


3-28-2018

# An Investigation of the Effects of Taking Remedial Math in College on Degree Attainment and College GPA Using Multiple Imputation and Propensity Score Matching

Meghan A. Clovis  
mclov001@fiu.edu

**DOI:** 10.25148/etd.FIDC006525

Follow this and additional works at: <https://digitalcommons.fiu.edu/etd>

 Part of the [Educational Assessment, Evaluation, and Research Commons](#), [Higher Education Commons](#), [Longitudinal Data Analysis and Time Series Commons](#), [Multivariate Analysis Commons](#), [Other Education Commons](#), and the [Secondary Education Commons](#)

---

## Recommended Citation

Clovis, Meghan A., "An Investigation of the Effects of Taking Remedial Math in College on Degree Attainment and College GPA Using Multiple Imputation and Propensity Score Matching" (2018). *FIU Electronic Theses and Dissertations*. 3573.  
<https://digitalcommons.fiu.edu/etd/3573>

This work is brought to you for free and open access by the University Graduate School at FIU Digital Commons. It has been accepted for inclusion in FIU Electronic Theses and Dissertations by an authorized administrator of FIU Digital Commons. For more information, please contact [dcc@fiu.edu](mailto:dcc@fiu.edu).

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

AN INVESTIGATION OF THE EFFECTS OF TAKING REMEDIAL MATH IN  
COLLEGE ON DEGREE ATTAINMENT AND COLLEGE GPA USING MULTIPLE  
IMPUTATION AND PROPENSITY SCORE MATCHING

A dissertation submitted in partial fulfillment of

the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

HIGHER EDUCATION

by

Meghan A. Clovis

2018

To: Dean Michael R. Heithaus  
College of Arts, Sciences and Education

This dissertation, written by Meghan A. Clovis, and entitled An Investigation of the Effects of Taking Remedial Math in College on Degree Attainment and College GPA Using Multiple Imputation and Propensity Score Matching, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this dissertation and recommend that it be approved.

---

George O'Brien

---

Norma Goonen

---

Benjamin Baez, Co-Major Professor

---

Mido Chang, Co-Major Professor

Date of Defense: March 28, 2018

The dissertation of Meghan A. Clovis is approved.

---

Dean Michael R. Heithaus  
College of Arts, Sciences and Education

---

Andrés G. Gil  
Vice President for Research and Economic Development  
and Dean of the University Graduate School

Florida International University, 2018

© Copyright 2018 by Meghan A. Clovis

All rights reserved.

## DEDICATION

To my mother, Kathy Clovis, and my late father, Rusty Clovis, for always believing in me. Your guidance, love, and support mean the world to me. To all of my students, past and present, who have overcome seemingly insurmountable obstacles to achieve your goals. Your commitment to your education and the sacrifices that you make motivate and inspire me every day. To my family and friends, who supported me during this exciting, exhausting, mind-numbing endeavor. Your words of encouragement and willingness to listen to me think out loud (or mumble nonsensically) were more helpful than my style manual or textbooks. To my dear friend, Jeffrey Miranda, for guiding me through this process and ensuring that I remained sane-ish until my defense was over.

I dedicate this dissertation to you.

## ACKNOWLEDGMENTS

I would like to express my appreciation for the countless hours of guidance, kindness, patience, and support of my Major Professor, Dr. Mido Chang. Dr. Chang gave of her time freely seven days a week and worked tirelessly to guide me through my research. I could not have done this without her.

I would also like to thank my Co-Major Professor, Dr. Benjamin Baez for his commitment to my education and success. Thank you to my committee members, Dr. Norma Goonen and Dr. George O'Brien, for their support, feedback, and professional insight during this process.

A special thank you to Maria Mendez and Kathleen Clovis for hours spent proofreading and editing this work.

ABSTRACT OF THE DISSERTATION

AN INVESTIGATION OF THE EFFECTS OF TAKING REMEDIAL MATH IN  
COLLEGE ON DEGREE ATTAINMENT AND COLLEGE GPA USING MULTIPLE  
IMPUTATION AND PROPENSITY SCORE MATCHING

by

Meghan A. Clovis

Florida International University, 2018

Miami, Florida

Professor Mido Chang, Co-Major Professor

Professor Benjamin Baez, Co-Major Professor

Enrollment in degree-granting postsecondary institutions in the U.S. is increasing, as are the numbers of students entering academically underprepared. Students in remedial mathematics represent the largest percentage of total enrollment in remedial courses, and national statistics indicate that less than half of these students pass all of the remedial math courses in which they enroll. In response to the low pass rates, numerous studies have been conducted into the use of alternative modes of instruction to increase passing rates. Despite myriad studies into course redesign, passing rates have seen no large-scale improvement. Lacking is a thorough investigation into preexisting differences between students who do and do not take remedial math.

My study examined the effect of taking remedial math courses in college on degree attainment and college GPA using a subsample of the Educational Longitudinal Study of 2002. This nonexperimental study examined preexisting differences between students who did and did not take remedial math. The study incorporated propensity

score matching, a statistical analysis not commonly used in educational research, to create comparison groups of matched students using multiple covariate measures. Missing value analyses and multiple imputation procedures were also incorporated as methods for identifying and handling missing data.

Analyses were conducted on both matched and unmatched groups, as well as on 12 multiply imputed data sets. Binary logistic regression analyses showed that preexisting differences between students on academic, nonacademic, and non-cognitive measures significantly predicted remedial math-taking in college. Binary logistic regression analyses also indicated that students who did not take remedial math courses in college were 1.5 times more likely to earn a degree than students who took remedial math. Linear regression analyses showed that taking remedial math had a significant negative effect on mean college GPA. Students who did not take remedial math had a higher mean GPA than students who did take remedial math. These results were consistent across unmatched groups, matched groups, and all 12 multiply imputed data sets.



## TABLE OF CONTENTS

CHAPTER	PAGE
I. INTRODUCTION .....	1
Statement of the Problem.....	4
Purpose of the Study .....	5
Sources of Data.....	6
Research Questions.....	6
Significance of the Study .....	7
II. LITERATURE REVIEW.....	9
Remedial/Developmental Education .....	9
Success in Developmental Courses .....	12
The Cost of Remedial Education .....	18
Increasing Student Success .....	20
Propensity Score Matching.....	30
Multiple Imputation .....	33
Propensity Score Matching and Multiple Imputation in Practice.....	37
Summary.....	47
III. METHODOLOGY .....	49
Research Questions.....	49
Sources of Data.....	50
Population .....	51
Variables and Sampling Method.....	60
General Research Methodology.....	63
Data Analysis .....	63
IV. RESULTS .....	65
Description of the Sample.....	65
Stage One: Analysis of the Initial Sample .....	68
Stage Two: Propensity Score Matching on the Original Sample .....	75
Stage Three: Missing Value Analysis and Multiple Imputation.....	82
Stage Four: Propensity Score Matching on All MI Sets.....	91
Summary of Results.....	100
V. DISCUSSION AND RECOMMENDATIONS.....	102
Overview of the Problem .....	102
Results.....	103
Limitations .....	107
Implications.....	109
Recommendations for Future Research .....	111
Summary .....	114

REFERENCES .....	116
APPENDICES .....	124
VITA.....	158

## LIST OF TABLES

TABLE	PAGE
1. Population Frequencies: Race, Sex, SES2 Quartiles .....	52
2. Population Frequencies: Race and Sex by Quartile Coding of SES2.....	53
3. Population Frequencies: Race by School Control .....	53
4. Population Frequencies: Race by School Urbanicity.....	54
5. Population Frequencies: Race by Family Composition.....	54
6. Population Frequencies: Educational Expectation by Concentration and Planning Program ( $N = 16,197$ ).....	56
7. Population Frequencies: Highest Math.....	57
8. Population Frequencies: Race by Highest Math Course.....	58
9. Population Frequencies: Highest Math Course by Quartile Coding of SES2 .....	58
10. Population: Known PSI Combination.....	59
11. PSI Combination by Remedial Mathematics Course(s) Taken .....	60
12. Sample: PSI Combination by Remedial Math and Degree Attainment.....	66
13. Descriptives: College GPA by PSI, Remedial, and Degree.....	67
14. Summary of Group Statistics for Continuous Covariates.....	70
15. <i>t</i> -test Summary: Remedial Math by Continuous Covariates.....	70
16. Summary of Group Statistics for Continuous Covariates Before and After Matching .....	76
17. Mean Difference Between Remedial Math Groups Before and After Matching ..	77
18. Logistic Regression Model 1 Summary Before and After Matching .....	79
19. Summary of Missing Value Analysis .....	83
20. Summary of Tabulated Patterns.....	84

21.	Estimated Means and Little's MCAR Test .....	85
22.	Pooled Covariate Estimates After Multiple Imputation.....	85
23.	Pooled Independent Samples <i>t</i> -Tests for 12 MI Sets .....	86
24.	Mean Difference Between Remedial Math Groups Before and After MI.....	86
25.	Summary of Group Statistics for Continuous Covariates Before and After Matching (First MI Set) .....	94
26.	Summary of Model 1 Before and After Matching for First MI Set.....	95

## LIST OF FIGURES

FIGURE	PAGE
1. Predicted Probability of Remedial Math Group Membership Before and After Matching in the Original Sample.....	75
2. Covariates Predicting Remedial Math Before and After Matching.....	78
3. Model 2: Predicted Probability of Degree Attainment Before and After Matching .....	80
4. Summary of Missing Values and Patterns .....	82
5. Predicted Probability of Remedial Math for Original Data and First MI Set .....	88
6. Predicted Probability of Degree Attainment: Original Sample and First MI Set ..	89
7. Predicted GPA by Remedial Math Group (Original and First MI Set) .....	91
8. Predicted Probability of Remedial Math Before and After Matching (Original and First MI Set) .....	93
9. Model 2: Predicted Probability of Degree Attainment Before and After Matching (Original and First MI) .....	97
10. Model 2: Predicted GPA by Remedial Math Group Before and After Matching (Original and First MI Set) .....	99

## CHAPTER I

### INTRODUCTION

As has always been the case, education is viewed, in large part, as the champion of social justice, the protector of freedom, and the path to the “American dream.” It is expected that education shall ensure equality, upward social mobility, and financial stability. Former President, Barak Obama, in his Address to the Joint Session of Congress in 2009, made this idea clear:

And so tonight, I ask every American to commit to at least one year or more of higher education or career training. This can be community college or a four-year school; vocational training or an apprenticeship. But whatever the training may be, every American will need to get more than a high school diploma. And dropping out of high school is no longer an option. It's not just quitting on yourself, it's quitting on your country - and this country needs and values the talents of every American.

From my perspective as an educator at one of the largest open-access public colleges in the U.S., however, I see the market-driven forces at work every day: students navigating the financial aid system, first generation students trying to learn the ropes, nontraditional students returning and trying to balance life, work, kids, and school, all in the hopes of making themselves more marketable in order to achieve the “American dream.” What American would want to let down his or her President and country by not getting more than a high school diploma? How can any American expect to make it in this world without a college education?

An unsurprising result of the importance placed on higher education is a rise in enrollment in degree-granting postsecondary institutions. Indeed, enrollment has risen by more than 30% since the year 2000. In 2015 there were an estimated 17 million undergraduates enrolled and this number is expected to rise to over 19 million by 2026. Despite this surge in enrollment, only 59% of full-time students attending public 4-year institutions complete a degree within six years of enrolling. Approximately 29% of full-time students attending public 2-year institutions complete a degree within three years (McFarland et al., 2017). Enrollment is increasing, but graduation rates are not keeping pace.

The number of students who are entering colleges and universities academically underprepared is also increasing. Underprepared students in college (those with deficiencies in basic reading, writing, and/or mathematics skills) is not a new concept. Higher education institutions have been accepting underprepared students for over 150 years and have continually developed services to meet the needs of diverse student populations with varying skill sets (Boylan, Bonham, & White, 1999; Casazza, 1999). The purpose of remedial (also known as developmental or college preparatory) coursework is not only to prepare students with the necessary skills to be successful in college-level courses, but also to “reduce disparities between disadvantaged and advantaged groups” that may exist beyond academic skill gaps (Bahr, 2007, p. 695).

Approximately 75% of all postsecondary institutions and 98% of 2-year institutions offer at least one level of remediation in reading, writing, and/or mathematics (NCES, 2010; Parsad & Lewis, 2003). Estimates of the numbers of students enrolling in these developmental courses range from 25% to as high as 75% of all incoming freshmen

(Aycaster, 2001; Bonham & Boylan, 2012; Donovan & Wheland, 2008; Parsad & Lewis, 2003). Additionally, more students require developmental math courses than reading and/or writing (Bahr, 2007; Bettinger & Long, 2005; Donovan & Wheland, 2008; Fike & Fike, 2012). In order to have any hope of attaining a degree, which is so highly sought-after, students must successfully negotiate these remedial courses.

Despite the increased need for, and prevalence of, remediation, many students do not successfully complete their coursework. In particular, less than 30% of students pass all remedial math courses in which they enroll (Attewell, Lavin, Domina, & Levey, 2006; Bahr, 2010). The lack of success of remedial math students has prompted numerous revisions to the teaching and learning process in these courses through massive redesign efforts. These redesigns have had mixed results (Bahr, 2007; Bettinger & Long, 2005; Bonham & Boylan, 2012; Illich, Hagan, & McCallister, 2004; Pretlow & Washington, 2011). Furthermore, as a result of the increasing demand, the cost of offering remediation in college is increasing. According to an August 2006 issue brief from the Alliance for Excellent Education, the estimated cost of remedial education in community colleges alone was \$3.7 billion annually. In essence, the need for remediation in college is rising—particularly mathematics remediation. The costs associated with offering remediation are also rising. Yet, the impact of remedial education programs on retention and success in college is questionable.

Years of redesign efforts have not improved success in remedial math on a large scale. So, what are we missing? The answer may not be in the effectiveness of the courses in addressing basic skill gaps alone, but in the characteristics of the students who are enrolled in these courses. Students taking remedial courses may be systematically



different from students who do not take remedial courses because preexisting differences in remedial students may impact their success no matter the structure of the course. Therefore, assessments of the effectiveness (or lack thereof) of remedial education that do not address differences in student populations may be inaccurate.

### **Statement of the Problem**

As a result of the increasing prevalence of remedial course offerings, the necessity, effectiveness, and costs of offering these courses continue to be topics of debate among policy makers and are areas of rising interest to researchers (Bahr, 2007; Bettinger & Long, 2005; Illich, Hagan, & McCallister, 2004; Pretlow & Washington, 2011). Critics of developmental education have raised many questions regarding the effectiveness and cost of remedial/developmental education programs. With the rising interest in research into these questions comes an increase in the need for institutional accountability and student data reporting. As Bettinger and Long (2005) point out, “Despite the growing numbers of underprepared students who enroll in remedial courses at community colleges each year, little is known about the causal effects of remediation on student outcomes” (p.17).

In particular, nonacademic and non-cognitive characteristics may influence student success (Aycaster, 2001). Alternative methods of instructional delivery have been, out of necessity, investigated in attempts to increase student success in remedial math (Bailey, 2009; National Governors Association, 2010). No single delivery method works for all students and because of the particularly high failure rate in remedial mathematics, special emphasis has been placed on the investigation of alternative modes of instruction. Course redesigns utilizing more than one teaching strategy have been

recommended and they include mastery learning, active learning, individualized instruction, supplemental instruction, and computer-assisted learning (Bonham & Boylan, 2012). Due to the diverse characteristics of adult learners, there may be particular characteristics that lend themselves to one instructional format over another. However, despite myriad studies into course redesigns that use behavioral and cognitive mathematics learning theories, there has been no significant improvement in completion for students in remedial math.

There is a gap in the literature in exploring characteristics of the students in remedial math courses. Thus, understanding which (if any) student characteristics may influence enrollment and success may be the key to understanding the impacts of remedial math coursework on overall retention and success. Additionally, educational research often does not lend itself to randomized experimental studies due to logistical and/or ethical reasons, and self-selection of students into programs may bias study results (Padgett et al., 2010). Statistical analysis in educational research is often hindered by techniques that assume the sample under investigation is the result of a randomized experiment. My study was undertaken to investigate the effect of taking remedial math courses in college on degree attainment and college GPA using propensity score matching and multiple imputation procedures in a nonexperimental study design.

### **Purpose of the Study**

This nonexperimental study examined (1) preexisting differences on multiple covariate measures between students who did and did not take remedial math in college; (2) the effect of those preexisting differences on participation in remedial math in college; and (3) the effect of taking remedial math on degree attainment and college

GPA. Propensity score matching was used to create comparison groups. Missing value analysis and multiple imputation procedures were utilized to handle missing data.

### **Sources of Data**

The data used in this study were obtained from the public use data file of Educational Longitudinal Study of 2002 (ELS:2002) sponsored by the National Center for Education Statistics. The ELS:2002 is a nationally representative longitudinal study of students' transition from high school into postsecondary education and the workforce. The National Center for Education Statistics (NCES) is a federal agency responsible for collecting and analyzing educational data. The NCES is part of the U.S. Department of Education and the Institute of Education Sciences and fulfills a Congressional mandate to collect, analyze, and report statistics on the condition of American education (NCES, About Us, n.d.). Background information of the ELS:2002 is provided in Appendix U.

### **Research Questions**

My study was guided by the following research questions:

1. Does taking remedial math in college have an effect on degree attainment?
2. Does taking remedial math in college have an effect on college GPA?
3. Are there preexisting differences between students who do and do not take remedial math courses in college?
4. Do preexisting differences predict participation in remedial math?
5. Does controlling for preexisting differences between remedial and non-remedial students produce different results in causal inference models?
6. Does propensity score matching improve the ability to control for preexisting differences in remedial and non-remedial students?

7. Do missing data have an impact on estimates of treatment effects?

### **Significance of the Study**

Given the lower success and completion rates of students in remedial math courses, investigations into the effect of remedial math-taking on degree attainment and GPA is warranted. Despite years of course redesign and the implementation of various instructional formats, student success rates have not been significantly altered on a large scale. If course redesign efforts do not effect a change in student retention and success, then it is plausible that nonacademic and/or non-cognitive characteristics may have an influence. Lacking in many studies is an investigation into preexisting differences in student populations that may influence outcomes regardless of whether or not students successfully complete remediation. If factors impacting student success are out of the control of the institution, then simply trying to remediate a skill deficiency will be inadequate in increasing retention and success.

There is a gap in the literature in examining factors outside of the control of the institution – preexisting differences – that may impact student success. There is also a gap with respect to using nonstandard statistical analyses in nonexperimental designs in educational research when it is not possible to randomly assign students to treatment and control groups (as is done in a randomized controlled trial). My study will add to the literature by conducting an investigation into factors that may impact enrollment and success in remedial math courses and the effect of taking remedial math on degree attainment and college GPA. The present study will also add to the literature by examining two methods of analysis not commonly used in nonexperimental educational research.

The study incorporated propensity score analysis, a statistical analysis not commonly used in education, that claims to mimic a randomized controlled trial. Students were measured on multiple academic and nonacademic variables. Propensity score matching was conducted to create comparison groups of matched students to attempt to obtain unbiased estimates of treatment effects of taking remedial math on degree attainment and college GPA. Missing data are a problem that often arises in longitudinal educational studies. Missing value analyses and multiple imputation procedures designed to handle missing data were also examined to address this problem.

## CHAPTER II

### LITERATURE REVIEW

In this chapter I will review relevant literature related to the variables under investigation in the study. Areas reviewed will include remedial/development mathematics education, course redesign efforts, factors that may impact student success, and statistical procedures that may improve analyses in nonexperimental studies.

#### **Remedial/Developmental Education**

Although researchers and policy-makers use the terms remedial and developmental interchangeably, a distinction does exist between remedial courses and developmental education programs. Remedial courses generally address specific skill gaps within the academic disciplines of reading, writing, and mathematics. Developmental education programs often encompass more than just remedial coursework and can include services such as orientations, advising, mentoring, individualized instruction, tutoring, instruction in study skills, and freshman experience courses (Boylan et al., 1999; Casazza, 1999; Saxon & Boylan, 2001). Distinguishing between solely remedial courses and developmental education programs is difficult, at best. Therefore, for the purposes of this review, I will use the terms interchangeably.

Remedial education faces much criticism. Many critics of remedial education contend that the primary responsibility of preparing students for college-level work rests with the high schools, though historically, postsecondary institutions have played a role, though often on an informal basis. The movement away from informal tutoring and college preparatory programs at the high school level and within higher education began

in the 1960s. Instead, more formal and structured programs offered at postsecondary institutions evolved to meet the increasing demand for remediation (Boylan et al., 1999).

On the academic front, there is no national standard for what defines “college-ready.” There is no national college readiness assessment. There is neither agreement on what constitutes acceptable remediation nor who should provide it. Many states have placed the responsibility of remediation with the community colleges and some with the high schools. Yet, there is no agreement on assessing college-readiness following the completion of remediation courses. Completion rates are also criticized; the more remedial courses students must take, the less likely they are to complete their degree or certificate. The only consensus is that the need for remedial education is high and remedial education is arguably unsuccessful.

**How many students take remedial courses?** The number of students requiring remedial education varies greatly by state, institution type, and area of remediation needed. These numbers also vary depending upon the source of the data. In 2009-10, it was estimated that 75.3% of all 4-year, and 99.6% of 2-year public institutions offered remedial education (NCES, 2010). On the basis of the sample data, the NCES reported that in 2007-08, 36% of all college freshman had taken at least one remedial course. If these data are subdivided by type of institution, the percentages of students who took a remedial course look slightly different: 42% at 2-year public institutions, 39% at 4-year, non-doctorate granting public institutions, and 24% at 4-year doctorate granting public institutions. The percentages of students taking remedial courses was largest (40%) for students aged 24 to 29, followed closely (38%) by students aged 30 or older (NCES, 2011, p. 70). According to the National Governors Association (2010), approximately

40% of students entering postsecondary education will require remediation and this number increases to 60% at the community college level.

The data is generally reported in aggregate, but this is misleading because it fails to show the differences that exist between various sectors of higher education. For example, the NCES' "Condition of Education 2004" reported that 28% of all entering freshman in the Fall of 2000 required some type of remediation, including both public and private institutions. Of these students, 22% required mathematics, 14% writing, and 11% reading remediation. However, if the data are separated by type of institution, the percentage (42%) is much higher for public two-year institutions than any other type. If the data are further divided by subject area, 35% of students at public two-year institutions require remediation in mathematics compared to only 16% at public four-year institutions.

Differences in type of remediation required also exist between states. The difficulty is that many states do not report these data individually and when they do, it is often not reported yearly, thus accurate and meaningful between-state comparisons are generally not possible. In Florida for example, the Office of Program Policy Analysis & Government Accountability (OPPAGA) reported that 55% of all students entering college in Florida in 2003-04 required remediation and of these students, 94% of them attended a community college for remediation. A total of 89% of these students needed remediation in math and 62% needed remediation in more than one area. Of the students who took remedial courses, only 52% of them finished their remedial course work (OPPAGA, 2007). Comparing the data from Florida to the NCES data illustrates the vast differences and inherent difficulties in making comparisons and generalizations when



data are reported in aggregate. There are no consistent data on the number of students who need remedial education, but one can confidently say that the number of students who are not college-ready is large.

### **Success in Developmental Courses**

Developmental education programs perform multiple functions; evaluating the effectiveness of such programs is a complex process. Among the many aspects of these programs, researchers often analyze metrics such as pass rates, retention rates, success in subsequent college-level courses, transfer rates, and degree completion. Success and retention rates are important metrics for evaluating any academic program in higher education, but the primary function of remedial coursework is to prepare students for college-level courses. Some research suggests that remedial students, as a whole, are less likely to complete their degree than students who place directly into college-level courses (Bailey, 2009; Brittenham et al., 2003). Because there is no standard for the reporting of student success data in remedial courses, determining the average pass rates for these courses is not possible. However, several analyses using large, longitudinal data sets that tracked remedial education students do exist, and these studies provide at least some information as to the success, retention, and completion rates of these students.

One such study was conducted by Attewell, Lavin, Domina, and Levey (2006), who analyzed a large subset of data from the National Educational Longitudinal Study (NELS), which tracked the progress of eighth grade students from 1988 until 2000. The analysis included students that had similar levels of academic preparedness, and compared the success of students who chose to enroll in remedial coursework to those who did not. Attewell et al. (2006) found that among the students who attended college,

40% enrolled in at least one remedial course, with mathematics remediation being the most common. The enrollment in remediation was higher for those students attending 2-year institutions, where 58% enrolled in at least one remedial course compared to 31% at non-selective 4-year institutions.

More startling than the high percentages of students enrolling in developmental courses were the percentages of students not passing their mathematics courses on the first attempt. Research suggests that students in remedial mathematics represent the largest percentage of total enrollment in remedial courses, and the study by Attewell et al. (2006) suggested that they have the lowest success rates. Attewell et al. concluded that most students passed all of their remedial courses in writing (68%) and reading (71%), but only 30% passed all of their remedial math courses. They also investigated how many students actually completed the remedial courses that they enrolled in and found that, within three years after beginning remedial coursework, 44% completed their remedial reading and 31% completed math. Of those students who placed into the lowest level for math or reading (generally three levels below college-ready), only 16% of remedial math students and 22% of remedial reading students finished their degree. Although these numbers appear dismal, the researchers found that, overall, most of these students took only one to two remedial courses, and most of them passed these courses within the first year. One possible problem with the results of this study may lie in the self-selection of courses by the students. It is possible that those students who chose not to enroll in remedial coursework were inherently more motivated and/or had higher levels of self-efficacy, which led to greater success. It is also possible that those not enrolling in remedial courses chose an alternate route such as vocational training.

Bettinger and Long (2005) attempted to control for preexisting differences in students in their study of a large longitudinal data set from colleges in Ohio. The researchers also compared students with equivalent placement test scores. The students attended different colleges with different placement test score cut-off points, but the students could not self-select. Thus, placement in remedial courses was a function of the institution attended. These students differed slightly in placement test scores, but were either just above or just below the cut-off score, depending upon the college attended. Among this population of students, the researchers found only slightly positive effects from taking remedial math courses, but no positive effects from remediation in English.

Possible limitations of Bettinger and Long's study are that the sample was limited to students (a) that were between the ages of 18 and 20; (b) who took the ACT; and (c) scored at the high end of the placement test cut-off point. Students between the ages of 18 and 20 are most likely direct entry students from high school and therefore they do not have a large time lapse between courses. Students who take the ACT most likely have an interest in going to college because this is a common college readiness assessment used by colleges and universities. Finally, students scoring at the high end of the placement scale could be more motivated to attempt college-level courses because they have higher self-efficacy.

Studies conducted in Florida by Calcagno and Long (as cited in Bailey, 2009), also focused on students scoring just above and just below the placement test cut-off score for remediation. Unlike Attewell et al.'s (2006) study, Calcagno and Long attempted to control for preexisting differences by comparing students who were required to remediate with those who were not required at institutions with the same cut-off

scores. Similar to Attewell et al.'s conclusions for two-year institutions, the researchers found that taking remedial courses in math and/or English had no effect on completing college-level courses or on college-completion rates. However, taking remedial courses in reading did have a negative effect on college-level course completion. Although Calcagno and Long's study attempted to control for the preexisting differences of the students, it narrowly focused on those students scoring just above and below the cut-off scores for placement in remediation. Most students do not score close to the cut-off point; therefore, Calcagno and Long's study cannot provide information about the majority of the students in remedial courses.

According to Bailey's (2009) assessment of the NELS data, less than 25% of students enrolled in remedial courses at the community college level complete a degree within eight years of beginning college, whereas 40% of those not in remedial courses graduated within eight years. Bailey also compared the success in college-level courses between remedial students and those placed directly into college-level courses and found that "among students with equally low assessment test scores, those who take developmental education do no better than those who enroll directly in college-level courses" (2009, p. 15).

Bailey's (2009) conclusion addresses a commonly cited problem with remedial education, which is the disconnect between high school proficiency required for graduation and what is really needed to be college-ready. Many colleges and universities have standards set by the state for what college-ready means, but these standards are not necessarily aligned with the learning outcomes of secondary institutions. This

misalignment may be a contributing factor to the relatively equal performance of students who take remedial courses and students who test college-ready.

In their study of remedial course-takers, Chen and Simone (2016) noted that 78% of students who took no high school math beyond algebra subsequently took remedial math courses. However, they also noted that even students with strong academic preparation took remedial courses (48% at two-year institutions). Moreover, students with moderate or strong academic preparation who did not complete remediation were less likely to earn college-level math credits and less likely to earn a degree within six years compared to non-remedial students.

As stated earlier, on average, pass rates for remedial math courses are far lower than the rates for reading and writing (Attewell et al., 2006; Bahr, 2007; Bonham & Boylan, 2012; Illich et al., 2004). However, those studies that have attempted to quantify the pass rates for students in developmental courses have not found highly significant correlations between pass rates in remedial courses and successful college-level course completion. Additionally, these studies did not attempt to account for the disparities potentially caused by preexisting differences in students. Despite the low pass rates in some subject areas, if successful remediation has little to no effect on success in college-level courses, might the preexisting differences play a larger role in success in both remedial and college-level coursework? Before research into the effectiveness of the preparatory aspect of developmental education can take place, low student success in remedial education necessitates examination. Given the low pass rates in many developmental courses, any effort in restructuring these programs must first include an investigation into factors that may be predictive of success.

Bahr (2007) identified two primary predictors of successful remediation as depth (degree of deficiency) and breadth (number of skill areas requiring remediation).

Students placing into the lowest levels of developmental math were far less likely to remediate successfully, and if these students also had deficiencies in English skills, this likelihood decreased (Bahr, 2010). Not only does the need for remediation in multiple areas have an additive negative effect on success, those students who need the most help are the least likely to successfully complete remediation (Bahr, 2007). The more levels of remediation and the more subject areas needed, the longer students spend attempting to complete a degree. Studies show that the longer students spend in remedial courses, the less likely they are to graduate (Attewell et al., 2006). Remediation is intended to function as a gateway to college-level courses and ultimately, to degree completion, but considering the high failure rates, remedial education is functioning as a gatekeeper.

