DETERMINANTS OF EDUCATIONAL SUCCESS IN SECONDARY AND POSTSECONDARY EDUCATION

by

JESSICA MUNZ SCHELD

A Dissertation Submitted to the

Graduate School-New Brunswick

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Economics

Written under the direction of

Carolyn Moehling

And approved by

New Brunswick, New Jersey

October 2016

ABSTRACT OF THE DISSERTATION

Determinants of Educational Success in Secondary and Postsecondary Education

By JESSICA MUNZ SCHELD

Dissertation Director:

Professor Carolyn Moehling

Expanding access to quality education is important to both policymakers and the nation as a whole. Although many policies have been designed to address inequality in educational opportunities, research does not always support their ability to intervene effectively. Two such policies intended to increase educational opportunities are introducing competition in public school districts and providing access to post-secondary schooling via community colleges. My dissertation examines the effects of these interventions with respect to a recent cohort of students in a period of economic instability.

In the first chapter of my dissertation, I examine the effect competition has on public schools in Massachusetts using publicly available data from the Massachusetts Department of Education. With eight years of data, from school year 2007-2008 to school year 2014-2015, I am able to construct a panel data set of public high schools. I define test scores as an outcome, using both statewide graduation requirements and the SAT. During this time frame, Massachusetts used the Massachusetts Comprehensive Assessment System (MCAS), which is comprised of three separate exams: Mathematics, English/Language Arts, and, starting in 2008, Science/Technology. As a measure of competitive influence, I construct the ratio of charter and private school enrollments to public school enrollments by district. Using a model which allows for

school district fixed effects, I find little evidence of positive effects of competition from charter schools on public school student test performance. Increased private school penetration has negative effects on Math MCAS scores and positive impacts on Math SAT scores, both of which are driven by the suburban districts. Because this model heavily relies on the within-district variation to identify the impact, as an alternative, I employ a control function approach. This method allows for more flexibility in the model and identifies the effect using heteroskedasticity in the error terms. Because alternative school locations are not random, I first estimate a Tobit model to predict the alternative enrollment ratio. Using this in the second stage together with the control variable generated from the error terms, I estimate the effect of alternative enrollment on public school exam scores. Results from this approach indicate a positive impact of alternative schools on public school test performance. As these two methods are quite different, it is not surprising that they produce different results. The main concern moving forward is identifying the appropriate model. At this point, all I can conclude is that the effect of competition is highly dependent upon model choice, which in itself is fruitful for the literature.

In my second and third chapters, I make use of the Education Longitudinal Study of 2002 (ELS: 2002) dataset, produced by the National Center for Education Statistics, to analyze the effect of attending two-year colleges on eventual educational attainment and labor market outcomes. The NCES sampled 750 high schools, and within each school, about 30 students were randomly chosen to participate. Students were surveyed beginning in 2002, as high school sophomores, through 2012, eight years after most graduated high school. Compared to students who began at four-year colleges, students who started at two-year colleges earned about thirty fewer postsecondary credits, one fewer year of education, and were twenty-five percentage points less likely to earn a bachelor's degree. Further, I examine whether the gap varies by racial, socio-economic, or academic differences. I find mixed results by racial groups, although Hispanic students seem to be most negatively impacted by starting at a two-year college, whereas Asian students do not seem to be impacted at all. Additionally, low income students who start at two-year colleges are less likely to earn a baccalaureate degree if they begin their postsecondary career at a two-year college instead of a four-year college. Finally, students with high school GPAs over 3.0 are disproportionately hurt by attending community colleges as they are less likely to earn a bachelor's degree, relative to their peers with high school GPAs above 3.0 who began at a four-year college.

Finally, I consider the effects of educational choices in the labor market. By 2012, almost a third of the ELS respondents had completed some college but had not earned a degree. I find no significant difference in wages or employment status in 2012 between high school graduates and students with some postsecondary attendance but no degree. Further, there are benefits to earning a certificate or associate degree over some two or four-year college credits. I find that males see negative labor market returns in terms of income when earning some college credits without earning a degree, relative to students with a high school degree or GED. Women see no impact from earning some credits in two-year colleges, but have positive returns to some four-year college credits relative to high school graduates. For both genders, the wage benefits of earning a bachelor's degree range from twenty-two to sixty-four percent by age twenty-six. The range of returns is larger than earlier studies, but still suggests that earning a bachelor's degree provides the greatest wage benefits, even in an unstable economy and among those early in their career.

Acknowledgments

Many people have contributed to this dissertation. I would first like to thank my advisor, Dr. Carolyn Moehling, for her unwavering support, encouragement, and guidance throughout my time at Rutgers. In addition to being an invaluable resource, Dr. Moehling helped shape my writing and presentation skills, providing many helpful suggestions for which I will always be grateful. Additionally, I am thankful for my committee members, Dr. Ira Gang, Dr. Roger Klein, and Dr. Paul Gottlieb, who all provided suggestions and advice that enriched my research.

I would additionally like to thank Dr. Anne Piehl, Dr. Rosanne Altshuler, Dr. Jennifer Hunt, and Dr. Hilary Sigman for their invaluable recommendations. I also could not have finished without the help of Dr. Thomas Prusa, who provided helpful advice both with respect to my research and teaching. Dr. Hugh Rockoff and Dr. Louise Russell helped develop my research by providing valuable research experience. Lastly, without the help of the graduate support staff, nothing would ever get done, so I must thank Linda, Janet, Donna, Debra, and Paula.

Graduate students who go through the program together form a special bond; we have helped each other through some of our toughest times. Georgia, sharing an office with you was such a blessing. I would also like to thank Jessica, Geoff, Han-Yen, Shuyang, Antonio, Dean, Samata, Freddy, and Walter. Finally, I would not have survived my first year without our study group; a special thank you to Roque, Atisha, and Vladimir for making that first year bearable.

Finally, I want to thank my family and friends, without whom I would not have been able to dedicate the time and energy to completing this task. I would especially like to thank my parents and my siblings, Andrew and Kaitlin, for their continued love and support. To my friends at Montgomery, especially Ilana, Maggie, and Dara thank you for the love, meals, and forced breaks from work. Completing a doctorate

truly takes a village, and I would not be here without the many wonderful people in my life who helped me along the way.

Dedication

To my parents, Patricia and Edward

Table of Contents

List of Tables

List of Tables Continued

List of Tables Continued

List of Figures

List of Figures Continued

1. Introduction

My research focuses on secondary and postsecondary schooling decisions, providing a unique look at students in a cohort of economic instability. Analyzing the impacts of education policies is crucial to creating an environment in which all students can prosper. With policies in place to provide both access to quality education and tools for success, students will be able to achieve both their academic and labor market goals. Students from disadvantaged backgrounds are disproportionately under-served with respect to quality education. Several policies are in place to help alleviate the educational gap for at-risk students. Two of these policies are the focus of my dissertation: school choice at the secondary level and community colleges at the postsecondary level. The goal of both school choice and increased college access goes further than equity of access; these policies aim to encourage equity of outcomes. First, I examine the effect of school choice on student achievement using publicly available data from the Department of Education in Massachusetts. I later make use of the restricted access Education Longitudinal Study of 2002 (ELS: 2002) to examine community college outcomes. First, I look at the impact of community colleges on baccalaureate completion, and then I consider labor market impacts of postsecondary educational outcomes.

In the first chapter, I examine the effect of competition generated by school choice within school districts in Massachusetts in the years 2008 through 2015. Massachusetts is an interesting state in which to study the effect of school choice, as there are many private and charter school options in both urban and suburban districts. While increased competition leads to lower prices in profit-maximizing industries, it is not clear that increased competition has a beneficial effect in the education sector. Competition among schools might raise average school quality and with it, academic achievement. Alternatively, increased competition could leave public schools with more difficult students and tighter budgets. School choice is a contentious issue; with funding decisions and tight budgets, how school choice impacts public school students is of great importance. If increased school choice has a positive effect on public school test scores, this would be an indication that schools are effectively re-optimizing their budgets and increasing their attention on academic outcomes. This would ultimately help both those who seek school choice as well as those students who remain, the best possible outcome.

I measure competition by determining the ratio of students attending alternative schools in district i to students attending the public school in district i. I measure achievement by test scores. Although there are certainly other interesting outcomes, test scores are tied to funding and scholarship decisions, and can mean free tuition to state schools for many students, especially those in more disadvantaged districts. During this time period, public school students in Massachusetts were required to pass an exit exam in high school in order to receive a diploma. These exams, called the Massachusetts Comprehensive Assessment System, or MCAS, are comprised of a Math, English, and Science component. I use the average passing rate, in addition to each specific exam passing rate as dependent variables. Alternatively, I consider average Math and Reading SAT scores as dependent variables. While ordinary least squares estimates suggest a positive effect of increased charter school enrollments relative to public school enrollments on average MCAS passing rates, the nature of the charter school movement combined with location decisions of private schools indicates that endogeneity is likely an issue. Using a school district fixed effects model, I find some evidence that increasing charter school penetration has a positive effect on exam scores. Increased private school penetration, on the other hand, leads to lower Math MCAS scores and higher Math SAT scores. Separating schools by urbanicity shows that in urban districts, increased charter penetration leads to higher Science MCAS scores, and that the private school impacts are driven by the suburban districts.

Fixed effects models, however, might not adequately control for the endogeneity of alternative school enrollments. Since there is not a lot of variation in alternative school enrollments or in the covariates between 2008 and 2015, even a small correlation between the error and demeaned enrollments might lead to a large bias in the fixed effects estimates. As an alternative, I employ a control function approach in which the first stage equation is a Tobit model, used to estimate the ratio of alternative enrollments relative to public enrollments. This allows for the estimation of a model in which there are many zeros, or districts in which there are no alternative schools. The control function approach then utilizes the heteroskedasticity of the errors to identify the model. These results indicate a positive impact of alternative schools on public school test performance, suggesting a positive effect of increasing schooling options on overall school quality, as measured by MCAS and SAT scores.

The results from the two econometric models produce different results, suggesting that the estimation of competitive effects is sensitive to modeling decisions. I believe it is important to account for the underlying first stage model in which districts without alternative schools are analyzed differently than those with school choice. However, there may still be school district fixed effects that are time invariant that are not accounted for within the control function approach. The fixed effects estimates do suggest positive charter school impacts on test scores, but they are not precisely estimated. Hence, I believe the control function approach is capturing the effect more precisely given the constraints of the data.

In my next two chapters, I use the restricted use version of the Education Longitudinal Study of 2002 (ELS: 2002) dataset to analyze the effect of attending two-year colleges on a variety of outcomes, including educational attainment and income. Community colleges are often seen as a strategy for increasing access to higher education for at-risk students and were most recently endorsed by President Obama as a way to make higher education affordable for all. Students who choose community college may be seeking a cheaper option on their path to an undergraduate degree. Alternatively, students may enroll in community colleges because their high school credentials do not support an institution with greater selectivity, but their desired careers require higher education. Community colleges have the potential to have a considerable impact on a diverse group of students, many of whom are minorities or are from lower socioeconomic backgrounds, both in terms of academic outcomes and labor market decisions.

The National Center for Education Statistics conducted a nationally representable survey of high school sophomores in 2002, with follow up surveys in 2004, 2006, and finally in 2012, eight years after most graduated high school. To estimate the impact of collegiate choice on attainment, I considered 7780 students who graduated high school in 2004, and pursued higher education in a public or non-profit private college. Compared to students who began at four-year colleges, students who started at two-year colleges earned about 30 fewer credits, one fewer year of education, and were twenty-five percentage points less likely to earn a bachelor's degree. Further, I examine whether this gap varies by race, socioeconomic status, or academic background. Of the racial subgroups, it seems that Hispanic students that begin at two-year colleges are much less likely to complete a bachelor's degree as compared to Hispanic students who begin at four-year colleges. On the other hand, postsecondary level does not impact Asian students' graduation rates. Additionally, low income students who start at two-year colleges are less likely to earn a baccalaureate degree than students of comparable socioeconomic background who enter four-year colleges. Finally, among students with high school GPAs of over 3.0, students who begin at a two-year college were significantly less likely to finish a bachelor's degree by 2012.

The finding that there is still a gap in attainment for those beginning their postsecondary career at community colleges does not necessarily mean that the policy prescription is to promote more four-year college entrants. Many students enter four-year programs and drop out; if the returns to a sub-baccalaureate degree outweigh those of some credit attainment, it might be preferable to encourage community college enrollment for some students. Although previous studies have generally shown that there is a positive effect of schooling on labor market outcomes, men and women in the ELS sample who only complete some college do not earn statistically different wages from high school graduates. By 2012 almost a third of individuals in the ELS sample had completed some college but had not earned a degree. Educational decisions are not random, however, and ordinary least squares estimates are likely biased.

Instrumental variables might seem to be an optimal method to account for selection bias, but identifying an instrument that impacts income only through educational choices is difficult. Alternatively, propensity score matching provides a more flexible model for estimating the impact of educational decisions on income.

Comparing traditional ordinary least squares estimates to instrumental variables and propensity score matching estimates provides evidence that there are moderate, but positive wage improvements for women who complete some college at a four-year institution but do not earn a degree compared to high school graduates. Men, however, have negative labor market returns due to not completing a degree relative to high school graduates. This finding contrasts the majority of previous studies which find positive returns to any postsecondary experience relative to high school graduates. Earning a bachelor's degree provides the greatest benefits, but the range of estimated returns is wider than previous studies find. Compared to the returns to a bachelor's degree, the returns to sub-baccalaureate degrees are lower, as expected. The relative returns, however, are smaller than those found in studies using data from earlier time periods. This suggests that either the returns to a bachelor's degree is growing faster than sub-baccalaureate degrees, or that sub-baccalaureate degrees are falling in desirability. However, earning a sub-baccalaureate degree has returns that outweigh no college or some college. Community colleges offer higher education opportunities to many who might not otherwise pursue it. However, if attendance alone does not increase wages, policies must be implemented to increase the incentive to complete a degree.

2. School Competition and the Performance of Public School Students

1 Introduction

Expanding access to quality education is important to both policymakers and the nation as a whole, as a highly-educated society is better able to sustain continued economic growth. “A Nation at Risk,” published by the United States Department of Education, brought attention to the problems the United States education system was facing in 1983, but equity among education is still sought after. One policy aimed at improving school quality is to increase competition by promoting school choice. However, while increased competition causes lower prices in profit-maximizing industries, it is not clear that increased competition has any effect on public schools. Innovation as a reaction to school choice may lead to better test scores in traditional public schools as they compete to retain students. A positive penetration might imply that by introducing competition, alternative schools force public schools to seek out efficiency; reallocate funds, focus more time or money on high achieving students or programs, or improvise methods to allow for student growth. Alternatively, increased competition could leave public schools with frustrated teachers, more difficult students, uninterested parents, increased scrutiny, and tighter budgets. In an ideal world, alternative schools would not only increase the educational gains of their own students, but induce public schools to invest in methods, curricular or otherwise, that would improve academic outcomes in their own schools. Thus, it is important to understand the various channels through which districts may respond to increased competition, as academic outcomes have long-lasting effects.

I look at the effect of competition on academic achievement among public school students at the high school level, and how this effect differs by the degree of urbanicity in Massachusetts between 2008 and 2015. I measure the degree of competition by calculating enrollments in district choice schools (private and charter schools) relative to the number of traditional public school enrollments. School choice is a broad term that includes voucher systems as well as private, charter, and magnet schools. In this paper, I focus on the effect that charter and private schools have, as Massachusetts did not have a voucher program during this time, and magnet schools tend to emphasize non-academic outcomes, such as arts or music.[1] Traditional public, charter public, and private schools are closely related, but not perfect substitutes. Private schools often come with entrance requirements and tuition payments, creating credit or academic constraints not all families can manage. Charter schools are arguably a closer substitute for traditional public schools, but have seen inconsistent results and are controversial in terms of funding. Massachusetts is an interesting setting in which competition may be studied: charter schools, especially in Boston, have been well-studied in their ability to produce MCAS scores significantly above that of traditional public schools. It is also home to some of the most prestigious private schools outside of New York City. I expect that if a competitive effect were to exist with respect to academic performance, it would be most easily seen in states where public schools are faced with high quality alternatives.

In order to test for the effect competition has on traditional public schools test scores, I first consider ordinary least squares. As the number and location of alternative schools is not exogenous, however, ordinary least squares estimates will be biased. To control for endogeneity, I consider a fixed effects model, as it is expected that each district will have unique features that are unobservable. For example, the quality and effectiveness of administration can significantly affect the academic progress a district makes, and these differences are accounted for with a unique intercept for each district. However, fixed effects models are identified off of within district variation over time. As this is a fairly short panel, and there was not a lot of variation

[1] There is one magnet school in Worcester: MA Academy of Math and Science at WPI that was included in Worcester Public School's data.

in enrollments, a concern is that the fixed effects model is not able to capture the penetration of alternative enrollments. Further, it is possible that the effect of school choice varies with time, creating heteroskedasticity in the error term. If this is the case, I am able to consider a control function model. Developed by Klein and Vella (2010), this is an alternative to instrumental variables when a valid instrument is not available. However, this model treats the dataset as if it is a cross-section, and does not account for the panel, or clustering the standard errors. In addition, estimating the model utilizes maximum likelihood techniques, which requires a more parsimonious model. This also creates issues when estimating effects of subsets of the data, such as the urban or suburban districts separately.

In the next section, I will discuss previous literature and its relation to my research, followed by the description of the model and results, and finally, the conclusion.

2 Literature Review

There are few studies examining the effect of school choice on students remaining in public schools, and none in Massachusetts, the state I will focus on in this study.[2] Private schools have historically offered alternatives to the traditional public schools but have restrictions of their own. Charter schools are a relatively new form of school choice that are closer in form and funding to traditional public schools. Both private and charter schools may induce competition for students through different avenues.

While private schools are becoming more representative of the national population, minority students are still under-represented. As of 2009-2010, White students

[2]Charter school students in Boston and Chicago are outperforming their traditional public school counterparts (Angrist et. al 2011, Hoxby 2005). The Center for Research on Educational Outcomes (CREDO 2013) recently published a study of Massachusetts and Boston charter schools and compared charter students to traditional public school students with similar demographic backgrounds. They also found positive effects of charter schools on their student population, although they found mixed results nationwide (CREDO 2009).

made up 72.6% of students in private schools, while African Americans only made up 9.2% of students and Hispanics 9.4% (NCES Table 9). On the other hand, in traditional public schools, these percentages were 54.1%, 16.2%, and 21.8% respectively, and in charter schools they were 37.0%, 29.9%, and 25.6% (National Alliance for Public Charter Schools 2010). Vouchers have enabled some affirmative action, but private schools still educate different racial populations. As of 2010, private schools made up about 25% of all schools, but only educated 10% of students. The majority are still Catholic, but between the late 1990's and school year 2011-2012, the percentage dropped from 54.5% to 42.9% (Council for American Private Education 2012). Public school students are likely to be different from private school students in both observable and unobservable ways; any comparison between them at face value is likely to be biased. Hoxby (1993) found evidence of improved public school academic performance in districts with high private school enrollments, but she analyzed a slightly different population than I will see in Massachusetts. It is plausible the number of students enrolled in private schools has an effect on public school academic achievement.

Charter schools, on the other hand, are a closer substitute as they are public schools; they cannot advocate for, or deny, any student a spot in their school, or charge tuition. However, one of the major departures charter schools make from traditional public schools is autonomy. This allows schools to enforce policies such as longer school days and school years, or stricter disciplinary and promotion policies. These differences might provide positive incentives for some students but crowd others out. Additionally, unlike administrators in the traditional public school system, charter school leaders are able to hire and fire as they see fit, as tenure is not granted.[3] While traditional public schools are held to the same academic standards, it takes much

[3] In exchange for the ability to make more decisions, charter schools are held to a stricter accountability standard. At the end of a charter school's contract (three-five years is most common), the Department of Education conducts an inspection whereby their charter is either extended or revoked (Mead and Rotherham 2007).

longer to close or reform a traditional public school than a charter school. States with stricter regulations regarding charter school renewals might be expected to have greater competitive effects via required high academic performance. Charter schools may additionally impose competition because of the way charter schools are funded. Charter school funding is dictated by the state and involves a formula in which the sending district of each student must allocate funds to the school the student actually attends. In Massachusetts, public school funding is made up of state, district, and a small amount of federal funding. Each student has a state determined "expenditure," or the cost of education, which is determined by district, special education status, and income levels. The charter school that educates the student will receive that aid out of the public school budget, but the sending district receives reimbursements as well. From the state, in the first year the sending district receives 100%, with 25% reimbursed in each of the next five years (Massachusetts Department of Education 2013a). The funding formula is intended to allow sending districts time to reallocate their budget and to alleviate some of the budgetary losses imposed by charter schools. Since a substantial amount of funding follows the student, traditional public schools have a financial incentive to compete for these students. A handful of seats does not make up an entire class, but a movement dispersed over various grades could cause a funding loss large enough to create significant budgetary concerns. On the other hand, it is possible that public schools would have a larger response if they were not compensated at all for the loss in student population. Finally, the competitive effect of charter schools might be diminished due to the charter school cap, which limits the number of students allowed to enroll in neighborhood charter schools.[4] It is plausible that in areas where the cap has been reached, the potential effect is greater than the observed effect. Many urban areas in Massachusetts have reached their cap, and

[4]Massachusetts has two types of charter school caps: (1) throughout the state, there can be a maximum of 120 charter schools (72 Commonwealth and 48 Horace Mann charter schools) and (2) a maximum of 9% of each districts' funding can be allocated to charter schools, and a total of 4% of students statewide may attend charter schools.

charter schools in these areas often have substantial waiting lists.[5] As removing the cap could more than double enrollments, its existence potentially distorts the true effect of charter school education. This might affect competition negatively, as the penetration charter schools could have as competitors to traditional public schools is diminished by the cap. Thus, the overall competitive effect from charter schools on traditional public schools is uncertain.

Dating back to Friedman in 1962, one of the main reasons for introducing school choice was to create competition in the public education sector as a means of improving academic standards and performance (Holmes et al. 2004). Holmes et al. show that charter schools increase competition for elementary schools for traditional public schools that are in close vicinity. This result is stronger the closer the charter school is to the traditional public school, and is robust across several specifications. The model used in this paper calculated the “cost” of attending a charter school, in terms of travel, to measure the effect of school choice on the traditional public schools. North Carolina, the state examined, currently does not impose caps on the number of charter schools or the percent of the student body allowed to enroll in charter schools. Thus, this significant penetration of competition may not be seen in states in which there are caps in place, or where growth is slower.[6]

Hoxby's research also suggests that school choice has a positive effect. Hoxby (2000) uses the number of rivers and streams in a metropolitan area as an instrument for the number of districts in that area. Natural landmarks such as waterways predict district formation, as before bridges were common, rivers created natural boundaries. Further, river location is certainly expected to be orthogonal to exam scores and other academic and non-academic outcomes. Districts such as Boston have

[5]For the school year 2013-2014, in Massachusetts, there were $40,376$ unique students on waiting lists for charter schools ($16,864$ of those in Boston alone), and $31,997$ students currently enrolled in charter schools (Massachusetts Department of Education 2013b, 2013c).

[6]Greene and Forster (2002) also consider a situation in which a district is faced with significant school choice. They also find positive effects of schools exposed to choice as compared to those which were not.

many surrounding districts which provide a greater degree of school choice than some other districts, with Miami being on the opposite end of the spectrum. Hoxby finds that metropolitan areas with greater choice have more productive public schools, via increased attainment and decreased spending. However, the instrumental variables approach hinges on a correct choice and measurement of the instrument. Rothstein (2005) refutes her work, suggesting that in fact, her estimates are overstated. He finds errors in both the measurement and coding that led to a failure to replicate her work; his findings with the corrected data suggest that the effect is much smaller than reported. Hoxby (2002) also reports positive results of increased school choice via charter schools in Michigan and Arizona and voucher programs in Milwaukee using a differences-in-differences approach. She concludes that the increase in academic achievement among public school students far outweighs any possible negative repercussions of increased school choice. However, the areas Hoxby (2002) studies had large percentages of students in charter schools, higher than the average nationwide, and certainly higher than Massachusetts, the context of this study. As I will show below, there are several ways in which districts with charter schools are significantly different from those without in Massachusetts, making a difference-in-differences approach in my context difficult.

Finally, Bohte (2004) uses publicly available data on high school students in Texas through the years 1996-2002. He first considers whether public school enrollments are significantly affected by the presence of charter schools, and finds that they are not.[7] Using the natural logarithm of enrollments to measure the effect of increased charter school penetration, Bohte finds that the traditional public schools' test scores increased with greater enrollments, and more-so among a subset of only low-income students. He concludes that charter schools, and competition in general, have positive effects on the public school students left behind. The exact channel

[7]This is also true in Massachusetts. Results are available upon request.

of this effect, possibly due to more "at-risk" students leaving, or increased funds devoted to the remaining population, is unclear. Additionally, while Bohte did account for time effects in his models, he did not explicitly address the endogeneity problem associated with where charter schools operate. In Texas, as in Massachusetts, charter schools are more often in areas in which there is high demand for alternatives. Since they are not randomly located across the state, treating increased enrollments in a district as exogenous is likely too strong an assumption.

The above studies have focused on areas with a large degree of school choice options, and have shown positive effects of competition. However, it is plausible that in areas with stricter caps or fewer choice options that competitive effects are smaller or negative. Zimmer and Buddin (2009) find that California charter schools impose little competitive pressure on public schools. Lubienski (2005) notes that when schools are forced to reallocate funds, it is more common to see more funding going towards areas such as marketing instead of areas that could directly benefit students. That is, the effect of an increase in school choice might not be an increase in academic performance, but rather an increase in mean academic ability of the student body through effective advertising. Mintrom (2000) finds that charter schools in Michigan had not in fact employed innovative techniques, and by and large had practices very similar to that of the public schools. Schools had increased their spending on areas such as advertising and research and development. Lubienski (2005) notes that while school marketing may indeed take money directly away from students, it may not be all negative. An increase in marketing indicates a school's transformation toward focusing on being consumer driven and making decisions based on what will attract families and students to their school. The overall effect of marketing is unclear as it takes money away from classroom practices, but may in fact attract higher ability students, which might have both individual and peer level impacts.

I contribute to the literature by analyzing the performance of Massachusetts public schools relative to the number of charter and private schools in close proximity. I measure alternative school penetration in a district by computing the ratio of alternative school enrollments to public school enrollments.[8] While school choice includes vouchers, pilot, charter, and private schools, in this study I consider only charter and private schools as a measure of school choice. Additionally, while other measures of school quality may be used (for example: future plans of graduates), I focus on standardized test scores here.

3 Data

I use publicly available district and school-level data from the Massachusetts Department of Education.[9] I was restricted by available data to studying school years 2007-2008 through 2014-2015.[10]

The main dependent variable in the model is a school's passing percentage on the Massachusetts Comprehensive Assessment System (MCAS). The MCAS was a response to the Education Reform Act in 1993. It is a statewide exam given to students in grades 3 - 8 and 10. As a graduation requirement in Massachusetts, public school students must pass a Math, English, and, starting in 2008, a Science/Technology exam in the tenth grade in order to graduate high school. Thus, the tenth grade exams are high stakes for both students and educators. I consider approximately 250 districts

[8] Alternate definitions are constructed for sensitivity analysis and the results do not change much. These alternate definitions of penetration are: number of schools and total enrollments (in log form)

[9] http://www.doe.mass.edu/ 9/11/13

[10] The school years 2007-2008 through 2013-2014 were used in some specifications in which grant data was utilized.

with high schools in each year.[11][12] On the MCAS, students receive a raw score that is translated into advanced, passing, needs improvement, or warning/failing. Students who receive passing or advanced scores are able to graduate. While all three are required for graduation, they are three very different exams. The Science exam passing percentages represent the highest exam performance on either the 9^{th} or 10^{th} grade Science exam (which can be Biology, Chemistry, Introductory Physics, or Technology/Engineering). On the other hand, the Math and English MCAS are proctored once in the 10^{th} grade. Also, the Science passing percentages tend to be the lowest, and is also the newest exam, so in calculating the average over the three, the overall average is significantly affected by the Science/Technology exam. Bohte (2004) averages test scores for schools in his study of Texas public schools' response to school choice, but the variable of interest is the percent of students passing all three exams, and this is not available in the Massachusetts data.