Remediation need in multiple areas is a strong predictor of success. However, there are other contributing factors to success in remedial coursework, including math anxiety, deficiencies in basic study skills, low self-efficacy, and poor attitude (Higbee & Thomas, 1999; Illich et al., 2004; Hammerman & Goldberg, 2003; Spradlin & Ackerman, 2010). Enrollment status (full-time/part-time), age, and ethnicity are also significant predictors of success in developmental math (Bettinger & Long, 2005; Penny & White, 1998). Students enrolled in developmental courses may also have the perception that they are in some way unfit to be in college and there may be a stigma associated with enrollment in remedial coursework, which can also have a negative impact on student success. If enrollment in remedial courses carries a stigma, students may be discouraged from enrolling altogether (Bettinger & Long, 2005; Hall & Ponton, 2005).

## **The Cost of Remedial Education**

At this point, we cannot say definitively just how many students take remedial courses or how successful those courses are at preparing students for college-level work. Furthermore, the impact on retention and degree completion is unclear. Thus, to calculate the actual monetary costs of remediation is problematic. Criticisms of the cost of offering developmental education programs come from legislators, educators, parents, students, colleges, and universities. Some argue that the need for remedial education would not be as prevalent if students learned what they needed while in high school. Taxpayers have already paid for this education once and are paying for it again when students are required to remediate at postsecondary institutions (Boylan et al., 1999; Saxon & Boylan, 2001).

Having a debate over the cost of offering developmental education is difficult because estimating the cost is problematic for the same reasons that evaluating program effectiveness is—lack of accurate and consistent data reporting. There is no standard reporting practice for colleges and universities. Each state, and in some cases each college within a state, reports data such as enrollments, costs, and revenue differently, making comparisons between states, and generalizations nationally, almost impossible. Adding to this difficulty is the fact that many states do not report this data at all, which seems odd given the ongoing debate about the costs and effectiveness of remediation.

Before one can estimate the cost of developmental education, one must determine the factors that contribute to the costs. Institutional costs may include faculty salaries, facility usage, tutoring services, learning support services, orientations, supplemental instruction sessions, advising, and mentoring. Several studies have attempted to estimate

the cost of remedial education using self-reported data from various states and higher education institutions. The most frequently cited investigation into these institutional costs is the 1998 study of Breneman and Haarlow (as cited in Bailey, 2009; Saxon & Boylan, 2001). Breneman and Haarlow estimated the average annual cost of developmental education to be over \$1 billion. Pretlow and Washington (2011) conducted a reevaluation of the investigation by Breneman and Haarlow using data from 2004-2005 and estimated the average cost at that time to be \$1.13 billion annually. However, Pretlow and Washington point out that, although their figure of \$1.13 billion seems to be an enormous amount of money, remedial education represented less than 1% of the total higher education revenues that year. Although Pretlow and Washington's estimate represents an increase of \$13 million, as a percentage of total revenue, the cost actually decreased by about 50% over that estimated by Breneman and Haarlow. According to an August 2006 issue brief from the Alliance for Excellent Education, the estimated cost of remedial education in community colleges alone was \$3.7 billion annually. Of this, the government's cost was \$978 million and tuition cost was \$283 million. However, no comparison was made between the total expenditures or revenues.

Total expenditures for remedial education do not necessarily accurately reflect the costs. Many services available to developmental students are also available to the rest of the student population. For example, tutoring labs are listed as an expenditure for some developmental education programs, but the labs are multipurpose facilities that provide services to all students at the college, thus the lab expense cannot be attributed solely to developmental courses. Faculty salaries fall in the same category as a multipurpose



expense. Faculty may teach college-level courses in addition to remedial courses. Thus, their salaries are not solely attributable to the cost of offering remedial courses.

Furthermore, investigations into the costs of offering remedial education do not take into account the revenues from tuition paid for these courses. Saxon and Boylan (2001) state that “remedial courses seldom cost institutions more than they generate in revenues” and within community colleges, “remedial courses typically generate more revenue than is spent in their delivery” (p. 8). Overall, remedial education accounts for only a fraction of the costs of higher education—it essentially pays for itself—and yet its cost is one of the major criticisms cited. With budget cuts growing larger each year, it is not surprising that potentially ineffective educational programs are in the spotlight when their perceived cost is high (Kelderman, 2011 as cited in Vaughan, Lalonde, & Jenkins-Guarnieri, 2014). Because the perceived cost of offering developmental education is one of the driving forces behind legislative changes that attempt to decrease or eliminate it, colleges and universities have undertaken numerous redesign efforts in an attempt to improve the success of these programs.

### **Increasing Student Success**

In answer to many of the criticisms of developmental education, institutions of higher education have increased their efforts to reform these programs in order to increase student success, decrease time in remediation, and improve accountability (Bailey, 2009). Several key components in these redesign efforts include more accurate assessment and placement measures, alternative modes of instruction, and supplemental instruction.

**Assessment and placement.** Program effectiveness, predictors of success, and the costs associated with developmental education are at the forefront of debate and of research interests. It is evident that higher education institutions cannot control the preexisting differences or academic preparedness of incoming students. Institutions must figure out ways to accurately assess and address these deficiencies. As with other aspects of developmental education, there is no consistency in defining college-ready. What constitutes college-ready depends on the state and, in some instances, the institution. Moreover, being college-ready does not simply mean that the student is academically prepared. College-readiness includes a student's ability to function in a new environment, to meet new and higher levels of expectations, and to be largely responsible for their own learning. Consequently, accurate assessment and placement is critical to student success and retention in remedial courses.

Administering some type of college-readiness diagnostic or college placement test is standard practice at many higher education institutions, but the assessments themselves vary widely across states and institutions (Bailey, 2009; Casazza, 1999). Common academic assessment and placement measures may include internal placement tests, AP scores, ACT scores, SAT scores, Accuplacer, Compass, and high school GPA (Bailey, 2009; Donovan & Wheland, 2008; Fike & Fike, 2012). College placement tests are often one-size-fits-all academic assessments with fixed cut-off scores that place students into the same level of remediation without taking into consideration varying degrees of skill gaps. Placement tests are ultimately high-stakes exams because students who perform poorly may be required to take one or more remedial courses in a given subject (Bailey, 2009). Aside from the lack of a common standard for academic preparedness and a wide

range of scores that fail to account for differences between students, college placement tests only assess specific content area knowledge. These tests do not assess the vast array of non-cognitive factors that can have an effect on student success; therefore, these assessments cannot accurately determine a student's college-readiness (Bailey, 2009; Boylan et al., 1999; Fike & Fike, 2012; Illich et al., 2004). Students may score low on these common placement exams for various reasons including

- a time lapse between high school graduation and college enrollment,
- a failure to learn the content in high school,
- a lack of knowledge about placement testing,
- test anxiety, and/or
- a lack of English skills necessary for understanding the questions.

These represent some of the most commonly cited factors that have the potential to influence a student's performance on a placement test (Bailey, 2009; Attewell et al., 2006). Some states, including California, Florida, and North Carolina, have started to move away from a common placement exam. Multiple measures are used to advise students and, in some cases, students are not required to remediate, even if multiple measures indicate they should (O'Connor, 2013; Xia, 2017; Zinshteyn, 2016).

However, once a student enrolls in a developmental education course(s), the learning needs of that student play a much larger role in their success. These different groups of students may not all need the same type of remediation and they may not all need the same amount, yet placement tests do not distinguish between these groups of students.

**Student population.** The population of students who enroll in remedial coursework is varied. However, Blacks and Hispanics are disproportionately represented in these courses (Chen & Simone, 2016). Students who enroll in developmental education courses are often labeled as nontraditional students. According to the NCES (2002), more than 75% of entering college students are nontraditional in some sense. The term nontraditional frequently describes a student who delays enrollment after finishing high school, is over the age of 25, attends part time, works full time, has dependents other than a spouse, is a single parent, or does not have a standard high school diploma (NCES, 2002). Students in developmental courses also frequently come from low-income families, and those from the lowest quartile of socioeconomic status are more likely to enroll in remedial courses. Additionally, nontraditional students who take remedial courses often come from urban high schools and they delay enrollment in college (Attewell et al., 2006; Boylan et al., 1999; Donovan & Wheland, 2008).

Many of the characteristics that define a nontraditional student are the same factors that may contribute to their lack of success. Nontraditional students have a variety of responsibilities and demands on their time. One of the most common characteristics of nontraditional students is that of a part-time, working student for whom any extra time spent in remediation represents an even greater delay and greater cost to graduation. Trying to balance school with work and/or a family will require sacrifices because of time constraints. More than likely, this student will sacrifice school over work and family. Although the prospect of a better job or higher salary is a motivating factor for many of these students to attend college, it is not an immediate result. Thus, sacrificing homework in order to pay the rent seems logical.

There is a steady increase in the number of reform (or redesign) movements at the state and national level in the area of developmental education in response to the criticisms of, and changes in, developmental education (Bailey, 2009; Bonham & Boylan, 2012; Boylan et al., 1999). In order to improve developmental education, institutions need to implement more comprehensive diagnostics that more accurately identify students' academic skill gaps as well as factors that may influence their success and retention. These assessments should inform not only their remedial course placement, but should consider the best setting in which each student will most likely be successful; the setting may need to be determined using other student characteristics. In addition to properly placing students, institutions should create more tailored interventions and alternative modes of course delivery (Fike & Fike, 2012). If college placement assessments were more holistic, the potential exists for institutions to greatly reduce or even eliminate the time students spend in developmental education programs. Alternative modes of delivery that account for the diverse student populations have the potential to increase student success. Just as a placement test should not be one-size-fits-all, neither should instruction if it is to be successful.

**Alternative modes of instruction in remedial math courses.** Students enrolled in remedial mathematics programs are largely unsuccessful. In addition to considering predictors of success and accurate placement, course design and delivery format are other areas that have been evaluated. Alternative modes of instruction exist and continue to evolve in an attempt to increase student success and retention in developmental education courses. The research indicates that students who enroll in a specific mode of instruction because it fits their needs have a higher probability of completion compared to students

who select courses based on the mode of instruction. Students that select courses based on what they thought would be easier or more beneficial in terms of delivery format are less successful, which suggests that students do not know how they learn best.

Institutions need to assist students in selecting a mode of instruction that best accommodates their outside commitments as well as one that best suits their learning style (Kinney & Robertson, 2003; Spradlin & Ackerman, 2010; Zavarella and Ignash, 2009).

***Computer-assisted learning.*** Computer-assisted learning is a term that encompasses any instruction that incorporates the use of computer technology, in whole or in part, within a course. Some studies indicate that computer-assisted learning is more effective than a traditional lecture class, but found mixed results with respect to remedial math courses. Students have to receive instruction in how to use the technology effectively and to its fullest potential; otherwise, there is no significant positive effect on student learning (Jacobson, 2006; Spradlin & Ackerman, 2010). The use of online learning management systems (LMS), which incorporate ebooks, tutorials, homework, and quizzes, can and do supplement the standard lecture class.

Common models of computer-assisted learning in which the technology plays a more prominent and direct role in instructional delivery include, hybrid (or blended), modular, and mastery-based. One mode of instruction that is computer-assisted is a hybrid class in which students receive part of their instruction in the form of lecture, and the other part outside of class through an online LMS. Hybrid courses have had mixed results; they have a higher withdrawal rate than standard lecture-based courses, but the retention rate increases if individualized tutoring supplements the lecture. Hybrid courses

may have higher withdrawal rates because students encounter challenges that they did not anticipate, such as the need to learn how to use the LMS and the increased amount of time spent on independent study (Zavarella & Ignash, 2009).

***Mastery-based learning and modular courses.*** Mastery-based learning is a method of instruction that allows students to work at their own pace. Mastery learning also includes comprehensive diagnostics to determine where students have skill gaps. Students spend time working only on those topics that they have deficiencies in. Mastery learning allows students to accelerate through a course by reducing the time spent reviewing topics that they already know. Often, courses employ an LMS that administers the diagnostic, provides students with an inventory of their skill gaps, and provides topic-based instruction for those deficiencies. Students prove mastery through frequent assessments and tests (Chickering & Gamson, 1987).

Bassett and Frost (2010) conducted a study in which they combined three developmental math courses into one course separated into 12 modules. The program provided multiple exit points depending on a student's educational and career goals. The course was a self-paced, mastery-based course that included on-demand, individualized assistance, study guides, and technology-driven instruction. The faculty did not lecture; instead they served as facilitators in the learning process, providing individualized tutoring and targeted feedback. Faculty spent less time preparing lectures and grading assessments, which gave them more time to dedicate to individualized instruction. Bassett and Frost compared students in the redesigned course to those in a standard lecture class using baseline data from posttest scores in both courses. Overall, redesign students increased their average posttest scores in all courses by 15 points and the pass

rates and retention rates increased. The costs to students decreased in that they only had to pay for one course instead of three. Because the courses were self-paced and no lecture was required, the institution was able to increase the class size to serve more students, and reduced institutional costs by 20%.

*Compressed or accelerated courses.* Mastery-based courses such as those designed by Bassett and Frost (2010) are also considered compressed courses. However, not all compressed or accelerated courses are mastery-based. Sheldon and Durdella (2010) assert that there has been little research examining the relationship between the length (number of weeks) of courses and course success in developmental education. Using historical enrollment data from a large, suburban community college in California, the authors attempted to study the correlation between course length and course success in developmental education while controlling for social and academic background.

For the purposes of Sheldon and Durdella's study, a compressed course was any course less than eight weeks in length. The study included compressed courses in the subjects of developmental English, reading, and math. The researchers found that shortened course length resulted in significant increases in student success across all categories of age, gender, and ethnicity, and across all disciplines. Students enrolled in compressed courses were more likely to succeed across all departments, with the highest successful course completion rates in the eight-week format in English. Students, regardless of age, race, or gender, were also more likely to successfully complete the compressed courses than their counterparts in courses taught at regular length.



***Concurrent enrollment.*** Concurrent enrollment refers to the practice of enrolling students in developmental education courses and in unrelated college-level courses simultaneously. It is a common practice in many higher education institutions, yet the research in this area is limited (Illich et al., 2004). Proponents of concurrent enrollment posit that a student's skill gaps in an unrelated discipline should not influence their success in an unrelated college-level course. For example, a student's skill gaps in mathematics should not influence their performance in an introductory history course. Illich et al. (2004) conducted an ex-post facto study using three years of data provided by the college's Office of Institutional Effectiveness. The researchers compared the passing rates of students concurrently enrolled in remedial and college-level courses to those of college-ready students.

Illich et al. (2004) found that at least 50% of the remedial students did not successfully complete their remedial course, and these students had significantly lower pass rates in their college-level course compared to the rest of the students. However, students who did pass their remedial courses had similar pass rates in the college-level courses to those students not in remedial courses, regardless of the subject of the college-level course. These results indicated that, for students successful in remedial courses, time to graduation can be decreased by allowing them to enroll concurrently in college-level courses. Nonetheless, 50% of the students in this study failed to pass their remedial courses and performed poorly in their concurrent college-level course. Thus, the issue of accurate and holistic placement and assessment still exists even within redesigned courses that show promising results.

On the basis of the literature review, large numbers of students are taking remedial mathematics courses in college. A large percentage of remedial math students do not pass these courses. Numerous methods of increasing student success in remedial math have been attempted and seem promising. Yet, overall, student success rates have not changed despite these efforts. Why? Remedial education serves diverse communities of students, including mainstream students from high school, students who have earned GEDs, employees coming back for degrees to increase their potential earnings, etc. I believe that the inability to increase overall student success, whether they are in traditional remedial classes or any of the number of redesigned courses, is due to student characteristics. Courses are redesigned and then the outcomes are compared for specific groups of students. Researchers investigate success based on gender, race/ethnicity, age, etc., but the design of the courses do not take these characteristics into account beforehand. It would behoove us to first better understand the characteristics of the diverse population of students in remedial math courses and what may or may not predict success. Then, we can design courses to better meet the needs of our students; i.e., we keep doing things backwards.

In order to better understand learning outcomes of remedial students, we need to compare groups of students with similar characteristics, but standard statistical analysis methods are not designed for comparing nonequivalent groups, such as those that often exist in educational research. Propensity score matching has been proposed as a method for addressing this problem.

## **Propensity Score Matching**

Studies in education often necessitate the analysis of nonequivalent groups of students. Though a randomized controlled trial is the gold standard in research, non-random assignment of participants to treatment and control groups as well as selection bias are common problems in educational research (Padgett, Salisbury, An, & Pascarella, 2010; Shadish & Steiner, 2010; Titus, 2007). When students are not randomly assigned to treatment and control groups, comparison groups cannot be assumed to be equivalent on any covariate measures that may affect the outcome under investigation because students in each group may be systematically different. Because of this, researchers need to account for this as a limitation when making causal inferences (Fan & Nowell, 2011). Longitudinal studies in particular may suffer from confounding because groups may have been different from the start and may “be subject to separate maturation processes or history effects” (Holmes, 2014, p.19).

One method that has been suggested to address the problem of comparing nonequivalent groups is propensity score matching (PSM). Supporters of PSM contend that it can mimic a randomized controlled trial by removing imbalance in covariate measures between comparison groups (Austin, 2011; Hill, 2004; Mitra & Reiter, 2016; Rosenbaum & Rubin, 1983). Propensity score matching originated in biomedical research and although its use in educational research is growing, the studies employing this method in education are limited (Fan & Nowell, 2011; Luellen, Shadish, & Clark, 2005; Padgett et al., 2010; Vaughan et al., 2014). According to Padgett et al. (2010),

In theory, propensity score methods account for observed factors that led students to participate in a program or not and, after accounting for these observed factors,

produce an estimated treatment effect that is closer to an estimated effect one would have produced had students been randomly assigned. (p. 30)

Propensity score matching is a three-step process. Step one involves carefully selecting covariates that are believed to cause differences in comparison groups; then, a “balancing score” is used to group participants so that members of each group have a similar balancing score. The balancing score, called the propensity score, is a function of the participants’ measures on multiple covariates and represents the conditional probability that a subject will be in the treatment or control group (Holmes, 2014). Propensity scores represent probabilities; thus they range from 0 to 1. In a randomized trial, for example, the propensity score is known to be .5 because participants are randomly chosen and have a 50% chance of assignment (Henderson & Chatfield, 2011; Luellen et al., 2005, Rosenbaum & Rubin, 1983; Shadish & Steiner, 2010). In a non-randomized study, the propensity scores are estimated using the covariate measures in the sample and are considered similar if the difference is within a predefined range (Shadish & Steiner, 2010). It is important to note that, although the treatment groups are known in a non-randomized experiment, each group has a non-zero propensity score because the scores are calculated based on the covariate measure; i.e., a control group member would still have *chance* of being in the treatment group based on their covariate measures (Guo & Fraser, 2010).

Step two of PSM consists of matching participants in each group based on how similar their propensity scores are. Step three involves the post-matching analysis. After forming subgroups of matched participants based on their propensity scores, comparisons between treated and untreated subgroups are conducted on outcome variables (Austin,

2011; Guo & Fraser, 2010; Padgett et al., 2010; Rosenbaum & Rubin, 1983; Vaughan et al., 2014).

There are four propensity score analysis methods: propensity score matching, stratification on the propensity score, inverse probability of treatment weighting, and covariate adjustment using the propensity score (Austin, 2011; Rosenbaum & Rubin, 1983). Propensity score matching is one of the most common methods and involves forming subgroups from the larger sample consisting of matched sets of treated and untreated participants whose propensity scores are similar. The most common method of matching is a one-to-one method in which each participant in one group is matched to a participant in the other group based on how close their propensity scores are, where closeness is predefined (Austin, 2011; Byun, Irvin, & Bell, 2014; Melguizo, Kienzl, & Alfonso, 2011).

There are four methods used for estimating the propensity score model that will classify the matched groups: logistic regression, discriminant function analysis, regression with a dummy dependent variable, and probit regression. Logistic regression is the most common method for estimating the propensity score when the dependent variable is dichotomous and the method of analysis of treatment effects is logistic regression (Holmes, 2014). The process is iterative, comparing participants' propensity scores and determining if matched sets still have systematic differences. After creating the model for subgroup assignment, the model is examined by comparing the difference in the covariate scores between the matched groups (Austin, 2011).

Although PSM is promising, it has limitations that must be considered. Calculation of the propensity score is based on observed variables. Thus, unobserved

variables could confound these results. PSM does not guarantee that every participant will be matched to another one. If there are two control participants with the same propensity score, for example, only one of them can be matched to someone in the treatment group in a one-to-one matching situation. Because the estimation of propensity scores relies on complete case analysis for all covariates used in the estimation process, PSM results in listwise deletion of cases with missing data. The problem is that deleted cases may be important to the study and deleting them listwise can result in the loss of crucial information and bias results. As such, PSM is often combined with missing data analysis and multiple imputation procedures (Hill, 2004; Mitra & Reiter, 2016).

### **Multiple Imputation**

Missing data in observational studies are not uncommon. Longitudinal studies and studies utilizing surveys in particular suffer from missing data (Carlin, Li, Greenwood, & Coffey, 2003). Common reasons for missing data are that respondents refused to answer a question or the question was not applicable to them. However, when data are missing, it is important to consider the cause of the missing data prior to analysis – are the data missing randomly, or is there a systematic reason for it? Unfortunately, the question is not always an easy one to answer. A number of techniques have been developed to investigate and handle missing values.

**Types of missing data.** Missing data are often described as missing completely at random (MCAR), missing at random (MAR), or missing not at random (MNAR). The ideal situation for a researcher is that data are MCAR. The MCAR data are missing data for variables that have no correlation with any of the other variables observed in a study and, on average, their values would be similar to those for cases without missing values.

As such, it is not likely that they would introduce bias during analysis. However, determining missing data to be MCAR is often impossible (Holmes, 2014; Manly & Wells, 2015; Meyers, Gamst, & Guarino, 2013).

MAR data tend to be missing for variables that are correlated with other observed variables and it is possible that bias may result during analysis, i.e. missing values are systematically related to (or dependent on) values for other observed variables. Although called random, they are actually not random at all. MNAR data are correlated with other variables and are the most likely to introduce bias during analysis. These missing values are a function of the variable itself and the missing values depend on unobserved variables. Meyers et al. (2013) use the example of low-income respondents to illustrate a case of data that are MNAR. When asked a question about their level of income, low-income respondents may refuse to answer, thus their missing value is a function of the question and not random.

Different types of missing data require different types of analyses. The researcher must investigate the patterns in, and impact of, the missing data in the analysis. If data are MCAR, then it may be possible to ignore these missing values. However, if data are MAR or MNAR, we cannot necessarily ignore them without introducing bias in our analysis. One method used to analyze missing data is Little's MCAR test. This test simultaneously compares differences between groups of missing and non-missing cases for every variable included in the test. If the test is significant at the .05 level, then data may be MAR or MNAR. The test does not definitively determine the type of missing data, nor does it determine which variable(s) contributed to the significant result, but it does demonstrate the possible existence of MAR and/or MNAR data (Myers et al., 2013).

**Handling missing data.** Several methods can be employed to handle missing data. The most basic method is listwise deletion – deleting all cases with missing values on any of the variables in the analysis. Although common, this method can result in a large reduction of the sample size because data with just one missing value across a multitude of variables will be removed from the investigation. In addition, if data are not MCAR, deleting cases with missing data can produce biased results in subsequent analyses (Carlin et al., 2007; Harel & Zhou, 2007).

Single imputation is an alternative to listwise deletion and involves assigning a value where one is missing, using some estimate. For example, the mean of all non-missing values may be used to replace any value that is missing. Single imputation is inherently biased and is not recommended (Carlin et al. 2003; Meyers et al., 2013). The most commonly employed and often recommended method to handle missing data is multiple imputation (Carlin, et al. 2003; Harel & Zhou, 2007; Miles, 2016; Reist & Larsen, 2012). Multiple imputation (MI) is similar to single imputation in that missing values are replaced with estimated values. The difference is that MI uses multiple variables to estimate missing data, creates multiple models for each estimate, and then combines the results of these models (Holmes, 2014). After MI, analyses are conducted on the pooled imputed data sets. Theoretically, the more models used to create the pooled estimates, the more valid the statistical analysis will be (Harel, 2007; Yuan, 2010). Consensus does not exist on the number of imputations one should use, but more recent studies suggest that the number of imputations should be greater than or equal to the percent of missing values (White, Royston, & Wood, 2011). Generating a larger



number of MI data sets may be made easier with the use of statistical software packages (Carlin et al., 2003; Miles, 2016).

Multiple imputation methods assume that the data are not MNAR; however, this assumption is not testable. Including more variables during the imputation modeling (that are correlated with the other variables) can make the case for the data being MAR more plausible (Harel & Zhou, 2007). Multiple imputation is appealing because it results in estimates of missing values and maintains the original sample size, but it is imperative that researchers investigate differences in the outcomes of statistical analyses both with and without the imputed values given the untestable assumption about the missing data mechanism (Meyers et al., 2013). Additionally, statistical packages that function to create MI data sets do not provide much flexibility in working with the combined results from MI. In many cases this is because there are not statistically valid methods for combining common statistics, such as those obtained in regression analyses (Miles, 2016; Mood, 2010). Although many studies cite “Rubin’s rule” for combining estimates, statisticians disagree on whether these estimates can be combined or, more importantly, interpreted so easily in the context of logistic regression (Mood, 2010). Indeed, if it were that simple, then all statistical packages that support MI would generate these pooled estimates, but they do not. Therefore, caution is warranted in interpreting results of analyses on MI data sets.

Manly and Wells (2015) have provided a guide for reporting the use of MI in higher education research. They cite the lack of accurate reporting of MI procedures specifically in higher education. The authors state that MI is considered to be the most appropriate way to handle missing data by many statisticians and it is generally

preferable to listwise deletion. However, there has been minimal implementation in education research. According to the authors, “The general point of MI is to produce valid statistical inferences, not particular values” (p. 399). In making valid inferences, it is just as important to adequately report the procedures that led to those inferences. The authors suggest that researchers using MI report the rates of missing data, reasons for missing values (if known), analysis uses to detect MNAR data, descriptions of the imputation model and variables used, the software used and the number of imputations created, statistics pooling procedures used and issues encountered, and notable imputation results (comparing observed and imputed values). Also included should be a discussion of any discrepancies seen in the analysis results.

### **Propensity Score Matching and Multiple Imputation in Practice**

Fan and Nowell (2011) provided a basic illustrative example of using PSM in educational research and echo others’ assertion that, “In educational research, experimental design provides the strongest foundation for making causal inferences. In reality, however, experimental design is often practically impossible for various reasons” (Fan & Nowell, 2011, p. 74). PSM may be used to make comparison groups more equivalent (statistically speaking), thus improving the foundation for making causal inferences. Fan and Nowell provide an explanation of the justification for PSM using the example of predicting a 12<sup>th</sup> grade math achievement score using school type (public vs. private) as a predictor. Socioeconomic status (SES) and parents’ expectations were posited to have an influence on school type. If this were true, then making causal inferences about 12<sup>th</sup> grade math scores by school type would have biased results.

Fan and Nowell (2011) conducted logistic regression with two covariates (SES and parent's expectations) predicting group membership (whether a student would go to a public or private school). The logistic regression model significantly predicted group membership, meaning that the groups were non-equivalent and the background characteristics of these students (SES and parent's expectations) were predictors of which school they went to. Therefore, matching these students based on covariates that predict group membership in order to make comparison groups more similar could improve causal inference on the outcome (12<sup>th</sup> grade math score). The researchers used stratification on the propensity score to match participants. However, they cautioned that generating propensity scores and matching participants is not sufficient for analysis.

To make the analysis more accurate, the matched groups first must be evaluated to ensure that they are equivalent on the covariates. Fan and Nowell (2011) chose a two-way ANOVA with each covariate (SES and parents' expectations) as a dependent variable and school type (public or private) and matched group membership as two categorical independent variables. If PSM was successful at removing imbalance, the ANOVA should indicate a nonsignificant outcome and the effect size should be essentially zero for SES and parents' expectations. The results of the ANOVA in this example met this criterion. After matching, the researchers tested for mean differences between groups (public vs. private) on the outcome variable of interest in their example (12<sup>th</sup> grade math score). The results indicated that students in private schools outperformed those in public schools on math achievement scores after matching – SES and parents' expectations were not factors in the outcome. Fan and Nowell (2011) identified limitations in PSM. Sample size was identified as a limiting factor because as

subgroups become smaller during matching, estimates become more unstable.

Additionally, PSM is largely dependent on the identification of appropriate covariates. In their example, it might have been possible that other covariates (aside from SES and parents' expectation) could have also been influencing school type. It is not possible to be certain that all relevant covariates have been included and unobserved variables can still bias results.

Melguizo, Kienzl, and Alfonso (2011) compared the success of community college transfer students to that of rising juniors already attending a 4-year postsecondary institution using postsecondary transcript data from the National Education Longitudinal Study of 1988. PSM was used to correct for selection bias and the authors asserted that they compared people who were very similar in observable characteristics, but followed different paths to a bachelor degree. The researchers limited their sample data to only those students who graduated from high school early or on time, which resulted in a sample size of 3,160. The researchers included covariates measuring background characteristics that took into account demographic variables, pre-college achievement, educational expectations, and financial aid/tuition costs.

Preliminary analyses found substantial preexisting differences between groups. However, linear regression (OLS) before matching found no significant differences between the two groups on college success. After matching students on the covariates, the balance of the matched groups was examined using standard *t*-tests. Post-matching analysis found similar results – no significant difference. Melguizo et al. (2011) asserted that OLS was sufficient enough to control for preexisting differences without matching, but they noted both analyses were subject to bias due to unobservable variables. The

researchers identified sample size as a limiting factor in their study, as a large number of cases were lost during the matching process. The authors cautioned that their results should not be interpreted as causal due to the limitations of PSM. Although Melguizo et al. included a number of covariate measures, because of missing values, their sample size was limited. They made no attempt to investigate or handle the missing values, which may have biased their results both before and after matching.

Giani, Alexander, and Reyes (2014) examined the impact of dual-credit coursework on postsecondary outcomes in Texas using PSM. A large sample of longitudinal data was obtained from a research center at the University of Texas at Austin. The researchers chose to select participants who had a number of similar characteristics, thus limiting their sample to similar student from the beginning, though it seems that only eight covariates were included in the subsequent analysis. After obtaining a matched sample of 31,432 students, they found that not all significant differences in covariates between groups were removed, but noted that the mean differences were greatly reduced. Pre- and post-matching analyses found no significant differences in outcomes. Given that the researchers included a small number of covariates, it is possible that their results were biased due to unobserved variables in this large sample.

Titus (2007) conducted an investigation of self-selection bias using PSM on a sample of 3,948 students from a large longitudinal study. Titus included a large number of covariates, including measures of background characteristics, income, and educational background. PSM was performed and there was a loss of 4% of the cases because of missing values on the covariates, and a loss of 13% of the cases who were not matched

during PSM. According to the author, PSM outperformed OLS as OLS did not detect substantial self-selection bias. The author noted, as have others, limitations associated with unobservable variables and loss of data during matching. Vaughan et al. (2014) also conducted an investigation of matching effectiveness by comparing PSM to a hierarchical propensity score method to match participants from a large scale data set. The authors examined the effect of a freshman year experience course on GPA and found that, after matching using eight covariates for PSM, the matched groups were still unbalanced. The alternative propensity score was used and seemed to have better results in terms of balancing matched groups. The authors made no causal inferences and only compared the result of the two matching models. Sample size was probably a limiting factor in this study as the sample size was only 684 students.

In their study of the causal effects of college attendance on participation, using PSM, Henderson and Chatfield (2011) found that even when observed covariates were reasonably well-balanced in matched groups, selection bias associated with unobservable variables was still a serious problem. They warned that one cannot interpret results from matching as unbiased causal inference. However, the authors did identify one positive outcome of failed matching analysis – PSM leads to a closer investigation of research designs methods.

Clark and Cundiff (2011) used PSM to assess the effect of a freshman seminar on retention rates and college GPA. Included were a number of covariates measuring demographic characteristics, personality traits, academic motivation, loneliness, depression, and educational commitment. Pre-matching analysis saw no significant difference in retention rates between groups, but there was a negative effect on GPA.

After matching, there was a significant increase in retention rate for the treatment group, but no significant difference in GPA. I found several items of concern when reading the Clark and Cundiff (2011) study. The sample was obtained by selecting students in certain psychology classes who either had taken or were currently taking the freshman experience course. Thus, the sample was already limited. The total initial sample size before matching was only 435 students. I also found myself confused in trying to determine just how many covariate measures there were in this study as I saw first 19, then 27, then 20, and then 25 covariates said to be included. Matching was reported to have resulted in a 51% loss of cases because of non-matching and resulted in a group of only 54 matched students. The authors did mention that they had a difficult time obtaining balance between matched groups on the covariates. This difficulty was likely caused by the small sample size. The authors mentioned a problem of missing values on the covariates. They mentioned imputing some missing values, but they did not report how many values were imputed or how. I believe that any results obtained in this study would be questionable.

Byun, Irvin, and Bell (2015) conducted an extensive study incorporating both PSM and multiple imputation (MI) using data from the 2002 longitudinal study conducted by the NCES. Byun et al. investigated the effects of taking advanced math courses in high school on math achievement and college enrollment. They also investigated how these effects varied by SES and race/ethnicity, but the authors cited limitations in prior research on math course taking, contending that it was unclear whether advanced math course taking had a causal effect on outcomes because there were systematic differences between students who do and do not take advanced math. As has

been discussed previously, preexisting differences result in selection bias and confound causal inferences on educational outcomes. The authors also asserted that most research on advanced math-taking focused on the average effect of treatment, assuming that all students should receive the same benefit from taking these courses, regardless of their backgrounds. They sought to address these two issues by using PSM to match students based on preexisting differences. As PSM relies on adequate selection of covariates, based on the literature, they included variables both at the student and school level. Student variables took into account measures of academic and non-academic characteristics such as previous math performance, motivation, educational expectations, and family support. Also included were school-level variables such as school control and urbanicity. After selecting a sample of 12,250 participants, but prior to conducting PSM, they performed multiple imputation for missing values on the covariates, creating five imputed data sets. The percent imputed values ranged from 0.2% to 30%. A preliminary analysis found that a number of the included covariates significantly predicted advanced math course taking.

After MI, PSM was used to generate five matched groups. Regression analyses were conducted before and after matching and then coefficients and standard errors were averaged over all MI data sets using Rubin's (1987) combining rules. Byun et al. (2015) found that regardless of analysis (before or after matching) advanced math-taking had a significant positive effect on math achievement and college enrollment. The authors found that this effect was largely due to preexisting differences between students rather than on advanced math-taking alone, and the effect of advanced math-taking was greatly reduced after matching.



Padgett et al. (2010) conducted an investigation into the outcomes of PSM compared to standard statistical analyses using a sample from a large longitudinal data set. The authors point out that, “a host of logistical and ethical issues often preclude educational researchers from randomly assigning college students to either participate in a given activity (the ‘treatment’ group) or abstain from that activity (the ‘comparison’ group)” (p. 30). Because of self-selection there may be unaccounted factors influencing that decision. When looking for treatment effects, this is a problem. In theory, propensity score methods account for observable variables that lead to the participation decision and, after accounting for those factors, can produce estimated treatment effects closer to what would be seen if random assignment were possible. The authors emphasized, as have many others, that the success of PSM largely depends upon the covariates chosen for inclusion. In comparing OLS and PSM analysis, the authors found that OLS and PSM produced similar results in treatment effect. However, PSM appeared to actually inflate the treatment effect. Padgett et al. (2010) suggested that using PSM is not always necessary when there is a large sample size and extensive covariates for inclusion in analyses. Similar to Henderson and Chatfield (2011), the authors cited one positive result of issues in conducting PSM – they lead to more investigations into research designs.

Several researchers have conducted extensive investigations of the claims that PSM can mimic a randomized experiment. Peikes, Moreno, and Orzol (2008) conducted a case study using data from a known randomized experiment and conducted PSM as if the data were from a nonexperimental study. They analyzed the data under what are considered to be ideal conditions for conducting PSM. The results of the investigation

showed the PSM produced incorrect impact estimates when compared with the randomized design. Hundreds of covariates known to be powerful predictors of the outcome under investigation were included in the analysis. PSM resulted in overestimates of treatment effects; i.e., PSM showed positive treatment effects when random assignment estimates showed negative effects. However, Peikes et al. (2008) cited major limitations in PSM citing results from their investigation:

- It requires knowledge of the treatment effects in an experimental design to assess whether PSM produces unbiased estimates;
- It requires correct covariate selection;
- It requires a very large sample size;
- It is labor-intensive and time-consuming;
- There is no way to determine in advance if PSM will work.

The authors asserted that it is possible that there could be a situation where no amount of model specification will produce good results using PSM. Hill, Weiss, and Zhai (2011) conducted a similar investigation using longitudinal data and also cited a number of limitations to PSM and cautioned that effective analysis requires that a lot of choices be made by the researcher such as:

- How to fit the model to the data;
- How to select the type of matching method that will be used;
- How to check for balance between matched groups on the covariates;
- How to decide if groups are balanced enough;
- How to analyze the results of the propensity score analysis.

Hill et al. (2011) cautioned that “these choices can have non-negligible impacts on the resulting estimates” (p. 504). In fact, Rosenbaum and Rubin (1985), the major contributors to the theory of PSM stated, “matching, when successful, is a persuasive method of adjusting for imbalances in observed covariates” (p. 33). In other words, if PSM works, it works well, but there is no way of knowing whether it will work, whether it actually accomplished the goal of creating statistically equivalent comparison groups, or if it produced unbiased estimates of treatment effects.

A recent working paper explicitly warns that propensity scores should not be used and asserts that the claims of Rosenbaum and Rubin are invalid. King and Nielsen (2016) assert that PSM often does the opposite of what it is supposed to do, i.e., it increases imbalance between groups and produces bias in estimates of treatment effects. The authors agreed with others in that PSM is becoming increasingly popular for preprocessing data to improve causal inferences in observational data. To investigate the results of PSM, the authors simulated 1,000 data sets from matched pair and completely randomized experiments, adding observations that caused the data sets to be imbalanced. PSM was applied and was able to sort out cases that were inserted in the simulation that had no similar treated cases. However, PSM could not separate the matched pair cases from the randomized cases; i.e., it produced results as if all data were from a completely randomized experiment, when in fact they were not. In addition, PSM resulted in more imbalance between groups than existed in the original data. The authors also reanalyzed multiple published studies that used PSM and conducted imbalance investigation using the original data sets. It was found that PSM results in one-to-one matching scenarios

produced worse imbalance than existed in the original data. The authors recommend that researchers abandon the use of PSM in favor of alternative methods.

### **Summary**

In Chapter II, I examined relevant literature with respect to the variables of interest in this study: remedial education, propensity score matching, and multiple imputation. Based on my review of the literature concerning remedial education, the fact that there is a sizeable number of students in remedial math courses and that a large percentage of them fail to complete seems undisputed. Additionally, the effectiveness of these programs cannot be measured solely by pass rates. Accurate measures of success must also include student success in college-level coursework in general, with special attention on level of achievement (GPA) and degree attainment. Furthermore, nonacademic and non-cognitive measures should be investigated.

Despite the many arguments against offering remedial education solely based on the costs associated with it, the costs are minimal compared to revenues from the offerings. Remedial education does not appear to be something that will be eliminated in higher education any time soon. However, promising practices and improvements in assessment, placement, and instruction may help institutions to better serve this ever-expanding population of students more successfully and more quickly. In order to achieve these improvements, a more targeted investigation into student characteristics and how they impact success is needed.

Research in higher education is often problematic due to non-randomized experiments, confounding variables, missing data, and insufficient methods of analysis. Propensity score matching has been suggested as a possible method for controlling for

preexisting differences between groups of students to create equivalent comparison groups. The research cited limitations and cautionary statements regarding the use of PSM. PSM is limited by sample size and missing data. Multiple imputation is recommended as a standard procedure for handling missing data prior to conducting PSM. In addition, sensitivity analyses and comparisons of models before and after matching is highly recommended due to PSM's sensitivity to covariate selection as well as unobservable differences. Finally, PSM is believed, by some, to cause more imbalance and/or produce biased estimates of treatment effects than standard statistical analyses; thus careful analysis before and after PSM is warranted. Therefore, PSM should be investigated as a possible method of controlling for selection bias, but post-matching results should be analyzed for remaining unbalance between groups. Additionally, estimates of treatment effects should be compared before and after matching to investigate any significant differences between results.

## CHAPTER III

### METHODOLOGY

Chapter III presents the purpose of the study and a refinement of the research questions. A description of the research design follows, along with a description of the sources of data, population, sampling method, and variables considered. Finally, summaries of the general research methodology and data analysis techniques are presented.

#### **Purpose of the Study**

The purpose of this nonexperimental study was to investigate: (1) preexisting differences between students who did and did not take remedial math in college; (2) the effect of those preexisting differences on remedial math-taking; (3) the effect of remedial math-taking on degree attainment and college GPA; and (4) alternative statistical analysis methods.

#### **Research Questions**

The study attempted to answer the following questions:

1. Does taking remedial math in college have an effect on degree attainment?
2. Does taking remedial math in college have an effect on college GPA?
3. Are there preexisting differences between students who do and do not take remedial math courses in college?
4. Do preexisting differences predict participation in remedial math?
5. Does controlling for preexisting differences between remedial and non-remedial students produce different results in causal inference models?

6. Does propensity score matching improve the ability to control for preexisting differences in remedial and non-remedial students?
7. Do missing data have an impact on estimates of treatment effects?

### **Sources of Data**

The data used in this study were obtained from the public use data file of Educational Longitudinal Study of 2002 (ELS:2002) sponsored by the National Center for Education Statistics. The National Center for Education Statistics (NCES) is a federal agency responsible for collecting and analyzing educational data. NCES is part of the U.S. Department of Education and the Institute of Education Sciences. “NCES fulfills a Congressional mandate to collect, collate, analyze, and report complete statistics on the condition of American education; conduct and publish reports; and review and report on education activities internationally” (NCES, About Us, n.d.).

*ELS:2002.* The Educational Longitudinal Study of 2002 concentrated on students' transition from high school into postsecondary education and the workforce. In particular, patterns of college access and persistence were investigated. The study was a nationally representative, longitudinal study of 10<sup>th</sup> grade students in 2002 and data collection involved six major waves:

- Base-year interview (2002)
- First follow-up interview (2004)
- High school transcript data collection (2005)
- Second follow-up interview (2006)
- Third follow-up interview (2012)
- Postsecondary transcript data collection (2013)

(Christopher, 2015). The data-collection method of the ELS:2002 involved a multistage probability sample design consisting of approximately 15,400 sophomores in the first wave and included postsecondary transcript data for 7,637 participants in 2013 (Bozick, Lauff, & Wirt, 2007).

### **Population**

The initial public use data file that I obtained through NCES contained 16,197 unweighted cases. Among sophomores in that sample, 56% were White, 13% were Black, and 15% were Hispanic. Asians and Hawaiian/Pacific Islanders accounted for 10%, and American Indians, Alaskan Natives, and students identifying as multiracial, comprised the remaining 6%. The sample was divided in half among male and female students. Students in the lowest and second socioeconomic status (SES) quartile groups each accounted for 24% of the sample. The third quartile groups contained 25% and the highest quartile was comprised of 28% of the students. Frequencies are presented in Table 1.



Table 1  
*Population Frequencies: Race, Sex, SES Quartiles*

	Frequency	Percent	Valid Percent
<b>Race</b>			
Amer. Indian/Alaska Native	140	.9	
Asian, Hawaii/Pac. Islander	1654	10.2	
Black or African American	2168	13.4	
Hispanic	2433	15.0	
More than one race	768	4.7	
White	9034	55.8	
Total	16197	100.0	
<b>Sex</b>			
Male	8090	49.9	
Female	8107	50.1	
Total	16197	100.0	
<b>Quartile coding of SES2</b>			
Lowest quartile	3600	22.2	23.6
Second quartile	3590	22.2	23.6
Third quartile	3753	23.2	24.6
Highest quartile	4301	26.6	28.2
Total	15244	94.1	100.0
Legitimate skip/NA	305	1.9	
Nonrespondent	648	4.0	
Total	16197	100.0	

Within their racial groups, White students comprised 15%, Blacks 33%, and Hispanics 50% of the lowest SES quartile; 23% of all males and 25% of females were in the lowest SES quartile group. Frequencies for each subgroup are presented in Table 2.

Table 2

*Population Frequencies: Race and Sex by Quartile Coding of SES2*

		Quartile coding of SES2 variable				Total
		Lowest quartile	Second quartile	Third quartile	Highest quartile	
Race	Amer. Indian/Alaska Native	45	35	32	18	130
	Asian, Hawaii/Pac. Islander	432	275	310	443	1460
	Black or African American	668	583	447	322	2020
	Hispanic	1003	510	402	302	2217
	More than one race	151	166	215	203	735
	White	1301	2021	2347	3013	8682
Total		3600	3590	3753	4301	15244
Sex	Male	1717	1792	1900	2169	7578
	Female	1883	1798	1853	2132	7666
Total		3600	3590	3753	4301	15244

Approximately 80% of the students attended a public school and 48% attend a suburban school. However, 48% of Black students and 47% of Hispanics attended urban schools compared to only 25% of Whites. Frequencies for school control and urbanicity by racial groups are presented in Tables 3 and 4, respectively.

Table 3

*Population Frequencies: Race by School Control*

		School control			Total
		Public	Catholic	Other private	
Race	Amer. Indian/Alaska Native	131	3	6	140
	Asian, Hawaii/Pac. Islander	1464	82	108	1654
	Black or African American	1919	167	82	2168
	Hispanic	2098	253	82	2433
	More than one race	604	80	84	768
	White	6549	1388	1097	9034
Total		12765	1973	1459	16197

Table 4  
*Frequencies: Race by School Urbanicity*

		School urbanicity			Total
		Urban	Suburban	Rural	
Race	Amer. Indian/Alaska Native	25	61	54	140
	Asian, Hawaii/Pac. Islander	736	824	94	1654
	Black or African American	1049	866	253	2168
	Hispanic	1155	1047	231	2433
	More than one race	244	406	118	768
	White	2277	4560	2197	9034
Total		5486	7764	2947	16197

Among the students, 23% lived in a single parent or guardian home and 77% lived with two parents or guardians. 46% of Black students lived in a single parent or guardian home, compared to 24% of all Hispanic and 17% of all Whites. Frequencies for family composition by racial groups presented in Tables 5.

Table 5  
*Population Frequencies: Race by Family Composition*

		Family composition			Total
		Two Parent/ Guardian	One Parent/ Guardian	Other	
Race	Amer. Indian/Alaska Native	88	48	4	140
	Asian, Hawaii/Pac. Islander	1365	269	20	1654
	Black or African American	1145	984	39	2168
	Hispanic	1807	596	30	2433
	More than one race	562	197	9	768
	White	7420	1551	63	9034
Total		12387	3645	165	16197

Forty-eight percent (N = 7,701) of students reported that they had been, or were currently, employed for pay not around the house. Of the 43% (N = 7,115) who reported they were currently working, they worked an average of 16 hours per week ( $M = 15.9$ ,

$SD = 11.5$ ). Although 79% of all sophomores expected to attend college, only 20% reported ever participating in a school program to help them plan or prepare for college and only 52% reported being in a high school program with a concentration considered to be academic (college preparatory). Frequencies for educational expectations by high school concentration and school program are presented in Table 6.

Table 6

*Population Frequencies: Educational Expectation by Concentration and Planning Program (N = 16,197)*

		High school Concentration				Ever in program to help plan for college		
		College				No	Yes	Total
		General	Prep	Vocational	Total			
Educational Expectation	Don't know	720	539	191	1450	1174	207	1381
	Less than high school graduation	71	40	17	128	96	22	118
	High school graduation or GED only	608	223	152	983	747	134	881
	Attend or complete 2-year college/school	409	223	248	880	676	149	825
	Attend college, 4-year degree incomplete	257	227	79	563	410	122	532
	Graduate from college	1910	3090	419	5419	3971	1115	5086
	Master's degree or equivalent	777	2162	214	3153	2208	762	2970
	PhD, MD, or other advanced degree	619	1879	170	2668	1815	703	2518
Total		5371	8383	1490	15244	11097	3214	14311

Based on the first follow-up survey in 2004, 47% of the students surveyed had taken at least one math course above algebra II (Table 7). Within racial groups, 34% of Hispanics and 36% of Blacks had taken a math course above algebra II, compared to 50% of Whites. Only 29% of students in the lowest quartile of SES, compared to 67% in the highest, had taken a course above algebra II. Frequencies of highest math course by race and by SES quartile are presented in Tables 8 and 9, respectively.

Table 7  
*Population Frequencies: Highest Math*

	Frequency	Percent	Valid Percent
No math course or math course is other	143	.9	1.0
Pre-algebra, general or consumer math	638	3.9	4.3
Algebra I	924	5.7	6.2
Geometry	1900	11.7	12.8
Algebra II	4272	26.4	28.9
Trigonometry, pre-calculus, or calculus	6925	42.8	46.8
Total	14802	91.4	100.0
Missing	128	.8	
Survey component legitimate skip/NA	136	.8	
Nonrespondent	1131	7.0	
Total	16197	100.0	

Table 8

*Population Frequencies: Race by Highest Math Course*

	Highest Math Course						Total
	None or other	Pre-alg., general or consumer math	Alg. I	Geom.	Alg. II	Trig, pre-calc, or calculus	
Am. Ind./Alaska Nat.	1	15	9	29	41	29	124
Asian, Hawaii/Pac. Isl.	12	38	51	122	305	984	1512
Black or African Am.	35	74	143	311	688	707	1958
Hispanic	27	136	211	379	687	736	2176
More than one race	5	38	48	82	194	302	669
White	63	337	462	977	2357	4167	8363
Total	143	638	924	1900	4272	6925	14802

Table 9

*Population Frequencies: Highest Math Course by Quartile Coding of SES2*

	Quartile coding of SES2 variable				Total
	Lowest quartile	Second quartile	Third quartile	Highest quartile	
No math course or math course is other	57	25	29	18	129
Pre-algebra, general or consumer math	233	185	105	39	562
Algebra I	357	261	168	74	860
Geometry	613	515	408	243	1779
Algebra II	1016	1041	1041	940	4038
Trigonometry, pre-calculus, or calculus	910	1211	1683	2722	6526
Total	3186	3238	3434	4036	13894

Postsecondary transcript data indicated, for respondents whose data were available, 46% attended strictly a 4-year institution compared to 24% attending a 2-year. Additionally, 13% began at a 2-year and then transferred to a 4-year PSI and 13% began at a 4-year and then transferred to a 2-year PSI (Table 10).

Table 10  
*Population: Known PSI Combination*

	Frequency	Percent	Valid Percent
4-year only	5280	32.6	46.1
2-year only	2762	17.1	24.1
Less-than-2-year only	118	.7	1.0
Less-than-2-year, then 2-year or 4-year	53	.3	.5
2-year, then 4-year	1465	9.0	12.8
2-year, then Less-than-2-year	168	1.0	1.5
4-year, then Less-than-2-year	87	.5	.8
4-year, then 2-year	1510	9.3	13.2
Total	11443	70.6	100.0
Missing	180	1.1	
Survey component legitimate skip/NA	3648	22.5	
Nonrespondent	926	5.7	
Total	16197	100.0	

Of the students attending only a 4-year PSI, 19% took remedial math courses in college compared to 50% attending only a 2-year, 46% who transferred from a 2-year to a 4-year, and 33% who transferred from a 4-year to a 2-year (Table 11).



Table 11

*PSI Combination by Remedial Mathematics Course(s) Taken*

	Took Remedial Math		Total
	No	Yes	
4-year only	4206	1007	5213
2-year only	1386	1367	2753
Less-than-2-year only	113	3	116
Less-than-2-year, then 2-year or 4-year	31	22	53
2-year, then 4-year	793	670	1463
2-year, then Less-than-2-year	81	85	166
4-year, then Less-than-2-year	57	30	87
4-year, then 2-year	1008	502	1510
Total	7675	3686	11361

**Variables and Sampling Method**

The main variables used to select a subsample of participants from the public use ELS:2002 data were postsecondary institution (PSI) attendance and taking remedial math course in college. The variable F3PSCHCOMBO was a composite variable and indicated the institution levels of PSI attended by each student. This variable was used to select an initial subsample and was also used as a covariate in analysis. The initial subsample taken on the basis of on known college attendance and known remedial math course taking resulted in a sample of 11,361 students. Students attending any combination that included a less-than-two-year institution accounted for a combined 4% of the total initial sample. These cases were removed from the investigation because of their small sample size, resulting in 10,939 students.

**Independent Variable.** The independent variable considered in the present study was remedial math-taking in college. The variable F3TZREMMTTOT indicated the total number of known remedial mathematics courses taken by students during their

undergraduate education, as of June 2013, and was taken from postsecondary transcript data. The number of remedial courses taken ranged from zero to five. This variable was transformed into a dichotomous variable consisting of two groups. Group 1 included those students who took zero known remedial math courses. Group 2 included those students who took between one and five known remedial math courses, inclusive.

**Dependent Variables.** The main dependent variables in the present study were college GPA and degree attainment as of June 2013. The variable F3TZGPAALL indicated the GPA at all PSIs attended as of June 2013 taken from transcript data. The variable F3TZANYDEGRE indicated whether the student received a known degree from any PSI as of June 2013 based on transcript data. Degree attainment and college GPA had 15 and 198 missing cases, respectively. Missing cases on the dependent variables were removed from the analysis, resulting in a final sample size of 10,736.

**Covariates.** Numerous covariates were considered in the current study to examine whether preexisting differences predict participation in remedial math and to examine the effects of participation on college GPA and degree attainment. Based on the literature, common student demographic variables included in this study were sex, race, socioeconomic status (SES), family composition, 2003-2004 working status, and native language. Student level academic variables at the high school level included highest math course taken in high school, high school program concentration, college planning program participation, base-year math IRT estimated number right, and base-year reading IRT. The math and reading scores are based on Item Response Theory (IRT), which uses patterns of correct, incorrect, and omitted answers to estimate a student's ability. By using the overall patterns of right and wrong responses, IRT can compensate for the

possibility of a student guessing items correctly, e.g. “If answers on several easy items are wrong, a correct difficult item is assumed, in effect, to have been guessed.” (NCES, 2004).

High school level variables considered were school control and urbanicity. Postsecondary level variables included school sector and level combination, timing of PSI enrollment, and student loan indicators. Two additional composite variables were included. The first was “postsecondary education pipeline.” This composite variable included a combination of high school level data such as students’ educational expectations and academic readiness, as well as postsecondary level data such as college entrance exam completion and college application. Students were considered to have completed the postsecondary pipeline if they had a 10<sup>th</sup> grade educational expectation of a Bachelor's degree or higher, were academically prepared, took a college entrance exam, and applied to a postsecondary institution. Students not meeting all of the conditions were considered to have partially completed the pipeline, and those who met none of these criteria were considered to have never entered the pipeline.

The second variable included was “high school attainment indicator.” The high school attainment indicator categorized high school attainment on the basis of graduation status, grades, and assessment scores. Students were coded as "successful graduates" if they received a high school diploma by August 2004, their 10th-grade academic GPA was at least 2.5, and their base year composite ELS assessment score was at or above one-fourth of a standard deviation below the mean. High school diploma recipients who did not meet all of the above criteria were considered "marginal graduates.” Non-completers and GED recipients were used as a reference category in this study. Nominal

variables with more than two levels were recoded using dummy variables (0 and 1) to create dichotomous variables, which were required for the regression analyses used in this study. A complete list of variable descriptions, as well as variables that were recoded, is presented in Appendix A.

### **General Research Methodology**

I conducted a nonexperimental investigation of a sample of 10,736 students from the ELS:2002 public use data with one independent variable: whether students took one or more remedial math courses during their undergraduate studies. I examined preexisting differences in students who did and did not take remedial math in college and whether those variables predicted participation in remedial math. I also investigated the effects of remedial math-taking (with and without the inclusion of covariates) on college GPA and degree attainment. This study was nonexperimental because I used an existing data set from the NCES, thus students were not randomly assigned to comparison groups. The variables of interest in this study were remedial math course taking, college degree attainment, and college GPA, thus sample participants were selected who had data values for all three of these variables.

### **Data Analysis**

All data analysis in this study was conducted using version 24 of the statistical software package, SPSS. Data analysis was a multi-stage process described below.

**Stage One.** I conducted preliminary analyses on the sample to investigate preexisting differences between students who did or did not take remedial math in college for each covariate used in this study. I used *t*-tests for differences in means for continuous variables and Chi-square tests for differences in frequencies for nominal

variables. Next, I conducted binary logistic regression with all covariates predicting remedial math-taking to determine if the covariates were significant predictors of the independent variable. Finally, I conducted logistic regression analyses predicting degree attainment and linear regression predicting college GPA using two models. Model 1 included remedial math-taking as the only predictor. Model 2 included remedial math plus all covariates as predictors.

**Stage Two.** I conducted propensity score matching to attempt to match participants using multiple covariates to create statistically equivalent comparison groups for subsequent analysis. I created a subsample of matched remedial math groups from the sample, including all covariates during the matching procedure. I investigated the balance of the matched groups on the covariates using *t*-tests, Chi square tests, and effect sizes, and then repeated the analyses from stage one using the matched sample. I then compared the results from stages one and two.

**Stage Three.** I investigated missing data for each covariate in the study using a missing value analysis procedure. I then conducted multiple imputation procedures to replace missing values on covariates, obtaining 12 multiply imputed complete case data sets. After conducting MI, I repeated the analyses conducted in stage one and compared the results to my previous models.

**Stage Four.** I conducted propensity score matching on each of the 12 multiply imputed data sets obtained in stage three. I then repeated the analyses conducted in stage one. Finally, I compared the outcomes from all four stages to investigate any differences in the resulting treatment effects.

## CHAPTER IV

### RESULTS

The results of the study are presented in this chapter. This chapter is divided into seven sections. The first section contains descriptive statistics of the sample. The second section presents the results of the stage one analysis, including preliminary and regression analyses on the initial sample. The third section presents stage two of the analysis including propensity score matching results and regression analyses on matched groups. The fourth section presents the results of the stage three analysis including a missing value analysis, multiple imputation procedures and results, and regression analyses on the imputed data sets. The fifth section presents the results of stage four, including the propensity score matching results and subsequent regression analyses on the multiply imputed data sets. The sixth concludes with a summary and comparison of the results from all stages of analysis.

#### **Description of the Sample**

**Independent and dependent variables.** A total of 10,736 unweighted participants were included in this study. In the final sample, 67% ( $N = 7,227$ ) of students had taken no remedial math and 33% ( $N = 3,509$ ) had taken at least one remedial math course during their undergraduate studies. Students attending only 4-year institutions accounted for 48% of the sample, compared to 24% attending only 2-year, 14% attending first a 2-year and then a 4-year, and 14% attending first a 4-year and then a 2-year institution (Table 12).

Of those students attending only a 4-year institution, 19% had taken at least one remedial math course compared to 51% of those attending only a 2-year PSI. Forty-six

percent of students who first attend a 2-year, and then a 4-year PSI had taken at least one remedial math course, compared to 33% who had first attended a 4-year and then a 2-year PSI. In the sample, 63% of students had attained a college degree as of June 2013 and 37% had not. Of those students attending only a 4-year institution, 32% had had attained a degree compared to 45% of those attending only a 2-year PSI. Only 10% of students who first attend a 2-year, and then a 4-year PSI had had attained a degree compared to 13% who had first attended a 4-year and then a 2-year PSI (Table 12).

Table 12

*Sample: PSI Combination by Remedial Math and Degree Attainment*

	Remedial math		Degree attained		Total
	No	Yes	No	Yes	
4-year only	4155	998	1284	3869	5153
2-year only	1283	1340	1817	806	2623
2-year, then 4-year	788	669	414	1043	1457
4-year, then 2-year	1001	502	502	1001	1503
Total	7227	3509	4017	6719	10736

The mean GPA for all participants at all known institutions attended was 2.72 ( $SD = .87$ ). Students attending only a 4-year institution had a higher mean GPA than all other PSI levels ( $M = 2.92$ ,  $SD = .76$ ) and students attending only a 2-year had the lowest ( $M = 2.29$ ,  $SD = 1.02$ ). Remedial math students had a lower mean GPA ( $M = 2.40$ ,  $SD = .85$ ) than non-remedial students ( $M = 2.87$ ,  $SD = .83$ ). Students who had attained a degree had a higher mean GPA ( $M = 3.10$ ,  $SD = .52$ ) than students who had not yet attained a degree ( $M = 2.08$ ,  $SD = .96$ ). Descriptive statistics for college GPA by PSI level, remedial math group, and degree attainment are presented in Table 13.