I consider dependent variables of the average passing rate, as well as average passing rates on Math, English, and Science/Technology exams separately. As alternative dependent variables, I consider average Reading and Math SAT scores.[13] These higher stakes exams are one indicator of college readiness, and a way a district might differentiate themselves from neighboring districts.

As the achievement outcomes occur in high school years, enrollment data is for

[11] In 2008, 2009, and 2011 there were 257 districts. In 2010, there were only 256, as Provincetown did not report MCAS scores. In 2012, there were 256 districts: Ayer and Shirley and Somerset and Berkley renamed their regional schools, and Provincetown did not report MCAS scores. In 2013, Harwich and Chatham merged to Monomoy Regional High School, and Southwick-Tolland renamed to include Granville, in addition to Provincetown not reporting MCAS scores. Finally, in 2015, Essex and North Shore Agricultural and Technical merged to Essex North Shore Agricultural and Technical. Thus, there are 257 districts in 2008, 2009, and 2011, 256 in 2010 and 2012, 255 in 2013 and 2014, and 254 in 2015. Note also that while the names change for some of these schools, they did not merge during this time period, so I treat them as the same school.

[12] A list of districts can be found at MassGIS Data 2015; the url is provided in the reference section.

[13] The percent of students who earn a 3 or higher on an AP exam was another plausible dependent variable. However, there are fewer observations for which AP scores are available, and the variable may also suffer from measurement error or selection bias. Schools did not always report scores, and students may take AP courses without taking the exam.

high school only. Enrollment data may incorporate one or more schools or towns, however. Many districts in Massachusetts are regional, including two or more towns. On the other hand, urban district enrollment data may include several schools.[14] To calculate charter and private school enrollments, schools were matched to the district in which they operated. I consider all choice schools operating in a district as being in close "vicinity." For example, all charter and private schools operating in Boston were considered to affect Boston Public Schools (Massachusetts Charter Public School Association, Massachusetts Department of Education).[15] About 90.4% of districts in Massachusetts have no charter schools, whereas about 64% have no private schools. There are some districts with one or the other, but many have both. It is very possible that the competition the affected districts face is masked by the majority of districts without alternative schools.

Once enrollment data for charter and private schools were found for each district, I construct the penetration variables. The measure of penetration is found by dividing choice (private or charter) enrollments by the public school enrollments in that year to arrive at the ratio of alternative school enrollments to public school enrollments per district.[16] As choice schools are unevenly distributed across the state, and specifically concentrated in urban locations, the percentage enrolled measures the effect while controlling for district size.[17]

[14] Note that in the construction of enrollment data, The Massachusetts Academy of Math and Science in Worcester, MA is a public school that accepts students on the basis of test scores, and thus was added to Worcester Public Schools enrollments (Massachusetts Department of Education).

[15] See Appendix A for a list of charter schools and their corresponding districts. I did not include two online schools that began operating in 2015: TEC Connections Academy Commonwealth Virtual School District (high school enrollment of 222), and Massachusetts Virtual Academy at Greenfield Commonwealth Virtual District (high school enrollment of 136).

[16] As a robustness check, I use the ratio of charter students to public students in consecutive years, and found only slightly more significant results.

[17] Alternate definitions were constructed for sensitivity analysis and the results were similar. These are: (1) ratio of alternative schools in the year prior to current public enrollments, (2) number of alternative schools and (3) total alternative enrollments (in log form). The third measure is calculated by adding 1 to each enrollment variable to allow for districts with no alternative schools to be coded as 0. Using subsequent years rather than contemporaneous is done to test the reaction time of public schools. The second alternative, number of schools, measures the difference in perceived threat due to two choice schools rather than one, which may play a significant role, as suggested by

Several factors are known to affect academic performance. Due to increased standards, such as adequate yearly progress and improved methods over time, I expect test performance to be an increasing function of time, and as such, yearly time effects are included. Class size may penetration the ability for individual students to learn, although reductions in class size do not necessarily affect achievement (Hoxby 1998). I include the student to teacher ratio, which includes aides and co-teachers, to account for the possible effects they might have on achievement outcomes.[18] This additionally considers the "ability" of the class, as classes with more students with disabilities will require more aides (Boozer and Rouse 2001). Attendance rates are included to control for the amount of time in the classroom, and the notion that attendance and achievement are positively related (Gottfried 2010).[19] Socioeconomic status and ethnicity also affect academic achievement (Sirin 2005), so I include variables for the percentage of African American, Hispanic, special education, and low-income students.[20] Note that special education consolidates a very heterogeneous group of students who may range greatly in capabilities. Finally, I include the natural log of public school enrollments in the year prior to account for larger districts having more students to educate but perhaps more budget flexibility.

I include the percent of administrators out of total full-time staff by district to capture varying budget constraints each district faces in allocations to administrators versus classroom expenditures. The administrative data was organized by choosing

---

Bresnahan and Reiss (1991). However, Massachusetts saw little expansion in high school charter or private schools over the time period, so it may not be the best measure of penetration due to lack of variation. Finally, total enrollments is a similar measure to the ratio in that it captures the amount of student loss, but does not control for differences in district sizes.

[18]Krueger (1997) finds positive effects of small class sizes on Kindergarten students, which was greater for minorities and low-income students. Jepsen and Rivkin (2002) also finds positive effects in elementary students, but that by expanding the small class size throughout the school lowered average teacher quality. Including a quadratic term does not change the results, and produces insignificant quadratic coefficients. This suggests that the driving effect is linear, and that there is not a peak or trough to consider with respect to optimal class size.

[19]It is possible that the attendance rate picks up stronger administrations that enforce strong attendance policies.

[20]These subpopulations are specifically studied in Massachusetts' reports of adequate yearly progress, and are targeted groups in test performance.

the titles that included "supervisor" or indicated a position of power, such as "principal" or "administrator". I summed these totals and calculated the percentage out of total full-time equivalents.[21] The effect of additional administrators is controversial, and could have a positive effect with strong administrators or a negative impact if the money is not being efficiently allocated (Chubb and Moe 1990, Meier, Wrinkle, and Polinard 2000). Average teacher age within the district was included to control for teacher experience. To construct average teacher age, I used the number of teachers in each of the seven age brackets. This composite variable was created by multiplying the mean age in each category by the district's total number of "full-time equivalents" in that category, summing these and dividing by the total number of full-time equivalents.[22] With many young teachers, the concern is inexperience in the classroom or a great deal of turnover. With many older teachers, the concern is a lack of freshness with respect to the curriculum, technology, or connections with students. While age may not perfectly predict teacher quality, it is expected that schools with averages on the low or high end of the spectrum might have lower quality teachers. The magnitude of the effect of teacher quality varies; Rivkin, Hanushek and Kain (2005) find that higher quality teachers, as measured by student performance, had more penetration on student achievement than reduced class size. Interestingly, they also found that teacher quality was difficult to predict and was not largely determined by education or experience.

A final important variable to control for is varying degrees of wealth across the state. I include the percent of funding attributed to grants, as lower income areas receive more aid. It is found by calculating the fraction of total expenditures

[21] The titles are: Superintendent, Assistant/Associate/Vice Superintendent, School Business Official, Other District Wide Administrators, Special Education Administrator, Principal/Headmaster(mistress)/Head of school, Deputy/Associate/Vice/Assistant Principal, Other School Administrator/Coordinator, and Supervisor/Director of Guidance, Pupil Personnel, Arts, Assessment, Curriculum, English Language Learners, English, Foreign Language, History/Social Studies, Library/Media, Math, Reading, Science, Technology and Professional Development (Massachusetts Department of Education).

[22] AvTeachAge $= \frac{(Less26*24+Btw2632*29+Btw3340*36.5+Btw4148*44.5+Btw4956*52.5+Btw5764*60.5+Over64*66)}{(Less26+Btw2632+Btw3340+Btw4148+Btw4956+Btw5764+Over64)}$

from "non-appropriated revenue sources from federal, state and private grants (Massachusetts Department of Education)." At this time, this variable is only available through 2014, and thus requires sacrificing a year of data. Alternatively, I consider Chapter 70 aid. In Massachusetts, each district has a foundation budget, which determines the amount the district is able to contribute to the school system.[23]. Once the foundation budget for the district is determined, the required local contribution is calculated by property values and income levels.[24] The gap between the foundation budget and the required local contribution is Chapter 70 aid. Across districts, aid is quite large, and thus in the dataset, Chapter 70 aid will always be reported in millions of dollars.

In Table 1, I present descriptive statistics. Average passing rates on the MCAS are driven down by the Science exam, which is a newer exam. It also has the largest variation in passing rates, which is likely due to the variety of subjects that may be tested. As a comparison, average Science MCAS exams among charter schools in this time period was 69.85% (standard deviation: 22.80), slightly below the statewide public average. For suburban charters, the average was 76.89% (18.43), and for urban charters, the average was 64.36% (24.39). In suburban public districts with charter schools, the average Science MCAS was 68.02% (17.36), and in urban public districts the average was 45.58% (14.92).

In contrast, charter schools outperform public schools statewide in terms of the English/Language Arts (ELA) and Math MCAS exam averages. Charter schools during this time period had an average ELS MCAS score of 89.84% (standard deviation: 14.28), with an average of 91.77% (13.88) in suburban districts and 88.30%

[23]It is calculated using district information such as the percent of special needs students, the percent of low income, and a wage adjustment factor (Massachusetts Department of Education 2015). The foundation budget quantifies, for an average district, "an adequate - but not excessive - level of funding (Edward Moscovitch, "Model School Budget". Cape Ann Economics, Rockport, Massachusetts, 1992, p1)."

[24]In 2011, this required contribution was determined by adding 0.3 percent of total property value to 1.4 percent of total income in the district. (Massachusetts Budget and Policy Center 2010).

(14.46) in urban districts. Average ELA MCAS exams among suburban public districts with charter schools was 84.14% (12.18) and among urban public districts was 73.08% (11.94). Finally, among Math MCAS, charter schools had an average score of 81.49% (standard deviation: 17.61), with an average of 83.62% (15.91) in suburban districts and 79.80% (18.73) in urban districts. In contract, suburban public districts with charter schools had average Math MCAS scores of 75.32% (14.15), whereas urban public districts had average Math MCAS scores of 61.17% (13.24). Thus, while suburban charters outperform urban charters on average, charters outperform public districts on average in both suburban and urban districts, and overall in terms of statewide testing in ELA and Math. This suggests that either charter schools are outperforming their public school counterparts, or they are effectively "cream-skimming" the better prepared students. In either scenario, public school districts might respond in a way that increases or decreases exam scores.

Across the state, the amount of competition varies substantially between districts with many alternative schools, and some with none. The difference between private and charter schools is large; the average district has 1.9% of students in charter schools and 12.0% of students in private schools. This is also visualized in the enrollment variables, which combine many districts with zero enrollments together with districts with positive enrollments. There is much variation in the demographic and income variables, implying vastly different needs across the state.

There is also quite a bit of variation in the subpopulations that might be considered higher risk. The average public district is comprised of 4% African American students, 8.6% Hispanic students, 17.2% of students requiring special education and 24.2% low income students, but these percentages vary; there are some districts that are almost entirely African American or low-income. Thus, it is important to include these variables to account for different needs by district.

Table 1: Descriptive Statistics of Dependent Variables in Massachusetts (Public Schools): 2008-2015

Variable	Observations	Mean	SD	Min	Max
Dependent Variables					
Average Passing Rate (%)	2,047	79.61	12.27	24.33	99.33
MCAS: Math (%)	2,047	79.59	12.34	9	100
MCAS: English/Language Arts (%)	2,047	87.38	10.51	37	100
MCAS: Science/Technology (%)	2,047	71.85	16.41	0	100
SAT: Math	1,966	518.13	45.98	382	717
SAT: Reading	1,966	506.53	43.86	368	661
Independent Variables					
Charter:Public Enroll (%)	2,047	1.885	7.663	0	71.06
Private:Public Enroll (%)	2,047	12.039	27.083	0	240.3846
Charter+Private:Public Enroll (%)	2,047	13.924	28.351	0	240.38
High School Enrollment (Public) [time t]	2,047	1105.635	1307.279	32	18521
High School Enrollment (Charter) [time t]	2,047	28.946	149.872	0	2676
High School Enrollment (Private) [time t]	2,047	153.03	379.054	0	4256
$ln(HSPublicEnroll_{t-1})$	2,047	6.752	0.671	3.93	9.84
% Administrators	2,047	6.010	1.685	2.99	17.81
Student:Teacher Ratio	2,047	13.486	1.95	5.6	23.2
% Teachers $<$ 26 years	2,047	4.755	3.003	0	41.9
% Teachers 26-32 years	2,047	18.168	6.645	0	51.6
% Teachers 33-40 years	2,047	20.311	4.848	4	37.3
% Teachers 41-48 years	2,047	20.014	4.978	0	39.3
% Teachers 49-56 years	2,047	21.456	5.961	0	45.1
% Teachers 57-64 years	2,047	14.098	4.412	0	36
% Teachers $>$ 64 years	2,047	1.197	1.192	0	16.6
Average Teacher Age	2,047	43.314	2.362	27.39	52.07
Attendance Rate	2,047	95.074	1.102	87.4	97.6
% African American	2,047	3.951	6.512	0	55.2
% Hispanic students	2,047	8.574	13.369	0	91.3
% Special Education	2,047	17.202	4.991	0	50.8
% Low Income	2,047	24.182	19.265	0	92.4
Chapter 70 Aid (millions)	2,047	15.152	30.443	0.25	301.59
% Grants	1,791	12.302	3.396	0.534	29.23

4 Methodology: OLS and Fixed Effects Model

I am interested in determining the effect of increased competition from choice schools on public school exam scores. The dependent variable, Y_{it} is either MCAS or SAT scores at a given time t for district i.[25] For the MCAS scores, I include average passing rates, and passing rates on the Math, English/Language Arts, and Science/Technology exams separately. The initial model is given below in equation

[25]District i may contain more than one high school. For example, Boston Public Schools are an aggregate for all Boston high schools, from English High School to Boston Latin Academy. On the other hand, Pentucket Regional High School includes students from Merrimac, West Newbury, and Groveland.

one, and estimated using ordinary least squares.

$$Y_{it} = \alpha_0 + \beta_1 ALT_{it} + \beta_2 X_{it} + \epsilon_{it} \tag{1}$$

The coefficient of interest, β_1, is shown in equation (1) and estimates the effect of increased alternative school penetration on public school test performance (ALT_{it}). However, the measure of alternative penetration is likely not exogenous, as the location of charter and private schools are not random. Districts with alternative schools are larger and serve more diverse and disadvantaged student populations than districts without, suggesting intrinsic differences between such districts. If alternative schools are more likely to exist in under-performing districts, the variable of interest, alternative enrollment, is endogenous, which will lead to inconsistent estimates.

To discern whether districts with charter or private schools are different than those without, I consider descriptive statistics in each. Districts with charter schools have, on average, significantly higher enrollments, a higher percentage of funding from grants, lower attendance rates, and higher percentages of African American, Hispanic, and low income students.[26] Districts with private schools also see, on average, higher enrollments, and greater percentages of African American and Hispanic students.[27] This suggests that there are intrinsic differences between districts with alternative schools and those without in terms of classroom needs, and these are important considerations in modeling decisions.

While I estimate the model above for reference, assuming uncorrelated error

[26] Districts with charter schools have an average of 2460 students, of which 9.1% are African American, 23.1% are Hispanic, 17.6% are special education students, and 42.0% are low-income. This is compared to districts without charter schools, with an average of 962 students, of which 3.4 % are African American, 7.0% are Hispanic, 17.2% are special education students, and 22.3% are low-income.

[27] Districts with private schools have an average of 1585 students, of which 5.5% are African American, 10.8% are Hispanic, 16.8% are special education students, and 26.2% are low-income. This is compared to districts without private schools, with an average of 837 students, of which 3.1% are African American, 7.3% are Hispanic, 17.4% are special education students, and 23.1% are low-income.

terms is likely too strong in this scenario as I risk incorrect standard errors and inefficient estimates. An error component model alleviates these concerns by removing time-invariant district-specific effects (α_i). Assuming an error term of the form: $e_{it} = \alpha_i + \epsilon_{it}$ and demeaning the variables leaves the model with an error (ϵ_{it}) uncorrelated with the explanatory variables as seen below in equations 2 and 3. A Hausman test indicates that a fixed effects model fits my data better than a random effects model and complete results are available upon request. [28] Further, in running a fixed effects regression, I note that in all cases, the errors (e_{it}) are correlated with the explanatory variables (X_{it}), although to varying degrees, as expected. This also indicates that I should not use a random effects model, in which the correlation is assumed to be zero. I cluster the standard errors by district in all specifications, as they allow the relaxation of the assumption that the error terms are identically distributed.

$$Y_{it} - \bar{Y}_i = \beta_1(ALT_{it} - A\bar{L}T_i) + \beta_2(X_{it} - \bar{X}_i) + (\alpha_i - \bar{\alpha}_i) + (\epsilon_{it} - \bar{\epsilon}_i) \tag{2}$$

That α_i does not depend on time allows the final equation to be of the form below, in equation (3).

$$Y_{it} - \bar{Y}_i = \beta_1(ALT_{it} - A\bar{L}T_i) + \beta_2(X_{it} - \bar{X}_i) + (\epsilon_{it} - \bar{\epsilon}_i) \tag{3}$$

4.1 The Effect of School Choice on Test Performance: Fixed Effects Model

Table 2 provides estimates of six specifications using average passing rates (AVPA) on the MCAS as the dependent variable. These specifications are: ordinary least squares (OLS) with charter and private enrollments separated and together, baseline fixed effects (FE) with charter and private enrollments separated and together, fixed effects

[28] Although the traditional Hausman test does not allow for clustered standard errors, the test statistic for the baseline model is 604.61. Using "xtoverid," which does allow for clustered standard errors, provides a Sargan-Hansen test statistic for the baseline model of 397.006.

with Chapter 70 Aid, fixed effects with grants, and fixed effects without vocational or agricultural schools. I provide different specifications to attempt to discern why the fixed effects model estimates vary so widely from the least squares estimates. I include Chapter 70 Aid in hopes of capturing the income effect. I include grant data to try to control for lower-income districts receiving more funding, and thus possibly offsetting some of the harm decreased enrollments might cause financially. Finally, I estimate the model in which all "vocational/technical" and "agricultural" schools are removed, as these schools typically focus on preparing students for training programs in specific fields and possibly put less emphasis on the MCAS and SAT. While they are still responsible for passing the MCAS exams, their academic goals likely differ from a typical public school student.

The standard errors are adjusted for clustering by district. In the ordinary least squares model, most variables are significant and have the expected sign. An increase in the ratio of charter enrollments has a positive effect on average passing rates, although the same effect is not seen when the measure of alternative school penetration is combined to include both charter and private school enrollments. Having older teachers, on average, has a negative impact on average passing rates. Higher attendance rates and larger districts have higher average pass rates, whereas the higher percentages of African Americans, Hispanics, low-income, or special education students lead to lower average passing rates.

Compared to the ordinary least squares models, the various fixed effects models show a break down in many of these effects, including the school choice penetration. Here, I see that an increase in the ratio of alternative to public enrollments has a negative effect, although it is no longer significant. An increase in the ratio of private schools to public schools has no significant effect on exam performance in any specification. Surprisingly, the estimated effect of the student to teacher ratio becomes positive, although not significantly so, and now implies that an increase in the ratio corresponds to an increase in average exam passing rates.[29] Further, the impact of district size switches signs, and significantly so. A one percentage point increase in the district enrollments leads to a 13-17% decline in average passing rates. The effect of attendance rates is still positive, but smaller in magnitude. Also surprisingly, the percentage of African American, Hispanic, and low income students seems to improve passing rates, contrary to much of the previous literature. While the impact of increased rates of special education students remains negative, and possibly picks up some of the overall negative effect of at-risk groups, the magnitudes do not compensate for the strong positive effect of increased Hispanic populations.

[29] Hanusheck (1986) reviews the education production function literature and suggests that while it might be expected that small student:teacher ratios would lead to academic gains, it does not seem to be the case across the board.

Table 2: Effect of Alternative Schools on Average Pass Rates

	OLS	OLS	FE$_{base}$	FE$_{base}$	FE$_{Ch70Aid}$	FE$_{Grants}$	FE$_{NoTech/Ag}$
Charter Ratio	0.078***		0.005				
	(0.029)		(0.055)				
Private Ratio	0.007		-0.011				
	(0.009)		(0.015)				
Alternative		0.013		-0.010	-0.010	-0.007	-0.012
Ratio		(0.009)		(0.014)	(0.015)	(0.015)	(0.014)
$ln(PubEnroll_{t-1})$	1.175**	1.117**	-13.034***	-13.059***	-13.716***	-13.750***	-17.027***
	(0.560)	(0.566)	(2.764)	(2.758)	(2.733)	(2.825)	(2.683)
% Admin	-0.010	-0.007	0.133	0.136	0.117	0.195	0.467**
	(0.166)	(0.167)	(0.230)	(0.229)	(0.234)	(0.229)	(0.191)
Student:	-0.349**	-0.330**	0.064	0.063	0.025	0.135	-0.186
Teacher	(0.165)	(0.167)	(0.186)	(0.186)	(0.184)	(0.182)	(0.179)
Average	-0.233*	-0.217*	-0.366**	-0.366**	-0.337*	-0.508**	-0.157
Teacher Age	(0.121)	(0.123)	(0.178)	(0.178)	(0.177)	(0.226)	(0.155)
Attendance	2.548***	2.499***	2.352***	2.352***	2.237***	2.425***	1.578***
Rate	(0.362)	(0.362)	(0.483)	(0.483)	(0.483)	(0.543)	(0.338)
% African	-0.170***	-0.163***	0.784*	0.785*	0.727	0.785*	0.587
American	(0.041)	(0.042)	(0.433)	(0.433)	(0.449)	(0.428)	(0.460)
% Hispanic	-0.294***	-0.289***	1.498***	1.504***	1.305***	1.577***	1.457***
	(0.032)	(0.033)	(0.149)	(0.148)	(0.167)	(0.154)	(0.138)
% SPED	-0.232***	-0.238***	-0.334***	-0.333***	-0.332***	-0.318***	-0.253***
	(0.072)	(0.071)	(0.054)	(0.054)	(0.055)	(0.054)	(0.060)
% Low Income	-0.221***	-0.221***	0.027**	0.027**	0.031***	0.023*	0.020*
	(0.019)	(0.019)	(0.011)	(0.011)	(0.011)	(0.012)	(0.011)
Chapter 70					0.186**		
					(0.075)		
% Grants						-0.151*	
						(0.084)	
Constant	-151.731***	-147.639***	-52.497	-52.457	-38.724	-49.029	41.613
	(35.406)	(35.317)	(54.021)	(54.005)	(54.219)	(59.102)	(39.241)
N	2047	2047	2047	2047	2047	1791	1808

OLS models include year dummies.
Standard errors clustered by district in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In Table 3A below, I present estimates using various dependent variables for the baseline fixed effects model. Thus, I replicate the fourth column from Table 2 for comparison. Still, there are very small negative effects of increasing the alternative enrollment ratio on public school test performance, as seen above, with the exception of the Math SAT. An increase in the alternative school ratio leads to a 0.09 point increase in average Math SAT scores. While this is a small value in terms of magnitude, that the effect is positive is encouraging. In Table 3B, I see that this positive effect is driven by private school enrollments; this means that districts with higher private school enrollments are more likely to see higher Math SAT scores. However, I also see that districts with greater private school enrollments decrease Math MCAS scores. Perhaps this is due to spillover effects in districts with private schools. Private schools are not required to administer the MCAS, which focuses on much different skills than does the SAT. Further, it is not surprising that the ELA and Reading exams did not produce significant results; many studies fail to find significant impacts on English and Reading scores. Most attribute the difficulty in estimating English effects to the vast array of ways in which students acquire literature skills, ranging from literacy at home to interest in classroom reading. The lack of evidence of penetration on the Science MCAS may be due to a number of things. The Science MCAS combines a range of possible exams, which means that schools can individually decide which course to offer freshman. Depending on the available resources, a school may decide to offer a number of courses, or only one. As there is more heterogeneity introduced, it is not surprising that it is difficult to determine an average effect.

Table 3A: Effect of Alternative Schools on MCAS/SAT Scores (Fixed Effects)

	MCAS				SAT	
	Average	Math	English	Science	Math	Reading
Alternate Schools Ratio	-0.010	-0.018	-0.003	-0.009	0.094*	0.033
	(0.014)	(0.012)	(0.018)	(0.022)	(0.050)	(0.040)
$ln(PubEnroll_{t-1})$	-13.059***	-5.903**	-19.372***	-14.103***	-6.598	-4.283
	(2.758)	(2.646)	(3.282)	(4.068)	(10.220)	(9.021)
% Administrators	0.136	-0.047	0.244	0.213	-1.056	-0.856
	(0.229)	(0.199)	(0.294)	(0.308)	(0.760)	(0.765)
Student:Teacher	0.063	0.022	0.122	0.049	1.319*	1.056
	(0.186)	(0.167)	(0.232)	(0.264)	(0.705)	(0.654)
Average Teacher Age	-0.366**	-0.265*	-0.629***	-0.219	-2.403***	-2.383***
	(0.178)	(0.135)	(0.237)	(0.226)	(0.548)	(0.543)
Attendance Rate	2.352***	1.861***	2.871***	2.276***	-1.013	-1.949
	(0.483)	(0.405)	(0.612)	(0.583)	(1.268)	(1.232)
% African American	0.785*	0.480	1.014**	0.866	-0.545	0.514
	(0.433)	(0.321)	(0.508)	(0.552)	(1.190)	(1.018)
% Hispanic	1.504***	0.630***	2.208***	1.671***	0.087	0.363
	(0.148)	(0.128)	(0.209)	(0.187)	(0.427)	(0.425)
% SPED	-0.333***	-0.440***	-0.290***	-0.269***	-0.022	0.041
	(0.054)	(0.075)	(0.063)	(0.076)	(0.134)	(0.140)
% Low Income	0.027**	0.007	0.029**	0.044***	0.148***	0.163***
	(0.011)	(0.016)	(0.014)	(0.013)	(0.044)	(0.041)
Constant	-52.457	-45.661	-49.263	-55.801	748.550***	804.664***
	(54.005)	(43.847)	(67.949)	(67.377)	(139.215)	(129.787)
N	Number of Observations: 2047.				Number of Observations: 1966.	

Standard errors clustered by district in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3B: Effect of Charter and Private Schools on MCAS/SAT Scores (Fixed Effects)

	MCAS				SAT	
	Average	Math	English	Science	Math	Reading
Charter Ratio	0.005	0.015	0.037	-0.037	0.126	0.265
	(0.055)	(0.035)	(0.089)	(0.068)	(0.210)	(0.207)
Private Ratio	-0.011	-0.021*	-0.006	-0.007	0.092*	0.015
	(0.015)	(0.013)	(0.019)	(0.023)	(0.053)	(0.041)
$ln(PublicEnroll_{t-1})$	-13.034***	-5.848**	-19.307***	-14.149***	-6.542	-3.874
	(2.764)	(2.651)	(3.289)	(4.071)	(10.236)	(9.075)
% Administrators	0.133	-0.054	0.236	0.219	-1.061	-0.897
	(0.230)	(0.200)	(0.296)	(0.309)	(0.760)	(0.763)
Student:Teacher	0.064	0.023	0.124	0.048	1.320*	1.064
	(0.186)	(0.167)	(0.232)	(0.264)	(0.706)	(0.653)
Average Teacher Age	-0.366**	-0.264*	-0.628***	-0.220	-2.402***	-2.376***
	(0.178)	(0.135)	(0.237)	(0.225)	(0.547)	(0.543)
Attendance Rate	2.352***	1.859***	2.869***	2.277***	-1.015	-1.965
	(0.483)	(0.405)	(0.612)	(0.583)	(1.268)	(1.229)
% African American	0.784*	0.478	1.012**	0.867	-0.546	0.510
	(0.433)	(0.321)	(0.508)	(0.553)	(1.190)	(1.008)
% Hispanic	1.498***	0.616***	2.191***	1.683***	0.073	0.263
	(0.149)	(0.130)	(0.210)	(0.192)	(0.441)	(0.431)
% SPED	-0.334***	-0.441***	-0.291***	-0.268***	-0.023	0.037
	(0.054)	(0.075)	(0.064)	(0.076)	(0.135)	(0.140)
% Low Income	0.027**	0.007	0.029**	0.044***	0.148***	0.163***
	(0.011)	(0.016)	(0.014)	(0.013)	(0.044)	(0.041)
Constant	-52.497	-45.750	-49.368	-55.728	748.457***	803.987***
	(54.021)	(43.841)	(67.980)	(67.420)	(139.285)	(129.794)
N	Number of Observations: 2047.				Number of Observations: 1966.	