Table 13

*Descriptives: College GPA by PSI, Remedial, and Degree*

	<i>N</i>	<i>M</i>	<i>SD</i>
Entire Sample	10736	2.72	.87
PSI Level			
4-year only	5153	2.92	.76
2-year only	2623	2.29	1.02
2-year, then 4-year	1457	2.80	.72
4-year, then 2-year	1503	2.70	.81
Remedial Math			
No	7227	2.87	.83
Yes	3509	2.40	.85
Degree Attainment			
No	4017	2.08	.96
Yes	6719	3.10	.52

**Covariates.** In the sample, 60% of the students were White, compared to 12% Black, 12% Hispanic, 11% Asian, and 5% other race or multiracial. Females accounted for 53% and males 47%. Eighty-four percent of students' native language was English and 16% was a language other than English. Twenty percent of the students lived in a single-parent or guardian household and 80% in a two-parent or guardian household. During the 2003-2004 school year (the students' senior year), 69% reported working during the year and 26% reported that they did not work. A summary of all descriptive statistics and frequencies is presented in Appendix B.

Seventy-four percent of students attend a public high school and 26% attended a Catholic or other private school. Thirty-five percent attended an urban, 48% a suburban, and 17% a rural high school. The mean SES for all participants was .20 ( $SD = .74$ ). SES was measured on a continuous scale ranging from -2.11 (lower SES) to 1.98 (higher SES). Students in a general high school program concentration accounted for 30% of the



sample compared to 61% in a college-preparatory (academic) program, and 8% in a vocational program. Fifty-four percent had taken a math course above algebra II (precalculus, trigonometry, or calculus), and 41% had taken algebra II or lower. Mean math IRT estimated number right was 41.18 ( $SD = 11.10$ ). Math IRT was measured on a continuous scale ranging from 12.52 (low) to 69.72 (high). Mean reading IRT estimated number right was 32.39 ( $SD = 9.04$ ). Reading IRT was measured on a continuous scale ranging from 10.20 (low) to 49.09 (high). Twenty-one percent reported participating in a special program to help plan for college and 69% did not participate in such a program. Forty-nine percent of the students were considered successful high school graduates, 39% were marginal graduates, and 3% were non-completers or had received a GED or other high school equivalency. Students who were considered to have completed the postsecondary education (PSE) pipeline accounted for 57% compared to 38% who had not completed the pipeline (Appendix B).

The first PSI attended by most students (76%) was a public institution compared to 24% attending a private PSI. Sixteen percent of students delayed postsecondary enrollment after high school and 72% had immediate enrollment in a PSI. Fifty-six percent reported having taken out student loans compared to 36% who reported not taking out loans (Appendix B).

### **Stage One: Analysis of the Initial Sample**

**Preliminary analyses.** I conducted preliminary analyses by investigating initial differences between students who did or did not take remedial math in college for each covariate in the study. I used independent sample  $t$ -tests for continuous covariates and Chi-square tests for nominal dichotomous covariates. Analyses were conducted on the

entire initial sample ( $N = 10,736$ ). However multiple covariates had missing values, thus not all analyses were conducted with complete case data sets.

**Continuous Covariates.** I conducted independent sample  $t$ -tests for equality of means to investigate differences in mean SES, math IRT, and reading IRT between the remedial math groups. Summaries of the group statistics and  $t$ -tests are provided in Tables 14 and 15 respectively. There was a significant difference in mean reading IRT between remedial math groups,  $t(10,611) = 31.65, p < .001$ . Cohen's  $d = .65$ , indicating a medium effect size. Students taking remedial math had a lower mean reading IRT ( $M = 28.57, SD = 8.60$ ) than students who did not take remedial math ( $M = 34.24, SD = 8.66$ ).

Levene's test for homogeneity of variance was violated for SES ( $F = 58.63, p < .001$ ) and for math IRT ( $F = 15.19, p < .001$ ). Equal variances were not assumed and the Cochran and Cox calculated test statistic with Satterthwaite-calculated degrees of freedom was used for both SES and math IRT (Hinkle, Wiersma, & Jurs, 2003). There was a significant difference in mean SES between remedial math groups,  $t(7,116.09) = 21.93, p < .001$ . Cohen's  $d = .46$ , indicating a medium effect size. Students taking remedial math had a lower mean SES ( $M = -.02, SD = .68$ ) than students who did not take remedial math ( $M = .30, SD = .75$ ). There was a significant difference in mean math IRT between remedial math groups,  $t(7,304.54) = 42.81, p < .001$ . Cohen's  $d = .89$ , indicating a large effect size. Students taking remedial math had a lower mean math IRT ( $M = 35.19, SD = 9.79$ ) than students who did not take remedial math ( $M = 44.08, SD = 10.52$ ).

Table 14

*Summary of Group Statistics for Continuous Covariates*

	Remedial math	N	Mean	SD	S.E. Mean
SES	No	6890	.30	.75	.01
	Yes	3320	-.02	.68	.01
	Total	10210			
Math IRT	No	7153	44.08	10.52	.12
	Yes	3460	35.19	9.79	.17
	Total	10613			
Reading IRT	No	7153	34.24	8.66	.10
	Yes	3460	28.57	8.60	.15
	Total	10613			

Table 15

*t-test Summary: Remedial Math by Continuous Covariates*

	Independent Samples Test								
	Levene's Test for Equality of Variances		t-test for Equality of Means				95% CI of the Difference		
	F	Sig.	t	df	Sig. <sup>b</sup>	Mean Diff.	S.E. Diff.	Lower	Upper
SES	56.63	.000	21.93 <sup>a</sup>	7116.09	.000	.326	.015	.297	.355
Math IRT	15.19	.000	42.81 <sup>a</sup>	7304.54	.000	8.896	.208	8.488	9.303
Reading IRT	1.45	.229	31.65	10611	.000	5.663	.179	5.312	6.014

a Levene's test was sig., equal variances not assumed

b Two-tailed

**Nominal Covariates.** I conducted Chi-square tests for differences between the remedial math groups on each nominal covariate. Note that nominal covariates with more than two levels were recoded as dichotomous using dummy variables (e.g. 0 = no, 1 = yes). Chi-square tests showed significant differences between remedial math groups on all nominal covariates with the exception of each level of high school urbanicity, students who first attended a 4-year and then a 2-year PSI, students in the Other race category,

sex, and students who were in a special program to help plan for college. Chi-Square test summaries are provided in Appendix C.

Preexisting differences between students who did and did not take remedial math in college were evident based on these initial comparisons. Next, I conducted a binary logistic regression analysis predicting remedial math course taking in college (using no remedial math as the reference category) including all covariates. Although several covariates were not significantly different between remedial math groups, they were included in the logistic regression analysis as they were correlated with the dependent variable. Public school control, general high school program, students of White race, vocational high school program, and 4-year, then 2-year PSI level combination were used as references categories for all regression analyses in this study.

***Binary logistic regression predicting remedial math.*** Due to missing values for multiple covariates, the initial sample size of 10,736 was reduced to 7,109 because of listwise deletion during the regression process. Results of the logistic analysis indicated that the model provided a statistically significant prediction of remedial math-taking,  $\chi^2(28, N = 7109) = 1853.55, p < .001$ . The predictive accuracy of the model was 76% (pseudo  $R^2 = .37$ ). Logistic regression coefficients, standard errors, and odds ratio summaries are presented in Appendix D. Preexisting differences on covariates were detected in remedial math groups and those combined covariates significantly predicted remedial math course taking. As such, the differences in remedial math groups could influence treatment effects of remedial math on the dependent variables. An analysis of the effect of remedial math-taking on degree attainment and college GPA was conducted

using two models. Model 1 included remedial math as the sole predictor of degree attainment and college GPA. Model 2 included remedial math plus all covariates.

**Analysis on the dependent variables.**

**Binary logistic regression Model 1.** Remedial math-taking was used in a binary logistic regression analysis predicting degree attainment. The logistic regression analysis indicated that the model provided a statistically significant prediction of degree attainment,  $\chi^2(1, N = 10,736) = 571.21, p < .001$ . Remedial math-taking significantly predicted degree attainment ( $b = 1.01, S.E. = .04, p < .001$ ). The predictive accuracy of the model was 65% (pseudo  $R^2 = .07$ ). The odds ratio for remedial math-taking was 2.7, indicating that, with no other information about the students, those who did not take remedial math were almost three times more likely to attain a degree than students who took remedial math.

**Binary logistic regression Model 2.** Remedial math-taking plus all covariates were used in a binary logistic regression analysis predicting degree attainment. This analysis had 3,627 missing cases on the covariates and these cases were automatically removed from the analysis. The logistic regression analysis indicated that the model provided a statistically significant prediction of degree attainment,  $\chi^2(29, N = 7,109) = 1968.27, p < .001$ . The predictive accuracy of the model was 77% (pseudo  $R^2 = .34$ ). With inclusion of the covariates, remedial math-taking was still a significant predictor of degree attainment ( $b = .38, S.E. = .07, p < .001$ ). The odds ratio for remedial math-taking was 1.5, indicating that, with all covariates included in the model, those who did not take remedial math were 1.5 times more likely to attain a

degree than students who took remedial math. Logistic regression coefficients, standard errors, and odds ratios for all variables in Model 2 are presented in Appendix E.

The predictive accuracy in Model 1 (remedial as the only predictor) was 65%. After inclusion of the covariates, the predictive accuracy was 77%. The odds ratio for remedial math-taking in Model 1 was 2.7. In Model 2 this ratio was 1.5, indicating that, after controlling for differences in the remedial groups on the covariates, students who did not take remedial math were 1.5 times more likely to earn a degree than students who took remedial math.

***Linear Regression Model 1.*** Remedial math course taking in college was used in a linear regression analysis to predict college GPA. Remedial math course taking significantly predicted college GPA ( $b = -.47, t = -27.38, p < .001$ ). Remedial math course taking explained a statistically significant proportion of variance in college GPA,  $R^2 = .07, F(1, 10,736) = 749.43, p < .001$ . Approximately 7% of the variance in college GPA was explained by remedial math-taking in college alone.

***Linear Regression Model 2.*** All covariates were entered in a linear regression analysis predicting college GPA. There were 3,627 missing cases on the covariates for this model and these were removed from the analysis. An investigation of the correlation matrix indicated that none of the covariates used in the model were too highly correlated with the dependent variable or with each other (correlations were .5 or smaller). The covariates explained a significant proportion of variance in college GPA,  $R^2 = .26, F(29, 7,109) = 84.49, p < .001$ . Approximately 26% of the variance in college GPA was explained by the combined covariates. Remedial math-taking had a significant negative impact on college GPA,  $b = -.13, t = -5.99, p < .001$ . Tests for multicollinearity indicated

that multicollinearity was present for the levels of high school attainment indicator (VIF > 10). This variable was separated into dichotomous variables using dummy coding. The multicollinearity was likely the result of redundancy. Three predictors (2-year, then 4-year PSI level combination, other race, and sex) had partial correlations that were slightly larger than their zero-order correlations, indicating the possibility of suppressor variables. A summary of all linear regression coefficients and part and partial correlations appears in Appendix F.

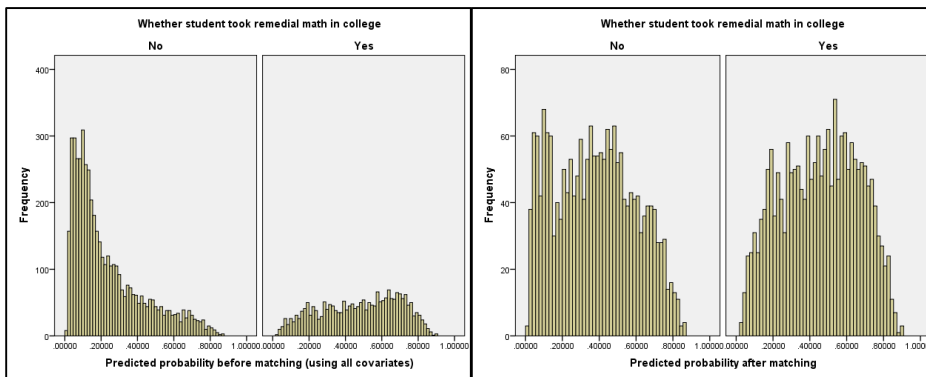
The total variance explained by remedial math in Model 1 was 7%. After inclusion of the covariates, the total explained variance was 26%. The linear regression analysis showed a significant negative effect of remedial math course taking on college GPA ( $b = -.47$ ) when it was the sole predictor in the model. When all other covariates were entered, the regression coefficient for remedial math course taking was  $-.13$ . In Model 1, the squared partial correlation for remedial math was  $.07$ . After entering all covariates, the squared partial correlation for remedial math was  $.01$ , indicating that, in the presence of all other variables in the model, remedial math-taking uniquely accounted for only one percent of the variance in GPA.

These results suggested that the effect of remedial math-taking on degree attainment and college GPA may have been the result of preexisting differences between students who did and did not take remedial math. Given the potential for biased coefficient estimates due to nonrandom group assignment and preexisting differences in the remedial math groups, propensity score matching was conducted on the original sample.

## Stage Two: Propensity Score Matching on the Original Sample

I conducted one-to-one propensity score matching (PSM) using all covariates in the model. Due to missing values on the covariates, the initial samples size available for matching was reduced from 10,736 to 7,109. One-to-one matching with a match tolerance of .1 resulted in 1,989 matches. Unmatched cases were removed from the analysis. Matching resulted in a total sample size of 3,978 divided equally between remedial math groups. Figure 1 shows side-by-side histograms of the predicted probability of remedial math group membership before and after matching. The predicted probability after matching is equivalent to the propensity scores used for the matching process.

Figure 1  
*Predicted Probability of Remedial Math Group Membership Before and After Matching in the Original Sample*



As can be seen from the histograms, the remedial math groups were more similar after matching. After matching, all previously run analyses in stage one were repeated using the matched sample.



**Preliminary analysis on matched sample.**

*Continuous covariates.* Preliminary analysis showed that there was a statistically significant difference between matched remedial math groups on mean reading IRT,  $t(3976) = 5.83, p < .001$ , and on mean SES,  $t(3976) = 2.17, p = .03$ . Levene’s test for homogeneity of variance was violated for math IRT ( $F = 21.64, p < .001$ ). The Cochran and Cox calculated test statistic with adjusted degrees of freedom showed a significant difference in mean math IRT between matched groups,  $t(3931) = 8.58, p < .001$ .

Summaries of group statistics for each variable before and after matching are provided in Table 16.

Table 16  
*Summary of Group Statistics for Continuous Covariates Before and After Matching*

	Remedial math	Before matching			After matching		
		Mean	SD	S.E. Mean	Mean	SD	S.E. Mean
SES	No	.30	.75	.01	.06	.70	.02
	Yes	-.02	.68	.01	.01	.69	.02
Math IRT	No	44.08	10.52	.12	38.96	10.87	.24
	Yes	35.19	9.79	.17	36.15	9.76	.22
Reading IRT	No	34.24	8.66	.10	31.03	9.13	.20
	Yes	28.57	8.60	.15	29.38	8.69	.19

Although statistically significant differences in means were found in the matched sample, this is not unusual due to the large sample size. A comparison of the mean difference and effect size for each covariate (Table 17) before and after matching indicates that the mean differences and effect sizes were greatly reduced after matching. The effect size for SES was reduced from .46 (medium) to .07 (minimal). The effect size for math IRT was reduced from .89 (large) to .27 (small). The effect size for reading IRT was reduced from .66 (medium) to .18 (small). PSM appeared to have created matched

groups that were more similar on the continuous covariate measures, effectively reducing the preexisting difference between the groups, though it did not eliminate all differences.

Table 17

*Mean Difference Between Remedial Math Groups Before and After Matching*

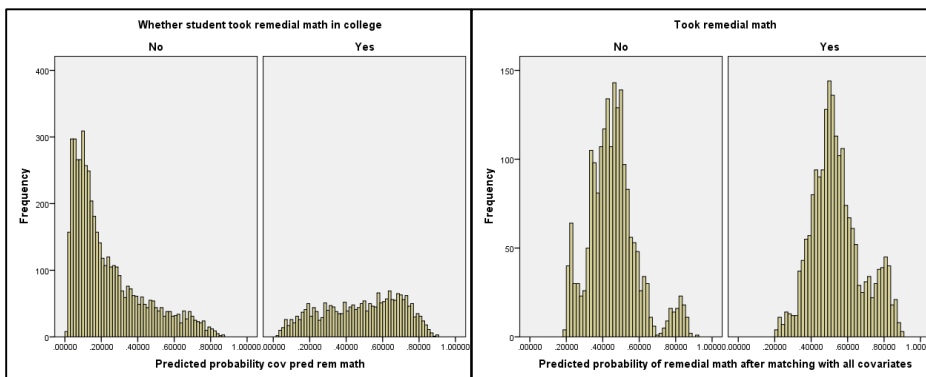
Covariate	Before matching			After matching		
	Mean Diff.	S.E. Diff.	Cohen's <i>d</i>	Mean Diff.	S.E. Diff.	Cohen's <i>d</i>
SES	.33	.01	.46	.05	.02	.07
Math IRT	8.90	.21	.89	2.81	.33	.27
Reading IRT	5.66	.18	.66	1.65	.28	.18

**Nominal covariates.** I conducted Chi-square tests of differences for nominal dichotomous covariates by matched remedial math groups. Tests showed significant differences for a number of covariates including high school control, high school urbanicity, highest high school math level, sector of first PSI attended, high school attainment indicator, PSI level combination (except 4-year, then 2-year), race (except Asian and other), completed PSE pipeline, whether English was the native language, and whether they took out student loans. A summary of chi-square values and significance levels for each covariate before and after matching appears in Appendix G. Although significant differences between matched groups were found, a comparison of the standardized residuals shows equal residuals (in absolute value) by level of dichotomous variable between the matched remedial math groups. A summary of the standardized residuals for each covariate before and after matching is provided in Appendix H.

**Binary logistic regression predicting remedial math.** A binary logistic regression was conducted with all covariates predicting remedial math-taking in the matched sample. The logistic analysis indicated that the model provided a statistically

significant prediction of remedial math-taking,  $\chi^2(28, N = 3978) = 346.59, p < .001$  (pseudo  $R^2 = .11$ ). Logistic regression coefficients, standard errors, and odds ratios for the matched sample are presented in Appendix I. These results indicated that there were still differences in remedial math groups after matching, but the combined explained variance was reduced from 32% before matching to 11% after matching, presumably because matched participants were more similar on the covariates, which was the goal of conducting PSM. Logistic regression resulted in ten significant regression coefficients before matching and six after matching. Figure 2 shows a comparison of the logistic regression predicted probability of being in a remedial math group with the inclusion of all covariates before and after matching, respectively.

Figure 2  
*Covariates Predicting Remedial Math Before and After Matching*



**Post-matching analysis on the dependent variables.**

***Post-matching binary logistic regression Model 1.*** Remedial math-taking was used in a binary logistic regression analysis predicting degree attainment in the matched sample. The logistic regression analysis indicated that the model provided a statistically significant prediction of degree attainment,  $\chi^2(1, N = 3,978) = 55.36, p < .001$  (pseudo  $R^2$

= .02). Remedial math-taking significantly predicted degree attainment in the matched sample ( $b = .48$ , S.E. = .07,  $p < .001$ ). Table 18 shows a comparison of the regression coefficients in Model 1 before and after matching for remedial math predicting degree attainment.

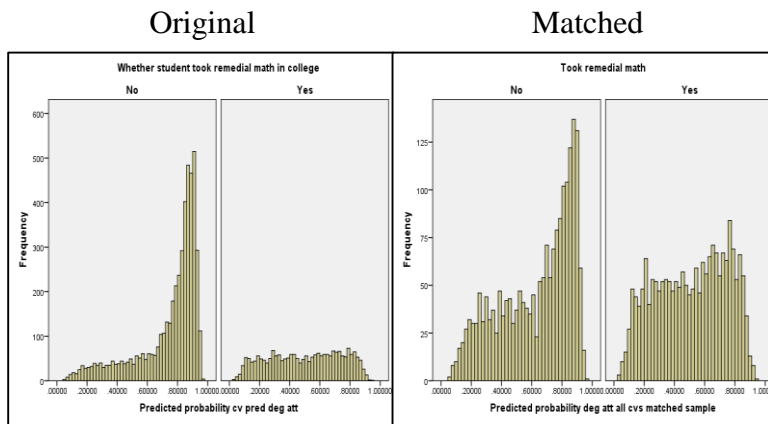
Table 18  
*Logistic Regression Model 1 Summary Before and After Matching*

		Coeff.	S.E	Sig	Odds Ratio	Pseudo $R^2$
Before	Remedial math	1.01	.04	.000	2.74	.07
	Constant	-.14	.03	.000	.87	
After	Remedial math	.48	.07	.000	1.61	.02
	Constant	.05	.05	.272	1.05	

Compared to Model 1 pre-matching, the odds ratio for remedial math decreased from 2.7 to 1.6. This indicated that, after controlling for differences in all other covariates in the model for the matched sample, students who did not take remedial were 1.6 times more likely to attain a degree than students who took remedial.

***Post-matching binary logistic regression Model 2.*** Remedial math-taking plus all covariates were used in a binary logistic regression analysis predicting degree attainment in the matched sample. The logistic regression analysis indicated that the model provided a statistically significant prediction of degree attainment,  $\chi^2(29, N = 3978) = 1033.15, p < .001$  (pseudo  $R^2 = .31$ ). Before matching, the predictive accuracy of Model 2 was 77%. After matching this was slightly reduced to 73%. Figure 3 shows a side-by-side comparison of histograms for the predicted probability of degree attainment before and after matching.

Figure 3  
*Model 2: Predicted Probability of Degree Attainment Before and After Matching*



The odds ratio for remedial math-taking in Model 2 before matching was 1.5. After matching, this ratio was still 1.5, indicating that after controlling for differences in all other covariates, students who did not take remedial math were 1.5 times more likely to earn a degree than students who took remedial math both before and after matching. Overall, the number of significant regression coefficients did not change in Model 2 after matching. A summary of Model 2 before and after matching is presented in Appendix J.

***Post-matching linear regression Model 1.*** Remedial math course taking was used in a linear regression analysis to predict college GPA in the matched sample. Remedial math course taking significantly predicted college GPA,  $b = -.23$ ,  $t = -8.23$ ,  $p < .001$ . Remedial math course taking explained a statistically significant proportion of variance in college GPA,  $R^2 = .02$ ,  $F(1, 3,978) = 67.68$ ,  $p < .001$ . Approximately 2% of the variance in college GPA was explained by remedial math-taking in college in the matched sample.

***Post-matching linear regression Model 2.*** All covariates in the matched sample were used in a linear regression analysis predicting college GPA. The covariates explained a significant proportion of variance in college GPA,  $R^2 = .20$ ,  $F(29, 3,978) = 34.01$ ,  $p < .001$ . Approximately 20% of the variance in college GPA was explained by the combined covariates. In the matched sample, the linear regression analysis showed a significant negative effect of remedial math course taking on college GPA ( $b = -.11$ ,  $t = -4.22$ ,  $p < .001$ ). When all other covariates were entered, the coefficient for remedial math course taking was reduced from  $-.23$  in Model 1 to  $-.11$  in Model 2 for the matched sample.

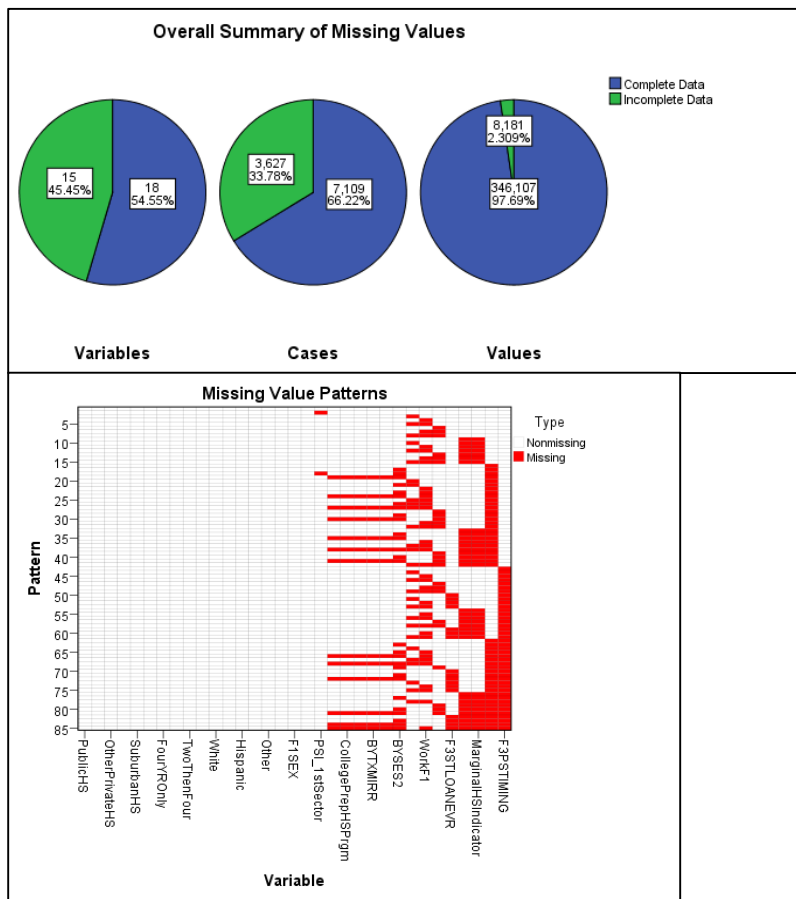
The total variance in college GPA explained by the covariates (Model 2) before matching was 26%. After matching, the explained variance was lower (20%). Before matching, the linear regression analysis (Model 2) showed a significant negative effect of remedial math course taking on college GPA ( $b = -.13$ ,  $t = -5.99$ ,  $p < .001$ ). After matching, remedial math still showed a significant negative effect ( $b = -.11$ ,  $t = -4.22$ ,  $p < .001$ ), but there was a slight decrease, from  $-.13$  to  $-.11$ . A summary of the linear regression coefficients before and after matching is presented in Appendix K.

Four predictors (other private high school, 2-year then 4-year PSI level combination, other race, and college planning program) had partial correlations that were larger than their zero-order correlations, indicating the possibility of suppressor variables. Compared to Model 2 before matching, the levels of high school attainment indicator still had elevated variance inflation factors, but the values were smaller after matching ( $VIF < 10$ ). A summary of the part and partial correlations is presented in Appendix L.

### Stage Three: Missing Value Analysis and Multiple Imputation

During the analyses in stages one and two, I noted varying amounts of missing data on multiple covariate measures included in the study. Because missing values were present, these cases were automatically deleted (listwise) during the preliminary analysis, during Model 2 construction using all covariates, and during propensity score matching. It is possible that the results obtained may have been influenced by the missing data. I proceeded to conduct a missing value analysis in SPSS. Figure 4 shows an overall summary of missing values along with the pattern of missing values.

Figure 4  
Summary of Missing Values and Patterns



It was determined that 15 variables had missing values and that the ratio of the number of missing values to the number of cases was 2.25 to 1. This large ratio implies that the listwise deletion methods used in stage one and two may have resulted in a loss of information. Simply deleting missing cases may produce biased results because cases with missing data may not be missing randomly. The percent of missing values ranged from .03% to 12.1% of the initial sample. Because some cases were missing values on multiple covariates, this resulted in a reduction in the valid sample size from 10,736 to 7,109 – a 34% loss of data. Table 19 summarizes the variables with missing values in decreasing order.

Table 19  
*Summary of Missing Value Analysis*

	Missing <i>N</i>	Percent	Valid <i>N</i>	Mean	<i>SD</i>
Timing of first postsecondary enrollment	1304	12.10%	9432		
Ever in program to help prepare for college	1088	10.10%	9648		
Marginal HS graduate	969	9.00%	9767		
Successful HS graduate	969	9.00%	9767		
Took out any student/PSE loans	919	8.60%	9817		
Completed PSE pipeline	629	5.90%	10107		
Worked during 03-04 school year	614	5.70%	10122		
Highest HS math level	545	5.10%	10191		
SES	526	4.90%	10210	.20	.74
Reading IRT estimated number right	123	1.10%	10613	32.39	9.04
Math IRT estimated number right	123	1.10%	10613	41.18	11.10
Vocational HS Program	123	1.10%	10613		
College Prep HS Program	123	1.10%	10613		
General HS Program	123	1.10%	10613		
Sector of 1st PSI attended	3	.03%	10733		

Minimum percentage of missing values for variable to be included: .01%



A summary of tabulated missing value patterns, along with the number of complete cases available for analysis if variables missing in those patterns were not used, is presented in Table 20.

Table 20  
*Summary of Tabulated Patterns*

										Dependent and Independent Variables					
										Degree Attain. <sup>d</sup>		Remedial math <sup>d</sup>			
Missing Patterns <sup>a</sup>										CC <sup>b</sup>	GPA <sup>c</sup>	No	Yes	No	Yes
N	SES	Help plan	Pipeline	High Math	Work	Successful	Marginal	Loans	Timing						
7109										7109	2.77	2360	4749	4889	2220
318			X							7427	2.58	154	164	190	128
384		X								7493	2.82	132	252	256	128
257	X	X								7750	2.70	84	173	165	92
212									X	7321	2.57	103	109	135	77
739								X	X	8060	2.30	464	275	472	267
246				X	X					7463	2.53	123	123	138	108
493						X	X			7602	2.91	138	355	384	109

Patterns with less than 1% cases (107 or fewer) are not displayed.

a Variables are sorted on missing patterns.

b Number of complete cases if variables missing in that pattern (marked with X) are not used.

c Means at each unique pattern

d Frequency distribution at each unique pattern

Little's MCAR test was significant,  $\chi^2 = 14.151, p = .007$ . Estimated means and Little's MCAR test are summarized in Table 21. These test results indicated that the missing values in the sample may not have been missing completely at random, and listwise deletion or single imputation methods are not recommended in this situation. Multiple imputation (MI) was used to handle the missing values. All variables were included during the MI. Recent studies suggest using as many imputations as the percent

of missing values. In this analysis, the highest percentage was 12.10%, thus 12 imputations were performed using MI procedures. After MI, there were 12 imputed data sets containing no missing values and the previous analysis conducted in stages one and two were repeated on all 12 sets.