Standard errors clustered by district in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

After estimating these models for the full sample, I consider urban districts separately from suburban districts to discern whether there are differential impacts. Several Massachusetts studies have shown that in urban areas, charter schools are outperforming their public counterparts (Angrist et al. 2011). The Boston Foundation (2009) showed that charter schools in Boston were able to eliminate half of the achievement gap in one year. While choice schools statewide are not showing positive effects, perhaps there is a difference between urban and suburban districts. Using the U.S. Census definition, I define urban areas to have populations greater than 50,000. Suburban areas, or urban clusters, have populations between 2,500 and 50,000. Rural areas have populations less than 2,500. Using the U.S. Census in 2010, I have classified each district as urban (3), suburban (2), or rural (1) (U.S. Census

Table GCT-PL2 2010). There are almost no districts in Massachusetts that meet the "rural" classification, so I group rural and suburban districts together. I then separated the data set into an "urban" sample and a "suburban" sample, and looked at choice effects separately.

Table 4: Effect of Alternative Schools on Achievement: Baseline Fixed Effects By Urbanicity

	Full Sample	Urban Sample	Suburban Sample
MCAS: Average Passing Rate	-0.01	0.277	-0.014
	(0.014)	(0.205)	(0.014)
MCAS: Math	-0.018	0.115	-0.020*
	(0.012)	(0.153)	(0.011)
MCAS: English/Language Arts	-0.003	0.272	-0.005
	(0.018)	(0.260)	(0.018)
MCAS: Science	-0.009	0.444*	-0.016
	(0.022)	(0.245)	(0.021)
N	2047	256	1791
SAT: Math	0.094*	0.422	0.092*
	(0.050)	(0.674)	(0.049)
SAT: Reading	0.033	0.565	0.026
	(0.040)	(0.682)	(0.039)
N	1966	245	1721

Standard errors clustered by district in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

In Table 4, a one percentage point increase in the ratio of alternative to public enrollments leads to an 0.01 percent decrease in average passing rates, although it is not significant. However, the sign is driven by the suburban sample. While not statistically significant, once these effects are broken into the urban and suburban samples, an increase in the alternative to public enrollments in urban locations has a positive effect, but an increase in the ratio of private to public schools has a negative effect.

The negative effect of increased alternative schooling on the Math MCAS is driven by the negative effect in suburban districts. As suburban districts are more likely to be impacted by increased private school enrollments, this implies a possible cream-skimming effect happening due to private schools in these districts. Interestingly, while there is no significant effect overall on the Science MCAS, a percentage point increase in alternative enrollments leads to a 0.44% increase in Science MCAS scores in urban districts. This suggests that urban districts with greater school choice induce higher MCAS scores on the Science exam in public schools. This may in part be due to urban districts starting with such low average Science passing rates, but it could also imply that schools are able to redistribute resources in a way that benefits students in Science courses.

The SAT tests material at a higher level than the MCAS, and is a large component of college acceptance. Public schools facing significant enrollment losses to charter schools may respond by spending additional resources on SAT preparation and boosting averages to better market themselves as compared to their competition. While I see no significant effect from increased alternative schooling on average Reading scores, there is a positive effect on Math SAT scores. A one percentage point increase in alternative enrollments leads to a 0.09 point increase in average Math SAT scores. This penetration is driven almost entirely by suburban schools. In light of the MCAS results above, this seems to suggest that suburban districts may be shifting resources from MCAS preparation to SAT preparation.

5 Methodology: Control Function Model

The issue of endogeneity is critical, and as it is often difficult to determine a valid instrument, it is useful to analyze alternative methods to correct for the biased estimates. With the fixed effects model, having a short panel with little within district variation proves to be problematic, as the identification comes from within-district variation alone. Even with time variant covariates, it can be difficult to identify an effect if the variation is small. Finally, using the differences in means model for fixed effects, while taking care of individual unobserved effects, may cause more measurement error than in levels, which leads to attenuation bias and potentially smaller estimates (Angrist and Pischke 2009). The bias, shown below, between the fixed effects estimator on the left and ordinary least squares on the right, may be worse with little variation. While one expects the numerator of the fixed effects bias to be smaller, it is also plausible that the denominator is smaller as well.

$$\frac{cov(\epsilon_{it}, ALT_{it} - A\bar{L}T_{it})}{var(ALT_{it} - A\bar{L}T_{it})} \geq \frac{cov(\alpha_i + \epsilon_{it}, ALT_{it})}{var(ALT_{it})}$$

If this is the case, I am able to consider a two step procedure to estimate the effect of competition on achievement. Using maximum likelihood estimation, I estimate equations (4) and (5) below. I first predict the ratio of alternative enrollments via a Tobit model, and then determine the effect of those on public school outcomes. Without a valid instrument, I use the heteroskedasticity in the error terms to identify the model. This is called the control, and in the model is ρ.

$$Y_{it} = \beta_0 + \beta_1 A\hat{L}T_{it} + \beta_2 X_{it} + \rho_{it} + u_{it} \tag{4}$$

$$ALT_{it} = \gamma_0 + \gamma_1 X_{it} + v_{it} \tag{5}$$

The derivation of the model and the construction of ρ are provided in Appendix E.[30] While this model allows for time varying district specific effects in addition to providing an alternative method for identifying the effect of competition, it does not control for the panel structure of the data and treats it as a cross-section. This could lead to overstating the effect. Further, the standard errors are not clustered, which will not impact the estimates but could lead to overstating the significance of the estimates.

[30]Further, I ran the model using simulated data and the results are presented in Appendix D. With close starting values, the program estimates the true model quite well.

5.1 The Effect of School Choice on Test Performance: Control Function Model

Using the above model, I estimate the effect of alternative schooling on public school test performance. While not presented, the reduced form model estimates the effect of competition separately for districts which have alternative schools from those that do not. Treating districts that do not have alternative schools differently from those that do controls for the fact that areas with alternative schools are not randomly chosen. The results are below in Table 5, and show the same baseline model as above, with the exception that the control variable is also included (ρ). Except for the model in which the Science MCAS is the dependent variable, ρ is significant. This implies that heteroskedasticity enters the model in a significant way, and further bolsters use of this methodology.

Table 5: Effect of Alternative Schools on Achievement: Control Function

	MCAS				SAT	
	Average	Math	English	Science	Math	Reading
Alternative Schools Ratio	0.023***	0.027***	0.011**	0.030**	0.126***	0.114***
	(0.008)	(0.008)	(0.005)	(0.015)	(0.043)	(0.043)
$ln(PubEnroll_{t-1})$	0.903***	1.269***	0.245**	1.202***	6.250***	3.279***
	(0.149)	(0.163)	(0.099)	(0.257)	(0.822)	(0.799)
% Administrators	0.306**	0.005	0.272***	0.490**	-4.967***	-4.469***
	(0.133)	(0.225)	(0.090)	(0.222)	(0.885)	(0.853)
Student:Teacher	-0.812***	-1.051***	-0.507***	-0.877***	0.285	0.235
	(0.143)	(0.158)	(0.094)	(0.251)	(0.719)	(0.728)
Average Teacher Age	-0.293**	-0.189	-0.151*	-0.601***	-0.831	-0.195
	(0.128)	(0.136)	(0.083)	(0.229)	(0.804)	(0.821)
Attendance Rate	2.026***	2.073***	1.559***	2.520***	5.266***	3.289***
	(0.207)	(0.213)	(0.144)	(0.331)	(1.128)	(1.141)
% African American	-0.486***	-0.043	-0.358***	-1.137***	-6.579***	-6.175***
	(0.155)	(0.174)	(0.115)	(0.251)	(0.794)	(0.821)
% Hispanic	-2.636***	-1.897***	-2.315***	-4.240***	-7.808***	-8.940***
	(0.280)	(0.333)	(0.187)	(0.420)	(1.226)	(1.175)
% SPED	-1.226***	-1.597***	-0.584***	-1.168***	-3.878***	-3.474***
	(0.159)	(0.164)	(0.116)	(0.268)	(0.853)	(0.844)
% Low Income	-6.829***	-8.118***	-3.538***	-7.565***	-15.070***	-14.807***
	(0.305)	(0.313)	(0.202)	(0.469)	(1.424)	(1.432)
Constant	-88.150***	-90.464***	-43.698***	-136.258***	63.527	238.517**
	(18.366)	(18.906)	(12.784)	(29.688)	(101.641)	(102.424)
Control	12.378***	19.398***	16.822***	5.714	-14.208***	-8.223**
	(4.198)	(4.226)	(4.175)	(4.130)	(3.680)	(3.812)
	Observations: 2047.				Observations: 1966.	

Year dummies included but not reported.
Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

In Table 5, I see that having greater relative enrollment in alternative schools leads to increased test scores in all exams. Of course, this model does not directly correlate to the fixed effects model, as the control function method does not account for the panel nature of the data. It is, however, similar to the ordinary least squares estimates, except that now, the model underlying alternative schools is corrected for via the Tobit model approximation in the first stage. All three models do include the same covariates, with year dummies included in the ordinary least squares and control function estimates.

Here, I see that a percentage point increase in the ratio of alternative enrollments to public enrollments leads to a 0.023% increase in average MCAS scores, a 0.027% increase in Math MCAS scores, 0.011% increase in ELA MCAS scores, and a 0.03%

increase in Science MCAS scores. Similarly, a percentage point increase in the ratio leads to a 12.6 point increase in Math SAT scores, and a 11.4 point increase in Reading SAT scores. While the magnitudes on the MCAS exam seem small, the point increases on the SAT scores are much larger than previously suggested.

A drawback to this model is that it relies on larger samples and thus, did not converge when estimating the effects when broken down into urban or suburban samples. However, these results do suggest that where the fixed effects models were not able to pin down an effect, the control function and maximum likelihood approach are.

6 Results and Conclusion

The effect of increased school choice on public school students is an ongoing concern. On one hand, school choice provides options for families who might seek alternatives to public school. However, more options leaves traditional public schools with fewer students and funds available. This might induce public schools to reevaluate budget decisions and optimize differently, leading to increased efficiency, or it might be detrimental to school policies and decisions. In this paper, I consider the effect of increased choice on test scores of public high school students. This is certainly not the only measure of competition, but as graduation and scholarship funds are determined by test scores, it is certainly an important facet to consider.

Ordinary least squares estimates suggest a positive effect of increased charter school penetration but negligible effects from private school enrollments on average passing rates of public school students on the MCAS. As private schools do not administer the MCAS exams, it is not surprising that their existence has negligible effects. However, the availability of alternative schooling is a consequence of poor historical results in public schools, as charter and private schools are more likely to open in areas with poor performing public schools. Thus, endogeneity is a concern.

To account for non-random locational decisions, fixed effects seemed to be an ideal model given the time series nature of the data. When considering all districts together, the effect of charter schools on MCAS passing rates is positive. On the other hand, private school enrollments have a smaller negative, although generally insignificant, effect. The finding that private school enrollment increases have little impact on public school performance is likely due to the fact that private schools are not as close of a substitute as compared to charter schools. Private schools have the ability to choose their student body and can do this on the basis of test scores or applications. In addition, private schools charge tuition; this is a substantial departure from public education. Separating schools by their urbanicity demonstrates that the positive effect seen on the Math SAT is being driven by suburban districts.

The overall lack of impact from competition in the fixed effects models could be a result of several factors. One, many charter schools in Massachusetts have been in operation for almost two decades, and private schools for much longer; it is possible that the competitive effect school choice imposes has dissipated over time. It is also likely that the competitive effect is spread across districts; many choice schools accept students from several neighboring districts and my measurement of "penetration" is probably not picking up the full effect. It is also possible that the charter school cap dampens the potential effect of charter schools. If all students on wait lists were granted seats in a charter school, there would be a significant difference in enrollments. It is also possible that the funding formula in its current state does not allow for true competition. While public schools do lose funds, they are still compensated for the loss in enrollment, yet no longer have to educate the student. Also, rather than through academic performance, it is possible the competitive effect is seen through other, non-academic outcome variables, such as effects on incarceration rates or lifetime earnings.

Finally, it is possible that controlling for district specific time invariant effects are not enough to estimate the effect of increased alternative school enrollments. With small variations over time in both the variable of interest and other covariates, it is possible that the bias on the estimate after differencing is worse than the bias on the estimate without. The combination of little growth in alternative schooling during this time period together with small variations in the covariate data may mean that fixed effects are not able to capture the true nature of the penetration.

If this is the case, modeling the effect of alternative schooling on public school test performance using a control function method is optimal. These results suggest a positive effect of increased enrollments in alternative schools, similar to the OLS results. This suggests that the fixed effects method in this case might not be sufficiently solving the endogeneity problem, and in fact, might make the bias worse. The control function model as is, requires additional robustness checks to ensure it is adequately accounting for the locational decisions of alternative schools and capturing the true penetration of greater school choice on public districts. However, the results are promising in that they show expected signs of other included covariates. Further, the fact that the coefficient on the "control" variable is incredibly significant in all but the Science exam model indicate that utilizing the heteroskedasticity of the model is indeed important.

Massachusetts differs from many of the states studied previously in demographic make-up, academic performance statewide, and median income. For example, aside from the urban areas of Massachusetts, students are primarily white. Education across the state is consistently ranked among the best in the United States (United States Education Dashboard). The overall population of Massachusetts is much smaller than many studied states, such as Texas, California, or Florida. Finally, per capita income in Massachusetts is high. These attributes enter into how a state responds to an influx in competition, and Massachusetts' schools may react differently than others. Finally, the political atmosphere plays a large role in both the creation and expansion of charter schools, and attitudes toward charter schools can be an important influence with respect to policy change or stagnation. Nevertheless, it is an interesting state in which to consider the effect of school choice especially as many studies have shown the school choice options for Massachusetts' students to be performing well (Angrist et. al 2011, CREDO 2013).

With many urban districts across the country failing to adequately prepare students for future endeavors after high school completion, it is crucial that the methods that induce positive change are identified and replicated throughout traditional public schools. Charter schools in Massachusetts have seen increasing demand through the years. With 40,376 unique students on wait lists for a seat at a charter school, it it imperative to determine why parents are so adamantly seeking this choice, and what traditional public schools can do to alleviate some of this demand (Massachusetts Department of Education 2013c). Introducing school choice should, in theory, at least maintain the status quo. A much better outcome, however, would be if increased choice improved outcomes of all students, not just the students able to utilize that choice. My results suggest that competitive effects are sensitive to modeling decisions. While alternative schooling locations are not exogenously determined, estimating the effect of choice is not a straightforward endeavor.

Appendices

Appendix A: Charter School Towns [31]

Charter School	Town
Abby Kelley Foster Charter Public	Worcester
Academy Of the Pacific Rim Charter Public	Hyde Park (Boston)
Advanced Math and Science Academy Charter	Marlborough
Amesbury Academy Charter Public	Amesbury
Atlantis Charter	Fall River
Berkshire Arts and Technology Charter Public	Adams
Boston Collegiate Charter	Dorchester (Boston)
Boston Day and Evening Academy Charter	Roxbury (Boston)
Boston Green Academy Horace Mann Charter School	South Boston (Boston)
Boston Preparatory Charter Public	Hyde Park (Boston)
City On A Hill Charter Public (Circuit Street and Dudley Square)	Roxbury (Boston)
City On A Hill Charter Public (New Bedford)	New Bedford
Codman Academy Charter Public	Dorchester (Boston)
Community Charter School of Cambridge	Cambridge
Edward M. Kennedy Academy for Health Careers (HCMS)	Boston
Four Rivers Charter Public	Greenfield
Foxborough Regional Charter	Foxborough
Francis W. Parker Charter Essential	Devens/Harvard
Global Learning Charter Public	New Bedford
Hampden Charter School of Science	Chicopee
Innovation Academy Charter	Tyngsborough
KIPP Academy Lynn Charter	Lynn
Lowell Middlesex Academy Charter	Lowell
Martha's Vineyard Charter	Martha's Vineyard
MATCH Charter Public School	Boston
Mystic Valley Regional Charter	Malden
New Leadership Charter	Springfield
New Liberty Charter	Salem
North Central Charter Essential	Fitchburg
Paulo Freire Social Justice Charter	Holyoke
Phoenix Academy Public Charter	Springfield
Phoenix Charter Academy	Chelsea
Pioneer Charter School of Science	Everett
Pioneer Charter School of Science II	Saugus
Pioneer Valley Chinese Immersion	Hadley
Pioneer Valley Performing Arts Charter Public	South Hadley
Prospect Hill Academy Charter	Cambridge
Rising Tide Charter Public	Plymouth
Sabis International Charter	Springfield
Salem Academy Charter	Salem
Salem Community Charter School	Salem
Spirit of Knowledge Charter School	Worcester
South Shore Charter Public	Norwell
Sturgis Charter Public	Hyannis/Barnstable
The Sizer School: A North Central Charter Essential	Fitchburg

[31] Prospect Hill Academy opened in Somerville, but the high school is located in Cambridge. (http://www.masscharterschools.org 2/4/13, http://www.doe.mass.edu 9/11/13).

Appendix B: Regional School Districts

Town	Corresponding District	Town	Corresponding District
Ashfield	Mohawk Trail Regional	Peru	Central Berkshire Regional
Barre	Quabbin Regional	Plainfield	Mohawk Trail Regional
Becket	Central Berkshire Regional	Plympton	Silver Lake Regional
Bernardston	Pioneer Valley Regional	Princeton	Wachusett Regional
Bolton	Nashoba Regional	Rochester	Old Rochester Regional
Brewster	Nauset Regional	Rowe	Mohawk Trail Regional
Buckland	Mohawk Trail Regional	Rowley	Trition Regional
Charlemont	Mohawk Trail Regional	Rutland	Wachusett Regional
Colrain	Mohawk Trail Regional	Salisbury	Trition Regional
Conway	Frontier Regional	Shelburne	Mohawk Trail Regional
Cummington	Central Berkshire Regional	Sterling	Wachusett Regional
Dalton	Central Berkshire Regional	Stockbridge	Berkshire Hills Regional
Deerfield	Frontier Regional	Stow	Nashoba Regional
Eastham	Nauset Regional	Sunderland	Frontier Regional
Halifax	Silver Lake Regional	Warren	Quaboag Regional
Hardwick	Quabbin Regional	Warwick	Pioneer Valley Regional
Hawley	Mohawk Trail Regional	Washington	Central Berkshire Regional
Heath	Mohawk Trail Regional	Wellfleet	Nauset Regional
Hinsdale	Central Berkshire Regional	West Brookfield	Quaboag Regional
Holden	Wachusett Regional	West Stockbridge	Berkshire Hills Regional
Housatonic	Berkshire Hills Regional	Whately	Frontier Regional
Hubbardston	Quabbin Regional	Williamstown	Mount Greylock Regional
Glendale	Berkshire Hills Regional	Windsor	Central Berkshire Regional
Great Barrington	Berkshire Hills Regional		
Kingston	Silver Lake Regional		
Interlaken	Berkshire Hills Regional		
Lancaster	Nashoba Regional		
Lanesborough	Mount Greylock Regional		
Leyden	Pioneer Valley Regional		
Marion	Old Rochester Regional		
Mattapoisett	Old Rochester Regional		
Newbury	Trition Regional		
New Ashford	Mount Greylock Regional		
New Braintree	Quabbin Regional		
Northfield	Pioneer Valley Regional		
Oakham	Quabbin Regional		
Orleans	Nauset Regional		
Paxton	Wachusett Regional		

Appendix C: Data Appendix

The administrative data was organized using the MA Department of Education and choosing the titles that included "supervisor" or indicated a position of power, such as "principal" or "administrator". I summed the total number of full-time equivalents in each position for each district and calculated the percentage out of total full-time equivalents. The titles are: Superintendent, Assistant/Associate/Vice Superintendent, School Business Official, Other District Wide Administrators, Special Education Administrator, Principal/Headmaster(mistress)/Head of school, Deputy/Associate/Vice/Assistant Principal, Other School Administrator/Coordinator, and Supervisor/Director of Guidance, Pupil Personnel, Arts, Assessment, Curriculum, English Language Learners, English, Foreign Language, History/Social Studies, Library/Media, Math, Reading, Science, Technology and Professional Development.[32]

As an alternative to using each percentage of teacher category, I calculated "Average Teacher Age." This variable was found by a weighted average of the percent of teachers in various age brackets by district. This composite variable was created by multiplying the mean age in each category by the district's total number of "full-time equivalents" in that category, summing these and dividing by the total number of full-time equivalents. That is, I found:

$$\begin{aligned} \text{AvTeachAge} = {} & (\%{<}26)*24 + (\%26\text{-}32)*29 + (\%33\text{-}40)*36.5 + (\%41\text{-}48)*44.5 \\ & + (\%49\text{-}56)*52.5 + (\%57\text{-}64)*60.5 + (\%{>}64)*66 \end{aligned}$$

Across districts, the average teacher age was 43.37 years old, with a standard deviation of 2.323. The lowest district's average was 37.14 years old and the highest 50.73 years old.

Percent grants is found by calculating the fraction of total expenditures from "non-appropriated revenue sources from federal, state and private grants."[33] At the time of this study, it was only available through 2013, and thus requires sacrificing a year of data. Alternatively, I considered Chapter 70 funding. In Massachusetts, each district has a foundation budget, which determines the amount the district is able to contribute to the school system. Once the foundation budget for the district is determined, the required local contribution is calculated by property values and income levels. The gap between the foundation budget and the required local contribution is Chapter 70 Aid. Across districts, aid is quite large, and thus in the dataset, Chapter 70 Aid will always be reported in millions of dollars.

[32] `http://www.doe.mass.edu/` 9/11/13.
[33] `http://profiles.doe.mass.edu/help/data.aspx` 10/13/13.

Appendix D: Data Simulation

Using 100 replications, I estimated the model $Y_1 = 5x_1 + 7x_2 + 2y_2 + u$, where $Y_2 = 3x_1 + 4x_2 + v$ if $Y_2 > 0$ and 0 otherwise. Using a random data simulator to generate x_1, x_2, v^*, and u^*, where v^*, u^* are the random components of the error, I estimate the above model with heteroskedastic errors formulated in the model. Specifically, I let $v = \sqrt{exp(x_1 + x_2)}v^*$ and $u = \sqrt{exp(2x_1 + x_2)}u^*$. Running 100 simulations provides estimates for the y_1 model shown below, with standard deviations also shown. That the estimates are so close to the truth provides evidence that the model is providing unbiased results given the constraints of a Tobit first stage and heteroskedastic errors.

	Coefficient	Standard Deviation
x_1	5.00006	0.0009
x_1	6.99997	0.0008
$Constant$	0.0001	0.0022
y_2	2.00003	0.0024

I additionally estimated $Y_1 = 2x_1 + 3x_2 - 2y_2 + u$ and $Y_2 = 2x_1 + 3x_2 + v$ when $Y_2 > 0$ and zero otherwise to ensure that the model retains its unbiased estimates when the coefficient of interest is negative. The results for the Y_1 model are below:

	Coefficient	Standard Deviation
x_1	1.9998	0.0009
x_1	3.0000	0.0009
$Constant$	-0.0001	0.0024
y_2	-1.9998	0.0026

Appendix E: Control Function Derivation

The main difficulty in this approach is the construction of ρ. The approach in Klein and Vella (2010) does not assume functional forms for ρ, nor does it make any distributional assumptions. However, this method requires a much larger sample size to work with in practice than is available here. Furthermore, there is an added complication in my application. Namely, many districts did not have any charter penetration in the years studied. To deal with these issues, I employ a parametric maximum likelihood approach. I assume a Tobit model for the reduced form equation with a potentially heteroskedastic error term. This parametric method will work in practice with much smaller sample sizes than the Klein-Vella approach while retaining heteroskedasticity as the basic identification strategy. Moreover, it will allow for the sizable fraction of districts with zero alternative schools. First, I will construct ρ, and then I will go through the derivation of the likelihood functions in two cases: (1) district i has alternative school enrollments and (2) district i does not have alternative school enrollments.

The estimation of the effect of competition requires the construction of ρ, a critical component to this model. To construct ρ, I first assume $\rho = \frac{s_u}{s_v}\hat{v}$ (Klein 2009), where the numerator is constructed from the outcome equation and the denominator from the reduced form equation. As the outcome equation involves an endogenous variable, determining s_u is more difficult than s_v.

To estimate s_v, I first run ordinary least squares on the reduced form equation, and get the residuals $\hat{v}$. Following Klein et al. (2009, 2010), I assume that $s_v^2 = e^{x\delta}$. Since $v = s_v v^*$, $v^2 = s_v^2 v^* = e^{x\delta} v^*$. This also implies that $ln(v^2) = x\delta + v^*$.[34] Thus, by regressing $ln(\hat{v}^2)$ on X, I am able to recover estimates $\hat{\delta}$, that allow the construction of $s_v = e^{\hat{\delta}X}$. This generates the denominator of the control (ρ).

To determine the numerator of ρ, s_u, more work is involved as the residuals,

[34] The exponential function is used to force positive values, which is required in order to use the natural logarithm.

$\hat{u}$, cannot be estimated via ordinary least squares alone. This is because there is an endogenous variable in the model. In this case, I use ordinary least squares estimates as starting values, similar to the above method, where I first regress Y_i on both the alternative penetration variable and X_i. By taking the residuals, $\hat{u}$, and regressing $ln(\hat{u}^2)$ on X, I obtain starting values of λ and subsequently estimate $s_u^2 = e^{x\lambda}$ using maximum likelihood estimation.

Combining the starting estimates for λ and δ with the estimates resulting from regressing Y_i on X_i, ALT_i, ρ and from regressing ALT_i on X_i, I obtain a full list of starting values. Since the number of starting values more than quadruples with each additional covariate, this model requires a trade-off with respect to efficiency as well as convergence. Using the likelihood functions provided below, I estimate equation 4 and 5 via maximum likelihood, maximizing with respect to β, γ and ρ. A significant coefficient on ρ implies the control belongs in the model, or in other words, that the heteroskedasticity enters the model in a significant way.

As mentioned above, in the reduced form equation, I estimate a Tobit model. That is, there are some districts ($d_1 = 1$) in which there are alternative school enrollments and some districts ($d_1 = 0$) in which there are not. I first consider the case in which there is positive enrollment ($d_1 = 1$), which will have a likelihood function as shown below.

$$f(Y_i, ALT_i|X_i) = f(ALT_i|X_i)f(Y_i|ALT_i, X_i) \tag{6}$$

This is equivalent to the log-likelihood below:

$$lnf(Y_i, ALT_i|X_i) = lnf(ALT_i|X_i) + lnf(Y_i|ALT_i, X_i) \tag{7}$$

Assuming $ALT_i \sim N(X_i\gamma, s_{v_i}^2)$, the first term, $lnf(ALT_i|X_i)$ is:

$$ln\Big(\frac{1}{\sqrt{2\pi}}\frac{1}{s_v}exp\Big(-\frac{1}{2}\frac{(ALT_i - X_i\gamma)^2}{s_v^2}\Big)\Big) = ln\frac{1}{\sqrt{2\pi}} + ln\frac{1}{s_v} - \frac{1}{2}\frac{(ALT_i - X_i\gamma)^2}{s_v^2} \tag{8}$$

This can be simplified to the log-likelihood function shown below, since $ALT_i = X_i\gamma + v_i$ (Equation 5), and $\frac{1}{\sqrt{2\pi}}$ will not affect the choices for maximization:

$$ln\frac{1}{s_v} - \frac{1}{2}\frac{(ALT_i - X_i\gamma)^2}{s_v^2} = -\frac{1}{2}\Big(ln(s_v^2) + \frac{v^2}{s_v^2}\Big) \tag{9}$$

Similarly, for the second term, $lnf(Y_i|ALT_i, X_i)$, I first note that:

$$E[Y_i|ALT_i, X_i] = \beta_0 + X_i\beta_1 + ALT_i\beta_2 + E[u_i|ALT_i, X_i] \tag{10}$$

Also, $Y_i|X_i, ALT_i$ is equivalent to $Y_i|X_i, v$. Then, assuming u and v are jointly normal, $E[u|v, X] = E[u|X] + \rho\frac{s_u}{s_v}(v - E[v|X]) = \rho\frac{s_u}{s_v}v$. This is due to the adjustment necessary when I "learn" information about v and X. If, for example, the variance of v is very large compared to u, $\frac{s_u}{s_v}$ will be very small, and will not affect the expectation of u by much. Similarly, if the actual value of v is higher than its expectation, this will adjust the expectation for u up. Finally, $var(Y_i|X_i, v) = (1 - \rho^2)s_u^2$. Note that as ρ increases, knowing v provides full information regarding u towards perfect correlation. If ρ is very small, this implies that knowing v gives very little information about u, and the variance rises. Thus,

$$Y_i|X_i, ALT_i \sim N\Big(\beta_0 + \beta_1 X_i + ALT_i\beta_2 + \rho\frac{s_u}{s_v}v, (1 - \rho^2)s_u^2\Big) \tag{11}$$

Then, $lnf(Y_i|X_i, ALT_i)$ is:

$$ln\Big(\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{(1 - \rho^2)s_u^2}}exp\Big(-\frac{1}{2}\frac{(Y_i - \beta_0 - X_i\beta_1 - ALT_i\beta_2 - \rho\frac{s_u}{s_v}v)^2}{s_v^2}\Big)\Big) \tag{12}$$

This simplifies to:

$$ln\Big(\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{(1-\rho^2)s_u^2}}exp\Big(-\frac{1}{2}\frac{(u-\rho\frac{s_u}{s_v}v)^2}{s_v^2}\Big)\Big) = ln\frac{1}{\sqrt{2\pi}}+ln\frac{1}{\sqrt{(1-\rho^2)s_u^2}}-\frac{1}{2}\frac{(u-\rho\frac{s_u}{s_v}v)^2}{s_v^2} \tag{13}$$

Or, similar to the above equation, this is equivalent to maximizing:

$$-\frac{1}{2}\Big(ln[(1-\rho^2)s_u^2]+\frac{(u-\rho\frac{s_u}{s_v}v)^2}{s_v^2}\Big) \tag{14}$$

Equation 14 will work when there are positive enrollments from alternative schools in district i. However, this only accounts for slightly over half of the districts. To estimate the model when there is no competition from alternative schools ($d_1 = 0$), I need to solve: $f(Y_i, ALT_i|X_i) = f(ALT_i|Y_i, X_i)f(Y_i|X_i)$, where the left-hand side in this setting is equivalent to $f(Y_i, 0|X_i)$. Thus, I have

$$f(Y_i, 0|X_i) = f(X_i\beta + u, 0|X_i) = Pr(ALT_i = 0|X_i\beta + u, X_i)f(X_i\beta + u|X_i) \tag{15}$$

I again take natural logarithms, so the equation I solve is: $lnf(X_i\beta+u, 0|X_i) = ln[Pr(ALT_i = 0|X_i\beta+u, X_i)]+lnf(X_i\beta+u|X_i)$. The second term, $lnf(X_i\beta+u|X_i)$, is the expression below, since I assume $Y_i|X_i \sim N(X_i\beta, s_u^2)$. Thus, I have:

$$ln\Big(\frac{1}{\sqrt{2\pi}}\frac{1}{s_u}exp\Big(-\frac{1}{2}\frac{(Y_i-X_i\beta)^2}{s_u^2}\Big)\Big) = ln\frac{1}{\sqrt{2\pi}}+ln\frac{1}{s_u}-\frac{1}{2}\frac{(Y_i-X_i\beta)^2}{s_u^2} = -\frac{1}{2}\Big(lns_u^2+\frac{u^2}{s_u^2}\Big) \tag{16}$$

The first term, $Pr(ALT_i = 0|X_i\beta + u, X_i)$ is equivalent to:

$$Prob(X_i\Pi + v \leq 0|X_i\beta + u, X_i) \tag{17}$$

Similar to the above, $E[v|X, u] = E[v|X] + \rho\frac{s_v}{s_u}(u - E[u(X)])$. Thus, since

$E[v|X] = E[u|X] = 0$, this simplifies to $\rho\frac{s_v}{s_u}u$. Similarly, $Var(v|u, X) = s_v^2(1-\rho^2)$, following the proof above. Then, assuming $v|X, u \sim N(\rho\frac{s_v}{s_u}, (1-\rho^2)s_u^2)$, I can re-scale the above to normalize the distribution:

$$Pr\Big(\frac{X_i\Pi + \rho\frac{s_v}{s_u}}{\sqrt{(1-\rho^2)s_u^2}} + \frac{v - \rho\frac{s_v}{s_u}}{\sqrt{(1-\rho^2)s_u^2}} \leq 0 \Big| X_i\beta + u, X_i\Big) \tag{18}$$

Let $Z \equiv \frac{v-\rho\frac{s_v}{s_u}}{\sqrt{(1-\rho^2)s_u^2}}$ and $Z \sim N(0,1)$. Then, I have:

$$Pr\Big(\frac{X_i\Pi + \rho\frac{s_v}{s_u}}{\sqrt{(1-\rho^2)s_u^2}} \leq -z \Big| X_i\beta + u, X_i\Big) \tag{19}$$

The above is equivalent to

$$1 - cdf\Big(\frac{X_i\Pi + \rho\frac{s_v}{s_u}}{\sqrt{(1-\rho^2)s_u^2}}\Big) \tag{20}$$

Thus, using Equation 14 for districts in which there are alternative enrollments and Equation 20 when there are not, I am able to estimate the effect of competition from alternative schools on public school test performance using the control function method.

3. The Path to a Bachelor's Degree: The Effect of Starting at a Community College

1 Introduction

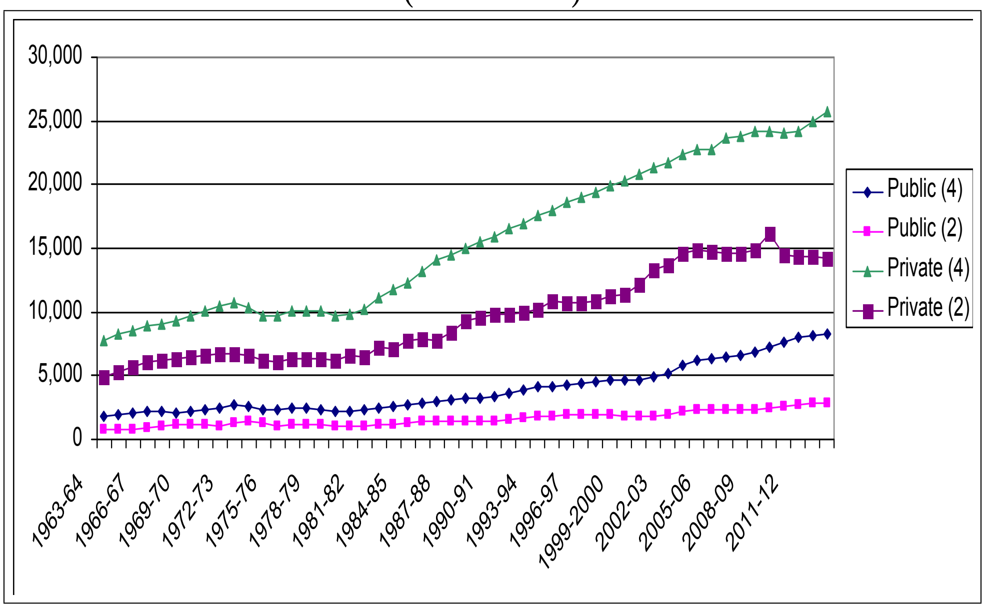
The role of community colleges has traditionally been to expand access to higher education. They are often located in high density areas and offer both traditional higher education coursework and job training programs. Most community colleges work with their surrounding neighborhoods, serving heterogeneous populations, to provide access to the higher education opportunities their community desires. Accordingly, community colleges are often key components in plans to expand access to higher education. President Obama recently endorsed tuition-free community colleges as a way to increase overall bachelor's degree attainment. In this paper, I consider the potential value of such policies by examining the effect of attendance at a community college on bachelor's degree attainment, and whether this effect varies by race, socio-economic status, and academic preparation.

The desire to increase access stems from the vast number of private and social gains to increased educational attainment. Unemployment rates are lower among workers with more education. In April 2015, unemployment rates among workers with at least a bachelor's degree was 2.7%; among workers with less than a high school diploma the rate was 8.6% (Bureau of Labor Statistics 2015). Further, as the fastest growing occupations require postsecondary education, having a degree allows for a greater selection of job opportunities (Carnevale, Smith, and Strohl 2014). In addition, earnings continue to increase with years of education. In 2014, median weekly earnings among workers with at least a bachelor's degree was $1193 compared to $488 among workers with less than a high school diploma. Median earnings among workers with some college or an associate degree in 2014 was $761, a smaller gap, but one that still amounts to a sizable yearly income gap (Bureau of Labor Statistics 2015). Further, I find in Chapter 4 that degree holders have higher wages and income. Greater educational attainment is also tied to better health outcomes and behaviors. People with greater education levels are less likely to smoke, drink excessively, or use illegal drugs, and are more likely to exercise regularly and access preventative health care such as flu shots and other vaccinations. This in turn corresponds to improved long-term health and longer life expectancies among more educated populations (Cutler and Lleras-Muney 2006).

Beyond the private returns to education, there are also social returns. Even with heavy subsidization of higher education, the government sees a positive rate of return to investment in college students on the order of 10% (Trostel 2010). Trostel (2010) estimates these returns based on lifetime tax revenue generated less government spending on higher education (ie. subsidized tuition and financial aid) and government assistance (ie. welfare, Medicaid). In terms of state income taxes, college graduates pay over double the amount a high school graduate pays over a lifetime, and over three times as much in federal income taxes. Further, college graduates are 21% as likely as comparable high school graduates to receive WIC (Trostel 2010). Citizens with postsecondary education are also less likely to be involved in crime, and are more likely to be informed citizens, participate in community service, and contribute to economic growth (Hanushek 1997; Levin 2005). In fact, an additional average year of schooling in a community lowers per-capita police expenditures by $170 (1996 dollars) (Psacharopoulos 2006). Lochner and Moretti (2004) find that increased schooling lowers the probability of incarceration and arrest. The social gains to a highly educated society have far reaching impacts for communities and for the nation as a whole. The combination of private and social returns to education explain the urgency with which American policymakers seek strategies to alleviate obstacles in students' paths through postsecondary education.

A growing number of students are considering community colleges as a viable pathway to a bachelor's degree. Costs at two-year colleges are much lower than four-year institutions. These differences have, in recent years, increased dramatically. Figure 1 shows differences in tuition plus required fees in terms of 2013-2014 dollars, for two and four-year institutions in both the public and private sectors (NCES Table 330.10).[1] Average annual tuition plus required fees for a full-time in-state community college student in 2013-2014 was $2,882 compared to $8,312 for public four-year colleges and $25,696 for private four-year colleges.

Figure 1: Average Tuition Rates in Postsecondary Institutions (1964-2014)



One important caveat to the above figure is that it does not show the differences accounting for financial aid. As the sticker price and actual price of college can vary quite drastically, net prices may be more relevant.[2] However, net price includes student loans, which impacts some students to a greater degree than others. Additionally, financial aid is only applicable to full-time students. Since many community

[1]All calculations assume in-state tuition.

[2]Net price is the total cost of attendance minus grant and scholarship aid from the federal government, state or local governments, or institutional sources. However, average net price by income level is calculated based on all students who received any type of Title IV aid, even those who received zero Title IV aid in the form of grants and received Title IV aid only in the form of work-study aid or loan aid.

college students attend part-time, the comparison is not straightforward. However, average net price for four-year public colleges was $12,410 in 2012, compared to $6,980 for public two-year colleges. This difference is slightly larger than the corresponding difference in tuition and required fees without accounting for aid (NCES Table 331.30). Although there are vast differences in two- and four-year college costs, some of this discrepancy is accounted for in differences in resources and facilities. Community colleges do not have the same resources that public four-year colleges do; in 2011-2012, in per-pupil terms, four-year institutions spent more than double that of two-year colleges on instruction and almost three times as much on academic support (NCES Table 334.10). Thus, the relative benefit of attending a two-year college in order to save money is still an open question.

Another important reason students attend community colleges is due to poor academic preparation in high school. Among students in the data analyzed in this paper, students beginning at two-year colleges had lower academic credentials when compared to students who began at four-year colleges. Average academic high school GPAs were more than 0.5 points lower among students entering two-year colleges, and the average student's standardized math score was almost 8 points lower. Students entering two-year colleges are often placed in remedial courses that do not count toward postsecondary degree attainment. The average student in the analyzed sample who began college at a two-year institution took two remedial courses. Even among students with high school GPAs above 3.0, the average student starting in a two-year college took one remedial course. While these courses may be critical to advancement in postsecondary education, they have explicit as well as implicit costs.

Community college enrollees are more likely to be racial and ethnic minorities, making them high priority for policymakers. In 2012, 15% of enrollees in public two-year colleges were African American and 20% were Hispanic. This is compared to 12%

and 13% enrolled in public four-year colleges (NCES Table 306.20).[3] Baccalaureate attainment also varies substantially by race. According to the U.S. Census, among 18-24 year olds in 2014, 4% of Hispanics and 5% of African Americans had earned at least a bachelor's degree, compared to 14% of non-Hispanic Whites and 18% of Asians (U.S. Census 2014).

In addition to minority students, community college students are also more likely to come from low-income households. In 2012, 80.7% of high-income high school graduates enrolled in college the following fall, compared to 50.9% of low-income students (NCES Table 302.30). Low-income students are also less likely to persist through college. Within five years of beginning postsecondary education, among students who started at two-year institutions, 11.7% from families with income levels below $25,000 had earned a bachelor's degree compared to 18.8% from families with incomes over $100,000. Among students who started at four-year institutions, the corresponding percentages of baccalaureate completion were 44.9% and 77.5% respectively (NCES Table 326.40).

Whether students seek out community colleges to save money or as a sole option after a poor high school record, two-year colleges are an important element of the higher education landscape. As earning a bachelor's degree is tied to social stratification, community colleges and their open access policies play a unique role in intergenerational mobility. An important contribution of this paper is the consideration of differences in baccalaureate attainment by race, socio-economic status, and academic background. It is imperative to identify the mechanisms that perpetuate the gap in baccalaureate attainment in order to inform policy decisions in the future.

[3] These were values of fall enrollments. Traditional students are defined as being between 18 and 24 years old.

2 Literature Review

Proponents of community colleges argue that two-year colleges increase access to education (Cohen and Brawer 2003, Medsker 1960). Many students enter community colleges without a desire to earn a bachelor's degree, and the purpose of community colleges within neighborhoods is usually more than simply a bridge between high school and a bachelor's degree (Hoachlandar, Sikora, Horn, and Carroll 2003). Nevertheless, the transfer function of community colleges has been an intrinsic component since their inception in the early twentieth century (Cohen and Brawer 2003). While it may not be the only mission of community colleges now, providing transfer options for students to baccalaureate programs is still an integral part of community college education.

However, many studies have found that community colleges "divert" students from continuing their education. Early studies considered community college effects directly by comparing attainment outcomes of students who began at two-year colleges to those who began at four-year colleges. These studies find differences in baccalaureate attainment between 10 and 20% (Anderson 1981, Dougherty 1992, Ganderton and Santos 1995, Nunley and Breneman 1988, Velez 1985, and Whitaker and Pascarella 1994). However, these studies did not account for selection bias: students who enter two-year colleges might be different in unobservable ways, such as motivation, resilience, or ability, compared to students who begin in four-year programs.

Various methods have been proposed to address the selection bias. Rouse (1995) utilizes an instrumental variable approach, using distance from a student's residence to a two-year and four-year institution and average in-state tuition as instruments for starting at a two or four-year college. She finds evidence of "democratization" in that students who attended two-year colleges had more total years of schooling and earned more bachelor's degrees than students who did not attend college. While she also finds evidence of diversion from baccalaureate completion among students who started at two-year colleges, these results were not significantly different from zero and were smaller in magnitude than the democratization effect. However, she did not control for a student's desire to earn a bachelor's degree. Leigh and Gill (2003) include the expectation of earning a bachelor's degree, using a "selection-on-observables" approach, and find that the inclusion of educational expectations mitigates the diversion effect.

Several studies have since used similar instrumental variable approaches and produce results that confirm the diversion effect of community college attendance even while controlling for selection bias (Long and Kurlaender 2009, Alfonso 2006, Christie and Hutcheson 2003, and Gonzalez and Hilmer 2006). Other models have also been suggested, such as a two-step choice model (Alfonso 2006), and propensity score matching (Doyle 2009, Melguizo, Kienzl, and Alfonso 2011, and Reynolds 2012). The problem with propensity score matching in this context is a small overlap between two-year and four-year college enrollees on observable factors, as well as a general issue of sensitivity to sample and included variables (Agodini and Dynarski 2004 and Smith and Todd 2004).

Additionally, several studies conducted on more narrowly defined samples and settings have seen similar results (Goodman, Hurwitz, and Smith 2015, Alba and Lavin 1981). Goodman, Hurwitz, and Smith (2015) employ a regression discontinuity method to analyze students just above and below an SAT threshold imposed in public Georgia colleges for four-year college admission. They find that students who were just above the threshold, and thus able to enroll in a public four-year college, were more likely to earn a bachelor's degree. Similarly, Alba and Lavin (1981) consider the effect of an experiment within the City University of New York, which in 1970 allowed all students access to a four-year college, although some had to complete two years at the community college level first. They find that, even with barriers to transfer taken away, students still were not pursuing bachelor's degrees at the same rate as students who began in four-year colleges.

The degree to which outside factors affect enrollment and persistence seems to vary widely by ethnicity (Freeman (2005), Perna (2000, 2007), Wood and Williams (2013), Nunez and Kim (2012), Schneider, Martinez, and Owens (2006), and Alon, Domina, and Tienda (2010)). Gonzalez and Hilmer (2006) find that Hispanic students are much more likely to enroll in two-year colleges, and policies that impact community colleges might have a disproportionately positive effect on this subgroup. However, it is also possible that Hispanics gravitate toward community colleges due to immigrant status, socio-economic status, and a mismatch of information regarding four-year colleges and costs (Bowen, Chingos and McPherson (2009), Schneider, Martinez, and Owens (2006)). Several papers (Olivas 1982, O'Connor 2009, Alon, Domina and Tienda, 2010) have indicated differences among subgroups in the ability to transfer status advantages. Further, O'Connor (2009) finds that among high income students, Hispanic students are at a disadvantage in starting at four-year programs, perhaps due to a language barrier. While 40.2% of full-time bachelor's degree-seeking African American students at four-year institutions completed their degree in six years (51.9% Hispanic), only 19.3% of those who enrolled in a post-secondary school with open admissions completed their degree in six years (30.3% Hispanic). To compare, the corresponding percentages among White students were 62.5% and 41.3%.[4] The disparity among college persistence by ethnicity alone makes this an important topic of study.

Finally, while the recognition of the importance of peer effects on student outcomes dates back to the Supreme Court decision in 1954 which overturned "separate but equal," peer impacts are often not directly included in analysis. The Coleman Report (1966) argues that student background and socio-economic status had a greater

[4] This data is from the 2006 cohort, and the percentages are more dire among males (NCES Table 326.10).

impact on outcomes than school funding. School quality, a difficult variable to measure, might be estimated using peer impacts. Students who are unsure of higher education might be persuaded to attempt a four-year degree if they attended a high school in which a large percentage of students pursue higher education. Jennings et. al (2015) conclude that disparities among high schools lead to vast differences in college attendance; that is, school quality makes a difference in determining educational outcomes. Angrist (2013) notes that individual exam performance was highly linked to peer performance, as was found in Sacerdote (2001), but that peer effects are difficult to disentangle. Smith and Stange (2015) find that peer quality, as measured by PSAT scores, explained about half of the gap in baccalaureate attainment among students who started in two-year colleges instead of four-year colleges. Including peer effects contributes to the growing literature on the importance of peer group and quality.

The data I use provide a unique look at students in a cohort of economic instability; students who graduated high school in 2004. With the Great Recession affecting unemployment rates from 2008 through 2012, the anticipated returns to a degree might have changed for some students throughout this period. The relevant policy implications are most importantly seen in the effects of diversion. If students who would have attended a four-year program are diverted to a community college and are economically hurt by this in the form of not attaining a bachelor's degree, then the expansion of community colleges and promotion of these programs as a low-cost alternative might not be in the best interest of college attenders. Further, if certain demographics of students are at a larger disadvantage than others, this study will address the degree to which interventions might be necessary.

3 Data

The National Center for Education Statistics collects data on students at various levels of education. The Education Longitudinal Study of 2002 (ELS:2002) sampled high school sophomores across the United States in the spring of 2002 and followed these students through 2013. The complete data set includes transcript files from high school as well as any higher education transcripts through 2013. It has a complex survey design, including a stratified two-stage sample of schools and students within them. In the first stage, approximately 750 high schools were drawn with probabilities inversely related to school size.[5] In the second stage, approximately 30 students in each school were sampled.[6] The survey had several intermediate follow-ups: spring 2004, 2006, and 2012.

As the ELS:2002 is survey data, one must be cautious in the treatment of missing data. Using multiple imputation methods would allow for the sample to remain nationally representative, as it is unlikely that data is missing at random.[7] However, once nonrespondents are removed, the missing data is less of an issue. Multiple imputation also limits the types of econometric models that can be estimated. Nevertheless, it is important to note that while the resulting sample analyzed is no longer nationally representative, using the "svy" command still allows for estimates to be weighted to ensure a nationally representative estimate.

The full sample consists of 16,200 students.[8] As I am concerned with the effect of collegiate choices on educational outcomes, I began by including only students who

[5]Private schools were over sampled to ensure adequately sized subsamples.

[6]Students from Hispanic and Asian populations were over sampled to ensure adequately sized subsamples.

[7]When data is not missing at random, list-wise deletion may lead to larger standard errors and biased estimates. Single imputation methods treat imputed values as known, which gives them more weight than perhaps should be awarded, and potentially results in more precise estimates than really exist. Multiple imputation, on the other hand, accounts for the between-imputation variability, and thus results in estimates that reflect the true degree of uncertainty among parameters.

[8]All reported sample sizes are rounded to the nearest 10, as required by the ELS restricted-use agreement with NCES.

were seniors in 2004 (a loss of 2180 students) and went on to pursue postsecondary education in a public or non-profit private college (a loss of 3850 students). I then removed all students without sample weights, as they would drop out of any analysis (a loss of 1610).[9] Finally, I removed students missing academic background variables: standardized math score (10), high school GPA (10), or postsecondary GPA (770). These modifications cut the sample to 7780 students.

Educational choice models are typically thought of in terms of a production function made up of several inputs: individual characteristics, family characteristics, school characteristics, and neighborhood characteristics (Hanushek 1986). Individual controls include gender, race (White, Asian, Black, Hispanic, and other), standardized math score (2004) and academic high school GPA (2004).[10] Math scores and GPA are thought to control for ability, and are used often in educational choice models when Armed Forces Qualification Test (AFQT) scores are unavailable. Additionally, I include whether the student intended on earning a bachelor's degree in 2004. Expectations are an important component in an educational choice model, as students make postsecondary decisions for a variety of reasons. I also include variables at the postsecondary level: (1) the number of remedial classes attended and (2) their GPA in the first year. Family controls include socio-economic status in 2004, which is a composite variable that equally weights parental education, occupation, and income.[11] Parental education and income levels are strongly correlated with student's education levels. School controls include whether the school is in an urban location,

---

[9] The NCES provides weights that combine the cross-sectional and longitudinal nature of the survey, and account for both differential nonresponse bias and over-sampling of some populations. Here, I use the weight f3f1tscpswt to use data from 2004 - 2012 together with the cohort of seniors in 2004, including transcripts in 2004 and 2012. This allows for a nationally representative sample of high school seniors in 2004. Missing weights (coded as -9) are due to missing transcript data, or if they did not respond in the 2006 or 2012 survey. Among students who did respond, extreme weights of zero are also dropped. These occur due to either a small probability of selection or a weight adjustment (NCES 2012a pg.77).

[10] The math exam was proctored by the NCES in each school. Scores were standardized across the senior 2004 sample to have a mean of 50 and standard deviation of 10.

[11] This variable utilizes the 1989 GSS occupational prestige scores rather than the 1961 Duncan SEI index.

the percent of students who are White in 2004, whether the high school is public, an interaction between urban and public schools, and the percent of the respondent's peers who started at a two or four-year college. Urban locations have more collegiate choices within a smaller radius, especially in terms of community colleges. I include the percent of White students as a control for diversity within schools. The peer variables are included to address potential peer impacts on student decisions. Especially if a student remains close to their peer group after high school graduation, their peers may influence collegiate persistence as well as postsecondary outcomes. Recent research suggests school quality may be an important determinant in attainment variation (Coleman et al. 1966, Hanushek 1986, Konstantopoulous 2006). Finally, I include geographic characteristics. Using the Bureau of Labor Statistics, I include per capita income at the county level to account for neighborhood differences. While I could have also include unemployment rates by county, this seems to affect the decision regarding enrollment in postsecondary education and not whether to enroll in a two or four-year college.[12] Census region dummies are also included in all models to control for differences by region.[13]

The dependent variables assessed are whether the student earned a bachelor's degree or higher, total postsecondary credits earned, and the total years of education as of June 2013. These variables were constructed using transcript data instead of self-reported data from the survey to increase reliability of results. Total credits allows for more flexibility, as the years of schooling variable is essentially degree attainment except among students who do not earn a degree. The group of "some college" students lumps all students who do not complete a degree into 12.5 years of education, which does not accurately capture variations in student experience.

[12]Bozick (2009), using the ELS:2002 data, found that students who graduate high school with more job opportunities not requiring higher education are more likely to enter the labor force immediately.

[13]While including state dummies might better control for state policy variations, the models failed to converge in these cases due to small sample sizes in some states. Combining small states (ie. Wyoming and Montana) was not enough to induce convergence.

However, credits do not necessarily translate into a degree, so there are trade-offs in using either variable.[14]

3.1 Descriptive Statistics

I begin by presenting descriptive statistics, weighted to account for varying probabilities of the likelihood to be sampled for the survey. As noted above, private schools and students of Asian and Hispanic decent were oversampled to create comparable subpopulations. This, in addition to nonresponse bias, is accounted for with probability weights, which ensure the means are representative of the population. Thus, the means presented below are estimated, with standard errors in parentheses below each mean.

[14] Total credits counts the number of normalized credits the respondent earned in any undergraduate institution by June 2013 (NCES Codebook). This accounts for differences across institutions. For example, Rutgers requires 120 degree credits for graduation, while the University of Washington requires 180 academic credits. Normalizing credits allows for equal comparison.