Table 21  
*Estimated Means and Little's MCAR Test*

	Estimated Means			
	GPA	SES	Math IRT	Reading IRT
All Values	2.72	0.20	41.18	32.39
EM	2.77	0.22	41.74	32.84

Little's MCAR test: Chi-Square = 14.151, DF = 4, Sig. = .007

### Post-Multiple Imputation Analysis.

*Preliminary analysis of the imputed data sets.* After imputation, I compared the frequencies and descriptive statistics for each of the 12 MI sets to the original sample. An investigation of the continuous variables showed that the descriptive statistics were similar before and after imputation. Table 22 summarizes these values for the original and pooled MI sets on the continuous covariates. A summary of the imputed nominal variables is presented in Appendix M.

Table 22  
*Pooled Covariate Estimates After Multiple Imputation*

		<i>N</i>	Range	Min	Max	<i>M</i>	SE	SD	Var.	FM	RI	RE
Original	SES	10210	4.09	-2.11	1.98	.20	.01	.74	.55			
	Math	10613	57.20	12.52	69.72	41.18	.11	11.10	123.26			
	Read	10613	38.62	10.46	49.09	32.39	.09	9.04	81.69			
Pooled	SES	10736	4.74	-2.27	2.48	.20	.01			.020	.020	.998
	Math	10736	58.66	11.06	69.72	41.16	.11			.003	.003	1.000
	Read	10736	42.52	8.09	50.61	32.36	.09			.018	.018	.998

Note. FM = fraction missing info., RI = relative increase variance, RE = relative efficiency

*Preliminary analysis of remedial groups on the imputed data sets.* I examined differences between remedial math groups in all 12 MI sets for the continuous covariates (SES, math IRT, reading IRT) using independent sample *t*-tests. Levene’s test for homogeneity of variance was significant for all three variables, thus the Cochran and Cox test statistic with adjusted degrees of freedom was used. Table 23 summarizes the pooled results.

Table 23  
*Pooled Independent Samples t-Tests for 12 MI Sets*

	<i>t</i> -test for Equality of Means			95% CI			
	<i>t</i>	df	Sig. <sup>b</sup>	Mean Diff.	S.E. Diff.	Lower	Upper
SES	22.35 <sup>a</sup>	14166	.000	.33	.01	.2988	.3563
Math IRT	42.97 <sup>a</sup>	238153	.000	8.88	.21	8.4795	9.2900
Reading IRT	31.745 <sup>a</sup>	46737	.000	5.65	.18	5.3038	6.0018

a Levene's Test was significant, equal variances not assumed

b Two-tailed

Test results showed significant differences in means between remedial math groups for all three variables. Table 24 summaries the differences for the original and pooled imputed sets.

Table 24  
*Mean Difference Between Remedial Math Groups Before and After MI*

Covariate	Before			After		
	Mean Diff.	S.E. Diff.	Cohen's <i>d</i>	Mean Diff.	S.E. Diff.	Cohen's <i>d</i>
SES	.33	.01	.46	.33	.01	.43
Math IRT	8.90	.21	.89	8.88	.21	.83
Reading IRT	5.66	.18	.66	5.65	.18	.61

A comparison of the means, standard errors, and Cohen’s *d* (Table 24) indicated that the imputed data sets had approximately the same values as did the original sample.

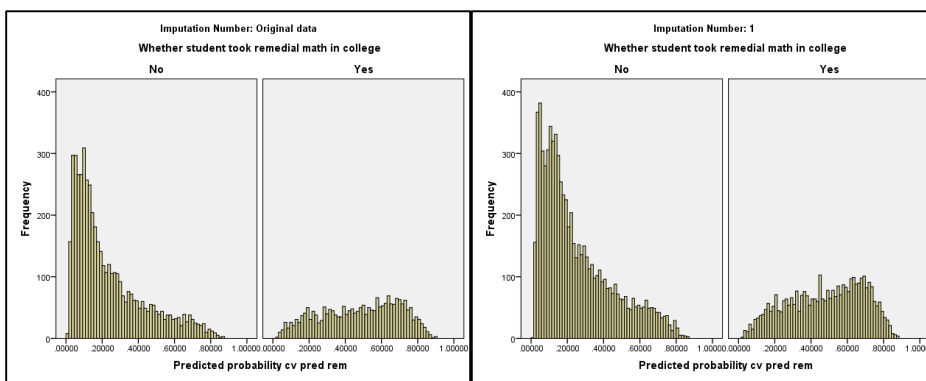
Next, I conducted chi-square tests of differences for all covariates that were imputed. Ten covariates with imputed values showed statistically significant differences between remedial math groups in all 12 imputed data sets. Because there is no agreed upon method of pooling Chi-square statistics for multiply imputed data, I considered each statistic separately and report the range of  $p$  values and statistics below. A summary of all Chi-square and  $p$  values for the 12 MI sets is presented in Appendix N.

College prep high school program, general high school program, highest high school math level, sector first postsecondary institution, successful graduate attainment indicator, marginal graduate attainment indicator, completed postsecondary education pipeline, worked during the 03-04 school year, and timing of postsecondary enrollment, were all significant,  $p < .001$ . Whether the student took out loans was significant for all imputed sets,  $.001 < p \leq .004$ . Chi-square values for all significant covariates ranged from 4.76 to 1,247.30. One covariate, whether the student was ever in a special high school program to help plan or prepare for college, was not significant,  $p > .05$ , for any of the 12 imputed data sets.

I conducted binary logistic regression on all MI sets with all covariates predicting remedial math-taking. All models were significant predictors of remedial math,  $2634.55 \leq \chi^2(28, N = 10736) \leq 2664.32, p < .001$ . This indicated that the inclusion of the covariates improved predication of remedial math-taking over the situation where no covariates are included. The predictive accuracy was 75% across all imputed sets with the inclusion of the covariates. The pseudo  $R^2$  was approximately .31 for all MI sets. In the original data set, the pseudo  $R^2$  was .37. Pooled logistic regression coefficients are summarized in Appendix O. These results indicated that there were preexisting

differences in remedial math groups in each MI data set. Figure 5 shows side-by-side histograms of the differences in group membership as predicted by inclusion of the covariates for the original sample and the first imputed data set. Histograms of the predicted probability of remedial math group membership for the other 11 MI sets had similar distributions and are not presented here.

Figure 5  
*Predicted Probability of Remedial Math for Original Data and First MI Set*



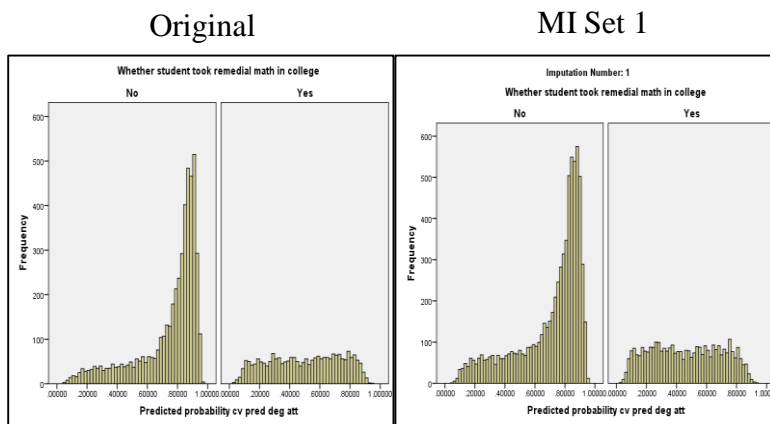
This demonstrates that after imputing missing values on the covariates to obtain complete cases data sets, the remedial math groups were relatively similar to the original sample groups when all covariates were used. Next, I re-ran Model 2 of the logistic and linear regressions. Model 1 was not repeated for the MI sets because Model 1 included no covariates (the source of the missing data that was imputed), thus the results would have been exactly the same.

**Logistic regression Model 2 on MI data sets.** I conducted a binary logistic regression with remedial math plus all covariates predicting degree attainment on the MI sets. There is no agreed upon method of pooling logistic regression statistics across models because logistic regression coefficients and odds ratios do not behave like linear regression statistics (Mood, 2010). Therefore, I considered each statistic separately and

report the range of  $p$  values and statistics below and I use this reporting method for the entirety of the MI analysis. All 12 models significantly predicted degree attainment,  $2975.57 \leq \chi^2(29, N = 10736) \leq 3017.08, p < .001$ . The predictive accuracy was 75% across all imputed sets with the inclusion of the covariates. The pseudo  $R^2$  was approximately .33 for all sets. Pooled logistic regression coefficients and odds ratios for all covariates are summarized in Appendix P. Remedial math-taking in college uniquely predict degree attainment and had an odds ratio of 1.44, indicating the students who did not take remedial math were about 1.5 times more likely to attain a degree than students who took remedial math. Figure 6 shows side-by-side histograms comparing the predicted probability of degree attainment with all covariates included for the original sample and first imputed data set. Histograms for the other 11 MI sets were similar and are not reported here.

Figure 6

*Predicted Probability of Degree Attainment: Original Sample and First MI Set*



In stages one and two of this investigation (pre-imputation), Model 2 with all covariates predicting degree attainment resulted in an odds ratio for remedial math-taking of 1.5 in both the unmatched and matched original sample. This indicated that students

who did not take remedial math were about 1.5 times more likely to attain a degree than students who took remedial math. The results in stages three and four (after MI) were similar, with an odds ratio of 1.4 for remedial math, even after imputing missing values for all covariates.

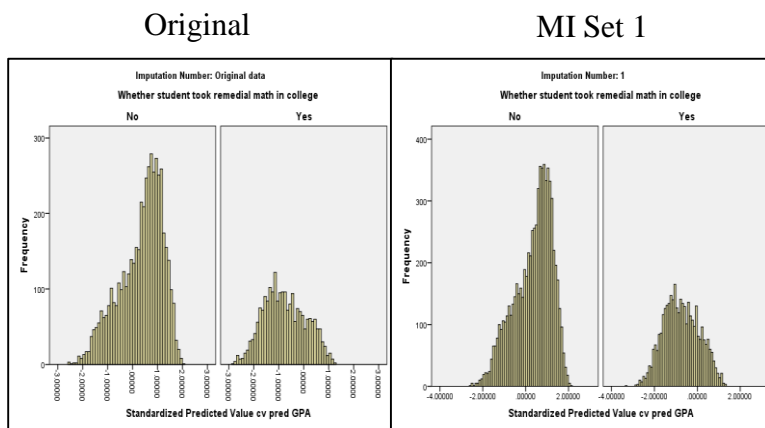
In stage one, the predictive accuracy of Model 2 was 77% for the original sample analysis. The analysis on the imputed sets also resulted in a significant model with 75% accuracy. In addition, the pseudo  $R^2$  in stage one was .34 for the original sample analysis. In stage two it was .33 for all imputed sets. Imputing the missing values of the covariates resulted in very little change in the statistics for analyses of the dependent variable of degree attainment.

***Linear regression Model 2 on MI sets.*** All covariates were used in a linear regression analysis predicting college GPA in the MI sets. The covariates explained a significant proportion of the variance in college GPA in all 12 imputed sets,  $R^2 \approx .25$ ,  $117.60 \leq F(29, 10,736) \leq 121.70$ ,  $p < .001$ . Approximately 25% of the variance in college GPA was explained by the combined covariates. In particular, for the pooled data, remedial math-taking showed a significant negative effect on college GPA in combination with the other covariates ( $b = -.14$ ,  $t = -7.62$ ,  $p < .001$ ). A summary of pooled estimates of coefficients can be found in Appendix Q.

In stage one (pre-imputation)  $R^2$  was .25 for the original sample. In stage three,  $R^2$  was .25. Imputing the missing values of the covariates resulted in no change in explained variance in GPA when all covariates were included in the model. In addition, in stage one, remedial math-taking in the original sample showed a significant negative effect on college GPA in combination with the other covariates ( $b = -.13$ ,  $t = -5.99$ ,  $p <$

.001). The results in stage three (presented above) were very similar ( $b = -.14$ ,  $t = -7.62$ ,  $p < .001$ ), even after imputing missing values for all covariates. Figure 7 shows side-by-side histograms for the original sample and first imputed MI set of the predicted value for college GPA for remedial math groups. Histograms for the other 11 MI sets were similar.

Figure 7  
*Predicted GPA by Remedial Math Group (Original and First MI Set)*



I next conducted one-to-one propensity score matching, this time using each of the 12 MI data sets to investigate differences in model 2 after controlling for preexisting differences in remedial math groups for the covariates.

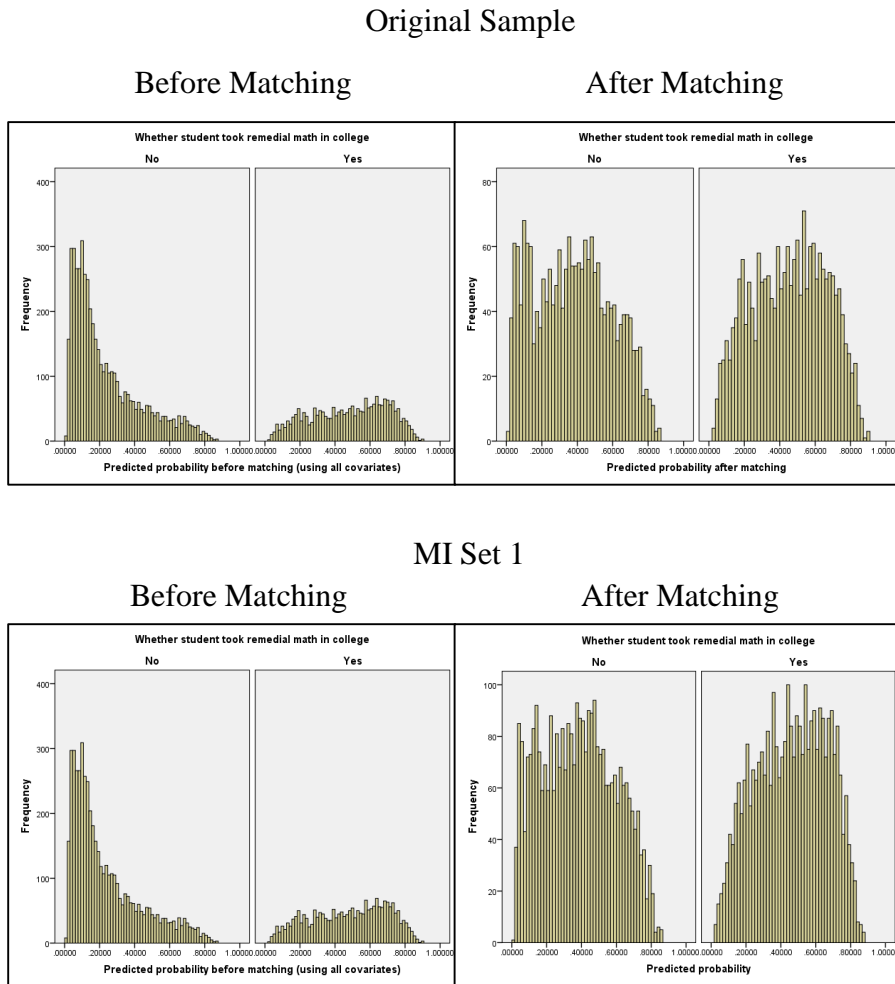
#### **Stage Four: Propensity Score Matching on All MI Sets**

There were no missing values because I used the imputed data sets, thus the matching was conducted on a complete case sample of 10,736 participants in each MI set. Because one-to-one matching was used, the maximum number of matches that could be made was 3,509 (the lesser of the two frequencies of students in the remedial math groups). Propensity score matching was conducted using a match tolerance of .1. The number of matches in all 12 imputed data sets ranged from 3,178 ( $N = 6,356$ ) to 3,216 ( $N$



= 6, 432). Matched samples were evenly divided between remedial math groups. Students who were not matched were removed from each sample. Figure 8 displays a comparison of side-by-side histograms of the predicted probability of remedial math group membership before and after matching for the original sample (stage one and two) and the first MI data set (stage three and four). As can be seen from the histograms, the remedial math groups were more similar after matching and the matching in the MI data has a similar histogram to the one obtained from the original sample. A comparison of the histograms for the other 11 imputed sets were similar and are not reported here.

Figure 8  
*Predicted Probability of Remedial Math Before and After Matching (Original and First MI Set)*



**Preliminary analysis after matching.** Preliminary analysis of the first imputed data set showed that there was a significant difference between matched remedial math groups on mean SES,  $t(6378) = 3.05, p = .002$ . Levene's test for homogeneity of variance was violated for math IRT ( $F = 43.05, p < .001$ ) and reading IRT ( $F = 19.29, p < .001$ ). Using the Cochran and Cox statistic with corrected degrees of freedom showed a significant difference between matched groups on mean math IRT,  $t(6296) = 10.93, p <$

.001, and mean reading IRT,  $t(6349) = 7.94, p < .001$ . Summaries of group statistics for each variable before and after matching for the first imputed data set are provided in Table 25.

Table 25  
*Summary of Group Statistics for Continuous Covariates Before and After Matching (First MI Set)*

	Remedial math	Before matching			After matching		
		Mean	SD	S.E. Mean	Mean	SD	S.E. Mean
SES	No	.30	.75	.01	.04	.70	.01
	Yes	-.02	.68	.01	-.01	.68	.01
Math IRT	No	44.08	10.52	.12	38.73	10.92	.19
	Yes	35.19	9.79	.17	35.90	9.74	.17
Reading IRT	No	34.24	8.66	.10	30.70	9.18	.16
	Yes	28.57	8.60	.15	28.93	8.58	.15

Chi-square tests of differences for nominal covariates by matched remedial math groups for the first imputed set showed significant differences on all nominal covariates with the exception of high school program, single parent home, sex, and timing of first postsecondary enrollment. A summary of chi-square values and significance levels for each covariate after matching appears in Appendix R. All preliminary analyses ( $t$ -tests and Chi-square tests) on the other 11 matched imputed data sets were similar to those presented above and are not reported here.

***Post-matching binary logistic regression predicting remedial math.*** I ran a binary logistic regression with all covariates predicting remedial math-taking in the matched sample for the first imputed data set. The logistic analysis indicated that the model provided a statistically significant prediction of remedial math-taking,

$\chi^2(28, N = 6380) = 560.43, p < .001$ . The pseudo  $R^2$  was .11. These results indicated

that there were still differences in remedial math groups after matching in the first MI set. These results are comparable to those obtained in the original sample (before imputation). Similar results were obtained for all of the other 11 MI sets and are not reported here.

**Post-matching analysis on the dependent variables for MI Sets.**

*Post-matching binary logistic regression Model 1.* Remedial math-taking was used in a binary logistic regression analysis predicting degree attainment in the matched sample for the first imputed set. The logistic regression analysis indicated that the model provided a statistically significant prediction of degree attainment,

$\chi^2(1, N = 6,380) = 81.67, p < .001$ . Remedial math-taking significantly predicted degree attainment in the matched sample ( $b = .46, S.E. = .05, p < .001$ ). Compared to Model 1 pre-matching, the regression coefficient was reduced from 1.01 to .46 and the odds ratio from 2.74 to 1.58. Similar results were obtained for the other 11 imputed data sets.

Table 26 shows a comparison of model 1 before and after matching for remedial math predicting degree attainment for the first imputed data set.

Table 26  
*Summary of Model 1 Before and After Matching for First MI Set*

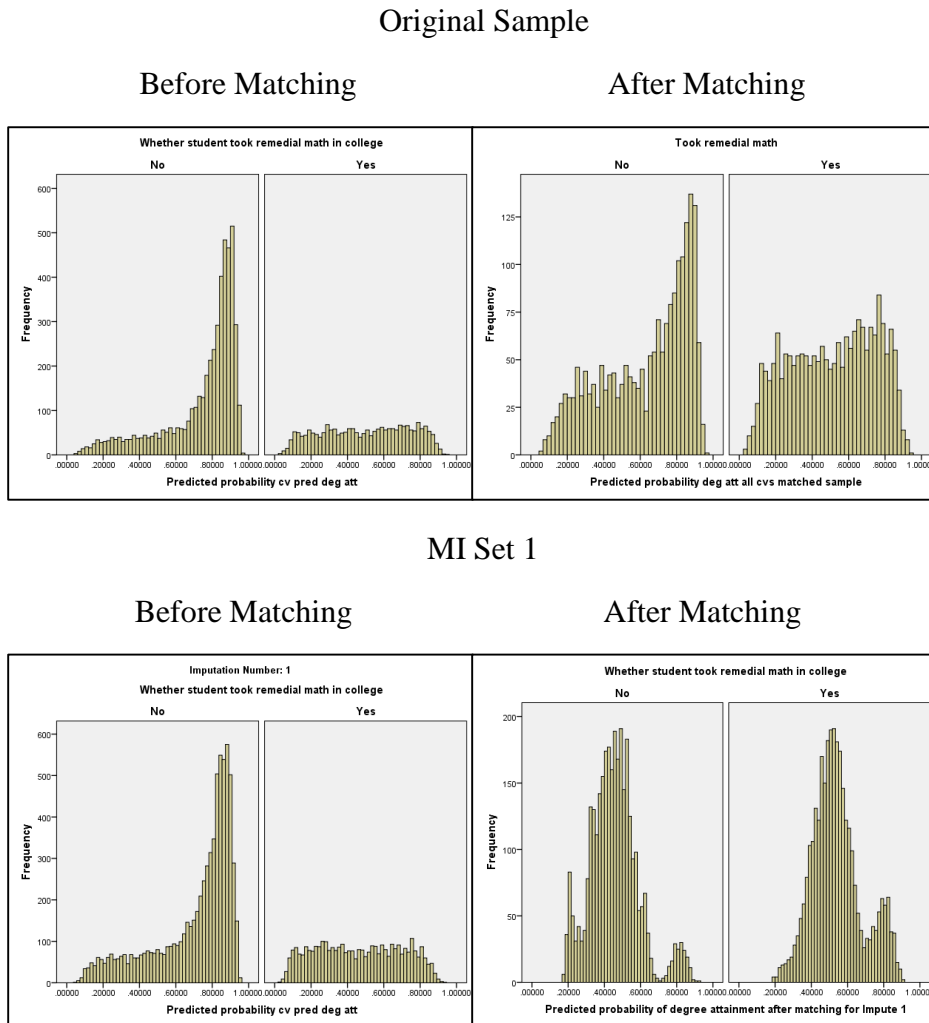
		Coeff.	S.E	Sig	Odds Ratio	Pseudo $R^2$
Before	Remedial math	1.01	.04	.000	2.74	.07
	Constant	-.14	.03	.000	.87	
After	Remedial math	.46	.05	.000	1.58	.02
	Constant	-.12	.04	.001	.89	

*Post-matching binary logistic regression Model 2.* Remedial math-taking plus all covariates were used in a binary logistic regression analysis predicting degree attainment in the matched sample for the first imputed data set. The logistic regression

analysis indicated that the model provided a statistically significant prediction of degree attainment,  $\chi^2(29, N = 6,380) = 1548.19, p < .001$ . The predictive accuracy of the model was 72% with the inclusion of the covariates. The pseudo  $R^2$  was .29. Logistic regression coefficients and odds ratios for all covariates before and after matching for the first imputed data set are summarized in Appendix S.

The odds ratio for remedial math, which is the independent variable of interest in this study, was 1.4, indicating that students who did not take remedial math were about 1.4 times more likely to attain a degree than students who took remedial math. Figure 9 shows a comparison of side-by-side histograms for the original sample and the first imputed data set of the predicted probability of degree attainment before and after matching. The results of this analysis and the histogram comparisons were similar for all of the other 11 imputed data sets and are not reported here.

Figure 9  
*Model 2: Predicted Probability of Degree Attainment Before and After Matching (Original and First MI)*



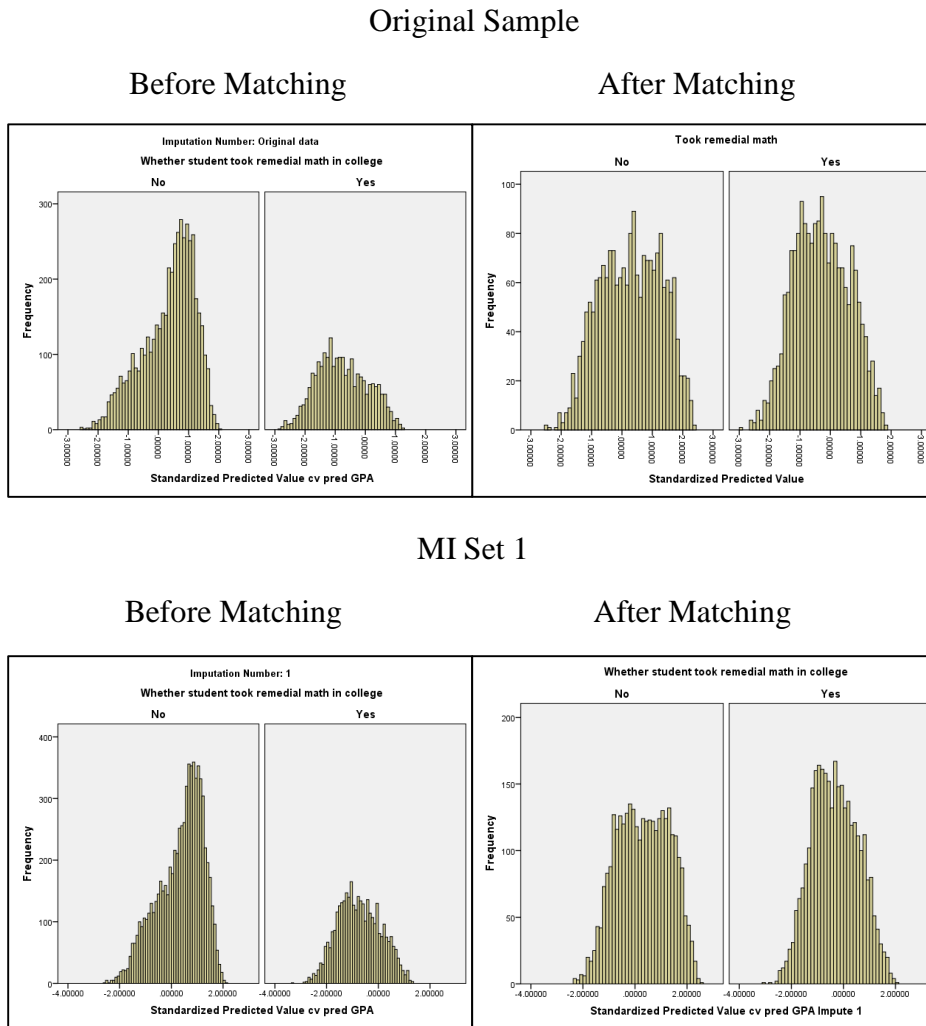
In stages one and two of this investigation (pre-imputation), Model 2 with all covariates predicating degree attainment resulted in an odds ratio of 1.5 for remedial math-taking. This indicated that students who did not take remedial math were about 1.5 times more likely to attain a degree than students who took remedial math. The results in stages three and four (after imputation) were similar, even after imputing missing values

for all covariates. In stages three and four, the odds ratio for the remedial math groups was 1.4 in both the unmatched and matched groups for the first imputed data set. All of the other 11 MI had similar results.

***Post-matching linear regression Model 1.*** Matched remedial math groups were used in a linear regression analysis to predict college GPA in the first imputed data set. Remedial math course taking explained a significant proportion of variance in college GPA,  $R^2 = .02$ ,  $F(1, 6380) = 119.99$ ,  $p < .001$ . Remedial math course taking significantly predicted college GPA,  $b = -.25$ ,  $t = -10.95$ ,  $p < .001$ . Approximately 2% of the variance in college GPA was explained by remedial math-taking in college for the first imputed data set.

***Post-matching linear regression Model 2.*** All covariates were used in a linear regression analysis predicting college GPA after matching in the first imputed sample. The covariates explained a significant proportion of variance in college GPA,  $R^2 = .19$ ,  $F(29, 6380) = 49.94$ ,  $p < .001$ . Approximately 19% of the variance in college GPA in the first imputed data set was explained by the combined covariates for matched remedial groups. Figure 10 shows a comparison of side-by-side histograms for the original sample and the first imputed data set of the predicted probability GPA value for remedial math groups before and after matching.

Figure 10  
*Model 2: Predicted GPA by Remedial Math Group Before and After Matching (Original and First MI Set)*



In the matched sample, the linear regression analysis showed a significant negative effect of remedial math course taking on college GPA ( $b = -.13$ ,  $t = -6.15$ ,  $p < .001$ ). A summary of regression coefficients for Model 2 before and after matching in the first imputed sample appears in Appendix T. The same analysis was conducted for the other 11 imputed data sets and the results of the analyses were similar to that for the first imputed set and are not reported here.



## Summary of Results

Preexisting differences on numerous covariates were detected in remedial math groups and these combined measures significantly predicted remedial math course taking in college. When no other information about the students was considered, analyses indicated that students who did not take remedial math were almost three times more likely to earn a college degree than students who took remedial math. After including the information about the students' characteristics (academic, non-academic, and non-cognitive measures), this predicted likelihood was reduced to one and a half. This indicated that the negative effect of remedial math-taking on degree attainment may have been influenced by these characteristics. However, even after controlling for preexisting differences, taking remedial math still showed a negative effect on degree attainment. Similarly, when used as the only predictor, taking remedial math showed a significant negative effect on college GPA. Students who did not take remedial math had a higher mean GPA than students who took remedial math courses. After all academic, non-academic, and non-cognitive variables were included, remedial math still had a negative effect on GPA, but it only uniquely accounted for a small proportion of the total variance in GPA.