Table 1: Individual Characteristics by Postsecondary Start

	Start 2	Start 4	Difference in Means
Individual Characteristics			
Sex (1: Male)	0.476	0.466	-0.010
	(0.012)	(0.010)	(0.016)
African American	0.134	0.119	-0.015
	(0.011)	(0.009)	(0.012)
Hispanic	0.196	0.082	-0.114***
	(0.015)	(0.007)	(0.015)
White	0.589	0.702	0.113***
	(0.017)	(0.012)	(0.018)
Asian	0.043	0.052	0.009
	(0.004)	(0.004)	(0.005)
Other Race	0.038	0.046	0.008
	(0.005)	(0.004)	(0.006)
Math Score (2004	47.944	55.730	7.786***
	(0.254)	(0.206)	(0.302)
High School GPA (2004)	2.470	3.089	0.619***
	(0.022)	(0.015)	(0.024)
First Year Postsecondary GPA	2.439	2.810	0.371***
	(0.029)	(0.018)	(0.034)
Expect Bachelor's Degree (2004)	0.611	0.935	0.324***
	(0.013)	(0.005)	(0.013)
% Peers Attending Two-Year College	34.990	22.593	-12.398***
	(0.659)	(0.641)	(0.774)
% Peers Attending Four-Year College	30.142	46.776	16.634***
	(0.764)	(0.886)	(1.014)
Number of Remedial Courses Taken	2.060	0.676	-1.384***
	(0.071)	(0.030)	(0.075)
N	2700	5080	

* 10%, ** 5%, *** 1%
Standard errors in parentheses

Above, I separate the sample into students who start at four-year colleges and students who start at two-year colleges. Women are more likely than men to enroll in college, though the percentage of women who enroll in two-year colleges is statistically equivalent to the percentage of women enrolling in four-year colleges. Among students who enroll in two-year colleges, African American and Hispanic students are over-represented relative to the full sample. On the other hand, students who start in four-year programs are more likely to be White or Asian. Additionally, there are significant differences in academic credentials, remedial course enrollment, and expectations between these groups.

Table 2: Family, School, and Geographic Characteristics by Postsecondary Start

	Start 2	Start 4	Difference in Means
Family Characteristics			
SES (Quartile) (2004)	2.387	3.012	0.625***
	(0.033)	(0.025)	(0.037)
Mother Education (Categorical)[a]	3.423	4.389	0.966***
	(0.049)	(0.047)	(0.063)
Father Education (Categorical)[a]	3.506	4.629	1.123***
	(0.057)	(0.052)	(0.067)
Income (Categorical) (2002)[b]	8.810	9.794	0.983***
	(0.073)	(0.047)	(0.080)
School Characteristics			
Average Math Score (2004)	49.348	51.841	2.493***
	(0.203)	(0.218)	(0.220)
Percent White (2004)	54.308	61.414	7.106***
	(1.381)	(1.153)	(1.361)
Urban	0.256	0.312	0.056***
	(0.016)	(0.013)	(0.021)
Public High School	0.943	0.870	-0.074***
	(0.005)	(0.007)	(0.008)
Urban×Public	0.230	0.233	0.003***
	(0.016)	(0.013)	(0.021)
Geographic Characteristics			
County: Per Capita Income [10,000's] (2004)	3.248	3.450	0.202***
	(0.043)	(0.039)	(0.043)
N	2700	5080	

* 10%, ** 5%, *** 1%
Standard errors in parentheses
a: Categories: 1) Less than HS, 2) HS/GED, 3) Some Two-Year, 4) Two-Year Degree
5) Some Four-Year, 6) BS, 7) MS, 8) PhD/Advanced Degree
b: Categories: 1) None, 2) $< \$1,000$, 3) \$1,001-5,000, 4) \$5,001-10,000, 5) \$10,001-15,000
6) \$15,001-20,000, 7) \$20,001-25,000, 8) \$25,001-35,000, 9) \$35,001-50,000
10) \$50,001-75,000, 11) \$75,001-100,000, 12) \$100,001-200,000, 13) $> \$200,001$

Table 2 describes family, school, and neighborhood characteristics. Students who start at four-year colleges are more likely to come from higher income families, more educated parents, and more well off counties. They are also more likely to attend urban schools with lower diversity among the student body, and they are less likely to attend public schools. That these characteristics are statistically different among the groups suggests that they are important to control for in models of postsecondary decisions.

Table 3: Postsecondary Characteristics by Postsecondary Start (Transcript Data)

	Start 2	Start 4	Significance
Total Credits	74.820	118.275	43.455***
	(1.532)	(0.935)	(1.699)
Years Schooling	13.559	15.439	1.880***
	(0.037)	(0.045)	(0.056)
At Least Bachelor's Degree (1: Yes)	0.203	0.695	0.492***
	(0.009)	(0.010)	(0.013)
Earned Credits at Two-year	47.376	8.130	-83.357***
	(0.935)	(0.396)	(1.319)
Earned Credits at Four-year	25.946	109.304	39.246***
	(1.049)	(0.938)	(0.984)
N	2700	5080	
* 10%, ** 5%, *** 1%			
Standard errors in parentheses			

Finally, differences among the dependent variables are shown in Table 3. Educational attainment varies significantly by whether a student starts at a two or four-year college. Students who begin at two-year colleges are much less likely to earn a bachelor's degree by 2013. About 70% of students who start at four-year colleges earned at least a bachelor's degree within nine years of high school graduation, but only 20% of those starting at two-year programs had earned at least a bachelor's degree, suggesting a large diversion effect due to community college attendance. Community college students earned, on average, 13.6 years of schooling compared to four-year college attenders who earned 15.4 years of schooling. Finally, total credits earned vary significantly by starting location. While community college students earned an average of 75 credits, four-year starters earned 118 credits.[15]

Collegiate attendance also varies by race, socio-economic status, and academic background. Summary statistics are provided in Appendix A (race), B (socioeconomic) and C (academic). Only 14% of African American students and 10% of Hispanics who began at two-year colleges had earned a bachelor's degree within nine years of high school graduation, compared to 24% of White students and 38% of Asian students. Additionally, students with lower high school academic achievement and students from low-income backgrounds were less likely to have earned a bachelor's

[15]Total credits are normalized by the NCES.

degree within nine years of high school graduation.

4 Empirical Methodology

Educational decisions have been long thought of as endogenous, depending on unobserved characteristics such as ability, motivation, and family background. Due to the probable endogeneity of attendance together with the nature of the data, care must be taken in modeling decisions. While I can control for family background, ability is still difficult to control for with GPA and math scores alone. Additionally, persistence and motivation are difficult to measure econometrically.

Although traditional ordinary least squares estimates do not account for selection bias, it is often difficult to determine a valid instrument, and these estimates provide a comparison to previous results. Equation (1) below shows the baseline model as estimated by ordinary least squares, where the control variables, X_i, include individual, family, school, and neighborhood characteristics.

$$Y_i = \beta' X_i + \rho S_i + \lambda E_i + \epsilon_i \tag{1}$$

In the above equation, Y_i denotes the outcome variable, which may be either continuous ((1) total credits or (2) years of schooling), or discrete ((3) bachelor's degree attainment). S_i denotes the starting location: either a student begins at a two-year college or a four-year college. Finally, E_i is an indicator variable that is 1 if the student expressed a desire to earn a bachelor's degree in 2004 and 0 otherwise.

Table 4 shows the results of the traditional ordinary least squares estimation under the baseline model for outcome variables (1) and (2). For baccalaureate attainment, I estimate the impact of two-year college attendance with a probit model. Starting at a two-year college has negative impacts on educational attainment with respect to each outcome variable. However, it is unlikely that students randomly select into community colleges, so accounting for endogeneity is imperative to obtain consistent estimates.

Table 4: Baseline Model

Dependent Variable	OLS Total Credits	OLS Years Schooling	Probit BA Attainment
Start at Two-Year College	-19.206***	-0.855***	-0.847***
	(1.929)	(0.060)	(0.050)
Math Score (2004)	0.573***	0.015***	0.011***
	(0.106)	(0.003)	(0.004)
High School Academic GPA (2004)	13.833***	0.663***	0.537***
	(1.381)	(0.047)	(0.048)
First Year Postsecondary GPA	15.292***	0.481***	0.421***
	(0.830)	(0.022)	(0.027)
Percent of Peers Attending Two-Year College	0.385***	0.012***	0.012***
	(0.064)	(0.002)	(0.002)
Percent of Peers Attending Four-Year College	0.202***	0.008***	0.008***
	(0.054)	(0.002)	(0.002)
Number of Remedial Courses Taken	2.420***	-0.034***	-0.041***
	(0.421)	(0.012)	(0.014)
Sex (1: Male)	-4.169***	-0.139***	-0.071
	(1.524)	(0.052)	(0.053)
Asian	11.039***	0.212**	0.206**
	(2.767)	(0.102)	(0.102)
African American	2.429	0.021	0.012
	(3.011)	(0.088)	(0.087)
Hispanic	-2.197	-0.017	-0.073
	(2.841)	(0.084)	(0.080)
Other Race	-0.742	-0.228*	-0.282**
	(3.791)	(0.119)	(0.115)
Expect Bachelor's Degree (2004)	21.640***	0.411***	0.579***
	(2.043)	(0.051)	(0.071)
SES Quartile 2	3.317	0.008	0.053
	(2.218)	(0.064)	(0.071)
SES Quartile 3	9.561***	0.306***	0.272***
	(2.452)	(0.079)	(0.078)
SES Quartile 4	11.728***	0.498***	0.404***
	(2.286)	(0.079)	(0.075)
School: Urban	-2.503	0.170	0.183*
	(2.712)	(0.118)	(0.106)
School: Public	-2.899	-0.031	0.018
	(2.286)	(0.096)	(0.094)
School: Urban×Public	6.614**	-0.183	-0.095
	(3.279)	(0.132)	(0.121)
School: Percent White	0.026	-0.0024**	-0.0015
	(0.037)	(0.001)	(0.001)
County: Per Capita Income (2004)	0.881	0.103***	0.090***
	(0.892)	(0.032)	(0.029)
Constant	-55.940***	9.670***	-4.485***
	(8.539)	(0.297)	(0.296)

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Number of Observations: 7780.
Census region dummies included but not reported.

Following Rouse (1995), I use distance to two-year and four-year public colleges and average state tuition to instrument for starting choice. The maintained assumption is that distance from residential zip code to the nearest two-year and four-year college is exogenous. This is a reasonable assumption to make since it is unlikely that a family makes the decision of where to live based on college proximity.[16] To calculate distances, I geocoded student residence zip codes and public two and four-year colleges with tuition information in 2004 from the IPEDS database.[17][18] To calculate the tuition variable, I used the in-state tuition rates provided by IPEDS, not including room or board, for public two-year and four-year colleges and averaged these values by state.[19] The descriptive statistics of these variables are below in Table 5. As above, these are all estimated means using the "svy" command in Stata to provide nationally representative estimates.

[16] As an alternative, I used distance between high school zip code and two-year and four-year colleges, and found essentially the same effect.

[17] In this year, there were about 20% of public or non-profit private colleges with missing tuition information. While many of these missing values seem to be departments of schools that are listed, further analysis should be done as a robustness check.

[18] There were approximately 830 missing student zip codes in 2004. 760 were replaced with student zip codes in 2002 (base year), 50 were replaced with student zip codes in 2006 (second follow-up), 10 were replaced with student zip codes in 2012 (third follow-up), and the remaining 20 zip codes were replaced using the school zip code in the base year. Several Hawaii zip codes were on islands without a corresponding school and thus school zip codes were used instead. Note that using base year school zip codes instead of student zip codes produced similar results. Once the zip codes were geocoded, I used the traveltime3 program in STATA to calculate distance in miles between two points via Google Distancematrix. Traveltime3 is a free alternative to ArcGIS software. It is also preferred to the geonear command, which provided the closest school using geodetic distances, although the two calculations are highly correlated.

[19] In Washington, D.C. there were no public two-year colleges in 2003-2004 with tuition data, and so in this case, the average of Maryland and Virginia was used.

Table 5: Descriptive Statistics: Instruments

	Start 2	Start 4	Difference in Means
Distance to Two-Year College	26.724	30.320	3.599
	(1.096)	(2.663)	(2.664)
Distance to Four-Year College	47.428	35.790	-11.634***
	(2.358)	(1.557)	(2.083)
Average Tuition (Two-Year)	6514.679	6622.36	107.680*
	(53.208)	(44.141)	(63.687)
Average Tuition (Four-Year)	8969.775	9004.623	34.849
	(58.298)	(53.241)	(73.181)
Tuition Ratio (Four-Year:Two-Year In-State Tuition)	139.913	137.106	-2.808***
	(0.736)	(0.482)	(0.776)
N	2700	5080	

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Of students who attend college, the distance to a two-year college is approximately 27 miles, which indicates a viable commuting distance on average. These students are also about 47 miles away from the closest four-year college. On the other hand, students who start at four-year colleges have a smaller differential in mileage between two and four-year colleges. The ratio of four to two-year college tuition is higher among students who start in two-year colleges. However, average tuition rates at two and four-year colleges are not significantly different among students.[20] Considering the first stage of the model, equation (2), determines the importance of distance and tuition on starting decisions.

$$S_i = \pi' X_i + \pi_2 D_{2i} + \gamma_2 (D_{2i})^2 + \pi_4 D_{4i} + \gamma_4 (D_{4i})^2 + \delta(\text{Tuition Ratio})_i + \eta_i \qquad (2)$$

As Table 6 suggests, students are more likely to start at a two-year college when two-year colleges are closer and four-year colleges are farther away. The ratio of tuition rates is significant as well. As the difference in tuition rates rises, students are more likely to start at a community college.[21]

[20] While net price might be more appropriate, these calculations are not available via IPEDS for the 2003-2004 school year.

[21] Since I have more than one exclusion restriction for the endogenous variable, I test for over identification in each model. All p-values are greater than 5%, indicating that there is no problem with the validity of instruments.

Table 6: First Stage Results (Average Marginal Effects)

Dependent Variable: Start at Two-Year College	IV1	IV2
Distance to Two-Year College	-0.001**	-0.002***
	(0.001)	(0.001)
(Distance to Two-Year College)^2	-9.73e-07	-6.19e-07
	(0.000)	(0.000)
Distance to Four-Year College	0.003***	0.003***
	(0.001)	(0.001)
(Distance to Four-Year College)^2	-9.12e-06***	-9.44e-06***
	(0.000)	(0.000)
Tuition Ratio (Four:Two)		0.002***
		(0.001)
Chi-Square	25.42	42.09

Number of Observations: 7780.
Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Other covariates: Gender, Race, Socio-economic status, Math Score, High School GPA BA Expectations, Per Capita Income, % White, % peers attending 2/4 year schools, Urban, Public, Urban×Public, Census region dummies.

As the endogenous variable in my model is binary, I set up an endogenous treatment model. When the outcome variable is continuous, it is inefficient to estimate the model using two stage least squares, but it is consistent. When the outcome variable is also binary, two stage least squares is no longer consistent. In each case, I choose to estimate using maximum likelihood. In the former case, I could additionally use optimal IV to estimate the effects. This would be more robust but less efficient than maximum likelihood estimation. In all models, I use the “svy” command in STATA, which accounts for the probability weights, clustering the standard errors by school, and stratification.[22] This model allows for an intercept effect, but not a slope effect because the coefficients are the same for students who start in two-year and four-year colleges. However, it is expected that the impact of the covariates on the outcome variables is the same whether a student starts at a two or four-year college.

Equation (3) indicates the first stage of the model, regardless of whether the outcome variable is continuous or binary. S_{2i} is a binary variable that is one for students who start at a two-year college and zero otherwise. X_{2i} includes variables

[22] If instead, I include the weights and cluster the standard errors by school, I arrive at almost identical results. The stratification is not accounted for, but this suggests the ability to use this method to perform traditional over-identification tests.

that might affect a student's decision to attend a two-year or four-year college. The instruments included are represented by Z_i.

$$S_{2i} = \begin{cases} 1 & \text{if } \pi' X_{2i} + \gamma Z_i + \eta_i > 0 \\ 0 & \text{Otherwise} \end{cases} \tag{3}$$

As mentioned above, when the outcome variable is also binary, two stage least squares is no longer consistent. In fact, incorrectly assuming either the outcome or endogenous variable is continuous may produce significant bias in the estimates (Freedman and Sekhon 2010). Thus, in the case of bachelor's degree completion being the outcome variable, I am restricted to using maximum likelihood methods for estimation. If starting at a two-year college and baccalaureate attainment are correlated, as might be expected, then a bivariate probit model is most fitting.[23]

An additional robustness check that I consider is whether the errors are heteroskedastic. If they are, using an endogenous treatment model might result in inconsistent estimates, as the treatment model would be mis-specified. Since the errors are typically assumed to be homogeneous, this is important to consider. To test the heteroskedasticity of the errors, I run a model that compares coefficients on a model where the errors are assumed homoskedastic to the model in which they are unrestricted. Doing this produces a chi-squared estimate of 59.33, suggesting that there is heteroskedasticity in the errors. To correct for this, I construct X^*, which divides each covariate in the treatment model by $s = e^{X_i \hat{\gamma}_i}$, where $\hat{\gamma}_i$ represents the estimates from the probit model allowing for heteroskedastic errors. Utilizing these adjusted variables allows for consistent estimation. The issue is maximum likelihood estimation requires many parameters to be simultaneously estimated, and this is difficult to manage with many included covariates. A two-step process would be preferred, but requires a standard error correction.

[23]If they were not correlated, I could estimate two probit models separately. However, starting at a two-year college and earning a bachelor's degree are negatively correlated (-0.422).

5 Results

In the tables below (Tables 7 and 8), I compare ordinary least squares estimates to two endogenous treatment models: one with distance to two and four-year colleges included in quadratic form as instruments, and one with distance as a quadratic in addition to the ratio of average state tuition rates as instruments. That is, I estimate the outcome equation (Equation (1)) in conjunction with the first stage estimation (Equation (3)).

In Table 7, I estimate the model with total postsecondary credits earned as the dependent variable. In Table 8, I use total years of education as the dependent variable.[24]

[24] Recall that years of education is categorical: 12.5: some college, 13: undergraduate certificate, 14: associate degree, 16: bachelor's degree or higher.

Table 7: Dependent Variable: Total Credits; Full Sample

	OLS	Endogenous Treatment	
		IV1	IV2
Start Two-Year College	-19.206***	-29.952***	-28.619***
	(1.929)	(4.764)	(4.796)
Math Score (2004)	0.573***	0.506***	0.515***
	(0.106)	(0.112)	(0.111)
High School Academic GPA (2004)	13.833***	11.878***	12.120***
	(1.381)	(1.576)	(1.608)
First Year Postsecondary GPA	15.292***	15.325***	15.323***
	(0.830)	(0.833)	(0.833)
Percent Peers Attending Two-Year College	0.385***	0.445***	0.437***
	(0.064)	(0.065)	(0.066)
Percent Peers Attending Four-Year College	0.202***	0.171***	0.174***
	(0.054)	(0.056)	(0.056)
Number of Remedial Courses Taken	2.420***	2.466***	2.460***
	(0.421)	(0.424)	(0.424)
Sex (1: Male)	-4.169***	-4.430***	-4.398***
	(1.524)	(1.520)	(1.520)
Asian	11.039***	11.218***	11.196***
	(2.767)	(2.798)	(2.794)
African American	2.429	1.330	1.467
	(3.011)	(2.977)	(2.992)
Hispanic	-2.197	-1.933	-1.965
	(2.841)	(2.845)	(2.845)
Other Race	-0.742	-1.692	-1.574
	(3.791)	(3.758)	(3.759)
Expect to Earn at Least a Bachelor's Degree (2004)	21.640***	19.023***	19.348***
	(2.043)	(2.370)	(2.366)
SES Quartile 2	3.317	3.277	3.282
	(2.218)	(2.243)	(2.239)
SES Quartile 3	9.561***	9.102***	9.159***
	(2.452)	(2.458)	(2.454)
SES Quartile 4	11.728***	10.827***	10.939***
	(2.286)	(2.313)	(2.311)
School: Urban	-2.503	-2.871	-2.825
	(2.712)	(2.705)	(2.708)
School: Public	-2.899	-3.406	-3.343
	(2.286)	(2.257)	(2.257)
School: Urban×Public	6.614**	6.635**	6.633**
	(3.279)	(3.275)	(3.275)
School: Percent White (2004)	0.026	0.033	0.032
	(0.037)	(0.037)	(0.037)
County: Per Capita Income (2004)	0.881	0.928	0.922
	(0.892)	(0.891)	(0.891)
Constant	-55.942***	-40.826***	-42.699***
	(8.539)	(10.480)	(10.442)

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Number of observations: 7780.
IV1: Instruments: distance, distance^2 to two, four year colleges.
IV2: Instruments: distance, distance^2 to two, four year colleges, ratio of average state tuitions.
Census region dummies included but not reported.

The average effect of starting at a two-year college on total credits is negative, conditional on the covariates. Specifically, students who start at two-year colleges earned about 29-30 fewer credits than students who started at four-year colleges.[25] This roughly translates to between eight or nine courses, or about a year of schooling. This estimate is larger than the ordinary least squares estimate, suggesting that the selection bias decreases the gap in attainment rather than increasing it, which is what is typically expected. However, many of the covariates that predict whether a student will complete a degree are included in the model, such as gender, race, and parental education. Thus, it seems that the model that does not correct for selection bias is over controlling for the effect of starting at a two-year college. Controlling for the non-random treatment into a two-year college increases the gap because it controls for unobservable differences between students who attend four and two-year colleges.

On the other hand, academic preparation positively affects earned credits. Interestingly, compared to White students, only Asian students earn more credits on average; all other racial subgroups had insignificant effects. Males earn about four fewer credits, all else equal. Desiring a bachelor's degree has a large positive impact on credit attainment, although it does not cancel out the negative effect of starting at a two-year college. The percent of peers attending a two-year college has more than twice the impact on earned credits as the percent of peers attending a four-year college. Finally, there are increasing impacts of socio-economic status; high-income students earn about 10 more credits than students in the lowest socio-economic quartile.

[25] Including only students who expect to earn at least a bachelor's degree increases the gap by about 3 credits. These results are available upon request.

Table 8: Dependent Variable: Years Schooling; Full Sample

	OLS	Endogenous Treatment	
		IV1	IV2
Start at Two-Year College	-0.855***	-1.215***	-1.214***
	(0.060)	(0.225)	(0.217)
Math Score (2004)	0.015***	0.013***	0.013***
	(0.003)	(0.004)	(0.004)
High School Academic GPA (2004)	0.663***	0.597***	0.597***
	(0.047)	(0.062)	(0.062)
First Year Postsecondary GPA	0.481***	0.482***	0.482***
	(0.022)	(0.022)	(0.022)
Percent of Peers Attending Two-Year College	0.012***	0.014***	0.014***
	(0.002)	(0.002)	(0.002)
Percent of Peers Attending Four-Year College	0.008***	0.006***	0.007***
	(0.002)	(0.002)	(0.002)
Number of Remedial Courses Taken	-0.034***	-0.032***	-0.033***
	(0.012)	(0.012)	(0.012)
Sex (1: Male)	-0.139***	-0.148***	-0.148***
	(0.052)	(0.052)	(0.052)
Asian	0.212**	0.218**	0.218**
	(0.102)	(0.102)	(0.102)
African American	0.021	-0.016	-0.016
	(0.088)	(0.089)	(0.089)
Hispanic	-0.017	-0.008	-0.008
	(0.084)	(0.083)	(0.083)
Other Race	-0.228*	-0.260**	-0.260**
	(0.119)	(0.119)	(0.119)
Expect to Earn at Least a Bachelor's Degree (2004)	0.411***	0.324***	0.324***
	(0.051)	(0.069)	(0.068)
SES Quartile 2	0.008	0.007	0.007
	(0.064)	(0.064)	(0.064)
SES Quartile 3	0.306***	0.290***	0.290***
	(0.079)	(0.080)	(0.080)
SES Quartile 4	0.498***	0.468***	0.468***
	(0.079)	(0.082)	(0.081)
School: Urban	0.170	0.158	0.158
	(0.118)	(0.118)	(0.118)
School: Public	-0.031	-0.048	-0.048
	(0.096)	(0.096)	(0.096)
School: Urban×Public	-0.183	-0.183	-0.183
	(0.132)	(0.132)	(0.132)
School: Percent White (2004)	-0.002**	-0.002*	-0.002*
	(0.001)	(0.001)	(0.001)
County: Per Capita Income (2004)	0.103***	0.105***	0.105***
	(0.032)	(0.032)	(0.032)
Constant	9.670***	10.177***	10.175***
	(0.297)	(0.438)	(0.427)

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Number of observations: 7780.
IV1: Instruments: distance, distance^2 to two, four year colleges.
IV2: Instruments: distance, distance^2 to two, four year colleges, ratio of average state tuitions.
Census region dummies included but not reported.

Table 8 indicates that starting at a two-year program results in about 1.2 fewer years of schooling, after controlling for observables.[26] The signs on the additional covariates correspond to what was seen in Table 7, except for remedial coursework. Here, the number of remedial courses has a negative impact on degree attainment, suggesting the requirement for remedial coursework imposes a barrier to completion.

[26]Including only students who expect to earn a bachelor's degree decreases this gap to about 0.6 years of school. Results are available upon request.

Table 9: Dependent Variable: BA; Full Sample

	Probit	Bivariate Probit	
		IV1	IV2
Start at Two-Year College	-0.847***	-0.938***	-0.892***
	(0.050)	(0.233)	(0.224)
Math Score (2004)	0.011***	0.011***	0.011***
	(0.004)	(0.004)	(0.004)
High School Academic GPA (2004)	0.537***	0.519***	0.529***
	(0.048)	(0.064)	(0.064)
First Year Postsecondary GPA	0.421***	0.421***	0.421***
	(0.027)	(0.027)	(0.027)
Percent of Peers Attending Two-Year College	0.012***	0.013***	0.012***
	(0.002)	(0.002)	(0.002)
Percent of Peers Attending Four-Year College	0.008***	0.007***	0.008***
	(0.002)	(0.002)	(0.002)
Number of Remedial Courses Taken	-0.041***	-0.041***	-0.041***
	(0.014)	(0.014)	(0.014)
Sex (1: Male)	-0.071	-0.073	-0.072
	(0.053)	(0.053)	(0.053)
Asian	0.206**	0.208**	0.207**
	(0.102)	(0.102)	(0.102)
African American	0.012	0.002	0.007
	(0.087)	(0.088)	(0.089)
Hispanic	-0.073	-0.070	-0.072
	(0.080)	(0.080)	(0.080)
Other Race	-0.282**	-0.290**	-0.286**
	(0.115)	(0.115)	(0.115)
Expect to Earn at Least a Bachelor's Degree (2004)	0.579***	0.555***	0.567***
	(0.071)	(0.090)	(0.089)
SES Quartile 2	0.053	0.053	0.053
	(0.071)	(0.071)	(0.071)
SES Quartile 3	0.272***	0.268***	0.270***
	(0.078)	(0.081)	(0.080)
SES Quartile 4	0.404***	0.396***	0.400***
	(0.075)	(0.079)	(0.079)
School: Urban	0.183*	0.179*	0.181*
	(0.106)	(0.106)	(0.106)
School: Public	0.018	0.014	0.016
	(0.094)	(0.095)	(0.095)
School: Urban×Public	-0.095	-0.094	-0.095
	(0.121)	(0.121)	(0.121)
School: Percent White (2004)	-0.002	-0.001	-0.001
	(0.001)	(0.001)	(0.001)
County: Per Capita Income (2004)	0.090***	0.091***	0.090***
	(0.029)	(0.029)	(0.029)
Constant	-4.485***	-4.350***	-4.420***
	(0.296)	(0.466)	(0.448)

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Number of observations: 7780.
IV1: Instruments: distance, distance^2 to two, four year colleges.
IV2: Instruments: distance, distance^2 to two, four year colleges, ratio of average state tuitions.
Census region dummies included but not reported.

Finally, in Table 9, I look to the effect of earning a bachelor's degree. I compare bivariate probit estimates to probit estimates, which do not account for endogeneity. Again, I see negative impacts on baccalaureate attainment and the same signs on other covariates. To address the magnitudes of the impact, however, I calculated the average marginal effects, which are shown in Table 10.

Table 10: Dependent Variable: BA; Average Marginal Effects

	Probit	Bivariate Probit	
		IV1	IV2
Start at Two-Year College	-0.232***	-0.259***	-0.245***
	(0.014)	(0.069)	(0.066)
Math Score (2004)	0.003***	0.003***	0.003***
	(0.001)	(0.001)	(0.001)
High School Academic GPA (2004)	0.128***	0.122***	0.125***
	(0.011)	(0.016)	(0.001)
First Year Postsecondary GPA	0.100***	0.099***	0.100***
	(0.006)	(0.006)	(0.006)
Percent of Peers Attending Two-Year College	0.003***	0.003***	0.003***
	(0.000)	(0.001)	(0.001)
Percent of Peers Attending Four-Year College	0.002***	0.002***	0.002***
	(0.000)	(0.000)	(0.000)
Number of Remedial Courses Taken	-0.010***	-0.010***	-0.010***
	(0.003)	(0.003)	(0.003)
Sex (1: Male)	-0.017	-0.017	-0.017
	(0.013)	(0.012)	(0.013)
Asian	0.049**	0.049**	0.049**
	(0.024)	(0.024)	(0.024)
African American	0.003	0.001	0.002
	(0.021)	(0.021)	(0.021)
Hispanic	-0.017	-0.017	-0.017
	(0.019)	(0.019)	(0.019)
Other Race	-0.068**	-0.069**	-0.069**
	(0.028)	(0.028)	(0.028)
Expect to Earn at Least a Bachelor's Degree (2004)	0.147***	0.140***	0.143***
	(0.019)	(0.025)	(0.025)
SES Quartile 2	0.013	0.013	0.013
	(0.018)	(0.018)	(0.018)
SES Quartile 3	0.068***	0.066***	0.067***
	(0.020)	(0.021)	(0.020)
SES Quartile 4	0.100***	0.097***	0.099***
	(0.019)	(0.021)	(0.020)
School: Urban	0.023*	0.022	0.023
	(0.014)	(0.014)	(0.014)
School: Public	-0.002	-0.003	-0.003
	(0.018)	(0.018)	(0.018)
School: Percent White (2004)	-0.0004	-0.0004	-0.0004
	(0.000)	(0.000)	(0.000)
County: Per Capita Income (2004)	0.021***	0.021***	0.021***
	(0.007)	(0.007)	(0.007)

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Number of observations: 7780.
IV1: Instruments: distance, distance^2 to two, four year colleges.
IV2: Instruments: distance, distance^2 to two, four year colleges, ratio of average state tuitions.
Census region dummies included but not reported.

The results from Table 10 suggest that students who start at two-year colleges are about 25-26 percentage points less likely to have earned a bachelor's degree within nine years of high school graduation. While the expectation of earning a baccalaureate degree counteracts this effect, it is not strong enough to outweigh the negative impact starting at a two-year college has on attainment.[27]

Finally, I considered whether these effects are being driven by public or private colleges. In the Appendix, Table D1 shows results for public colleges only. Here, I see that students who start at two-year public colleges earn about 20 fewer credits, 1.2 fewer years of schooling, and are 20 percentage points less likely to earn a bachelor's degree as compared to their peers who start at four-year colleges. It is interesting that the effect on years of schooling remains the same while the effect of starting at a public two-year college on total credits and baccalaureate attainment are less drastic.

In all of the tables above, I see a negative impact of starting at a two-year college on educational attainment. Students who start at two-year colleges earned about 28 fewer credits or 1.2 fewer years of schooling, and are about 25 percentage points less likely to earn a bachelor's degree. In the next section, I will consider the impact of two-year college attendance on attainment for several subgroups of particular interest: racial, socio-economic, and academic subgroups.

5.1 Subgroups

By looking at effects by subgroup, I can discern whether the impacts differ by racial groups, socio-economic status, and academic background. In the tables below, I consider the impact by subgroup in the full sample as well as the sample of public colleges only. All estimates below indicate results in which the endogenous treatment model was utilized, with distance to two and four year colleges in quadratic form and the ratio of average state tuitions at two and four year colleges included as instruments for starting location.

[27] Running the models with only bachelor's degree aspirants increases the gap by about 0.7 percentage points. Results available upon request.

Table 11: Subgroup - Race

	White	Asian	African American	Hispanic	Other
Dependent Variable: Total Credits					
Start Two-Year College	-31.310***	-9.551	-12.928	-16.953*	-5.827
	(5.600)	(10.866)	(16.613)	(8.730)	(18.545)
Public Schools Only: Start Two-Year	-23.170***	6.290	-7.461	-17.580*	0.187
	(5.600)	(32.709)	(16.781)	(10.174)	(20.680)
Dependent Variable: Years Schooling					
Start Two-Year College	-1.289***	-0.158	-1.605***	-1.125***	-0.993
	(0.280)	(1.242)	(0.503)	(0.362)	(0.656)
Public Schools Only: Start Two-Year	-1.290***	-0.841*	-1.544***	-1.352***	-0.708
	(0.211)	(0.476)	(0.492)	(0.269)	(0.913)
Dependent Variable: Bachelor's Degree Attainment (Reported: Average Marginal Effects)					
Start Two-Year College	-0.279***	–	-0.466	-0.083	-0.150
	(0.081)	–	(0.496)	(0.132)	(0.236)
Public Schools Only: Start Two-Year	-0.246**	0.047	-0.561***	-0.131	0.147
	(0.108)	(0.227)	(0.094)	(0.159)	(0.246)
N (All Schools)	4880	860	810	850	380
N (Public Schools Only)	3710	710	650	730	290

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Instruments: distance, distance^2 (two/four year colleges), ratio of average state tuitions.

Table 11 depicts impacts by racial subgroups. Overall, White students drive the results in all models because their sample size is significantly larger than the other subgroups. In terms of total credits, Hispanic students earn about 17 fewer credits when they start at two-year colleges, but other subgroups are not significantly affected. African American students are disproportionately impacted in terms of total years of schooling. They earned about 1.6 fewer years of schooling compared to African American students who started at four-year colleges. In addition, Hispanic students earned about 1.1 fewer years of schooling by starting at two-year colleges, relative to those who started at four-year colleges. The results for public schools only are generally smaller in magnitude, except with respect to Hispanic students, but have the same implications. As far as bachelor's degree attainment, among public schools African American students seem to be significantly negatively impacted. However, this does not translate to all schools.[28]

[28]The full model for Asian students did not converge so I can not speak to the effect two-year college attendance has on Asian subgroups specifically.

Additionally, I considered the "cmp" command in STATA, which allows for the estimation as above but additionally allows for the treatment effect to be interacted with the subgroup variable of interest. Using this with the same covariates as presented above, I see similar results, presented in Appendix E. The differences in the magnitudes can be attributed to the restriction that, with interaction terms, I am forcing the additional covariates to impact racial subgroups in the same manner, which the separate subgroup analysis did not require. Asian students were the least likely to be negatively impacted by starting at a community college, relative to White students, as they see positive results upon starting at a four year college, and no significant effect on starting at a two year college. Hispanics, on the other hand, are 36 percentage points more likely to earn a bachelor's degree if they start at a four year college and 84 percentage points less likely to earn one by starting at a two year college relative to White students. The gap for African Americans is less drastic, but significant. African American students who start at a four-year college are 21 percentage points more likely to graduate with at least a bachelor's degree, and those who start at a two-year college are 38 percentage points less likely to earn a bachelor's degree nine years after high school. Total credits and years of schooling seem to be driven by the negative impact on starting at a two-year college for most racial subgroups.

Table 12: Subgroup-Socioeconomic Quartile

	Low SES	Middle-Low SES	Middle-High SES	High SES
Dependent Variable: Total Credits				
Start Two-Year College	-82.575***	-12.457	-32.521**	-35.342***
	(25.096)	(7.750)	(12.800)	(6.925)
Public Schools Only: Start Two-Year	-34.770	-11.435	-33.523*	-25.453***
	(21.307)	(7.685)	(19.622)	(7.774)
Dependent Variable: Years Schooling				
	Low SES	Middle-Low SES	Middle-High SES	High SES
Start Two-Year College	-1.513***	-0.971***	-1.093***	-1.315*
	(0.549)	(0.212)	(0.348)	(0.680)
Public Schools Only: Start Two-Year	-1.434***	-1.029***	-1.044***	-1.575***
	(0.264)	(0.187)	(0.261)	(0.596)
Dependent Variable: Bachelor's Degree Attainment (Reported: Average Marginal Effects)				
	Low SES	Middle-Low SES	Middle-High SES	High SES
Start Two-Year College	-0.549***	-0.256***	-0.175	-0.277***
	(0.085)	(0.093)	(0.117)	(0.107)
Public Schools Only: Start Two-Year	-0.521***	-0.220*	-0.126	-0.313**
	(0.107)	(0.112)	(0.138)	(0.156)
N (All Schools)	1180	1610	2050	2940
N (Public Schools Only)	1060	1370	1640	2030

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Instruments: distance, distance^2 (two/four year colleges), ratio of average state tuitions.

In Table 12, I consider the effect of two-year college enrollment by socio-economic status. Students from the lowest socio-economic quartile are much more negatively impacted than other income groups when considering all public and non-profit private schools by starting at a two-year college. The effect dissipates in the next quartile, and appears again for students in middle and high-income groups, creating an upside-down "U" shaped effect, although the magnitudes between middle-high and high-income groups are not statistically different. When looking at only public schools, it is students of higher income categories that are more significantly impacted by starting at a two-year college. This is perhaps due to less financial aid to students in higher income brackets. With respect to years of schooling, I see large negative impacts among low-income students, but also among high-income students. Relative to students who start in four-year colleges, students from high-income families that attend community colleges earn 1.6 fewer years of schooling. Finally, I see a similar pattern in baccalaureate attainment. Low-income students who attended two-year colleges were more than 50 percentage points less likely to earn a bachelor's degree than if they had started in a four-year college. While high-income students are also negatively impacted, the effect is significantly smaller.

Similarly, in considering the model with interactions rather than subgroups, I again see that it is students from lower socioeconomic backgrounds that are more negatively impacted by starting at a two year college.[29] First of all, only students from higher socioeconomic statuses are significantly impacted by starting at a four year college; they are 29 to 44 percentage points more likely to earn a bachelor's degree, earn between 9 and 11 more credits, and have 0.3 to 0.46 more years of schooling, while students from lower socioeconomic statuses see no significant effect. Further, by starting at a two year college, students from lower socioeconomic backgrounds are between 82 and 84 percentage points less likely to earn a bachelor's degree, compared to a disadvantage of between 38 and 49 percentage points for students of higher socioeconomic backgrounds. All students who start at two-year colleges seem to earn fewer credits, but the negative gap is smaller among higher-income students. The gap in years schooling increases with income, but due to the positive impact of starting at four-year colleges, the overall gaps still are largest among lower-income students.

[29] Tables are presented in Appendix E, Table E2.

Table 13: Subgroup - Academic Background

	Low GPA	High GPA[a]	Low Grad	High Grad[b]
Dependent Variable: Total Credits				
Start Two-Year College	-21.276***	-25.948***	-18.494***	-29.774***
	(6.700)	(5.613)	(6.803)	(4.993)
Public Schools Only: Start Two-Year	-20.200***	-21.001***	-16.592***	-23.036***
	(6.486)	(5.716)	(6.264)	(5.229)
Dependent Variable: Years Schooling				
	Low GPA	High GPA[a]	Low Grad	High Grad[b]
Start Two-Year College	-1.046***	-1.643***	-0.911***	-1.482***
	(0.152)	(0.262)	(0.150)	(0.304)
Public Schools Only: Start Two-Year	-1.008***	-1.724***	-0.875***	-1.531***
	(0.122)	(0.330)	(0.123)	(0.334)
Dependent Variable: Bachelor's Degree Attainment (Reported: Average Marginal Effects)				
	Low GPA	High GPA[a]	Low Grad	High Grad[b]
Start Two-Year College	-0.194**	-0.400***	-0.133	-0.328***
	(0.086)	(0.071)	(0.100)	(0.069)
Public Schools Only: Start Two-Year	-0.143*	-0.408***	-0.052	-0.354***
	(0.081)	(0.089)	(0.084)	(0.076)
N (All Schools)	3930	3850	3190	4580
N (Public Schools Only)	3370	2720	2780	3310

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Instruments: distance, distance^2 (two/four year colleges), ratio of average state tuitions.
a) High GPA (academic, high school) is above a 3.0.
b) High Graduate (high school) is a student who: (1): earned diploma before 8/04, (2): 10^{th} grade academic GPA > 2.5, and (3): base year ELS composite assessment score $\geq$ 0.25 s.d. below mean.

Finally, I considered the impact of collegiate attendance by academic background. If students seek community colleges to save money before earning a bachelor's degree, better prepared students should not have as much difficulty. However, students with higher academic credentials were more impacted by attending two-year colleges with respect to total credit attainment. Students with high GPAs who started at two-year colleges earned about 26 fewer credits than students with high GPAs who started at four-year colleges. Students with low GPAs who started at two-year colleges earned 21 fewer credits than students who started at four-year colleges.

I see similar impacts in total years of schooling and bachelor's degree attainment. Students with high GPAs earned about 1.6 fewer years of schooling if they started at a two-year rather than a four-year school. Finally, students with higher academic backgrounds who started at a two-year school were 40 percentage points less likely to earn a bachelor's degree within nine years of high school graduation than their peers who started at four-year institutions. In comparison, among students with lower academic credentials, students who started at two-year colleges were only 19 percentage points less likely to earn a bachelor's degree compared to students who started at four-year colleges.[30]

Further, I compare the above model to one in which interaction terms are included, allowing the treatment effect, starting at a two year college, to be interacted with having a high GPA.[31] Credit earning does not differ significantly among students who begin at a two year college by their GPA, but it is significantly lower for students who start at four year colleges. The almost seven credit differential suggests that students with higher GPAs are either taking fewer classes or being more prudent with class choices in order to graduate more quickly. Further, students who start at a four year college with a high GPA are 37 percentage points more likely to earn a bachelor's degree, compared to 23 percentage points for those with lower high school GPAs. This seems to be the driver for the gap in degree attainment by academic background, as the decline in bachelor's degree attainment by 64 to 65 percentage points for students who start at two year colleges is not statistically different by GPA status. In other words, the overall gap for students with high GPAs between those who start at two and those who start at four year schools is driven by the significant difference among those who start at four-year colleges.

That students with promising academic backgrounds who start at two-year colleges are as likely to earn a bachelor's degree as those without suggests an institutional issue. It is possible that some high performing students are unable to optimally match with colleges of a caliber that match their skills (Roderick, Coca, and Nagaoka 2011). Once improperly matched at a lower-ranked school, lower peer quality might impact

[30]In considering only students who aspire to a bachelor's degree, the gap in credit earning increases, while the gaps on baccalaureate completion and years of schooling decrease. I also replaced the percent of White students in a school with average math scores to account for school quality, and found statistically equivalent results.

[31]Results are presented in Appendix E, table E3.

student performance and persistence.

6 Conclusion

Access to higher education via community colleges is an important research topic for a number of reasons. It is one way to improve equality among students; with more choices, students are better able to identify their optimal path after high school. From the above evidence, it seems students who start at two-year colleges are likely to earn fewer years of education, indicating a diversion effect. If the productivity of higher education is not contingent upon educational attainment, then it may be optimal for some students to not go on to earn a bachelor's degree. However, as the labor market changes and more positions require additional education or training, this could have serious implications for students who start at two-year colleges, especially those that express interest in earning a baccalaureate degree. It may be that community colleges are not serving as a stepping stone to the bachelor's degree that so many students seek.

Men, specifically, are less likely in all models to earn credits, years of schooling, or complete a bachelor's degree. This may in part be due to higher baseline wages men in this sample see. As I show in Chapter 4, male high school graduates have significantly higher wages than female high school graduates. Thus, it is possible that men are making educational decisions based on their available job opportunities, which may differ substantially from those available to women.

Asian students seem to perform better relative to White students, and where they start their postsecondary career does not seem to impact their achievement in the long term. While it did seem that Hispanics were less likely to earn a bachelor's degree relative to Whites, this effect was not seen in years of schooling or credits when subsamples are used. Using the model in which the treatment is interacted however, shows negative effects for Hispanic students. The impact on African American students is stronger in the interaction model, and although African American students are less likely to earn credits, the negative impact of community college attendance on baccalaureate attendance is smaller than most other racial subgroups. Interestingly, African Americans in this sample enroll in two and four-year colleges at statistically equivalent rates.

Lower income students seem at a large disadvantage among student groups either with the subgroup analysis or the interaction model. Policy interventions for lower income students might serve to aid in the transition to college and degree attainment. It is also plausible that race interacted with income has differential effects.

Most strikingly, students with better academic credentials who start at two-year schools earn fewer credits, attain fewer years of schooling, and are less likely to earn a baccalaureate degree than their peers who start at four-year colleges. While it is plausible that these results are driven by high ability students from low quality high schools, this does not seem to be the case. Comparing average math scores in schools in which students with high GPAs did earn a bachelor's degree to average math scores in schools in which students with high GPAs did not earn a bachelor's degree shows very little difference between the two groups. However, this result does seem to be driven by high achieving students entering four year programs, as they are much more likely to graduate with a bachelor's degree. That is, for students with higher academic backgrounds in high school, it is more important that they enter a four year program in order to see baccalaureate attainment goals satisfied, but they are no more likely to falter at a two year college than students with lower academic credentials.

In addition to differential effects by demographic and academic characteristics, several variables were important predictors in postsecondary attainment. Higher academic achievement in the form of math scores, high school GPA, and postsecondary GPA were significant predictors of increased total credits, years of schooling, and the likelihood of earning a bachelor's degree. The number of remedial courses has a positive impact on credit attainment, but negative impacts on total years of schooling or baccalaureate attainment. This suggests that remedial coursework imposes barriers to students in terms of degree attainment. Peer impacts were positive and significant, albeit small in magnitude. A percentage point increase in the fraction of peers attending two-year colleges has almost double the impact of a percentage point increase in peers attending four-year colleges on total credit attainment, and has one and a half times the impact on years of schooling and bachelor's degree attainment. Finally, expecting a bachelor's degree had a large positive impact on academic attainment, as expected.

The results from this paper bolster previous research that suggests starting at a two-year college is a deterrent for baccalaureate attainment. Further, the negative impacts are more starkly found among high achievers. That is, students who might be choosing a community college to save money might end up worse off than if they had gone to a four-year college. This has timely policy implications; if students are not able to successfully transfer to a four-year program and earn a bachelor's degree, then free tuition policies that have been proposed might not have the desired implications. If the goal of community colleges is to increase access to a population of students who might not otherwise pursue education, they are indeed succeeding. If, however, the goal has shifted to inducing more bachelor's degrees, then the objective has yet to be realized.

Appendices

Appendix A: Racial Descriptive Statistics

Table A1: Descriptive Statistics by Postsecondary Start and Race

	White		Asian		African American		Hispanic		Other	
	Start 2	Start 4	Start 2	Start 4	Start 2	Start 4	Start 2	Start 4	Start 2	Start 4
Individual Characteristics										
Sex (1: Male)	0.480	0.474	0.553	0.498	0.412	0.455	0.487	0.398	0.498	0.452
	(0.015)	(0.012)	(0.031)	(0.030)	(0.035)	(0.027)	(0.030)	(0.032)	(0.056)	(0.039)
ELS Math Score (2004)	49.910	57.191	51.326	59.185	42.050	48.491	45.643	52.375	46.333	54.182
	(0.283)	(0.205)	(0.990)	(0.642)	(0.532)	(0.505)	(0.459)	(0.592)	(1.053)	(0.699)
High School GPA (2004)	2.613	3.185	2.656	3.262	2.067	2.607	2.296	2.950	2.358	2.921
	(0.025)	(0.014)	(0.057)	(0.029)	(0.048)	(0.039)	(0.048)	(0.050)	(0.067)	(0.062)
First Year Postsecondary GPA	2.553	2.907	2.653	2.952	2.108	2.350	2.295	2.624	2.341	2.684
	(0.035)	(0.018)	(0.097)	(0.046)	(0.081)	(0.055)	(0.056)	(0.059)	(0.136)	(0.093)
Expect Earn $\geq$ Bachelor's (2004)	0.631	0.941	0.703	0.968	0.611	0.934	0.535	0.891	0.605	0.888
	(0.016)	(0.006)	(0.041)	(0.009)	(0.032)	(0.014)	(0.029)	(0.017)	(0.052)	(0.032)
% Peers Attending Two-Year	34.788	22.093	34.640	24.754	32.704	23.330	37.503	23.821	33.638	23.681
	(0.716)	(0.724)	(1.776)	(1.293)	(1.507)	(1.144)	(1.529)	(1.407)	(2.108)	(1.574)
% Peers Attending Four-Year	31.529	48.201	33.117	47.547	29.632	41.552	25.211	42.724	32.498	44.834
	(0.892)	(0.961)	(2.159)	(2.320)	(1.389)	(1.876)	(1.596)	(2.036)	(2.542)	(2.213)
Number of Remedial Courses	1.552	0.473	1.803	0.584	3.450	1.455	2.701	1.387	2.003	0.599
	(0.077)	(0.025)	(0.197)	(0.068)	(0.226)	(0.120)	(0.153)	(0.155)	(0.242)	(0.100)
N	1530	3350	290	570	290	520	450	400	130	250
Standard errors in parentheses										

Table A2: Descriptive Statistics by Postsecondary Start and Race

	White		Asian		African American		Hispanic		Other	
	Start 2	Start 4	Start 2	Start 4	Start 2	Start 4	Start 2	Start 4	Start 2	Start 4
Family Characteristics										
Socio-economic status (2004)^{a}	0.044	0.432	-0.099	0.273	-0.242	0.037	-0.432	-0.084	-0.151	0.235
	(0.021)	(0.019)	(0.062)	(0.048)	(0.045)	(0.037)	(0.054)	(0.052)	(0.069)	(0.059)
Mother Education (Categorical)^{b}	(0.058)	(0.054)	(0.160)	(0.117)	(0.107)	(0.108)	(0.121)	(0.151)	(0.202)	(0.161)
Father Education (Categorical)^{b}	3.701	4.802	3.913	4.986	3.484	4.093	2.868	3.762	3.379	4.510
	(0.065)	(0.056)	(0.196)	(0.129)	(0.148)	(0.121)	(0.128)	(0.148)	(0.211)	(0.201)
Income (Categorical)^{c} (2002)	9.372	10.176	8.494	9.153	7.637	8.712	8.022	8.726	8.669	9.375
	(0.064)	(0.050)	(0.178)	(0.174)	(0.175)	(0.125)	(0.189)	(0.148)	(0.277)	(0.193)
School Characteristics										
Percent White [0,1]	0.712	0.731	0.272	0.345	0.309	0.311	0.270	0.301	0.472	0.475
	(0.011)	(0.009)	(0.029)	(0.028)	(0.025)	(0.024)	(0.022)	(0.022)	(0.042)	(0.028)
Urban	0.154	0.245	0.401	0.466	0.448	0.500	0.422	0.468	0.151	0.410
	(0.015)	(0.014)	(0.048)	(0.044)	(0.045)	(0.037)	(0.047)	(0.044)	(0.037)	(0.047)
Public High School	0.933	0.857	0.944	0.887	0.983	0.941	0.953	0.854	0.915	0.891
	(0.007)	(0.008)	(0.015)	(0.025)	(0.004)	(0.010)	(0.010)	(0.020)	(0.023)	(0.016)
Urban×Public	0.127	0.160	0.372	0.375	0.436	0.460	0.391	0.380	0.117	0.342
	(0.015)	(0.014)	(0.047)	(0.044)	(0.046)	(0.038)	(0.047)	(0.046)	(0.034)	(0.047)
Geographic Characteristics										
County: Per Capita Income (2004)	3.173	3.418	3.723	3.914	3.418	3.526	3.256	3.221	3.223	3.621
	(0.041)	(0.043)	(0.118)	(0.120)	(0.123)	(0.090)	(0.087)	(0.106)	(0.111)	(0.100)
N	1530	3350	290	570	290	520	450	400	130	250

Standard errors in parentheses

a: [range: -2,2]

b: Categories: 1) Less than HS, 2) HS/GED, 3) Some Two-Year, 4) Two-Year Degree, 5) Some Four-Year, 6) BS, 7) MS, 8) PhD

c: Categories: 1) None, 2) < \$1,000, 3) \$1,001-5,000, 4) \$5,001-10,000, 5) \$10,001-15,000, 6) \$15,001-20,000, 7) \$20,001-25,000, 8) \$25,001-35,000, 9) \$35,001-50,000, 10) \$50,001-75,000, 11) \$75,001-100,000, 12) \$100,001-200,000, 13) > \$200,001

Table A3: Descriptive Statistics by Postsecondary Start and Race (Transcript Data)

	White		Asian		African American		Hispanic		Other	
	Start 2	Start 4	Start 2	Start 4	Start 2	Start 4	Start 2	Start 4	Start 2	Start 4
Total Credits	80.292	120.854	103.564	133.764	62.270	104.833	60.365	112.989	76.461	105.483
	(1.986)	(1.127)	(5.404)	(2.431)	(3.946)	(3.272)	(3.088)	(3.735)	(7.001)	(5.222)
Years Schooling	13.738	15.615	14.212	16.005	13.238	14.616	13.132	15.091	13.385	14.870
	(0.048)	(0.054)	(0.159)	(0.113)	(0.091)	(0.124)	(0.064)	(0.151)	(0.129)	(0.169)
Earned at least BA	0.243	0.738	0.384	0.821	0.137	0.498	0.101	0.615	0.147	0.555
	(0.012)	(0.011)	(0.042)	(0.024)	(0.022)	(0.031)	(0.015)	(0.033)	(0.029)	(0.044)
N	1530	3350	290	570	290	520	450	400	130	250

Standard errors in parentheses

Appendix B: Socioeconomic Status Descriptive Statistics

Table B1: Descriptive Statistics by Postsecondary Start and Socioeconomic Status

	SES:1		SES:2		SES:3		SES:4	
	Start 2	Start 4	Start 2	Start 4	Start 2	Start 4	Start 2	Start 4
Individual Characteristics								
Sex (1: Male)	0.434	0.381	0.470	0.465	0.478	0.467	0.539	0.488
	(0.024)	(0.027)	(0.022)	(0.022)	(0.023)	(0.018)	(0.025)	(0.014)
African American	0.168	0.229	0.141	0.163	0.121	0.116	0.096	0.072
	(0.019)	(0.026)	(0.016)	(0.018)	(0.019)	(0.013)	(0.019)	(0.008)
Hispanic	0.357	0.229	0.170	0.094	0.143	0.061	0.087	0.050
	(0.037)	(0.027)	(0.018)	(0.015)	(0.017)	(0.009)	(0.015)	(0.006)
White	0.380	0.398	0.621	0.637	0.662	0.736	0.728	0.788
	(0.031)	(0.030)	(0.022)	(0.024)	(0.023)	(0.018)	(0.026)	(0.012)
Asian	0.051	0.085	0.032	0.049	0.038	0.038	0.053	0.054
	(0.010)	(0.011)	(0.005)	(0.007)	(0.006)	(0.005)	(0.008)	(0.006)
Other Race	0.045	0.059	0.037	0.057	0.035	0.048	0.036	0.036
	(0.011)	(0.014)	(0.008)	(0.011)	(0.007)	(0.008)	(0.009)	(0.005)
ELS Math Score (2004)	45.183	50.748	47.286	53.485	49.042	55.344	51.155	58.296
	(0.378)	(0.482)	(0.391)	(0.411)	(0.410)	(0.312)	(0.484)	(0.236)
High School GPA (2004)	2.341	2.881	2.456	2.994	2.538	3.081	2.574	3.192
	(0.039)	(0.040)	(0.034)	(0.033)	(0.035)	(0.021)	(0.043)	(0.019)
First Year Postsecondary GPA	2.361	2.575	2.427	2.700	2.407	2.778	2.604	2.943
	(0.060)	(0.058)	(0.050)	(0.038)	(0.060)	(0.032)	(0.048)	(0.022)
Expect Earn $\geq$ Bachelor's (2004)	0.520	0.872	0.590	0.901	0.643	0.944	0.725	0.961
	(0.023)	(0.019)	(0.020)	(0.014)	(0.021)	(0.008)	(0.024)	(0.005)
% Peers Attending Two-Year	35.507	25.427	34.366	22.870	35.451	22.478	34.598	21.790
	(1.160)	(1.013)	(0.807)	(0.923)	(0.852)	(0.777)	(0.978)	(0.802)
% Peers Attending Four-Year	25.083	36.925	29.537	41.815	31.583	46.453	35.902	51.790
	(1.085)	(1.359)	(0.978)	(1.103)	(0.967)	(1.075)	(1.167)	(1.129)
Number of Remedial Courses	2.393	1.213	2.018	0.910	2.090	0.739	1.635	0.388
	(0.109)	(0.097)	(0.118)	(0.070)	(0.135)	(0.055)	(0.111)	(0.026)
N	650	530	750	860	720	1330	580	2360