Given the potential for biased estimates of treatment effects in this non-experimental study, propensity score matching was conducted in an attempt to create statistically equivalent comparison groups. Students were matched based on the similarity of the various characteristics included previously, and then the effects of remedial math-taking were re-analyzed. After matching, it was noted that the groups still showed differences for many of the variables included, but the groups were found to be

more similar on these measures than they were before matching and the effect sizes of the differences were greatly reduced. However, the analyses of the effects of remedial math-taking for the matched groups (both with and without the inclusion of covariates) on degree attainment and GPA did not produce results that differed significantly from those found before conducting propensity score matching.

Due to a number of missing values for the covariates, it was possible that the previous results obtained in this study were influenced by the missing data. A missing value analysis indicated that the data might not have been missing randomly. As such, multiple imputation was used to handle the missing data and 12 complete case data sets were created. A subsequent repeat of the earlier analyses, both with and without matching, on all 12 data sets did not produce any treatment effects that were significantly different from those found before imputing the missing data.

All models indicated that students who did not take remedial math in college were about 1.5 times more likely to attain a degree and had a higher mean GPA than students who took remedial math, even after using propensity score matching and handling missing data using multiple imputation. Much of the variation in degree attainment and GPA was attributable to differences in student characteristics, but controlling for these differences did not completely remove the negative impact of taking remedial math in college.

## CHAPTER V

### DISCUSSION AND RECOMMENDATIONS

The purpose of this nonexperimental study was to investigate preexisting differences between students who did and did not take remedial math in college; the effect of those preexisting differences on remedial math-taking; and the effect of remedial math-taking on degree attainment and college GPA. This study also investigated whether multiple imputation procedures and propensity score matching would improve estimates of treatment effects. The sample in this study consisted of a subsample of 10,736 participants from the ELS:2002 public use data file. Multiple academic, non-academic, and non-cognitive covariates were included in the analyses.

#### **Overview of the Problem**

National statistics indicate that the number of students entering college academically underprepared in mathematics is increasing. The success rates of students in remedial math courses in college is less than 50%. Years of research have been conducted on alternative modes of instruction in an effort to improve the passing rates of remedial math students, yet there has been no large scale increase in the passing rates. Given that redesigning courses has not improved the pass rates, it is plausible that student characteristics may be a factor in student success in remedial education. If remedial math-taking affects college GPA and degree attainment, then by extension, students' characteristics may be predictors of college GPA and degree attainment. This led me to ask, are there preexisting differences in students who do and do not take remedial math? And if so, is it possible to control for the differences when investigating the effects of remedial math-taking on GPA and degree attainment?

The research questions that were addressed in this study were as follows.

1. Does taking remedial math in college have an effect on degree attainment?
2. Does taking remedial math in college have an effect on college GPA?
3. Are there preexisting differences between students who do and do not take remedial math courses in college?
4. Do preexisting differences predict participation in remedial math?
5. Does controlling for preexisting differences between remedial and non-remedial students produce different results in causal inference models?
6. Does propensity score matching improve the ability to control for preexisting differences in remedial and non-remedial students?
7. Do missing data have an impact on estimates of treatment effects?

## **Results**

### **Does taking remedial math in college have an effect on degree attainment?**

Logistic regression analyses showed that remedial math-taking significantly predicted degree attainment. Students who did not take remedial math were almost three times more likely to attain a college degree than students who took remedial math.

**Does taking remedial math in college have an effect on college GPA?** Linear regression analysis showed that remedial math-taking had a significant negative effect on GPA. Approximately 7% of the variance in college GPA was explained by remedial math.

**Are there preexisting differences between students who do and do not take remedial math courses in college?** Independent samples *t*-tests showed significant differences between remedial and non-remedial math students on SES, math IRT, and

reading IRT. Students who did not take remedial math had higher mean SES, math IRT estimated number right, and reading IRT estimated number right than students who took remedial math. Chi-square tests showed significant differences between remedial math groups for most covariate measures including high school (HS) control, HS program concentration, highest HS math class taken, sector of PSI attended, HS attainment indicator, PSI level combination, race, family composition, postsecondary education pipeline completion, working status, whether English was the native language, timing of PSI enrollment, and taking out student loans. Based on these results, it was evident that preexisting differences did exist between remedial math groups.

**Do preexisting differences predict participation in remedial math?** Binary logistic regression with all covariates included resulted in a model that significantly predicted remedial math-taking in college. Preexisting differences did exist and they predicted participation in remedial math in college.

**Does controlling for preexisting differences between remedial and non-remedial students produce different results in causal inference models?** Including all covariates in regression analyses provided significant models predicting both degree attainment and college GPA. However, after controlling for the preexisting differences, results indicated that students who did not take remedial math were 1.5 times more likely to attain a degree, compared to three times more likely when remedial math was the only predictor. In addition, approximately 26% of the variance in GPA was explained by the combined covariates, compared to 7% when remedial math was the only predictor. Including the covariates appeared to reduce the effect of remedial math-taking on degree attainment and college GPA, though it did not remove it as a predictor entirely.

**Does propensity score matching improve the ability to control for preexisting differences in remedial and non-remedial students?** Comparison groups of matched remedial and non-remedial students were formed using propensity score matching in an attempt to further control for the preexisting differences in groups. Statistically significant preexisting differences were still found between the matched groups. However, the effect sizes of SES, math IRT, and reading IRT were all greatly reduced compared to the unmatched sample. Similarly, significant differences were found for a number of the nominal covariates, yet standardized residual measures showed equivalent residuals across groups. In addition, remedial math-taking was still a significant predictor of degree attainment and showed a significant negative effect on college GPA, even after matching participants.

A comparison of the odds ratios before and after matching showed the likelihood that a non-remedial student would attain a degree compared to a remedial student decreased from 2.7 to 1.6. This was interesting because in this case, remedial math was used as the only predictor of degree attainment. Therefore, matching participants seemed to control for the preexisting differences. However, the results after matching were similar to those obtained from the standard regression analyses including all covariates as predictors in the unmatched sample. Therefore, matching did not improve the results for predicting degree attainment.

Similar results were found when comparing the regression analyses predicting GPA; however, there was a decrease (from 7% to 2%) in the total amount of variance in GPA that was explained by remedial math-taking alone. This made sense because, assuming that PSM did its job, students were more similar on covariate measures, thus

the explained variance in GPA by remedial math-taking was reduced, though still significant. Just as in the case of degree attainment, when all covariates were included in the model, remedial math-taking still showed a negative effect on degree attainment and the regression coefficients before and after matching were very similar. As was noted by other authors, it seemed that, although matching helped control for preexisting differences, it did not improve the results over and above those obtained using the standard regression analyses conducted on the unmatched sample using all covariates as predictors (Byun et al., 2015; Melguizo et al., 2011, Padgett et al., 2010).

**Do missing data have an impact on estimates of treatment effects?** To investigate the effect of missing data, multiple imputation procedures were used to create 12 complete case data sets. All previous analyses were repeated on the multiply imputed data sets, including propensity score matching. Although a missing value analysis indicated that missing data might have produced biased estimates of treatment effects, regression analyses pre- and post-matching on the multiply imputed data sets showed similar results to the previous analyses on the original data set.

Overall, propensity score matching and multiple imputation did not improve the ability to control for preexisting differences between remedial and non-remedial students, nor did they result in any noticeably different estimates of treatment effects. All models controlling for preexisting differences indicated that students who do not take remedial math were more likely to attain a degree and have higher GPAs than students who did take remedial math.

## **Limitations**

**Data.** This study was limited by the data used. I utilized the public use data file available from the ELS:2002, thus finding non-academic measures to include in the study was more difficult due to a number of restricted-use variables. Some student-level variables of interest, such as actual high school GPA, were restricted. The independent variable in this study was remedial math-taking. This variable was based on transcript data that reported the numbers of remedial math courses a student took. I did not consider how many of those courses the student actually passed. In addition, as discussed in the literature review, there is a distinction between remedial math courses and developmental math courses. Generally, remedial math focuses mainly on remediating a skill deficiency, whereas developmental education courses often encompass much more. The variable used in this study did not distinguish between these two types of courses. Therefore, it was unknown what other resources and/or instruction these students may have received.

An investigation in this area is warranted because the goal of developmental education is to focus on the whole student and to consider many areas that may need improvement, such as time management, organization, study skills, etc. It may be that the educational outcomes of students in “remedial” math are different from those of students in “developmental” math. Additionally, no information was known about how these students were placed into remediation or whether or not they were required to complete it. Not all institutions use placement tests nor do they all require students to complete remediation. Further investigation is needed into the differences in outcomes



between students who self-select into remedial courses and those who are required by the institution to complete remediation.

**Analysis.** This study was limited by the inability to achieve balance between matched groups of participants. Although I was able to reduce the magnitude of the differences, I was not able to remove them entirely by using propensity score matching. Therefore, I was not able to create statistically equivalent comparison groups, which was the goal of conducting propensity score matching. As PSM relies heavily on selection of adequate covariates, it is plausible that not all relevant covariate measures were included in the analysis; however, there is no way of knowing whether adequate covariates were obtained in advance. Furthermore, before imputation, missing data resulted in a 34% decrease in my sample size during regression analysis including all covariates. In addition, PSM also resulted in a large reduction in my sample size. This reduction was approximately 63% before imputation and 40% after imputation.

The matching process I used for this study included all levels of postsecondary institutions attended. Therefore, during the matching process, no distinction was made between students who attended 2-year colleges and those who attended 4-year institutions. As students are more likely to take their remedial coursework at 2-year colleges, this was a limiting factor in my study. In addition, although the PSI level combination did distinguish between 2- and 4-year institutions, not all 4-year PSIs are the same. For example, a current trend seen in higher education is to transform 2-year colleges into 4-year colleges by having 2-year PSIs offer a limited number of bachelor's degrees. However, these 4-year colleges still function largely as community colleges,

which mainly confer associates degrees. No distinction was made between the aforementioned 4-year colleges and traditional universities.

Multiple imputation and propensity score matching themselves were limiting factors. Both of these techniques were time-consuming. Running a single round of 12 MI data sets took approximately one and a half hours each time. In addition, systematically adding and removing covariate measures to the PSM analysis and repeating regression analyses in an attempt to balance matched groups took weeks. Given that these methods did not provide significantly different estimates of treatment effects, the length of time it took to conduct these analyses is a deterrent to using them again.

Finally, there is no agreement among statisticians on an adequate way to pool the results of multiply imputed data set analyses. Therefore, each regression analysis had to be evaluated separately, which was also time-consuming and did not provide a comparison of results across the 12 multiply imputed data sets.

### **Implications**

The results of this study were consistent with prior research in that it was found that students who take remedial math are less likely to graduate and more likely to have lower academic achievement. However, it was also shown that preexisting differences in students were significant predictors of remedial math-taking in college. In particular, non-academic characteristics such as educational expectations, socioeconomic status, race, and delaying postsecondary enrollment can affect students' likelihood of being in remedial math from the start. As Anyon (2014) points out, "We have been counting on education to solve the problems of unemployment, joblessness, and poverty for many

years. But education did not cause these problems, and education cannot solve them” (p. 5). Despite this, educators and institutions of higher learning have spent years and countless dollars to increase access, enrollment, retention and success rates of academically underprepared students. Yet the student base in higher education does not change. The multitude of factors influencing the success of college students, which are often beyond the control or influence of postsecondary educational institutions, have not been adequately considered or addressed. Instead, what we see is an ever-increasing role of government in shaping practice and policy based on the belief that it is the instructors or the course designs that are the problem.

No Child Left Behind and Race to the Top are examples of the federal government inserting itself in K-12 education and they have redefined what constitutes education as well as its purpose. The result is a shift to high-stakes testing and merit pay (Jackson, 2011). This shift has occurred both in K-12 and in higher education. The outcome is a cookie cutter system that produces a generation of students who know how to take a test and who fear numbers because everything is a grade, everything must be quantified, and if it cannot be quantified it is not worth evaluating. Unfortunately, those student characteristics that cannot be graded (motivation, expectation, etc.) are the very ones that can greatly influence educational outcomes that are quantified, measured, and funded (or defunded). What is needed is a shift in focus.

This study found that a number of non-academic student characteristics predict participation in remedial math in college. Therefore, continually redesigning courses to try to increase passing rates may not be sufficient to effect a change in these numbers. Additionally, remedial math had a negative impact on degree attainment and GPA, but, is

this really all that surprising? If students enter college underprepared, and their characteristics have contributed to their lack of preparation, then it is logical that they might have lower levels of academic performance in college, despite remediating their basic skills gaps. Redesigning courses based on student characteristics may improve their passing rates in remedial math, but that might not improve their overall academic performance in college. Making up for years of skills gaps in one or two semesters does not seem plausible to me. By extension, students' rates of degree completion are a factor of their level of academic preparedness. Successfully negotiating one or more remedial math courses cannot be expected to change years of student development that occurred prior to college enrollment.

It is my belief that higher education needs to accept the fact that not all students belong in college as it currently exists. The higher education system was designed and remains structured for the "traditional" college student. It was not designed for students with low academic preparation, low SES, first-generation students, etc. A student who has had 12 or 13 years to accumulate math deficiencies, among other things, is expected to correct those deficiencies and elevate their academic preparation to the level of a higher performing student; yet remedial courses are often only one or two semesters. It is not surprising then that we, as educators, cannot "fix" the remedial math problem in higher education because low math performance is not the sole contributing factor to lack of success.

### **Recommendations for Future Research**

This study attempted to use propensity score matching because it is purportedly a superior method for creating comparison groups in nonexperimental studies. However,

aside from being very time-consuming, it did not produce any noticeably improved results. Rather than attempting to create equivalent groups of remedial and non-remedial students, future researchers should consider within group comparisons. What makes one group of remedial students different from another group of remedial students? Why do some remedial students complete their degrees when others do not? For example, aside from the fact that low income students are overrepresented in remedial courses, as are minority students, what other characteristics do these students have in common that may impact their success in college? Can any of the non-academic negative predictors of success be influenced through developmental education programs?

As mentioned previously, most students requiring remediation do not complete these courses at traditional universities. Future research in this area should focus on students attending 2-year institutions. As a number of community colleges have been transformed into 4-year colleges, efforts should be made to distinguish between these types of institutions (if they are included in the research). Furthermore, these investigations should gather more detailed information about the types of remedial courses students take. Are the courses strictly for skills-gap remediation? Are the courses more in line with the developmental education concept, which includes instruction in areas such as study skills, time-managements, and the like? A greater effort to make this distinction should be made.

**Recommendations for Florida.** This study used what is deemed a nationally representative sample of students. Florida schools, in particular, have a diverse student population that is not necessarily mirrored in these national data sets. For example, according to a 2016 report based on the Department of Education's Integrated

Postsecondary Education Data System, at Miami Dade College (one of the largest open access colleges in the U.S.) the student population is largely Hispanic (68%) and White students make up only 6% of the population. At Florida International University, the student population is 63% Hispanic and 12% White (Student Diversity, 2016). The population used in my study (considered nationally representative) was 56% White and 15% Hispanic. Caution is warranted when making policy decisions based on studies incorporating supposedly nationally representative samples that may not be representative of populations in specific states or even individual schools within a state. Future research in Florida, specifically, should involve in-state data collection on a large scale, rather than using national data sets.

Furthermore, as was discussed previously, not all students are required to complete remediation, thus not accounting for this in estimating treatment effects of remediation on learning outcomes may bias results. The passage of Senate Bill 1720 in the State of Florida in 2013 made the placement and remediation of entering students all the more difficult for postsecondary institutions. Students who graduate from a public Florida high school (retroactive to 2002), or those who are active military, are declared college-ready and public colleges/universities cannot require these students to take a common placement exam, nor can they require students to enroll in remedial courses. Students who qualify for this exemption may opt to take a common placement test. However, even if the results indicate that a student needs remediation, the college cannot force them to remediate (FS 1008.30). The result is that many students who qualify for this exemption enter directly into gateway courses such as freshman composition and intermediate algebra, whether or not they are prepared.

Future research in Florida should investigate preexisting differences in remedial students and the effects of remediation on degree attainment and GPA, but should be conducted in the context of SB1720. This research should attempt to track those students who chose not to remediate (those who qualified for exemption per SB1720), and compare their outcomes to those of students who did elect to remediate. As self-selection bias is considered a confounding factor in data analysis, this would be an informative investigation in Florida. Similar investigations in other states with diverse student populations that are not accurately represented in national studies should also be considered.

### **Summary**

This study analyzed the effects of remedial math-taking on degree attainment and college GPA in a sample of 10,736 participants from the ELS:2002 data. Four stages of analyses were conducted: (1) preliminary analysis of preexisting differences in remedial math groups and the effects of remedial math-taking on degree attainment and college GPA; (2) propensity score matching on the original sample followed by a repeat of the previous analyses; (3) a missing value analysis and multiple imputation of missing data, followed by a repeat of the analyses done in stage one; and (4) propensity score matching on the multiply imputed data sets, followed by analyses and a comparison of models from all four stages. Remedial math groups were found to have significant preexisting differences on many covariates included in the study and these differences significantly predicted participation in remedial math.

Logistic regression analyses, both with and without the inclusion of covariates, indicated that remedial math-taking in college significantly predicted degree attainment.

After controlling for preexisting differences, analyses indicated that students who did not take remedial math were approximately 1.5 times more likely to attain a degree. Furthermore, linear regression analyses, both with and without inclusion of the covariates, indicated that remedial math-taking had a significant negative effect on college GPA. Students who did not take remedial math had higher mean GPAs than students who took remedial math. Moreover, estimates of treatment effects after multiple imputation and propensity score matching were similar to the estimates obtained from the analyses on the original sample. Neither multiple imputation procedures nor propensity score matching resulted in an improvement in the ability to control for significant preexisting differences between remedial math groups.



## REFERENCES

- Alliance for Excellent Education. (2006). *Paying double: Inadequate high schools and community college remediation*. Issue Brief. Washington, DC.
- Anyon, J. (2014). *Radical possibilities. Public policy, urban education, and a new social movement* (2nd ed.). New York, NY: Routledge.
- Attewell, P., Lavin, D., Domina, T., & Levey, T. (2006). New evidence on college remediation. *Journal of Higher Education*, 77, 886-924.
- Austin, G. A. (2010). Administrative challenges and rewards of online learning in a rural community college: Reflections of a distance learning administrator. *New Directions for Community Colleges*, 2010(150), 27-36. doi:10.1002/cc.402
- Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research*, 46, 399-424. doi:10.1080/00273171.2011.568786
- Aycaster, P. W. (2001). Factors impacting success in community college developmental mathematics courses and subsequent courses. *Community College Journal of Research and Practice*, 25, 406-416.
- Bahr, P. R. (2007). Double jeopardy: Testing the effects of multiple basic skill deficiencies on successful remediation. *Research in Higher Education*, 48, 695-725. doi:10.1007/s11162-006-9047-y
- Bahr, P. (2008). Does mathematics remediation work?: A comparative analysis of academic attainment among community college students. *Research in Higher Education*, 49, 420-450. doi: 10.1007/s11162-008-9089-4
- Bahr, P. R. (2010). Preparing the underprepared: An analysis of racial disparities in postsecondary mathematics remediation. *The Journal of Higher Education*, 81, 209-237.
- Bailey, T. (2009). Challenge and opportunity: Rethinking the role and function of developmental education in community college. *New Directions for Community Colleges*, 2009(145), 11-30. doi:10.1002/cc.352
- Bassett, M. J., & Frost, B. (2010). Smart math: Removing roadblocks to college success. *Community College Journal of Research & Practice*, 34, 869-873. doi:10.1080/10668926.2010.509232

- Bettinger, E., & Long, B. T. (2005). Remediation at the community college: Student participation and outcomes. *New Directions for Community Colleges*, 2005(129), 17-26.
- Bonham, B. S., & Boylan, H. R. (2012). Developmental mathematics: Challenges, promising practices, and recent initiatives. *Journal of Developmental Education*, 36(2), 14-21.
- Boylan, H. R., Bonham, B. S., & White, S. R. (1999). Developmental and remedial education in postsecondary education. *New Directions for Higher Education*, 1999(108), 87-101.
- Bozick, R., Lauff, E., & Wirt, J. (2007). Education longitudinal study of 2002 (ELS: 2002): A first look at the initial postsecondary experiences of the high school sophomore class of 2002. *National Center for Education Statistics*.
- Brittenham, R., Cook, R., Hall, J. B., Whitesell, P. M., Ruhl-Smith, C., Shafii-Mousavi, M., Showater, J., Smith, K., & White, K. (2003). Connections: An integrated community of learners. *Journal of Developmental Education*, 27(1), 18-27.
- Byun, S. Y., Irvin, M. J., & Bell, B. A. (2015). Advanced math course taking: effects on math achievement and college enrollment. *The Journal of Experimental Education*, 83, 439-468.
- Carlin, J. B., Li, N., Greenwood, P., & Coffey, C. (2003). Tools for analyzing multiple imputed datasets. *The Stata Journal*, 3(3), 226-244.
- Casazza, M. E. (1999). Who are we and where did we come from? *Journal of Developmental Education*, 23(1), 2.
- Chickering, A. W., & Gamson, Z. F. (1987). Seven principles for good practice in undergraduate education. *AAHE Bulletin*, 39(7), 3-7.
- Chen, X & Simone S. (2016). Remedial Coursetaking at US Public 2-and 4-Year Institutions: Scope, Experiences, and Outcomes. Statistical Analysis Report. NCES 2016-405. *National Center for Education Statistics*.
- Christopher, Elise, M. (2015). Education Longitudinal Study of 2002 (ELS: 2002) Postsecondary Education Transcript Study: Addendum to the Public-Use File. *National Center for Education Statistics, Institute of Education Sciences, US Department of Education*. Washington, DC. Retrieved from [https://nces.ed.gov/surveys/els2002/pdf/ELSPETS-PUF-MAY\\_Final.pdf](https://nces.ed.gov/surveys/els2002/pdf/ELSPETS-PUF-MAY_Final.pdf)

- Clark, M., & Cundiff, N. (2011). Assessing the effectiveness of a college freshman seminar using propensity score adjustments. *Research in Higher Education, 52*, 616-639. doi:10.1007/s11162-010-9208-x
- Donovan, W. J., & Wheland, E. R. (2008). Placement tools for developmental mathematics and intermediate algebra. *Journal of Developmental Education, 32*(2), 2-11.
- Fan, X., & Nowell, D. L. (2011). Using propensity score matching in educational research. *Gifted Child Quarterly, 55*, 74-79. doi:10.1177/0016986210390635
- Fike, D. S., & Fike, R. (2012). The consequences of delayed enrollment in developmental mathematics. *Journal of Developmental Education, 35*(3), 2-10.
- Fla. Stat. § 1008.02 (2013). Retrieved from [http://www.leg.state.fl.us/statutes/index.cfm?App\\_mode=Display\\_Statute&URL=1000-1099/1008/1008.html](http://www.leg.state.fl.us/statutes/index.cfm?App_mode=Display_Statute&URL=1000-1099/1008/1008.html)
- Giani, M., Alexander, C., & Reyes, P. (2014). Exploring variation in the impact of dual-credit coursework on postsecondary outcomes: A quasi-experimental analysis of Texas students. *The High School Journal, 97*(4), 200-218.
- Guo, S., & Fraser, M. W. (2010). *Propensity Score Analysis: Statistical Methods and Applications* (Vol. 12). Los Angeles: CA. SAGE Publications.
- Hall, J., & Ponton, M. K. (2005). Mathematics self-efficacy of college freshman. *Journal of Developmental Education, 28*(3), 26-33.
- Hammerman, N., & Goldberg, R. (2003). Strategies for developmental mathematics at the college level. *Mathematics and Computer Education, 37*, 79-95.
- Harel, O. (2007). Inferences on missing information under multiple imputation and two-stage multiple imputation. *Statistical Methodology, 4*(1), 75-89.
- Harel, O., & Zhou, X. H. (2007). Multiple imputation: review of theory, implementation and software. *Statistics in medicine, 26*(16), 3057-3077.
- Henderson, J., & Chatfield, S. (2011). Who matches? Propensity scores and bias in the causal effects of education on participation. *Journal of Politics, 73*, 646-658. doi:10.1017/S0022381611000351
- Higbee, J. L., & Thomas, P. V. (1999). Affective and cognitive factors related to mathematics achievement. *Journal of Developmental Education, 23*(1), 8.

- Hill, J. (2004). *Reducing bias in treatment effect estimation in observational studies suffering from missing data* (Working Paper 04-01). Columbia University Academic Commons, <https://doi.org/10.7916/D8B85G11>.
- Hill, J., Weiss, C., & Zhai, F. (2011). Challenges with propensity score strategies in a high-dimensional setting and a potential alternative. *Multivariate Behavioral Research*, *46*(3), 477-513.
- Hinkle, D. E., Wiersma, W., & Jurs, S. G. (2003). *Applied statistics for the behavioral sciences* (5<sup>th</sup> ed.). Boston, MA: Houghton Mifflin Company.
- Holmes, W. M. (2014). *Using propensity scores in quasi-experimental designs*. Los Angeles, CA: SAGE Publications.
- Illich, P. A., Hagan, C., & McCallister, L. (2004). Performance in college-level courses among students concurrently enrolled in remedial courses: Policy implications. *Community College Journal of Research & Practice*, *28*, 435-453. doi: 10.1080/10668920490444463
- Jackson, S. (2011). Commentary on the rhetoric of reform. A twenty-year retrospective. In Saltman, K. J. & Gabbard, D. A. (Eds.), *Education as enforcement. The militarization and corporatization of schools* (2nd ed., pp. 247-260). New York, NY: Routledge.
- Jacobson, E. (2006). Computer homework effectiveness in developmental mathematics. *Journal of Developmental Education*, *29*(3), 2-8.
- King, G., & Nielsen, R. (2016). *Why propensity scores should not be used for matching*. Working paper. Retrieved from <https://gking.harvard.edu/files/gking/files/psnot.pdf>.
- Kinney, D. P., & Robertson, D. F. (2003). Technology makes possible new models for delivering developmental mathematics instruction. *Mathematics and Computer Education*, *37*, 315-328.
- Luellen, J. K., Shadish, W. R., & Clark, M. H. (2005). Propensity scores: An introduction and experimental test. *Evaluation Review*, *29*, 530-558. doi:10.1177/0193841X05275596
- Manly, C. A., & Wells, R. S. (2015). Reporting the use of multiple imputation for missing data in higher education research. *Research in Higher Education*, *56*(4), 397-409.