Standard errors in parentheses

Table B2: Descriptive Statistics by Postsecondary Start and Socioeconomic Status

	SES:1		SES:2		SES:3		SES:4	
	Start 2	Start 4	Start 2	Start 4	Start 2	Start 4	Start 2	Start 4
Family Characteristics								
Socio-economic status (2004)^a	-0.890	-0.856	-0.302	-0.287	0.196	0.216	0.856	0.981
	(0.018)	(0.013)	(0.006)	(0.005)	(0.008)	(0.006)	(0.016)	(0.011)
Mother Education (Categorical)^b	1.852	1.971	3.002	3.043	3.946	4.130	5.449	5.793
	(0.044)	(0.057)	(0.060)	(0.060)	(0.070)	(0.062)	(0.087)	(0.041)
Father Education (Categorical)^b	1.904	1.939	2.917	3.156	4.097	4.380	5.727	6.155
	(0.049)	(0.064)	(0.065)	(0.065)	(0.093)	(0.058)	(0.079)	(0.043)
Income (Categorical)^c (2002)	6.671	6.775	8.707	8.801	9.703	9.925	10.655	10.948
	(0.124)	(0.111)	(0.074)	(0.066)	(0.072)	(0.053)	(0.065)	(0.041)
School Characteristics								
Percent White (2004) [0,1]	0.420	0.449	0.590	0.596	0.577	0.648	0.596	0.644
	(0.023)	(0.024)	(0.016)	(0.019)	(0.018)	(0.015)	(0.018)	(0.013)
Urban	0.328	0.391	0.216	0.299	0.240	0.284	0.239	0.316
	(0.035)	(0.029)	(0.020)	(0.025)	(0.022)	(0.019)	(0.027)	(0.020)
Public High School	0.985	0.951	0.955	0.926	0.923	0.879	0.899	0.817
	(0.004)	(0.008)	(0.006)	(0.007)	(0.010)	(0.009)	(0.012)	(0.013)
Urban×Public	0.321	0.365	0.195	0.253	0.204	0.211	0.196	0.205
	(0.035)	(0.029)	(0.020)	(0.025)	(0.022)	(0.018)	(0.027)	(0.019)
Geographic Characteristics								
County: Per Capita Income (2004)	3.132	3.214	3.185	3.315	3.273	3.452	3.461	3.571
	(0.062)	(0.059)	(0.045)	(0.047)	(0.057)	(0.053)	(0.073)	(0.046)
N	650	530	750	860	720	1330	580	2360

Standard errors in parentheses

a: [range: -2,2]

b: Categories: 1) Less than HS, 2) HS/GED, 3) Some Two-Year, 4) Two-Year Degree, 5) Some Four-Year, 6) BS, 7) MS, 8) PhD

c: Categories: 1) None, 2) < $1,000, 3) $1,001-5,000, 4) $5,001-10,000, 5) $10,001-15,000, 6) $15,001-20,000, 7) $20,001-25,000, 8) $25,001-35,000, 9) $35,001-50,000, 10) $50,001-75,000, 11) $75,001-100,000, 12) $100,001-200,000, 13) > $200,001

Table B3: Descriptive Statistics by Postsecondary Start and Socioeconomic Status (Transcript Data)

	SES:1		SES:2		SES:3		SES:4	
	Start 2	Start 4	Start 2	Start 4	Start 2	Start 4	Start 2	Start 4
Total Credits	61.695	101.627	68.780	109.788	81.367	118.419	92.562	126.343
	(2.799)	(3.241)	(2.323)	(2.401)	(2.982)	(1.833)	(3.292)	(1.324)
Years Schooling	13.200	14.640	13.357	14.919	13.736	15.415	14.101	15.897
	(0.060)	(0.124)	(0.048)	(0.085)	(0.078)	(0.078)	(0.096)	(0.065)
Earned at least BA	0.110	0.517	0.153	0.595	0.251	0.689	0.340	0.791
	(0.015)	(0.030)	(0.013)	(0.021)	(0.021)	(0.017)	(0.024)	(0.012)
N	650	530	750	860	720	1330	580	2360

Standard errors in parentheses

Appendix C: Academic Background Descriptive Statistics

Table C1: Descriptive Statistics by Postsecondary Start and Academic Background

	Low GPA		High GPA		Low Graduate		High Graduate	
	Start 2	Start 4	Start 2	Start 4	Start 2	Start 4	Start 2	Start 4
Individual Characteristics								
Sex (1: Male)	0.498	0.564	0.404	0.400	0.491	0.529	0.445	0.438
	(0.014)	(0.016)	(0.023)	(0.012)	(0.015)	(0.018)	(0.018)	(0.011)
African American	0.163	0.212	0.039	0.056	0.180	0.245	0.039	0.062
	(0.014)	(0.016)	(0.009)	(0.006)	(0.015)	(0.019)	(0.008)	(0.006)
Hispanic	0.221	0.093	0.113	0.074	0.231	0.108	0.123	0.069
	(0.017)	(0.011)	(0.021)	(0.007)	(0.019)	(0.013)	(0.016)	(0.006)
White	0.541	0.605	0.747	0.767	0.503	0.539	0.769	0.774
	(0.019)	(0.019)	(0.024)	(0.012)	(0.020)	(0.022)	(0.019)	(0.011)
Asian	0.035	0.036	0.070	0.063	0.043	0.041	0.043	0.057
	(0.004)	(0.004)	(0.011)	(0.006)	(0.005)	(0.005)	(0.007)	(0.005)
Other Race	0.040	0.054	0.031	0.040	0.044	0.066	0.027	0.037
	(0.005)	(0.007)	(0.007)	(0.005)	(0.006)	(0.009)	(0.005)	(0.004)
ELS Math Score (2004)	46.196	51.319	53.696	58.717	44.868	49.633	54.390	58.459
	(0.266)	(0.325)	(0.382)	(0.193)	(0.272)	(0.371)	(0.266)	(0.175)
High School GPA (2004)	2.183	2.468	3.416	3.509	2.169	2.526	3.101	3.341
	(0.017)	(0.015)	(0.013)	(0.008)	(0.021)	(0.028)	(0.021)	(0.011)
First Year Postsecondary GPA	2.230	2.340	3.128	3.128	2.215	2.427	2.910	2.982
	(0.034)	(0.030)	(0.038)	(0.015)	(0.035)	(0.035)	(0.037)	(0.017)
Expect Earn $\geq$ Bachelor's (2004)	0.555	0.878	0.798	0.974	0.539	0.855	0.764	0.971
	(0.015)	(0.010)	(0.023)	(0.003)	(0.014)	(0.013)	(0.020)	(0.004)
% Peers Attending Two-Year	34.313	21.007	37.220	23.666	34.278	21.908	36.483	22.899
	(0.666)	(0.702)	(1.043)	(0.734)	(0.747)	(0.883)	(0.836)	(0.686)
% Peers Attending Four-Year	31.004	47.072	27.306	46.575	30.379	45.183	29.645	47.488
	(0.781)	(1.108)	(1.065)	(0.933)	(0.836)	(1.283)	(0.945)	(0.906)
Number of Remedial Courses	2.390	1.168	0.973	0.343	2.591	1.368	0.948	0.366
	(0.082)	(0.059)	(0.088)	(0.023)	(0.085)	(0.075)	(0.072)	(0.023)
N	1990	1940	710	3140	1740	1460	960	3620

Standard errors in parentheses

Table C2: Descriptive Statistics by Postsecondary Start and Academic Background

	Low GPA		High GPA		Low Graduate		High Graduate	
	Start 2	Start 4	Start 2	Start 4	Start 2	Start 4	Start 2	Start 4
Family Characteristics								
Socio-economic status (2004)[a]	-0.145	0.199	0.043	0.412	-0.176	0.137	0.056	0.410
	(0.023)	(0.027)	(0.035)	(0.019)	(0.027)	(0.029)	(0.026)	(0.019)
Mother Education (Categorical)[b]	3.339	4.116	3.701	4.574	3.281	4.040	3.720	4.545
	(0.054)	(0.071)	(0.089)	(0.051)	(0.062)	(0.081)	(0.077)	(0.050)
Father Education (Categorical)[b]	3.413	4.263	3.813	4.877	3.359	4.187	3.814	4.827
	(0.063)	(0.076)	(0.100)	(0.058)	(0.069)	(0.082)	(0.082)	(0.056)
Income (Categorical)[c] (2002)	8.695	9.563	9.190	9.950	8.598	9.281	9.256	10.023
	(0.079)	(0.076)	(0.116)	(0.051)	(0.091)	(0.086)	(0.088)	(0.049)
School Characteristics								
Percent White (2004) [0,1]	0.519	0.555	0.624	0.654	0.500	0.529	0.634	0.652
	(0.014)	(0.016)	(0.021)	(0.011)	(0.015)	(0.018)	(0.017)	(0.011)
Urban	0.270	0.369	0.212	0.274	0.284	0.394	0.197	0.276
	(0.017)	(0.018)	(0.026)	(0.015)	(0.020)	(0.022)	(0.019)	(0.015)
Public High School	0.948	0.878	0.929	0.864	0.950	0.889	0.931	0.861
	(0.005)	(0.009)	(0.011)	(0.009)	(0.005)	(0.009)	(0.009)	(0.008)
Urban×Public	0.241	0.295	0.194	0.192	0.258	0.328	0.172	0.191
	(0.017)	(0.018)	(0.026)	(0.014)	(0.020)	(0.022)	(0.018)	(0.014)
Geographic Characteristics								
County: Per Capita Income (2004)	3.284	3.493	3.128	3.421	3.296	3.442	3.148	3.454
	(0.046)	(0.051)	(0.056)	(0.040)	(0.051)	(0.054)	(0.046)	(0.043)
N	1990	1940	710	3140	1740	1460	960	3620

Standard errors in parentheses

a: [range: -2,2]

b: Categories: 1) Less than HS, 2) HS/GED, 3) Some Two-Year, 4) Two-Year Degree, 5) Some Four-Year, 6) BS, 7) MS, 8) PhD

c: Categories: 1) None, 2) < \$1,000, 3) \$1,001-5,000, 4) \$5,001-10,000, 5) \$10,001-15,000, 6) \$15,001-20,000, 7) \$20,001-25,000, 8) \$25,001-35,000, 9) \$35,001-50,000, 10) \$50,001-75,000, 11) \$75,001-100,000, 12) \$100,001-200,000, 13) > \$200,001

Table C3: Descriptive Statistics by Postsecondary Start and Academic Background (Transcript Data)

	Low GPA		High GPA		Low Graduate		High Graduate	
	Start 2	Start 4	Start 2	Start 4	Start 2	Start 4	Start 2	Start 4
Total Credits	64.289	101.703	109.481	129.495	62.360	101.490	100.934	125.787
	(1.592)	(1.787)	(3.181)	(1.074)	(1.638)	(1.956)	(2.463)	(1.072)
Years Schooling	13.236	14.520	14.623	16.062	13.217	14.513	14.277	15.854
	(0.034)	(0.066)	(0.093)	(0.053)	(0.036)	(0.072)	(0.070)	(0.049)
Earned at least BA	0.127	0.490	0.456	0.834	0.121	0.483	0.375	0.790
	(0.009)	(0.017)	(0.023)	(0.010)	(0.009)	(0.018)	(0.018)	(0.010)
N	1990	1940	710	3140	1740	1460	960	3620

Standard errors in parentheses

Appendix D: Public School IV Results

Table D1: The Effect of Starting at a Community College (IV2): Public Schools Only

	Total Credits	Years Schooling	BA (AME)
Start Two-Year College	-20.344***	-1.226***	-0.203**
	(4.778)	(0.142)	(0.085)
Math Score (2004)	0.703***	0.011***	0.003***
	(0.120)	(0.004)	(0.001)
High School Academic GPA (2004)	15.807***	0.593***	0.132***
	(1.686)	(0.055)	(0.019)
First Year Postsecondary GPA	15.552***	0.451***	0.097***
	(0.898)	(0.023)	(0.007)
Percent Peers Attending Two-Year College	0.388***	0.014***	0.003***
	(0.072)	(0.002)	(0.001)
Percent Peers Attending Four-Year College	0.191***	0.004**	0.002***
	(0.063)	(0.002)	(0.001)
Number of Remedial Courses Taken	2.608***	-0.031**	-0.010***
	(0.448)	(0.012)	(0.004)
Sex (1: Male)	-4.955***	-0.119**	-0.015
	(1.691)	(0.057)	(0.014)
Asian	12.666***	0.261**	0.056**
	(2.946)	(0.109)	(0.026)
African American	5.167	0.030	0.015
	(3.321)	(0.092)	(0.024)
Hispanic	-1.526	-0.018	-0.028
	(2.970)	(0.090)	(0.022)
Other Race	-0.711	-0.228*	-0.068**
	(4.255)	(0.137)	(0.032)
Expect to Earn at Least a Bachelor's Degree (2004)	21.108***	0.329***	0.161***
	(2.333)	(0.058)	(0.029)
SES Quartile 2	2.050	0.004	0.009
	(2.319)	(0.067)	(0.018)
SES Quartile 3	8.259***	0.261***	0.070***
	(2.518)	(0.084)	(0.022)
SES Quartile 4	11.958***	0.445***	0.097***
	(2.397)	(0.085)	(0.023)
School: Urban	2.506	0.152	0.019
	(3.571)	(0.143)	(0.015)
School: Public	-0.992	-0.106	-0.010
	(2.947)	(0.121)	(0.023)
School: Urban×Public	0.807	-0.158	–
	(4.135)	(0.158)	–
School: Percent White (2004)	6.125	-0.111	-0.0004
	(4.396)	(0.125)	(0.000)
County: Per Capita Income (2004)	1.808*	0.114***	0.025***
	(0.995)	(0.032)	(0.008)
Constant	-72.322***	10.357***	–
	(11.093)	(0.374)	–

Number of observations: 7780.
Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Appendix E: CMP Results

Table E1: The Effect of Starting at a Community College (IV2): Race

	Total Credits	Years Schooling	BA (AME)
White: Start Four-Year	–	–	0.303***
	–	–	(0.099)
Asian: Start Four-Year	7.883**	0.197*	0.451***
	(3.074)	(0.117)	(0.135)
African American: Start Four-Year	4.523	-0.187	0.205*
	(3.468)	(0.124)	(0.112)
Hispanic: Start Four-Year	4.819	0.038	0.362***
	(3.486)	(0.130)	(0.126)
Other: Start Four-Year	-5.548	-0.401***	-0.031
	(4.369)	(0.143)	(0.141)
White: Start Two-Year	-27.219***	-1.306***	-0.628***
	(4.886)	(0.218)	(0.152)
Asian: Start Two-Year	-18.044***	-1.249***	-0.335
	(6.460)	(0.267)	(0.214)
African American: Start Two-Year	-34.174***	-0.902***	-0.383*
	(6.437)	(0.236)	(0.203)
Hispanic: Start Two-Year	-38.905***	-1.354***	-0.840***
	(5.754)	(0.236)	(0.160)
Other: Start Two-Year	-16.000*	-0.920***	-0.792***
	(9.048)	(0.272)	(0.238)

Number of observations: 7780.
Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
IV2: Instruments: distance, distance^2 to two/four year colleges, ratio of average state tuitions.

Table E2: The Effect of Starting at a Community College (IV2): SES

	Total Credits	Years Schooling	BA (AME)
SES 1: Start Four-Year	–	–	0.055
	–	–	(0.121)
SES 2: Start Four-Year	3.116	0.011	0.151
	(2.218)	(0.064)	(0.106)
SES 3: Start Four-Year	9.042***	0.292***	0.291***
	(2.446)	(0.080)	(0.093)
SES 4: Start Four-Year	10.927***	0.463***	0.437***
	(2.285)	(0.081)	(0.093)
SES 1: Start Two-Year	-29.920***	-1.073***	-0.819***
	(5.641)	(0.234)	(0.150)
SES 2: Start Two-Year	-31.464***	-1.206***	-0.835***
	(5.265)	(0.219)	(0.145)
SES 3: Start Two-Year	-27.111***	-1.257***	-0.491***
	(5.445)	(0.228)	(0.162)
SES 4: Start Two-Year	-24.173***	-1.379***	-0.378**
	(5.397)	(0.245)	(0.173)

Number of observations: 7780.
Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
IV2: Instruments: distance, distance^2 to two/four year colleges, ratio of average state tuitions.

Table E3: The Effect of Starting at a Community College (IV2): Academic

	Total Credits	Years Schooling	BA (AME)
Low GPA: Start Four-Year	–	–	0.232**
	–	–	(0.093)
High GPA: Start Four-Year	-6.756***	0.397***	0.374***
	(2.485)	(0.085)	(0.102)
Low GPA: Start Two-Year	-31.525***	-1.121***	-0.652***
	(4.997)	(0.204)	(0.146)
High GPA: Start Two-Year	-27.297***	-1.105***	-0.637***
	(5.608)	(0.231)	(0.163)

Number of observations: 7780.
Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
IV2: Instruments: distance, distance^2 to two/four year colleges, ratio of average state tuitions.
a) High GPA (high school) is above a 3.0.

4. Education Decisions and Labor Market Outcomes

1 Introduction

Median income levels have been consistently higher among individuals with more education, leading to greater investments in education over the years. As of 2012, individuals with undergraduate degrees had median income levels more than double those without a high school diploma (NCES Table 502.30). However, the wage benefits of postsecondary education to those who do not complete a baccalaureate degree are less studied but increasingly important as the labor market shifts to support a new era. With 21.2% of citizens 25 years or older with some postsecondary experience in 2014, and 37.2% with at least an associate degree, job requirements are changing, and the relative return to a degree might not be what it was 20 or 30 years ago (U.S. Census Table S1501). In the last forty years, the number of associate degrees has more than tripled (NCES Table 318.10). Thus, it is now more urgent than ever to understand the effect these degrees have on labor market outcomes.

The Bureau of Labor Statistics calculates projected growth rates among different occupations by education level requirements. Job growth between 2012 and 2022 among those with associate degrees is expected to rise by 17.6%. As the second fastest growing category among educational credentials, this suggests an importance in understanding returns (BLS Employment Projections - 2012-2022 Table 7). Further, while much job growth will occur in occupations in which higher education is required, job growth only accounts for about one-third of expected new jobs; replacement due to retirees makes up the rest. Such jobs are mostly in sectors requiring little to no higher education; twenty-two of the thirty occupations expected to see the highest growth require no postsecondary education.[1] On the other hand, growth

[1]Home heath aides are expected to grow by 48.5% and personal care aides by 48.8%, neither of

within sectors requiring a bachelor's degree is expected to be significantly smaller.[2] Of the thirty occupations projected to have the largest number of openings by 2022 due to growth and replacement, only six require a bachelor's degree or higher (BLS Employment Projections - 2012-2022 Table 8).

Finally, sub-baccalaureate programs are especially popular among minority groups, as both African American and Hispanic students are more likely to enroll in two-year colleges. Here, both associate degrees, which typically require two years of college, and undergraduate certificates, typically requiring one year, are offered. Associate degrees in liberal arts and health are common, from nursing to philosophy. Undergraduate certificates tend to focus more on job training; some examples of undergraduate certificates include law enforcement, medical assistants, and automotive technology. Students also have the ability to transfer to four-year colleges. Between 2000 and 2013, the number of associate degrees conferred to Whites has increased by 51%, to African Americans by 126%, and to Hispanics by 206% (NCES Table 321.20). Further, undergraduate certificates conferred to Whites have increased by 55%, African Americans by 82%, and Hispanics by 129% (NCES Table 320.20).

Using data from the National Center for Education Statistics Education Longitudinal Study of 2002 (ELS:2002), I consider the difference in wages by educational attainment among individuals surveyed in 2012. At this time, individuals were about 27 years old. There are 5100 males and 5560 females that are at least high school graduates and have at most a baccalaureate degree with positive wages within the ELS:2002 sample. As seen below in Figures 1 and 2, median wages increase with higher educational attainment for both females and males. Not only do those with a high school diploma have the lowest wages, but the distribution of wages is tighter than other categories.

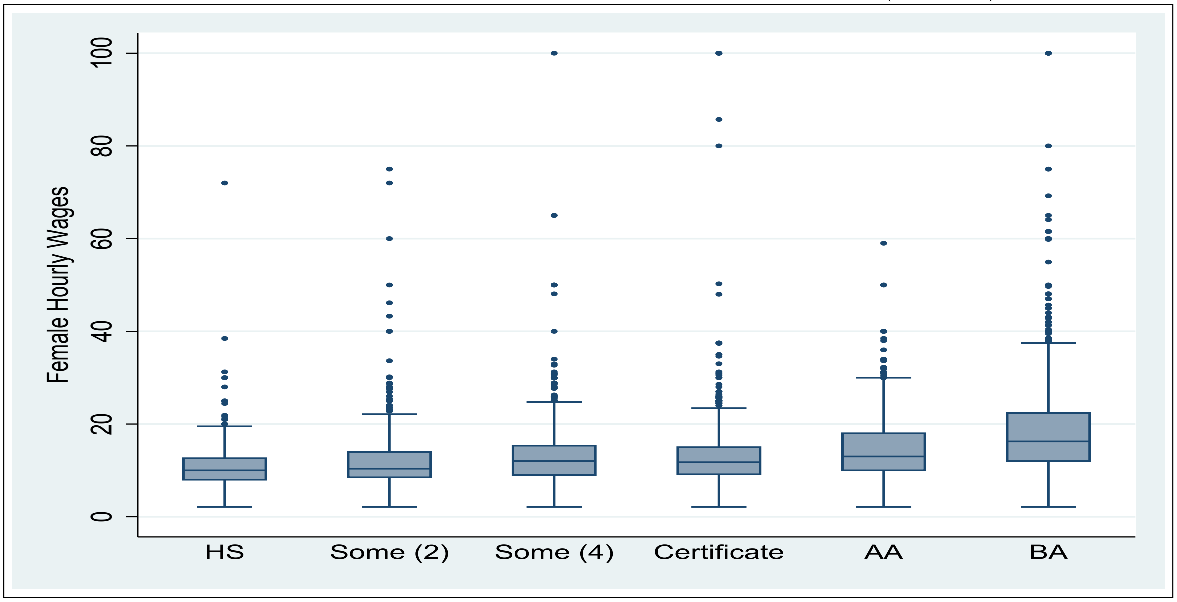
---

which require a high school diploma (BLS Employment Projections - 2012-2022 Table 8).

[2]For example, accountants are expected to grow by 13.1%, and elementary school teachers by 12.3% (BLS Employment Projections - 2012-2022 Table 8).

For women, there are returns to those who start at four-year colleges, but the distribution of wages for women earning some two-year college credits looks quite similar to that of high school graduates.

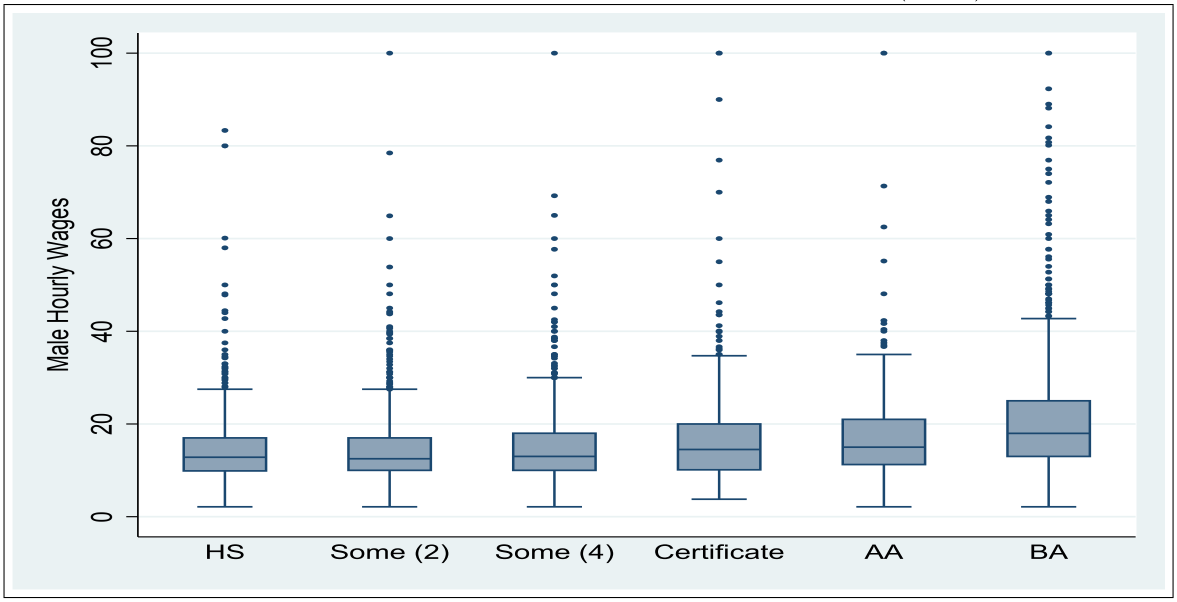
Figure 1: Hourly Wages by Educational Attainment (Female)



For males, it is not clear that high school graduates' median wages are different from those who attend some college. This may partially explain the rate at which males in this sample enroll in postsecondary institutions and why their completion rates are consistently lower than that of similar females.

With both genders, the highest wage benefits lie in earning a baccalaureate degree, which is consistent with previous literature. Also, the distribution of wages for women in each educational category is more tightly centered around the median than the distribution of wages for men.

Figure 2: Hourly Wages by Educational Attainment (Male)



The question I answer in this paper is whether there are financial returns to attending college or earning a degree among sub-baccalaureate earners, as compared to students who complete high school and do not attempt any postsecondary credential. Using the Education Longitudinal Study of 2002, I consider individuals at the beginning of their careers. It is possible that students enter college and make connections which then lead to job prospects, even if they do not finish a degree. On the other hand, attending and not earning a degree may send a negative signal to future employers. If benefits lie in degree completion, then more must be done to help students attain these goals. Especially if free tuition becomes a reality, it is imperative that we understand the differential impacts of college attendance versus degree completion. In the next section I review the literature to date. I will then go through the data and methodology, the results, and a discussion and policy implications to conclude.

2 Literature Review

Previous literature has typically focused on the returns to bachelor's degrees. Some studies include controls for an associate degree, but do not differentiate between associate degrees and postsecondary certificates, or between these credentials and some college but no degree. Typically this is due to small sample sizes and variation in the meaning of an undergraduate certificate. Further, among those studies that consider some college separately, it is rare to see two and four-year college credits considered separately. Thus, most research that considers the returns to education combines a very heterogeneous group into the "some college" category, imposing a restriction on the returns to education for students with some college to have the same returns as those who earn either a certificate or associate degree. Especially if wage benefits apply to degree earners only, disaggregating this group is important. Nationally, 21.2% of individuals over 25 have some postsecondary experience only, so there are many people for which the returns may be overstated (U.S. Census Table S1501, 2014). For many years, it has been suggested that simply attending college has a positive impact on earnings, but it is possible this is no longer the case.