- McFarland, J., Hussar, B., de Brey, C., Snyder, T., Wang, X., Wilkinson-Flicker, S., ... & Bullock Mann, F. (2017). The Condition of Education 2017. NCES 2017-144. *National Center for Education Statistics*. Retrieved from <https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2017144>
- Melguizo, T., Kienzl, G. S., & Alfonso, M. (2011). Comparing the educational attainment of community college transfer students and four-year college rising juniors using propensity score matching methods. *Journal of Higher Education*, 82, 265-291.
- Meyers, L. S., Gamst, G., & Guarino, A. J. (2013). *Applied multivariate research: Design and interpretation* (2<sup>nd</sup> ed.). Los Angeles, CA: SAGE Publications.
- Miles, A. (2016). Obtaining predictions from models fit to multiply imputed data. *Sociological Methods & Research*, 45(1), 175-185.
- Mitra, R., & Reiter, J. P. (2016). A comparison of two methods of estimating propensity scores after multiple imputation. *Statistical methods in medical research*, 25(1), 188-204.
- Mood, C. (2010). Logistic regression: Why we cannot do what we think we can do, and what we can do about it. *European sociological review*, 26(1), 67-82.
- National Center for Education Statistics [NCES] (n.d.). *About Us*. Washington, DC: U.S. Department of Education. Retrieved from <https://nces.ed.gov/about/>
- National Center for Education Statistics [NCES] (n.d.). *Educational Longitudinal Study of 2002 (ELS:2002). Historical Background*. Washington, DC: U.S. Department of Education. Retrieved from <https://nces.ed.gov/surveys/els2002/overview2.asp>
- National Center for Education Statistics [NCES] (n.d.). *Educational Longitudinal Study of 2002 (ELS:2002). Policy and Research Issues: Introduction*. Washington, DC: U.S. Department of Education. Retrieved from <https://nces.ed.gov/surveys/els2002/policy.asp>
- National Center for Education Statistics [NCES] (n.d.). *Educational Longitudinal Study of 2002 (ELS:2002). Policy and Research Issues: Post-High School Transition*. Washington, DC: U.S. Department of Education. Retrieved from <https://nces.ed.gov/surveys/els2002/policy2.asp>
- National Center for Education Statistics [NCES] (n.d.). *Educational Longitudinal Study of 2002 (ELS:2002). Purpose*. Washington, DC: U.S. Department of Education. Retrieved from <https://nces.ed.gov/surveys/els2002/index.asp>

- National Center for Education Statistics [NCES] (n.d.). *National Educational Longitudinal Study of 1988 (NELS:88). Overview*. Washington, DC: U.S. Department of Education. Retrieved from <https://nces.ed.gov/surveys/nels88/>
- National Center for Education Statistics [NCES] (2002). *The condition of education 2002*. Washington, DC: U.S. Department of Education. Retrieved from <http://nces.ed.gov/pubs2002/2002025.pdf>
- National Center for Education Statistics [NCES] (2004). *The condition of education 2004*. Washington, DC: U.S. Department of Education. Retrieved from <https://nces.ed.gov/pubs2004/2004077.pdf>
- National Center for Education Statistics [NCES] (2010). *The digest of education statistics 2010*. Washington, DC: U.S. Department of Education. Retrieved from [http://nces.ed.gov/programs/digest/d10/tables/dt10\\_340.asp](http://nces.ed.gov/programs/digest/d10/tables/dt10_340.asp)
- National Center for Education Statistics [NCES] (2011). *The condition of education 2011*. Washington, DC: U.S. Department of Education. Retrieved from <https://nces.ed.gov/pubs2011/2011033.pdf>
- National Governors Association (2010). *Return on investment: Strategies for improving remedial education*. Washington, DC. Retrieved from <http://www.nga.org/files/live/sites/NGA/files/pdf/C2CBriefingPaperRemedialEd.pdf>
- Obama, B. (2009). Remarks of President Barack Obama—Address to joint session of Congress, February 24, 2009. Retrieved from <https://obamawhitehouse.archives.gov/the-press-office/remarks-president-barack-obama-address-joint-session-congress>
- O'Connor, J. (2013, October 28). Why Remedial Classes Are No Longer Required at Florida Colleges. *StateImpact Florida*. Retrieved from <https://stateimpact.npr.org/florida/2013/10/28/why-remedial-classes-are-no-longer-required-at-florida-colleges/>
- Office of Program Policy Analysis & Government Accountability (2007). Half of college students needing remediation drop out; Remediation completers do almost as well as other students. Tallahassee, FL: OPPAGA Report No. 07-31.
- Padgett, R. D., Salisbury, M. H., An, B. P., & Pascarella, E. T. (2010). Required, practical, or unnecessary? An examination and demonstration of propensity score matching using longitudinal secondary data. *New Directions for Institutional Research*, 2010, 29-42. doi:10.1002/ir.370

- Parsad, B., & Lewis, L. (2003). *Remedial education at degree-granting postsecondary institutions in fall 2000* (NCES Publication No. 2004-010). Washington, DC: National Center for Educational Statistics, U.S. Department of Education.
- Peikes, D. N., Moreno, L., & Orzol, S. M. (2008). Propensity score matching: A note of caution for evaluators of social programs. *The American Statistician*, *62*(3), 222-231.
- Penny, M. D., & White, J. G. (1998). Developmental mathematics students' performance: Impact of faculty and student characteristics. *Journal of Developmental Education*, *22*(2), 2.
- Pretlow III, J., & Washington, H.D. (2011), Cost of developmental education: An update to Breneman and Harlow. *Journal of Developmental Education*, *35*(1), 2-12.
- Reist, B. M., & Larsen, M. D. (2012). Post-Imputation Calibration Under Rubin's Multiple Imputation Variance Estimator. In *Section on Survey Research Methods, Joint Statistical Meeting*.
- Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41-55.
- Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*, *39*(1), 33-38.
- Saxon, D., & Boylan, H. R. (2001). The cost of remedial education in higher education. *Journal of Developmental Education*, *25*(2), 2.
- Shadish, W. R., & Steiner, P. M. (2010). A primer on propensity score analysis. *Newborn and Infant Nursing Reviews*, *10*(1), 19-26.
- Sheldon, C. Q., & Durdella, N. R. (2010). Success rates for students taking compressed and regular length developmental courses in the community college. *Community College Journal of Research & Practice*, *34*, 39-54. doi: 10.1080/10668920903385806
- Spradlin, K., & Ackerman, B. (2010). The effectiveness of computer-assisted instruction in developmental mathematics. *Journal of Developmental Education*, *34*(2), 12-42.
- Student Diversity at More Than 4,600 Institutions (2016). *The Chronicle of Higher Education*. Retrieved from <https://www.chronicle.com/interactives/student-diversity-2016>

- Titus, M. A. (2007). Detecting selection bias, using propensity score matching, and estimating treatment effects: An application to the private returns to a master's degree. *Research in Higher Education*, 48(4), 487-521.
- Vaughan, A., Lalonde, T., & Jenkins-Guarnieri, M. (2014). Assessing student achievement in large-scale educational programs using hierarchical propensity scores. *Research in Higher Education*, 55, 564-580. doi:10.1007/s11162-014-9329-8
- White, I. R., Royston, P., & Wood, A. M. (2011). Multiple imputation using chained equations: issues and guidance for practice. *Statistics in medicine*, 30(4), 377-399.
- Xia, R. (2017, August 3). Cal State will no longer require placement exams and remedial classes for freshmen. *Los Angeles Times*. Retrieved from <http://www.latimes.com/local/lanow/la-me-cal-state-remedial-requirements-20170803-story.html>
- Yuan, Y. C. (2010). Multiple imputation for missing data: Concepts and new development (Version 9.0). *SAS Institute Inc., Rockville, MD*, 49, 1-11.
- Zavarella, C. A., & Ignash, J. M. (2009). Instructional delivery in developmental mathematics: Impact on retention. *Journal of Developmental Education*, 32(3), 2-13.
- Zinshteyn, M. (2016 March 9). Getting rid of placement exams. *The Atlantic*. Retrieved from <https://www.theatlantic.com/education/archive/2016/03/the-problem-with-college-placement-exams/472944/>



## APPENDICIES

## Appendix A

### *Description of Variables Used in the Study*

Variable	Description
F3TZGPAALL “College GPA”	Transcript: GPA at all known postsecondary institutions attended
F3TZANYDEGRE “Degree Attainment”	Transcript: Any known degree attained as of June 2013 0 = No, 1 = Yes
F3TZREMMTTOT “Remedial Math”	Total number of known remedial mathematics courses taken by the student during their undergraduate education, as of June 2013 (range 0 to 5)  rem_math (recoded): Whether student took remedial math in college (0 = No, 1 = Yes)
BYSCTRL	High school control: Public (reference category), Catholic, Other private  Recoded as dichotomous using dummy variables (0 = No, 1 = Yes)
BYURBAN	High school urbanicity: Urban, suburban, rural (reference category)  Recoded as dichotomous using dummy variables (0 = No, 1 = Yes)
BYSCHPRG	Composite: High school program reported by student: General, College preparatory-academic, Vocational-including technical/business (reference category)  Recoded as dichotomous using dummy variables (0 = No, 1 = Yes)
F1HIMATH	Highest math course taken of a half year or more: No math course or math course is other, Pre-algebra, general or consumer math, Algebra I, Geometry, Algebra II, Trigonometry, pre-calculus, or calculus  HighMath (recoded): 0 = Basic math (Algebra II or lower), 1 = Advanced math (Trigonometry, pre-calculus, or calculus)
F3TZPS1SEC	Indicated sector (by level) of first attended postsecondary institution: 4-year public, 4-year private, not-for-profit, 4-year private, for-profit, 2-year public, 2-year private, for-profit, Less than 2-year private, for-profit, Other  PSI_1stSector (recoded): Dichotomous (0 = Public, 1 = Private)

Variable	Description
F2HSATTM	High school attainment indicator (academic risk): Successful graduate (hs diploma recipient), Marginal graduate (hs diploma recipient), Completer (GED/equivalency/cert of att), Non-completer  Recoded dichotomous (0 = No, 1 = Yes). Completer/Non-completer used as reference category
F3TZSCHCOMBO	Indicated the postsecondary institution level change from the first known postsecondary institution to second known postsecondary institution: 4-year only 2-year only 2-year, then 4-year 4-year, then 2-year (reference category) Less-than-2-year only (omitted) Less-than-2-year, then 2-year or 4-year (omitted) 2-year, then Less-than-2-year (omitted) 4-year, then Less-than-2-year (omitted)  Recoded as dichotomous using dummy variables (0 = No, 1 = Yes)
F1RACE	Composite: F1 student's race/ethnicity  Recoded into dichotomous using dummy variables (0 = No, 1 = Yes): White, Black, Hispanic, Asian, Other)
F1FCOMP	Family composition: Mother and father, Mother and male guardian, Father and female guardian, Two guardians, Mother only, Father only, Female guardian only, Male guardian only, Lives with student less than half time  OneParentHouse (Recoded): 1 = Single-Parent Household (One parent or guardian) 0 = Other household composition
F2PSPPLN	Composite: Postsecondary education pipeline. Still in high school as of F2 interview, never entered (no PS attendance), partially completed, completed/no PS attendance, completed/1 <sup>st</sup> attend <4yr, completed/1 <sup>st</sup> attend 4yr  Recoded (PipeCompl) to dichotomous variable: 0 = No (Partial or none), 1 = Yes (completed)
F1WRKHRS	F1 hours worked per week during 03-04 school year. Ordinal variable.  Recoded (WorkF1) to dichotomous variable: 0 = No (did not work), 1 = Yes (worked)

Variable	Description
F1SEX	F1 sex-composite. Recoded 0 = Male, 1 = Female
F1STLANG	Composite: F1 whether English is student's native language
F3PSTIMING	Timing of first postsecondary enrollment: no enrollment (reference category), delayed, immediate.
F3STLOANEVR	Whether respondent took out any student/PSE loans. 0 = No, 1 = Yes
BYS33L	Indicated whether the student had ever been in a special high school program to help them plan or prepare for college as self-reported on the base year student questionnaire. 0 = No, 1 = Yes
BYSES2	Socio-economic status composite, v.2 (SES): Continuous variable comparable to the NLS-72/HS&B/NELS:88 composite variable and was constructed from parent questionnaire data when available and student substitutions when not. SES was based on five equally weighted, standardized composite variables: father's/guardian's education, mother's/guardian's education, family income, father's/guardian's occupation, and mother's/guardian's occupation.
BYTXMIRR	Math IRT (Item-response theory) estimated number right Continuous variable. An estimate of the number of items students would have answered correctly had they responded to all 72 items in the ELS:2002 math item pool. The ability estimates and item parameters derived from the IRT calibration can be used to calculate each student's probability of a correct answer for each of the items in the pool. These probabilities are summed to produce the IRT estimated number right score. The score is not an integer because it is a sum of probabilities, not a count of right and wrong answers.
BYTXRIRR	Reading IRT (Item-response theory) estimated number right. Continuous variable. An estimate of the number of items students would have answered correctly had they responded to all 51 items in the ELS:2002 reading item pool. The ability estimates and item parameters derived from the IRT calibration can be used to calculate each student's probability of a correct answer for each of the items in the pool. These probabilities are summed to produce the IRT-estimated number right score. The score is not an integer because it is a sum of probabilities, not a count of right and wrong answers.

## Appendix B

### *Descriptive Statistics/Frequencies for All Variables in Original Sample*

Categorical Variables	Frequency	Percent
Remedial Math (independent)		
No	7227	67.3
Yes	3509	32.7
Total	10736	100.0
Degree Attainment (dependent)		
No	4017	37.4
Yes	6719	62.6
Total	10736	100.0
HS Control		
Public (reference)	7910	73.7
Catholic	1676	15.6
Other Private	1150	10.7
Total	10736	100.0
HS Urbanicity		
Urban	3721	34.7
Suburban	5169	48.1
Rural (reference)	1846	17.2
Total	10736	100.0
HS Program		
General HS Program	3236	30.1
College Prep HS Program	6569	61.2
Vocational (reference)	808	7.5
Missing cases	123	1.1
Total	10736	100.0
Highest HS math level		
Basic	4393	40.9
Advanced	5798	54.0
Missing cases	545	5.1
Total	10736	100.0
Sector of 1st PSI attended		
Public	8165	76.1
Private	2568	23.9
Missing cases	3	0.0
Total	10736	100.0
HS Attainment indicator		
Successful graduate	5257	49.0
Marginal graduate	4172	38.9
Other (reference)	338	3.2

Categorical Variables	Frequency	Percent
Missing	969	9.0
Total	10736	100.0
PSI level combo		
4-year only	5153	48.0
2-year only	2623	24.4
2-year, then 4-year	1457	13.6
4-year, then 2-year (reference)	1503	14.0
Total	10736	100.0
Race		
White (reference)	6399	59.6
Black	1250	11.6
Hispanic	1331	12.4
Asian	1214	11.3
Other	542	5.1
Total	10736	100.0
Single parent/guardian household		
No	8605	80.2
Yes	2131	19.8
Total	10736	100.0
Completed PSE pipeline		
No	4036	37.6
Yes	6071	56.5
Missing	629	5.9
Total	10736	100.0
Worked during 03-04 school year		
No	2754	25.7
Yes	7368	68.6
Missing	614	5.7
Total	10736	100.0
Sex		
Male	5026	46.8
Female	5710	53.2
Total	10736	100.0
English is student's native language		
No	1751	16.3
Yes	8985	83.7
Total	10736	100.0
Timing of first PSE		
Delayed	1660	15.5
Immediate	7772	72.4

Categorical Variables	Frequency	Percent				
Missing cases	1304	12.1				
Total	10736	100.0				
Took out any student loans						
No	3809	35.5				
Yes	6008	56.0				
Missing cases	919	8.6				
Total	10736	100.0				
Ever in program to help plan for college						
No	7430	69.2				
Yes	2218	20.7				
Missing cases	1088	10.1				
Total	10736	100.0				
Continuous Variables	N	Missing	Mean	SD	Min.	Max.
College GPA (dependent)	10736	0	2.72	.87	.00	4.00
SES	10210	526	.20	.74	-2.11	1.98
Math IRT	10613	123	41.18	11.10	12.52	69.72
Reading IRT	10613	123	32.39	9.04	10.46	49.09

## Appendix C

### *Chi-Square Tests of Differences Between Remedial Math Groups on Nominal Covariates in the Original Sample*

Variable	$\chi^2$ <sup>a</sup>	Sig. <sup>b</sup>
Public HS	161.10	***
Catholic HS	40.88	***
Other private HS	111.72	***
Urban	40.22	
Suburban	11.09	
Rural	8.92	
General	121.46	***
College Prep	168.69	***
Vocational	21.64	***
Highest HS math level	1151.56	***
Sector 1st PSI	451.10	***
Successful Grad	1126.50	***
Marginal Grad	990.08	***
Attended 4-yr PSI	798.72	***
Attended 2-yr PSI	534.24	***
Attended 2-yr, then 4-yr PSI	134.15	***
Attended 4-yr, then 2-yr	.41	
White	138.32	***
Black	123.8	***
Hispanic	149.7	***
Asian	40.37	***
Other	0.69	
Single parent/guardian household	57.11	***
Completed PSE pipeline	1174.62	***
Worked 03-04 school year	6.8	**
Sex	3.55	
English is native language	7.66	**
Timing of 1 <sup>st</sup> PSI enrollment	168.97	***
Took out student loans	10.3	**
Program to help plan for college	0.03	

*Note.* 0 cells have expected count less than 5.

a df = 1

b Asymptotic significance (2-sided)

\* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$



Appendix D

*Logistic Regression: Covariates Predicting Remedial Math Original Sample (N = 7,109)*

Variable	Coeff.	S.E.	Sig.	Odds Ratio
HS control is Catholic	.071	.093		1.074
HS control is other private	.368	.118	**	1.445
HS is urban	-.116	.095		.890
HS is suburban	-.070	.083		.932
General HS Program	-.196	.116		.822
College Prep HS Program	-.120	.114		.887
Highest HS math level	.703	.069	***	2.021
Sector of 1st PSI attended	.907	.086	***	2.476
Successful HS graduate	.342	.180		1.408
Marginal HS graduate	-.061	.170		.941
Attended only 4-yr PSI	.629	.087	***	1.876
Attended only 2-yr PSI	.408	.108	***	1.504
Attended 2-yr, then 4-yr PSI	.086	.106		1.090
Black	.076	.102		1.078
Hispanic	-.313	.104	**	.731
Asian	.101	.128		1.106
Other	.076	.140		1.079
Lives in single parent/guardian household	-.050	.076		.951
Completed PSE pipeline	.487	.081	***	1.628
Worked during 03-04 school year	-.067	.070		.936
Sex	-.028	.062		.972
English is student's native language	-.164	.106		.849
Timing of 1 <sup>st</sup> PSE	-.196	.085	*	.822
Took out student loans	-.124	.065		.883
Ever in program to help plan for college	-.096	.071		.908
SES	-.113	.048	*	.893
Math IRT	-.052	.004	***	.949
Reading IRT	.009	.005		1.009
Constant	-.931	.475		.394
* $p < .05$ , ** $p < .01$ , *** $p < .001$				
Pseudo $R^2$				.373

## Appendix E

### *Logistic Regression Model 2: Covariates Predicting Degree Attainment in Original Sample (N = 7,109)*

Variable	Coeff.	S.E.	Sig.	Odds Ratio
Whether student took remedial math in college	.381	.068	***	1.463
HS control is Catholic	-.258	.092	**	.773
HS control is other private	-.266	.110	*	.766
HS is urban	.089	.095		1.093
HS is suburban	-.031	.083		.970
General HS Program	.028	.118		1.029
College Prep HS Program	.004	.115		1.004
Highest HS math level	-.412	.071	***	.662
Sector of 1st PSI attended	-.193	.078	*	.824
Successful HS graduate	-1.343	.215	***	.261
Marginal HS graduate	-.695	.208	**	.499
Attended only 4-yr PSI	-.420	.085	***	.657
Attended only 2-yr PSI	.388	.106	***	1.474
Attended 2-yr, then 4-yr PSI	-.775	.112	***	.461
Black	.380	.102	***	1.462
Hispanic	.114	.107		1.121
Asian	-.239	.128		.787
Other	.423	.134	**	1.526
Single parent/guardian household	.065	.076		1.067
Completed PSE pipeline	-.378	.082	***	.685
Worked during 03-04 school year	-.051	.069		.950
Sex	-.446	.061	***	.640
English is student's native language	.109	.107		1.115
Timing of first PSE	-.792	.086	***	.453
Took out student loans	-.500	.064	***	.606
Ever in program to help plan for college	.064	.071		1.066
SES	.267	.048	***	1.306
Math IRT	.005	.004		1.005
Reading IRT	-.017	.005	**	.983
Constant	3.082	.489	***	21.807
<i>*p &lt; .05, **p &lt; .01, ***p &lt; .001</i>				
Pseudo R <sup>2</sup>				.336

Appendix F

*Summary of Linear Regression: All Covariates Predicting College GPA (N = 7,109)*

Variable						Correlations			VIF
	Coeff.	S.E.	$\beta$	$t$	Sig.	Pearson $r$	Partial	Part	
Constant	1.982	.114		17.32	***				
Whether student took remedial math in college	-.127	.021	-.07	-5.99	***	-.261	-.071	-.061	1.312
HS Control (public as reference)									
HS control is Catholic	-.019	.026	-.01	-.73		.045	-.009	-.007	1.286
HS control is other private	.080	.031	.03	2.59	*	.091	.031	.026	1.191
HS Urbanicity (rural as reference)									
HS is urban	-.035	.028	-.02	-1.24		-.035	-.015	-.013	2.376
HS is suburban	.008	.025	.01	.33		.033	.004	.003	2.063
HS Program (vocational as reference)									
General HS Program	.022	.036	.01	.62		-.110	.007	.006	3.674
College Prep HS Program	.060	.035	.03	1.69		.158	.020	.017	3.914
Highest HS math level	.144	.022	.09	6.54	***	.308	.077	.067	1.598
Sector of 1st PSI attended	.162	.022	.08	7.40	***	.151	.088	.076	1.194
HS Attainment indicator (other as reference)									
Successful HS graduate	.370	.061	.22	6.10	***	.403	.072	.063	12.022
Marginal HS graduate	.071	.058	.04	1.22		-.376	.014	.012	10.756
PSI level combo (4-yr, then 2-year as reference)									
Attended only 4-yr PSI	.143	.026	.09	5.56	***	.225	.066	.057	2.220
Attended only 2-yr PSI	-.008	.034	.00	-.24		-.270	-.003	-.002	2.624
Attended 2-yr, then 4-yr PSI	.258	.033	.11	7.73	***	.022	.091	.079	1.781
Race (white as reference)									
Black	-.357	.032	-.13	-11.33	***	-.210	-.133	-.116	1.291
Hispanic	-.072	.032	-.03	-2.25	*	-.086	-.027	-.023	1.435

Variable						Correlations			VIF
	Coeff.	S.E.	$\beta$	$t$	Sig.	Pearson $r$	Partial	Part	
Asian	-.030	.036	-.01	-.84		.037	-.010	-.009	1.628
Other	-.169	.041	-.04	-4.09	***	-.036	-.049	-.042	1.048
Lives in single parent/guardian household	.010	.023	.01	.43		-.082	.005	.004	1.083
Completed PSE pipeline	.058	.026	.03	2.20	*	.326	.026	.023	2.198
Whether worked during 03-04 school year	-.021	.020	-.01	-1.08		-.015	-.013	-.011	1.042
F1 sex-composite	.219	.018	.13	12.21	***	.122	.144	.125	1.073
English is student's native language	-.028	.032	-.01	-.88		.030	-.010	-.009	1.730
Timing of first postsecondary enrollment	-.024	.028	-.01	-.86		.169	-.010	-.009	1.299
Whether R took out any student/PSE loans	.049	.019	.03	2.60	**	.077	.031	.027	1.124
Ever in program to help prepare for college	-.035	.021	-.02	-1.70		-.026	-.020	-.017	1.029
SES	.035	.014	.03	2.56	*	.212	.030	.026	1.424
Math IRT estimated number right	.007	.001	.10	5.54	***	.364	.066	.057	2.843
Reading IRT estimated number right	.000	.001	.00	-.22		.328	-.003	-.002	2.416

\* $p < .05$  (2-tail), \*\* $p < .01$ , \*\*\* $p < .001$

$R^2$  .257

Note. VIF = Variance inflation factor

## Appendix G

### *Chi-Square Tests Between Remedial Math Groups and Nominal Covariates Before and After Matching*

Variable	Before matching		After matching	
	$\chi^{2a}$	Sig. <sup>b</sup>	$\chi^{2a}$	Sig. <sup>b</sup>
Public HS	161.10	***	169.24	***
Catholic HS	40.88	***	99.61	***
Other private HS	111.72	***	60.01	***
Urban	40.22		40.22	***
Suburban	11.09		11.09	**
Rural	8.92		8.92	**
General	121.46	***	1.35	
College Prep	168.69	***	2.07	
Vocational	21.64	***	.30	
Highest math	1151.56	***	48.37	***
Sector 1st PSI	451.10	***	56.25	***
Successful Grad	1126.50	***	51.59	***
Marginal Grad	990.08	***	46.31	***
Attended 4-yr PSI	798.72	***	38.89	***
Attended 2-yr PSI	534.24	***	11.28	**
Attended 2-yr, then 4-yr PSI	134.15	***	10.92	**
Attended 4-yr, then 2-yr	.41		.32	
White	138.32	***	25.63	***
Black	123.8	***	6.55	*
Hispanic	149.7	***	25.72	***
Asian	40.37	***	.03	
Other	0.69		.20	
Lives with single parent	57.11	***	1.79	
Completed PSE pipeline	1174.62	***	39.56	***
Worked 03-04 school year	6.8	**	2.57	
Sex	3.55		.08	
English is native language	7.66	**	7.35	**
Timing PSI enrollment	168.97	***	.34	
Took out student loans	10.3	**	7.91	**
Program help plan	0.03		.01	

*Note.* 0 cells have expected count less than 5.

a df = 1

b Asymptotic significance (2-sided)

\* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$

## Appendix H

### *Standardized Residuals of Dichotomous Variables Before and After Matching for Original Sample*

Variable		Stdzd. residual	Remedial Math	
			No	Yes
HS control is public	No	Before	6.2	-8.9
		After	-8.6	8.6
	Yes	Before	-3.7	5.3
		After	3.3	-3.3
HS control is Catholic	No	Before	-1.4	2.1
		After	2.1	-2.1
	Yes	Before	3.4	-4.8
		After	-6.7	6.7
HS control is other private	No	Before	-2.0	2.8
		After	1.1	-1.1
	Yes	Before	5.7	-8.2
		After	-5.4	5.4
HS Urban	No	Before	-.4	.6
		After	2.4	-2.4
	Yes	Before	.5	-.8
		After	-3.8	3.8
HS Suburban	No	Before	.1	-.1
		After	-1.7	1.7
	Yes	Before	-.1	.1
		After	1.7	-1.7
HS Rural	No	Before	.3	-.4
		After	-1.0	1.0
	Yes	Before	-.6	.9
		After	1.9	-1.9
General HS Program	No	Before	3.5	-5.0
		After	.5	-.5
	Yes	Before	-5.2	7.5
		After	-.7	.7
College Prep HS Program	No	Before	-5.8	8.4
		After	-.8	.8
	Yes	Before	4.6	-6.6
		After	.7	-.7
Vocational HS Program	No	Before	.7	-1.1
		After	.1	-.1
	Yes	Before	-2.6	3.7
		After	-.4	.4
Highest HS math level	Basic	Before	-14.5	21.1
		After	-3.2	3.2
	Advanced	Before	12.6	-18.4
		After	3.8	-3.8
Sector of 1st PSI	Public	Before	-5.9	8.5

Variable	Stdzd. residual	Remedial Math		
		No	Yes	
Private	After	-2.3	2.3	
	Before	10.6	-15.2	
	After	4.8	-4.8	
Successful HS graduate	No	Before	-14.2	20.1
		After	-3.3	3.3
	Yes	Before	13.2	-18.6
		After	3.8	-3.8
Marginal HS graduate	No	Before	11.9	-16.8
		After	3.5	-3.5
	Yes	Before	-13.7	19.4
		After	-3.3	3.3
4-yr PSI	No	Before	-11.2	16.1
		After	-2.7	2.7
	Yes	Before	11.7	-16.7
		After	3.5	-3.5
2-yr PSI	No	Before	6.5	-9.4
		After	1.3	-1.3
	Yes	Before	-11.5	16.5
		After	-2.0	2.0
2-yr, then 4-yr PSI	No	Before	2.4	-3.5
		After	1.0	-1.0
	Yes	Before	-6.2	8.8
		After	-2.1	2.1
4-yr, then 2-yr PSI	No	Before	.1	-.2
		After	.2	-.2
	Yes	Before	-.3	.5
		After	-.4	.4
White	No	Before	-5.2	7.4
		After	-2.8	2.8
	Yes	Before	4.3	-6.1
		After	2.3	-2.3
Black	Not	Before	2.2	-3.1
		After	.7	-.7
	Yes	Before	-6.0	8.6
		After	-1.7	1.7
Hispanic	No	Before	2.5	-3.5
		After	1.3	-1.3
	Yes	Before	-6.5	9.4
		After	-3.3	3.3
Asian	No	Before	-1.2	1.8
		After	.0	.0
	Yes	Before	3.4	-4.9
		After	.1	-.1
Other	Before	.1	-.2	
	After	-.1	.1	

Variable		Stdzd. residual	Remedial Math	
			No	Yes
	Yes	Before	-5	.7
		After	.3	-.3
Single parent household	No	Before	1.9	-2.8
		After	.4	-.4
	Yes	Before	-3.9	5.6
		After	-.8	.8
Completed PSE pipeline	No	Before	-15.1	21.9
		After	-3.0	3.0
	Yes	Before	12.3	-17.8
		After	3.2	-3.2
Worked 03-04	No	Before	1.3	-1.8
		After	-1.0	1.0
	Yes	Before	-.8	1.1
		After	.6	-.6
Sex	Male	Before	.8	-1.1
		After	.2	-.2
	Female	Before	-.7	1.1
		After	-.1	.1
English is native language	No	Before	-1.4	2.1
		After	-1.8	1.8
	Yes	Before	.6	-.9
		After	.7	-.7
Timing 1st PSI enrollment	Delayed	Before	-6.7	9.7
		After	-.4	.4
	Immediate	Before	3.1	-4.5
		After	.2	-.2
Took out loans	No	Before	-1.4	2.1
		After	-1.6	1.6
	Yes	Before	1.1	-1.6
		After	1.2	-1.2
Program to help plan	No	Before	.0	-.1
		After	.0	.0
	Yes	Before	-.1	.1
		After	.0	.0



Appendix I

*Logistic Regression: Covariates Predicting Remedial Math After Matching (N = 3,978)*

Variable	Coeff.	S.E.	Sig.	Odds Ratio
Catholic HS	-1.266	.139	***	.282
Other private HS	-1.545	.207	***	.213
Urban	-.087	.105		.917
Suburban	.016	.088		1.016
General	.020	.123		1.020
College Prep	.059	.121		1.061
Highest math	.247	.077	**	1.280
Sector 1st PSI	.585	.095	***	1.796
Successful Grad	.167	.193		1.182
Marginal Grad	.109	.180		1.115
Attended 4-yr PSI	.202	.101	*	1.223
Attended 2-yr PSI	.143	.119		1.153
Attended 2-yr, then 4-yr PSI	-.033	.121		.967
Black	-.207	.113		.813
Hispanic	-.366	.119	**	.694
Asian	-.136	.145		.873
Other	-.014	.159		.986
Single parent household	-.016	.083		.984
Completed PSE pipeline	.172	.090		1.188
Worked 03-04 school year	.028	.080		1.028
Sex	-.039	.069		.962
English is native language	.055	.122		1.056
Timing PSI enrollment	-.074	.091		.929
Took out student loans	-.019	.073		.981
Program help plan	-.057	.079		.945
SES	-.013	.054		.987
Math IRT	-.009	.005		.991
Reading IRT	.000	.006		1.000
Constant	2.645	.566	***	14.086
* $p < .05$ , ** $p < .01$ , *** $p < .001$				
Pseudo $R^2$				.111

Appendix J

*Summary of Logistic Regression Model 2 Predicting Degree Attainment Before and After Matching*

Variable	Before Matching				After Matching			
	Coeff.	S.E.	Sig.	Odds Ratio	Coeff.	S.E.	Sig.	Odds Ratio
Remedial math in college	.38	.07	***	1.46	.41	.08	***	1.50
HS control is Catholic	-.26	.09	**	.77	-.17	.14		.84
HS control is other private	-.27	.11	*	.77	-.37	.19	*	.69
HS is urban	.09	.10		1.09	.07	.12		1.07
HS is suburban	-.03	.08		.97	-.01	.10		1.00
General HS Program	.03	.12		1.03	-.07	.14		.93
College Prep HS Program	.00	.12		1.00	-.06	.13		.94
Highest HS math level	-.41	.07	***	.66	-.46	.08	***	.63
Sector of 1st PSI attended	-.19	.08	*	.82	-.13	.10		.88
Successful HS graduate	-1.34	.22	***	.26	-1.19	.24	***	.31
Marginal HS graduate	-.70	.21	**	.50	-.64	.23	**	.53
Attended only 4-yr PSI	-.42	.09	***	.66	-.47	.11	***	.63
Attended only 2-yr PSI	.39	.11	***	1.47	.30	.13	*	1.35
Attended 2-yr, then 4-yr PSI	-.78	.11	***	.46	-.91	.13	***	.40
Black	.38	.10	***	1.46	.34	.12	**	1.40
Hispanic	.11	.11		1.12	.07	.13		1.07
Asian	-.24	.13		.79	-.23	.17		.80
Other	.42	.13	**	1.53	.34	.17	*	1.40
Lives in single parent/guardian household	.07	.08		1.07	.13	.09		1.13
Completed PSE pipeline	-.38	.08	***	.69	-.45	.10	***	.64
Whether worked during 03-04 school year	-.05	.07		.95	-.15	.09		.86
Sex	-.45	.06	***	.64	-.37	.08	***	.69
English is student's native language	.11	.11		1.12	.10	.14		1.10

	Before Matching				After Matching			
Timing of first postsecondary enrollment	-.79	.09	***	.45	-.73	.10	***	.48
Whether R took out any student/PSE loans	-.50	.06	***	.61	-.53	.08	***	.59
Ever in program to help prepare for college	.06	.07		1.07	.15	.09		1.16
SES	.27	.05	***	1.31	.22	.06	***	1.25
Math IRT estimated number right	.01	.00		1.01	.01	.01		1.01
Reading IRT estimated number right	-.02	.01	**	.98	-.02	.01	***	.98
Constant	3.08	.49	***	21.81	3.34	.62	***	28.17
* $p < .05$ , ** $p < .01$ , *** $p < .001$								
Pseudo $R^2$	.336				.307			

Appendix K

*Summary of Linear Regression Model 2 Predicting GPA Before and After Matching*

Variable	Before Matching					After Matching				
	Coeff.	S.E.	$\beta$	$t$	Sig.	Coeff.	S.E.	$\beta$	$t$	Sig.
Constant	1.982	.114		17.32	***	2.035	.149		13.64	***
Remedial math in college	-.127	.021	-.07	-5.99	***	-.111	.026	-.06	-4.22	***
HS control is Catholic	-.019	.026	-.01	-.73		-.116	.048	-.04	-2.40	*
HS control is other private	.080	.031	.03	2.59	*	.130	.066	.03	1.96	
HS is urban	-.035	.028	-.02	-1.24		-.052	.040	-.03	-1.30	
HS is suburban	.008	.025	.01	.33		.007	.034	.00	.20	
General HS Program	.022	.036	.01	.62		.075	.047	.04	1.59	
College Prep HS Program	.060	.035	.03	1.69		.159	.047	.09	3.41	**
Highest HS math level	.144	.022	.09	6.54	***	.168	.030	.09	5.71	***
Sector of 1st PSI attended	.162	.022	.08	7.40	***	.182	.035	.08	5.17	***
Successful HS graduate	.370	.061	.22	6.10	***	.303	.075	.17	4.06	***
Marginal HS graduate	.071	.058	.04	1.22		.018	.070	.01	.26	
Attended only 4-yr PSI	.143	.026	.09	5.56	***	.164	.038	.09	4.29	***
Attended only 2-yr PSI	-.008	.034	.00	-.24		.026	.045	.01	.57	
Attended 2-yr, then 4-yr PSI	.258	.033	.11	7.73	***	.326	.046	.14	7.12	***
Black	-.357	.032	-.13	-11.33	***	-.375	.043	-.15	-8.67	***
Hispanic	-.072	.032	-.03	-2.25	*	-.065	.045	-.03	-1.46	
Asian	-.030	.036	-.01	-.84		-.023	.055	-.01	-.41	
Other	-.169	.041	-.04	-4.09	***	-.156	.061	-.04	-2.58	*
Single parent household	.010	.023	.01	.43		.011	.032	.01	.36	
Completed PSE pipeline	.058	.026	.03	2.20	*	.070	.034	.04	2.04	*
Worked 03-04 school year	-.021	.020	-.01	-1.08		-.010	.030	-.01	-.34	
Sex	.219	.018	.13	12.21	***	.184	.026	.10	6.99	***
English is student's native language	-.028	.032	-.01	-.88		.006	.046	.00	.13	

	Before Matching					After Matching				
Timing of 1st PSE	-.024	.028	-.01	-.86		-.046	.035	-.02	-1.34	
Took out any student loans	.049	.019	.03	2.60	**	.071	.028	.04	2.55	*
Ever in program to help plan	-.035	.021	-.02	-1.70		-.052	.030	-.03	-1.73	
SES	.035	.014	.03	2.56	*	.010	.021	.01	.47	
Math IRT	.007	.001	.10	5.54	***	.007	.002	.08	3.67	***
Reading IRT	.000	.001	.00	-.22		-.003	.002	-.03	-1.27	
<i>*p &lt; .05, **p &lt; .01, ***p &lt; .001</i>										
<i>R</i> <sup>2</sup>	.257					.200				

Appendix L

*Summary of Correlations: Linear Regression Model 2 After Matching*

	Correlations			VIF
	Zero-order	Partial	Part	
Took remedial math	-.129	-.067	-.060	1.090
Catholic HS	-.068	-.038	-.034	1.199
Other private HS	.006	.031	.028	1.078
Urban	-.102	-.021	-.018	2.116
Suburban	.061	.003	.003	1.819
General	-.075	.025	.023	3.222
College Prep	.130	.054	.049	3.404
Highest math	.231	.090	.081	1.340
Sector 1st PSI	.101	.082	.074	1.181
Successful Grad	.329	.065	.058	8.638
Marginal Grad	-.303	.004	.004	7.679
Attended 4-yr PSI	.153	.068	.061	2.130
Attended 2-yr PSI	-.216	.009	.008	2.778
Attended 2-yr, then 4-yr PSI	.092	.113	.101	1.866
Black	-.196	-.137	-.123	1.377
Hispanic	-.065	-.023	-.021	1.532
Asian	.017	-.007	-.006	1.539
Other	-.027	-.041	-.037	1.058
Lives with single parent	-.063	.006	.005	1.091
Completed PSE pipeline	.245	.032	.029	1.855
Worked 03-04 school year	.027	-.005	-.005	1.043
Sex	.110	.111	.099	1.074
English is native language	.041	.002	.002	1.723
Timing PSI enrollment	.114	-.021	-.019	1.275
Took out student loans	.116	.041	.036	1.161
Program help plan	-.025	-.027	-.025	1.035
SES	.120	.007	.007	1.286
Math IRT	.282	.058	.052	2.394
Reading IRT	.246	-.020	-.018	2.245

*Note.* VIF = Variance inflation factor

## Appendix M

### *Frequencies for Original and Pooled MI Variables*

Categorical Variables			Freq.	Percent	Valid Percent
<b>Timing of 1<sup>st</sup> PSE</b>					
Original	Valid	Delayed	1660	15.5	17.6
		Immediate	7772	72.4	82.4
		Total	9432	87.9	100
	Missing	Missing	1304	12.1	
		Total	10736	100	
Pooled	Valid	Delayed	1976.5	18.4	
		Immediate	8759.5	81.6	
		Total	10736	100	
<b>Program to help plan for college</b>					
Original	Valid	No	7430	69.2	77
		Yes	2218	20.7	23
		Total	9648	89.9	100
	Missing	Missing	1088	10.1	
		Total	10736	100	
Pooled	Valid	No	8261.8	77.0	
		Yes	2474.2	23.0	
		Total	10736	100	
<b>Marginal HS graduate</b>					
Original	Valid	No	5595	52.1	57.3
		Yes	4172	38.9	42.7
		Total	9767	91	100
	Missing	Missing	969	9	
		Total	10736	100	
Pooled	Valid	No	6126.3	57.1	
		Yes	4609.7	42.9	
		Total	10736	100	
<b>Successful HS graduate</b>					
Original	Valid	No	4510	42	46.2
		Yes	5257	49	53.8
		Total	9767	91	100
	Missing	Missing	969	9	
		Total	10736	100	
Pooled	Valid	No	4900.1	45.6	
		Yes	5835.9	54.4	
		Total	10736	100.0	
<b>Took out student loans</b>					

Categorical Variables			Freq.	Percent	Valid Percent
Original	Valid	No	3809	35.5	38.8
		Yes	6008	56	61.2
		Total	9817	91.4	100
		Missing	919	8.6	
		Total	10736	100	
Pooled	Valid	No	4214.7	39.3	
		Yes	6521.3	60.7	
		Total	10736	100.0	
Completed PSE pipeline					
Original	Valid	No	4036	37.6	39.9
		Yes	6071	56.5	60.1
		Total	10107	94.1	100
		Missing	629	5.9	
		Total	10736	100	
Pooled	Valid	No	4395.2	40.9	
		Yes	6340.8	59.1	
		Total	10736	100.0	
Worked during 03-04 school year					
Original	Valid	No	2754	25.7	27.2
		Yes	7368	68.6	72.8
		Total	10122	94.3	100
		Missing	614	5.7	
		Total	10736	100	
Pooled	Valid	No	2919.1	27.2	
		Yes	7816.9	72.8	
		Total	10736	100.0	
Highest HS math level					
Original	Valid	Basic	4393	40.9	43.1
		Advanced	5798	54	56.9
		Total	10191	94.9	100
		Missing	545	5.1	
		Total	10736	100	
Pooled	Valid	Basic	4728.5	44.0	
		Advanced	6007.5	56.0	
		Total	10736	100.0	
Vocational HS Program					
Original	Valid	No	9805	91.3	92.4
		Yes	808	7.5	7.6
		Total	10613	98.9	100



Categorical Variables			Freq.	Percent	Valid Percent
		Missing	123	1.1	
		Total	10736	100	
Pooled	Valid	No	9864.7	91.9	
		Yes	871.3	8.1	
		Total	10736	100.0	
College Prep HS Program					
Original	Valid	No	4044	37.7	38.1
		Yes	6569	61.2	61.9
		Total	10613	98.9	100
		Missing	123	1.1	
		Total	10736	100	
Pooled	Valid	No	4106	38.2	
		Yes	6630	61.8	
		Total	10736	100.0	
General HS Program					
Original	Valid	No	7377	68.7	69.5
		Yes	3236	30.1	30.5
		Total	10613	98.9	100
		Missing	123	1.1	
		Total	10736	100	
Pooled	Valid	No	7435.4	69.3	
		Yes	3300.6	30.7	
		Total	10736	100.0	
Sector of 1st PSI attended					
Original	Valid	Public	8165	76.1	76.1
		Private	2568	23.9	23.9
		Total	10733	100	100
		Missing	3	0	
		Total	10736	100	
Pooled	Valid	Public	8166.5	76.1	
		Private	2569.5	23.9	
		Total	10736	100.0	

## Appendix N

*Summary of Pearson Chi-Square Values for All Covariates with Imputed Values (df = 1)*

Variable	Imputation Number												
	1	2	3	4	5	6	7	8	9	10	11	12	
Cllge Prep	Val.	165.43	159.57	170.72	168.74	170.69	168.39	169.54	169.69	165.94	172.71	160.74	161.03
	Sig.	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
Gen. Prg	Val.	128.53	131.19	120.96	127.06	119.69	123.19	118.41	115.26	115.27	124.13	119.34	122.76
	Sig.	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
High Math	Val.	1170.42	1190.45	1206.41	1205.41	1189.09	1195.80	1197.44	1205.52	1232.57	1185.78	1184.79	1189.66
	Sig.	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
Sector PSI	Val.	451.07	450.67	449.57	452.17	451.07	450.67	450.12	451.62	451.62	450.12	450.12	451.62
	Sig.	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
Successful	Val.	1156.22	942.34	944.25	1163.74	1138.92	1152.18	1149.71	1162.72	940.98	1164.48	931.87	1156.43
	Sig.	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
Marginal	Val.	1028.10	810.81	798.40	1032.28	1011.12	1018.43	1028.37	1024.53	795.24	809.26	802.39	1013.84
	Sig.	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
Pipeline	Val.	1204.82	1218.35	1247.30	1219.59	1221.98	1217.07	1232.15	1213.51	1223.61	1212.71	1214.00	1208.76
	Sig.	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
Worked	Val.	4.76	6.49	7.21	11.26	7.19	7.25	6.65	6.13	7.17	8.35	5.21	6.92
	Sig.	.029	.011	.007	.001	.007	.007	.010	.013	.007	.004	.022	.009
Timing	Val.	181.91	165.82	171.33	170.30	175.35	164.64	184.62	174.82	197.88	161.03	188.53	163.93
	Sig.	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
Loans	Val.	12.92	12.65	13.24	8.12	14.71	12.02	14.17	14.72	13.75	14.83	13.45	9.37
	Sig.	.000	.000	.000	.004	.000	.001	.000	.000	.000	.000	.000	.002
Prg plan	Val.	.12	.36	.05	.36	.06	.26	.28	.06	.96	.00	.62	.57
	Sig.	.733	.546	.816	.549	.809	.611	.598	.806	.327	.947	.431	.452

Val. = Pearson Chi-Square (0 cells have expected count less than 5.). *N* = 10,736

## Appendix O

### *Pooled Logistic Regression Coefficients for Covariates Predicting Remedial Math (12 Imputed Data Sets)*

	Coeff.	S.E.	Sig.	Odds Ratio	Fraction Missing Info.	Relative Increase Variance	Relative Efficiency
HS control is Catholic	.006	.076		1.006	.010	.010	.999
HS control is other private	.342	.093	***	1.407	.018	.018	.998
HS is urban	-.076	.075		.927	.005	.005	1.000
HS is suburban	-.104	.067		.901	.002	.002	1.000
General HS Program	-.206	.101	*	.814	.253	.324	.979
College Prep HS Program	-.162	.098		.850	.229	.285	.981
Highest HS math level	.700	.057	***	2.015	.070	.075	.994
Sector of 1st PSI attended	1.024	.070	***	2.784	.006	.006	1.000
Successful HS graduate	.276	.130	*	1.318	.107	.118	.991
Marginal HS graduate	-.122	.121		.885	.103	.113	.992
Attended only 4-yr PSI	.588	.072	***	1.800	.002	.002	1.000
Attended only 2-yr PSI	.369	.086	***	1.446	.011	.011	.999
Attended 2-yr, then 4-yr PSI	.039	.087		1.039	.002	.002	1.000
Black	-.121	.079		.886	.006	.006	1.000
Hispanic	-.267	.082	**	.766	.004	.004	1.000
Asian	.206	.098	*	1.228	.006	.006	.999
Other	-.061	.108		.941	.007	.007	.999
Single parent household	-.041	.059		.960	.004	.004	1.000
Completed PSE pipeline	.485	.065	***	1.624	.049	.051	.996
Worked 03-04 school year	-.060	.057		.942	.081	.087	.993
Sex	.022	.049		1.022	.010	.010	.999
English is native language	-.092	.082		.912	.007	.007	.999
Timing of 1 <sup>st</sup> PSE	-.205	.070	**	.815	.183	.217	.985
Took out any student loans	-.082	.055		.921	.134	.151	.989
Program to help plan	-.075	.060		.927	.113	.125	.991
SES	-.104	.039	**	.901	.040	.042	.997
Math IRT	-.045	.004	***	.956	.030	.030	.998
Reading IRT	.008	.004		1.008	.078	.084	.994
Constant	-.905	.380	*	.405	.050	.052	.996

\* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$

Appendix P

*Pooled Logistic Regression Coefficients for Covariates Predicting Degree Attainment (12 Imputed Data Sets)*

	Coeff.	S.E.	Sig.	Odds Ratio	Fraction Missing Info.	Relative Increase Variance	Relative Efficiency
Remedial math	.364	.05	***	1.439	.015	.015	.999
HS control is Catholic	-.194	.075	*	.824	.014	.014	.999
HS control is other private	-.141	.086	**	.868	.030	.031	.997
HS is urban	-.014	.075		.986	.006	.006	.999
HS is suburban	-.058	.066		.944	.004	.004	1.000
General HS Program	.023	.104		1.023	.279	.369	.977
College Prep HS Program	-.005	.101		.995	.276	.363	.978
Highest HS math level	-.342	.060	***	.711	.140	.158	.988
Sector of 1st PSI attended	-.251	.062	***	.778	.009	.009	.999
Successful HS graduate	-.935	.214	***	.393	.617	1.460	.951
Marginal HS graduate	-.330	.205		.719	.619	1.471	.951
Attended only 4-yr PSI	-.340	.070	***	.712	.007	.007	.999
Attended only 2-yr PSI	.444	.084	***	1.559	.011	.012	.999
Attended 2-yr, then 4-yr PSI	-.763	.090	***	.466	.003	.003	1.000
Black	.469	.080	***	1.599	.017	.017	.999
Hispanic	.228	.083	*	1.257	.009	.009	.999
Asian	-.077	.096		.926	.011	.011	.999
Other	.354	.105	*	1.424	.005	.006	1.000
Single parent household	.075	.060		1.078	.011	.011	.999
Completed PSE pipeline	-.334	.068	***	.716	.131	.147	.989
Worked 03-04 school year	.028	.057		1.029	.097	.106	.992
Sex	-.446	.049	***	.640	.029	.030	.998
English is native language	.158	.082		1.172	.008	.009	.999
Timing of 1 <sup>st</sup> PSE	-.795	.073	***	.451	.233	.292	.981
Took out any student loans	-.467	.052	***	.627	.065	.069	.995
Program to help plan	.027	.061		1.028	.178	.209	.985
SES	.274	.039	***	1.315	.092	.100	.992
Math IRT	.007	.004	**	1.007	.121	.134	.990
Reading IRT	-.016	.004	***	.984	.076	.081	.994
Constant	2.004	.407	***	7.416	.169	.198	.986

\* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$

Note.  $N = 10,736$  for all imputed sets

## Appendix Q

### *Summary of Pooled Estimates of Unstandardized Coefficients for Linear Regression with All Covariates Predicting College GPA (All 12 Imputed Data Sets)*

Variable	Coeff.	S.E.	<i>t</i>	Sig.	FMI	RIV	RE
(Constant)	1.991	.110	18.15	***	.35	.50	.97
Remedial math	-.137	.018	-7.62	***	.03	.04	1.00
HS control is Catholic	-.013	.023	-.56		.02	.02	1.00
HS control is other private	.074	.026	2.81	**	.04	.04	1.00
HS is urban	-.014	.024	-.57		.01	.01	1.00
HS is suburban	.011	.021	.50		.01	.01	1.00
General HS Program	.011	.032	.35		.17	.20	.99
College Prep HS Program	.036	.031	1.16		.17	.19	.99
Highest HS math level	.130	.019	6.67	***	.10	.11	.99
Sector of 1st PSI attended	.164	.019	8.64	***	.01	.01	1.00
Successful HS graduate	.373	.062	5.97	***	.56	1.14	.96
Marginal HS graduate	.118	.055	2.17	*	.47	.83	.96
Attended only 4-yr PSI	.124	.023	5.52	***	.01	.01	1.00
Attended only 2-yr PSI	-.039	.029	-1.34		.03	.03	1.00
Attended 2-yr, then 4-yr PSI	.259	.029	8.95	***	.01	.01	1.00
Black	-.367	.026	-14.07	***	.01	.01	1.00
Hispanic	-.101	.027	-3.79	***	.00	.00	1.00
Asian	-.034	.030	-1.16		.01	.01	1.00
Other	-.159	.034	-4.64	***	.00	.00	1.00
Single parent household	-.016	.019	-.81		.01	.01	1.00
Completed PSE pipeline	.079	.025	3.22	**	.23	.29	.98
Worked 03-04 school year	-.031	.018	-1.77		.09	.09	.99
Sex	.219	.016	13.95	***	.07	.07	.99
English is native language	-.031	.026	-1.21		.01	.01	1.00
Timing of 1st PSE	-.037	.026	-1.40		.34	.49	.97
Took out any student loans	.041	.017	2.36	*	.15	.18	.99
Program to help plan	-.042	.019	-2.22	*	.15	.17	.99
SES	.042	.012	3.57	***	.02	.02	1.00
Math IRT	.008	.001	6.44	***	.20	.24	.98
Reading IRT	.000	.001	.07		.30	.40	.98

\* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$

Note. FMI = fraction missing info., RIV = relative increase variance, RE = relative efficiency

## Appendix R

### *Chi-Square Tests: Matched Remedial Math Groups by All Nominal Covariates (Impute 1)*

Variable	$\chi^2$ <sup>a</sup>	Sig. <sup>b</sup>
Public	268.08	***
Catholic	155.68	***
Other Private	97.41	***
Suburban	14.11	***
Rural	14.65	***
General	2.36	
College Prep	1.06	
Vocational	.03	
Highest HS Math	71.10	***
Sector of 1st PSI	90.74	***
Successful	90.55	***
Marginal	78.81	***
4-year only	78.76	***
2-year only	28.51	***
2-year, then 4-year	13.14	***
4-year, then 2-year	1.15	
White	36.73	***
Black	12.69	***
Hispanic	32.05	***
Asian	1.07	
Other	.08	
Single parent household	1.16	
Completed PSE pipeline	65.33	***
Worked 03-04	5.02	*
Sex	.00	
English is native language	9.58	**
Timing of 1st PSE	1.93	
Took out student loans	10.89	**
Program to plan for college	.50	.479

0 cells have expected count less than 5.

a df = 1

b Asymptotic Significance (2-sided)

\* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$

## Appendix S

### *Summary of Logistic Regression Model 2 Predicting Degree Attainment Before and After Matching (Impute 1)*

Variable	Before Matching				After Matching			
	Coeff.	S.E.	Sig.	OR	Coeff.	S.E.	Sig.	OR
Remedial math	.381	.068	***	1.463	.341	.059	***	1.406
HS control is Catholic	-.258	.092	**	.773	-.070	.108		.932
HS control is other private	-.266	.110	*	.766	-.212	.146		.809
HS is urban	.089	.095		1.093	-.021	.090		.979
HS is suburban	-.031	.083		.970	-.070	.077		.932
General HS Program	.028	.118		1.029	.044	.102		1.045
College Prep HS Program	.004	.115		1.004	.003	.101		1.003
Highest HS math level	-.412	.071	***	.662	-.326	.065	***	.722
Sector of 1st PSI attended	-.193	.078	*	.824	-.210	.083	*	.811
Successful HS graduate	-1.343	.215	***	.261	-.984	.165	***	.374
Marginal HS graduate	-.695	.208	**	.499	-.430	.156	**	.650
Attended only 4-yr PSI	-.420	.085	***	.657	-.320	.087	***	.726
Attended only 2-yr PSI	.388	.106	***	1.474	.372	.099	***	1.450
Attended 2-yr, then 4-yr PSI	-.775	.112	***	.461	-.872	.106	***	.418
Black	.380	.102	***	1.462	.466	.095	***	1.593
Hispanic	.114	.107		1.121	.224	.099	*	1.251
Asian	-.239	.128		.787	.006	.122		1.006
Other	.423	.134	**	1.526	.342	.128	**	1.408
Single parent household	.065	.076		1.067	.128	.071		1.137
Completed PSE pipeline	-.378	.082	***	.685	-.360	.074	***	.698
Worked school year	-.051	.069		.950	-.079	.068		.924
Sex	-.446	.061	***	.640	-.385	.059	***	.680
English is native language	.109	.107		1.115	.171	.100		1.186
Timing of 1st PSE	-.792	.086	***	.453	-.775	.074	***	.461
Took out student loans	-.500	.064	***	.606	-.506	.062	***	.603
Program to help	.064	.071		1.066	.038	.068		1.039
SES	.267	.048	***	1.306	.183	.047	***	1.200
Math IRT	.005	.004		1.005	.006	.004		1.006
Reading IRT	-.017	.005	**	.983	-.017	.005	***	.983
Constant	3.082	.489	***	21.807	2.065	.470	***	7.883

\* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$

Pseudo $R^2$	.336	.307
--------------	------	------

Appendix T

Summary of Linear Regression Model 2 Predicting GPA Before and After Matching (MI I)

Variable	Before Matching					After Matching				
	<i>b</i>	S.E.	$\beta$	<i>t</i>	Sig.	<i>b</i>	S.E.	$\beta$	<i>t</i>	Sig.
Constant	2.02	.09		22.36	***	2.09	.12		18.04	***
Remedial math	-.13	.02	-.07	-7.60	***	-.13	.02	-.07	-6.15	***
HS control is Catholic	-.01	.02	-.01	-.53		-.07	.04	-.02	-1.68	
HS control is other private	.07	.03	.03	2.73	**	.14	.05	.03	2.55	*
HS is urban	-.01	.02	-.01	-.60		-.02	.03	-.01	-.53	
HS is suburban	.01	.02	.01	.53		.02	.03	.01	.54	
General HS Program	.00	.03	.00	-.06		-.02	.04	-.01	-.45	
College Prep HS Program	.02	.03	.01	.84		.04	.04	.02	1.00	
Highest HS math level	.13	.02	.07	6.98	***	.15	.02	.08	6.10	***
Sector of 1st PSI attended	.16	.02	.08	8.68	***	.17	.03	.07	5.81	***
Successful HS graduate	.41	.05	.23	8.97	***	.39	.06	.21	7.05	***
Marginal HS graduate	.14	.04	.08	3.19	**	.13	.05	.07	2.56	*
Only 4-yr PSI	.12	.02	.07	5.50	***	.12	.03	.06	3.57	***
Only 2-yr PSI	-.04	.03	-.02	-1.28		-.01	.04	-.01	-.29	
2-yr, then 4-yr PSI	.26	.03	.10	8.99	***	.30	.04	.12	7.80	***
Black	-.37	.03	-.14	-14.10	***	-.37	.04	-.14	-10.67	***
Hispanic	-.10	.03	-.04	-3.84	***	-.13	.04	-.05	-3.74	***
Asian	-.04	.03	-.01	-1.26		-.05	.04	-.02	-1.13	
Other	-.16	.03	-.04	-4.61	***	-.15	.05	-.04	-3.19	**
Single parent household	-.02	.02	-.01	-.82		-.03	.03	-.01	-.96	
Completed PSE pipeline	.09	.02	.05	4.00	***	.10	.03	.05	3.42	**
Worked 03-04	-.03	.02	-.02	-1.91		-.02	.03	-.01	-.77	
Sex	.22	.02	.12	14.32	***	.20	.02	.11	9.27	***
English is native language	-.03	.03	-.01	-1.31		-.04	.04	-.02	-1.09	
Timing of 1st PSE	-.05	.02	-.02	-2.28	*	-.07	.03	-.03	-2.46	*
Took out student loans	.04	.02	.03	2.78	**	.06	.02	.03	2.43	*
Program to help plan	-.05	.02	-.02	-2.59	*	-.05	.03	-.02	-1.81	
SES	.04	.01	.04	3.71	***	.03	.02	.02	1.48	
Math IRT	.01	.00	.10	6.93	***	.01	.00	.09	5.29	***
Reading IRT	.00	.00	.00	-.10		.00	.00	-.02	-1.23	
* <i>p</i> < .05, ** <i>p</i> < .01, *** <i>p</i> < .001										
<i>R</i> <sup>2</sup>										
	.247					.186				



## Appendix U

### Background to ELS:2002

The ELS:2002 is the fourth school-based longitudinal study in a series of NCES studies and it builds upon three prior NCES longitudinal studies, including the National Longitudinal Study of the High School Class of 1972 (NLS-72), 1980 High School and Beyond (HS&B), and the National Education Longitudinal Study of 1988 (NELS:88).

**NLS-72.** In 1972, NLS-72 began with a sample of over 21,000 high school seniors and included longitudinal data linking educational experiences with postsecondary education enrollment and attainment. The NLS-72 placed great emphasis on low-income areas and schools with significant minority enrollments. The study consisted of four follow-up data collections waves in 1973, 1974, 1979, and 1986, as well as a postsecondary education transcript study in 1984 (NCES, Historical Background, n.d.).

**HS&B.** High School and Beyond included two cohorts: one cohort comparable to the seniors in NLS-72, and a second cohort that included high school sophomores. The HS&B study investigated the relationship between the second cohort's early high school experiences and their subsequent educational experiences in high school and beyond. The study included an investigation into such factors as family, community, school and classroom and whether these factors had an impact on student learning. The HS&B study addressed postsecondary issues such as student enrollment patterns, degree completion, and the impacts of financial aid on enrollment, persistence, and completion (NCES, Historical Background, n.d.).

**NELS:88.** The National Education Longitudinal Study of 1988 began with a sample of 24,599 eighth graders. Subsamples of base year participants and nonparticipants were resurveyed in 1990, 1992, 1994, and 2000. Postsecondary transcripts were collected in 2000. The NELS:88 focused on “trend data about critical transitions experienced by students as they leave middle or junior high school, and progress through high school and into postsecondary institutions or the work force” (NCES, NELS:88, n.d.).

## VITA

MEGHAN A. CLOVIS

Born, Miami, Florida

1998-2002

B.S., Mathematics  
University of Miami  
Coral Gables, Florida

2002-2004

M.A., Mathematics  
University of Miami  
Coral Gables, Florida

2005-2008

Lecturer  
University of Miami  
Coral Gables, Florida

2008-Present

Associate Professor, Senior  
Miami Dade College  
Homestead, Florida

2013

League for Innovation: John & Suanne Roueche  
Excellence Award

2014

Mardee S. Jenrette Endowed Teaching  
Chair of Excellence Award

2015-2018

Doctoral Candidate  
Florida International University  
Miami, Florida

## PRESENTATIONS AND PUBLICATIONS

Archbold, A., Clovis, M. (2010). *Project 2 + 6 prep express*. Presentation at the Miami Dade College Provost's Interdisciplinary Teleconference: Spotlight on Research, Miami, Florida.

Brown, A., Clovis, M., Coronel, A., Garman, S., and Pena, J. (November, 2013). *Different students, different solutions: Accelerated and modularized course design for developmental education*. Presentation at the Florida Developmental Education Association Conference, Miami, Florida.

Clovis, M. (2011). *Super prep express: Project 4 + 4 + 8*. Presentation at the Miami Dade College Annual Math Faculty Retreat, Miami, Florida.

Clovis, M. (2012). *Accelerated developmental mathematics courses*. Presentation to the Miami Dade College District Board of Trustees, Miami, Florida.

Clovis, M. and Coronel, A. (October, 2014). *Developmental education reform: Innovations in developmental education at Miami Dade College*. Presentation at the Association of Florida Colleges: Developmental Education Summit, Destin, Florida, Miami, Florida.

Clovis, M. and Villalobos, S. (2017). *Virtual college MAT1033 intervention pilot*. Presentation at the Miami Dade College 2017 Fall Convocation, Miami, Florida.