Of the studies that do disaggregate some college, many use data that are several decades old, and do not account for selection bias that is inherent in education research. However, several studies have found positive returns to both some postsecondary experience, as well as associate degree attainment (Bailey et. al (2004), Kane and Rouse (1995), Grubb (1993, 1995, 1997, 2002), Gill and Leigh (1997, 2003), and Marcotte et al. (2005), and Monk-Turner (1994)). Using the National Longitudinal Survey of Youth of 1979 (NLSY), both Gill and Leigh (1997) and Kane and Rouse (1995) find positive and significant effects of earning an associate degree. Kane and Rouse (1995) find benefits of 21% for women and 23% for men to earning an associate degree, and about 40% to earning a bachelor's degree. Earnings benefits to some college were markedly smaller, between 4-9% for men and as much as 6% for women. Gill and Leigh (1997) find similar magnitudes among continuing students in their study, though they find between 8-15% improvements to wages for some college credits for men, and higher estimates for women except among bachelor's degree earners. Kane and Rouse (1995), using the National Longitudinal Survey of 1972 (NLS-72), see much smaller estimates for men; 4% for some two-year credits or an associate degree, and 26% for a bachelor's degree. However, their results for women are similar, 6-7% for some college, 26% for an associate degree, and 33% for a bachelor's degree. Marcotte et al. (2005) use the National Education Longitudinal Study of 1988 (NELS:88) to study a cohort of students enrolled in the 1990s, and find positive effects of associate degrees on earnings of between 17 - 40%, though they do not find much in terms of earned credits or to those students who earn undergraduate certificates.

On the other hand, Grubb (1995) finds negligible certificate effects for both men and women using the NLS-72. Further, he finds negative effects on earnings for men with vocational or academic associate degrees, and positive effects for women only from a vocational associate degree. There were mixed effects among those students earning some college credits but no degree for both men and women. With the SIPP data, however, Grubb (1995) finds more similar effects to the above literature on earnings, with stronger vocational impacts for men, but otherwise higher impacts for women. Grubb (1997) expanded the SIPP data analysis to compare earnings over time; between 1984-1990, returns to an associate degree fall for both genders, returns to vocational certificates fall for men but increase for women, and returns to some college all but disappear for women and decline for men. He concludes that the returns to sub-baccalaureate degrees are quite variable. These potentially negative returns were also found in Anderson (1984) and Breneman and Nelson (1981).

Bailey et. al (2004) support Grubb (2002)'s claims. Using the NELS-88 and High School and Beyond (HS&B) surveys, they find that later cohorts saw returns to some college education and that returns to undergraduate certificates were only seen among women. Both associate and bachelor degree earners saw significant increases, though the associate degree benefits seem to lie with occupational, rather than academic, majors. Monk-Turner (1994) further shows that the wage benefit to starting at a four-year college outweighs the cost-savings of a two-year college among 27 year olds in the Parnes National Longitudinal Study. Overall, these studies generally agree that there are returns to an associate or bachelor's degree, but identify more variation among some credit earning and certificates or diplomas.

Some more recent studies have used panel data and fixed effects models to determine the impact of community college degrees, and show strong positive labor market effects due to undergraduate certificates and associate degrees. Among students in Kentucky, Jepson et al (2014) find positive effects of community college degrees on labor market outcomes; females see increases of almost 40% upon earning an associate degree or diploma, and men see increases of about 18-20%. They consider two recent cohorts of students, those that started college during the 2002-2003 school year, and those that started college during the 2003-2004 school year. Dadgar and Trimble (2015), also using a fixed effects model and a cohort of students from the 2001-2002 year, consider returns to sub-baccalaureate credentials with respect to a Washington state community college. They find increases to credentials are greater for women, although there was more variation in returns by major rather than degree. Liu et. al (2015) study North Carolina community college entrants in 2002-2003, and find strong returns to associate and bachelor's degrees, and small returns for students with some postsecondary credits, certificates, or diplomas. These studies often utilize a specific group of students, those returning to higher education after being in the workforce, to identify returns to education. While these results provide an interesting contribution to the literature, in addition to students returning to school, these studies also require a specific type of dataset and are typically focused on one school or location.

Identification is an ongoing issue in this literature, as it is likely that there are both measurement error concerns as well as omitted variables in the wage equation. However, without a policy change, natural experiment, or panel data, researchers are left with identifying an instrument or utilizing matching methods in order to address endogeneity. Determining an instrument that impacts income only through educational attainment is a challenge. Using birth order, family's educational hopes for their children, and financial shocks as instruments, Blundell et al. (2004) find stronger returns to higher education as compared to their least squares estimates, though they were imprecisely estimated. Others have considered distance to college. However, identifying a strong instrument remains a unsolved problem.

Matching methods are an alternative method to identify the causal effect of education on income. Blundell et al. (2004) use propensity score matching to identify the effect of education on earnings in a similar comparison study of traditional regression analysis, instrumental variables, and matching methods. They find significant returns to various education levels among the different estimation methods. Brand and Xie (2010), using the Wisconsin Longitudinal Study and the NLSY 1979, find that students who are least likely to obtain a college degree are the most likely to benefit from a diploma, again using propensity score matching. As there are drawbacks to using matching methods as well, there are few studies employing these methods currently. First, conditional independence is required; once the matching has taken place, the covariates included should have similar means and variance between the treatment and control. The second assumption is overlap; there must be a positive probability of any observation being in either the treatment or control. This is often difficult to obtain in observational studies.

Earlier studies have generally shown that there is a positive effect of postsecondary schooling on labor market outcomes. However, many of these studies are several decades old, and it is unclear whether the effect of attending a community college has changed over time. Further, failing to account for selection bias of attainment choices potentially biases the results. This paper will contribute to the literature in several ways. First, I look to new data and a new cohort of job seekers in a rapidly changing labor market. These students are at the beginning of their careers, which has different implications with respect to the returns to education as compared to more experienced workers. With student loan debt a very real concern for many students, community colleges and the degrees within might be more enticing in terms of labor market outcomes among recent graduates. Second, I look to the effect of community college certificates separately from associate degrees since recent literature suggests the earnings gap might vary by type of degree earned. Further, I look to compare these students to those who did not earn a postsecondary credential. I present ordinary least squares estimates in conjunction with estimates that account for endogeneity and compare results in the sections that follow.

3 Data

The Education Longitudinal Study of 2002 (ELS:2002) data set is a stratified two-stage sample of schools, and students within them, beginning in the spring of 2002. The survey had several follow-ups: spring 2004, 2006, and 2012. In the first stage, approximately 750 high schools were drawn with probabilities inversely related to school size. In the second stage, approximately 30 high school sophomores in each school were sampled; students from Hispanic and Asian populations were over sampled to ensure adequately sized subsamples. As I am concerned with the effect of collegiate choices on labor market outcomes, I begin by including only students who were high school seniors in 2004 and had earned a high school credential by 2012. This reduced my sample from 16,200 to 13,910. Further, I include only individuals with reported hourly wages or income data in the third follow-up. This reduced the sample to

11,550. I then consider only students who earned at most a bachelor's degree, which reduced the sample to 10,410. I do not include students pursuing graduate degrees, as they would be too close to the beginning of their career to feasibly measure educational returns. I drop 500 students who attended for-profit schools, as such schools have different incentives and possibly outcomes.[3] I drop 30 students who were missing postsecondary information, 500 students with missing ELS proctored math scores, and 510 respondents who indicated they were self-employed. Thus, the final sample is 8930 respondents.[4] These respondents range in age from 26-30 at the time of the last questionnaire, with the majority being 27. Thus, these respondents are early in their career and will lack significant labor market experience that is often included in labor market research.

The NCES survey asked respondents for annual earnings in 2011 before taxes and deductions. Total income, on the other hand, is comprised of the respondent's earnings in 2011 and her spouse's earnings (if applicable). Hourly wages at the respondent's current or most recent job were also asked within the survey, although to preserve anonymity, these values were bottom coded at \$2.13 and top coded at \$100. The labor market variables I consider are (1): a binary variable indicating whether the respondent is employed full-time, (2): a categorical variable indicating the respondent's employment status (0: out of labor market, 1: unemployed, 2: working part-time, 3: working full-time), (3): the natural log of hourly wages, (4): the natural log of respondent income, and (5): the natural log of total income.[5] In the analysis

[3]As a robustness check, leaving these students in does not affect the results. Including for-profit schools implies slightly smaller magnitudes in some cases, though not statistically different from those presented.

[4]All reported sample sizes are rounded to the nearest 10, as required by the ELS restricted-use agreement with NCES. The NCES provides weights that combine the cross-sectional and longitudinal nature of the survey, which accounts for both nonresponse bias and over-sampling of some populations. Here, I use the weight f3f1tpnlwt to use data from 2004 - 2012. This does not include transcripts, but the results with the transcript weight were similar.

[5]There were some missing observations in the wage data. I consider two possibilities and show that they produce similar results: (1) replace missing values with zero's and (2) recode missing hourly wages using earned income divided by 2080 (40 hours per week/52 weeks) for those indicating positive earned income. Individuals who reported a "per day" wage were coded as missing (-9) due

below, I focus on hourly wages and respondent's income as the dependent variables.

I consider six mutually exclusive groups of students; students who complete high school but no college, students who start at a two-year college and do not earn a degree, students who start at a four-year college and do not earn a degree, students who earn a certificate, students who earn an associate degree, and students who earn a bachelor's degree. Below, in Table 1, I present descriptive statistics of all income variables; the top panel includes observations with zero wages or incomes while the bottom panel excludes such observations.

Table 1: Income Variables

	HS	Some (2)	Some (4)	Cert	AA	BA
Hourly Wages (1)	12.12	11.99	13.26	14.32	14.75	17.88
	(0.342)	(0.203)	(0.306)	(0.502)	(0.401)	(0.227)
Hourly Wages (2)	12.71	12.69	14.00	15.03	15.34	18.69
	(0.333)	(0.203)	(0.290)	(0.495)	(0.387)	(0.226)
Respondent Earnings (2011)	20995.59	21259.17	24184.41	23863.37	25257.64	33665.01
	(719.285)	(563.768)	(695.588)	(815.135)	(848.922)	(692.069)
Total Earnings (2011)	33944.99	33927.64	36656.33	37926.71	41352.98	50609.51
	(1023.609)	(878.781)	(1045.851)	(1252.333)	(1392.403)	(980.769)
Employment	0.64	0.66	0.68	0.72	0.72	0.79
	(0.018)	(0.013)	(0.016)	(0.018)	(0.018)	(0.009)
N	920	1610	1310	810	800	3480
Hourly Wages (+)	13.96	13.87	14.66	15.60	16.11	19.30
	(0.361)	(0.202)	(0.309)	(0.557)	(0.326)	(0.235)
N	830	1440	1200	760	740	3250
Respondent Earnings (2011) (+)	25488.39	24808.54	26704.60	25689.87	28009.06	35604.80
	(733.912)	(631.034)	(726.257)	(790.282)	(860.833)	(729.902)
N	750	1370	1160	730	730	3270
Total Earnings (2011) (+)	38113.00	37518.64	39061.67	39340.23	43852.23	52824.31
	(1171.295)	(990.935)	(1146.101)	(1298.719)	(1434.511)	(1041.709)
N	820	1460	1210	770	770	3330

Standard errors in parentheses
(1): Hourly wages calculated such that missing observations are dropped.
(2): Hourly wages calculated such that missing observations replaced with respondent income/2080.
+: Only positive values included

The raw data do not reveal significant differences between students who finish high school and those who earn some two-year college credits. However, there are significant wage differences between students who have some two-year college experience and those that have some four-year college experience, suggesting a possible

to lack of information regarding hours per day worked.

difference in quality of education or selection of students. Further, there are moderate differences between some four-year college experience and earning a certificate, and significant differences to earning an associate degree. Differences between earning a certificate or associate degree appear small.

There are many reasons why we might be interested in wage variations by gender, so in Tables 2 and 3, I consider income variables separately for women and men.

Table 2: Income Variables - Female

	HS	Some (2)	Some (4)	Cert	AA	BA
Hourly Wages (1)	9.36	10.59	12.52	12.46	14.13	17.08
	(0.396)	(0.245)	(0.336)	(0.330)	(0.361)	(0.258)
Hourly Wages (2)	9.68	10.88	13.00	12.99	14.39	17.65
	(0.392)	(0.246)	(0.333)	(0.298)	(0.349)	(0.250)
Respondent Earnings (2011)	12791.59	16373.24	21669.22	19312.18	21108.72	30284.36
	(802.146)	(809.333)	(899.591)	(707.817)	(884.828)	(570.299)
Total Earnings (2011)	33016.74	32739.67	38819.51	36736.03	41398.65	52461.20
	(1816.979)	(1351.264)	(1691.348)	(1491.091)	(1845.520)	(1133.431)
Employment	0.44	0.58	0.65	0.65	0.65	0.76
	(0.030)	(0.019)	(0.023)	(0.025)	(0.025)	(0.013)
N	370	790	630	480	470	1890
Hourly Wages (+)	11.32	12.21	13.65	13.60	15.26	18.20
	(0.479)	(0.240)	(0.361)	(0.313)	(0.327)	(0.257)
N	320	710	580	450	440	1780
Respondent Earnings (2011) (+)	18188.65	20567.91	24239.54	21053.00	24059.65	32037.95
	(903.141)	(890.033)	(859.537)	(704.131)	(890.187)	(550.409)
N	270	630	540	420	420	1780
Total Earnings (2011) (+)	38014.81	36699.05	40253.89	37528.10	43560.68	54011.46
	(2369.037)	(1507.060)	(1815.092)	(1602.355)	(1881.174)	(1194.082)
N	330	710	590	450	460	1820

Standard errors in parentheses
(1): Hourly wages calculated such that missing observations are dropped.
(2): Hourly wages calculated such that missing observations replaced with respondent income/2080.
+: Only positive values included

For women, there are significant differences between those who finish high school and those who attempt two or four-year credits. On average, a woman who completes high school is earning about a dollar less per hour by not going to a two-year college, and over three dollars less by not attending a four-year college. Additionally, there seem to be significant differences between those who attempt two and four-year credits, and those who earn certificates or an associate degree. However, there is only some evidence of differences between attempting some four-year credits and earning a certificate or associate degree. In all cases, there is a significant increase to earning a bachelor's degree.

Table 3: Income Variables - Male

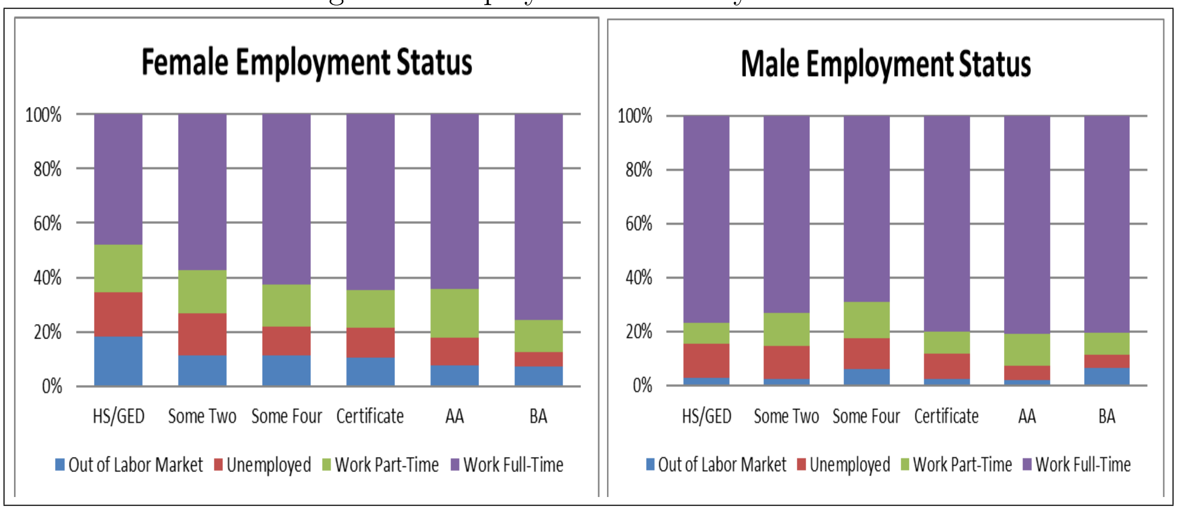
	HS	Some (2)	Some (4)	Cert	AA	BA
Hourly Wages (1)	13.82	13.39	13.88	16.80	15.60	18.77
	(0.468)	(0.327)	(0.503)	(1.067)	(0.786)	(0.364)
Hourly Wages (2)	14.58	14.49	14.85	17.74	16.67	19.87
	(0.455)	(0.320)	(0.483)	(1.042)	(0.730)	(0.363)
Respondent Earnings (2011)	26052.19	26113.81	26302.72	29933.72	31002.62	37483.82
	(965.423)	(809.275)	(1059.239)	(1445.570)	(1501.306)	(1334.795)
Total Earnings (2011)	34517.12	35108.01	34834.49	39514.84	41289.75	48517.82
	(1258.708)	(1152.305)	(1412.747)	(1957.199)	(1851.157)	(1522.010)
Employment	0.76	0.75	0.70	0.80	0.82	0.82
	(0.020)	(0.017)	(0.023)	(0.024)	(0.025)	(0.013)
N	550	820	680	330	330	1600
Hourly Wages (+)	15.24	15.39	15.52	18.24	17.30	20.58
	(0.483)	(0.319)	(0.525)	(1.155)	(0.617)	(0.379)
N	510	730	620	310	290	1470
Respondent Earnings (2011) (+)	29054.22	28703.20	28791.03	31789.04	33502.70	39737.72
	(974.797)	(919.080)	(1131.517)	(1357.918)	(1538.323)	(1436.424)
N	480	730	620	310	310	1490
Total Earnings (2011) (+)	38160.96	38271.37	38052.57	41723.84	44257.77	51448.77
	(1318.303)	(1335.828)	(1568.684)	(1950.512)	(1988.336)	(1641.528)
N	490	750	620	310	310	1510

Standard errors in parentheses
(1): Hourly wages calculated such that missing observations are dropped.
(2): Hourly wages calculated such that missing observations replaced with respondent income/2080.
+: Only positive values included

Aside from the large difference between those that earn a bachelor's degree and those that do not, significant wage differences only exist between students who attempt some four-year credits and those who earn a certificate or an associate degree for men. Especially interesting is the small gap between earning a certificate and a bachelor's degree for men, though there is more variation in the earnings variables for those who earn a certificate. These wage differences suggest the importance of completing a degree might vary by gender.

In Figure 3, I consider differences in employment status by gender and degree. For both men and women, the percent of those working full-time increases with degree attainment. Interestingly for men, however, the percent of high school graduates working full-time is slightly higher than those with some college but no degree. For both genders, the fraction of unemployed decreases with attainment, as previous research has suggested.

Figure 3: Employment Status by Gender



Of course, wages and the probability of being in the workforce are affected by many characteristics. Below, I present descriptive statistics for several individual, background, school, and geographic characteristics by educational attainment.

Table 4: Individual Characteristics

	HS	Some (2)	Some (4)	Cert	AA	BA
Gender (1: Male)	0.616	0.504	0.536	0.437	0.417	0.465
	(0.021)	(0.017)	(0.018)	(0.022)	(0.022)	(0.012)
African American	0.143	0.142	0.170	0.195	0.121	0.085
	(0.015)	(0.012)	(0.016)	(0.020)	(0.016)	(0.008)
Hispanic	0.166	0.248	0.111	0.153	0.140	0.085
	(0.016)	(0.019)	(0.012)	(0.017)	(0.018)	(0.007)
White	0.604	0.532	0.633	0.571	0.655	0.732
	(0.021)	(0.020)	(0.020)	(0.023)	(0.024)	(0.012)
Asian	0.017	0.030	0.035	0.044	0.028	0.059
	(0.004)	(0.004)	(0.004)	(0.006)	(0.006)	(0.005)
Other	0.070	0.048	0.051	0.037	0.057	0.040
	(0.010)	(0.007)	(0.008)	(0.008)	(0.010)	(0.005)
ELS Proctored Math Score (2004)	42.556	46.476	51.865	45.830	49.093	56.085
	(0.342)	(0.289)	(0.342)	(0.423)	(0.379)	(0.205)
Number of Carnegie Units	23.983	24.418	25.685	24.976	25.809	26.325
	(0.227)	(0.179)	(0.168)	(0.218)	(0.188)	(0.147)
Academic GPA (High School)	2.092	2.244	2.711	2.433	2.679	3.155
	(0.028)	(0.022)	(0.027)	(0.029)	(0.030)	(0.014)
BA Expectations (2004)	0.222	0.543	0.838	0.525	0.685	0.940
	(0.014)	(0.015)	(0.014)	(0.022)	(0.022)	(0.005)
Married (2012) (1: Yes)	0.316	0.266	0.263	0.305	0.329	0.278
	(0.018)	(0.015)	(0.017)	(0.019)	(0.020)	(0.011)
Children (2012) (1: Yes)	0.437	0.348	0.260	0.378	0.303	0.116
	(0.021)	(0.014)	(0.016)	(0.020)	(0.020)	(0.008)
Worked in 2006 (1: Yes)	0.440	0.111	0.035	0.091	0.051	0.009
	(0.020)	(0.010)	(0.007)	(0.012)	(0.010)	(0.002)
Job Training (1: Yes)	0.302	0.308	0.324	0.354	0.437	0.471
	(0.018)	(0.015)	(0.016)	(0.020)	(0.022)	(0.011)
N	920	1610	1310	810	800	3480

Standard errors in parentheses

Individual characteristics expected to affect wages and employment opportunities include race, academic ability, whether the individual is married or has children, and work experience. Over 60% of those students who do not attempt any postsecondary experience are male, possibly due to differences in job opportunities by gender. Hispanic students make up almost 25% of students who earn some two-year college credits, compared with 15% of those who earn a certificate, 14% of those earning an associate degree, and 9% of those earning a bachelor's degree. Student ELS proctored math scores, academic coursework, and academic high school GPAs do rise with college participation, which is expected.[6] Individuals with no postsecondary experience

[6] Carnegie units measure academic coursework. A student who follows a traditional college preparatory track in high school would earn approximately six Carnegie units per school year.

are much more likely to have children by 2012, although they are not necessarily more likely to be married. About 44% of students with no postsecondary experience were working in 2006, compared to only 1% of those who earned a bachelor's degree.[7] I additionally include whether the respondent had on the job training. Base year urbanicity is included to control for cost of living adjustments.

In Table 5 below, I consider background characteristics.

Table 5: Background Characteristics

	HS	Some (2)	Some (4)	Cert	AA	BA
Parental Education						
Less than High School	0.095	0.079	0.039	0.082	0.034	0.019
	(0.011)	(0.010)	(0.007)	(0.013)	(0.007)	(0.003)
High School/GED	0.352	0.215	0.182	0.239	0.264	0.130
	(0.016)	(0.013)	(0.014)	(0.017)	(0.020)	(0.008)
Some Two-Year College	0.152	0.147	0.092	0.138	0.126	0.079
	(0.013)	(0.011)	(0.010)	(0.013)	(0.013)	(0.006)
Some Four-Year College	0.099	0.117	0.119	0.108	0.073	0.092
	(0.011)	(0.009)	(0.011)	(0.013)	(0.011)	(0.006)
AA/Certificate	0.128	0.165	0.139	0.163	0.163	0.110
	(0.012)	(0.012)	(0.013)	(0.015)	(0.016)	(0.007)
BA	0.123	0.171	0.260	0.161	0.219	0.299
	(0.013)	(0.012)	(0.015)	(0.015)	(0.018)	(0.010)
Greater than BA	0.052	0.105	0.170	0.110	0.121	0.270
	(0.008)	(0.011)	(0.013)	(0.014)	(0.013)	(0.010)
Parental Income						
Low Income ($< \$35,000$)	0.489	0.385	0.300	0.392	0.305	0.186
	(0.019)	(0.016)	(0.016)	(0.020)	(0.020)	(0.009)
Middle Income ($\$35,000 - 75,000$)	0.401	0.436	0.423	0.417	0.457	0.408
	(0.018)	(0.015)	(0.017)	(0.020)	(0.022)	(0.012)
High Income ($> \$75,000$)	0.109	0.179	0.277	0.191	0.238	0.406
	(0.012)	(0.013)	(0.015)	(0.015)	(0.018)	(0.012)
N	920	1610	1310	810	800	3480

Standard errors in parentheses

Students who start their postsecondary career at a four-year college are more likely to have parents with at least a bachelor's degree. Students who have the least amount of education are also most likely to have parents with low educational attainment. Additionally, income brackets follow expected distributions according

[7]The experience variable (worked in 2006) is generated using several questions in the 2006 survey, which asked students if they were working in the beginning half of 2006. If a student was working at least three of the six months in the first half of 2006, I considered them working. While not an exact measure of experience, this survey does not ask for actual experience. Further, as these students are all the same age and many work part-time through higher education, potential experience is difficult to measure.

to what students' educational attainment is, with the exception of those students earning some four-year college credits looking more similar to the income distribution of students with a bachelor's degree.

Below, in Table 6, I consider characteristics of the respondent's high school and geographic characteristics.

Table 6: School Characteristics

	HS	Some (2)	Some (4)	Cert	AA	BA
School and Geographic Characteristics						
School: Average Math Score (2004)	48.042	48.950	50.542	49.190	49.662	52.009
	(0.204)	(0.212)	(0.252)	(0.231)	(0.258)	(0.202)
School: Percent White	56.222	50.381	57.811	54.843	59.461	62.676
	(1.578)	(1.474)	(1.542)	(1.714)	(1.733)	(1.078)
School: % Peers Attending Two-Year Schools	28.499	32.299	20.688	28.206	28.509	25.213
	(0.789)	(0.752)	(0.708)	(0.912)	(0.890)	(0.729)
School: % Peers Attending Four-Year Schools	31.147	30.025	44.221	34.551	35.241	44.756
	(0.914)	(0.832)	(1.088)	(0.982)	(1.077)	(0.927)
School: Attended Public High School	0.985	0.956	0.904	0.942	0.932	0.867
	(0.003)	(0.005)	(0.007)	(0.006)	(0.008)	(0.007)
School: Rural	0.237	0.201	0.189	0.213	0.236	0.198
	(0.015)	(0.015)	(0.017)	(0.016)	(0.022)	(0.013)
School: Suburban	0.508	0.525	0.462	0.509	0.545	0.509
	(0.020)	(0.018)	(0.018)	(0.022)	(0.024)	(0.015)
School: Urban	0.256	0.274	0.348	0.277	0.218	0.293
	(0.020)	(0.018)	(0.017)	(0.021)	(0.018)	(0.013)
County: Per Capita Income (2004)	3.220	3.265	3.344	3.237	3.301	3.462
	(0.044)	(0.043)	(0.044)	(0.055)	(0.052)	(0.039)
County: Unemployment Rate (2004)	5.677	6.057	5.571	5.844	5.657	5.510
	(0.071)	(0.171)	(0.061)	(0.146)	(0.086)	(0.055)
N	920	1610	1310	810	800	3480
Standard errors in parentheses						

Students who start at four-year colleges but do not finish a degree come from schools that look quite similar to students with bachelor's degrees. However, students who do finish a bachelor's degree are more likely to have come from suburban schools that have a higher average math score among students sampled. Bachelor's degree earners are the least likely to have attended a public high school and attend high schools with the largest percent of White students. The Bureau of Labor Statistics (BLS) county level data are used to capture neighborhood effects and include unemployment rates (2004) and per capita income (2004).[8]

[8]Bozick (2009), using the ELS:2002 data, finds that students who graduate high school with more

Finally, for respondents who attended college, I look at the characteristics of their postsecondary experiences in Table 7.

Table 7: Postsecondary Characteristics

	Some (2)	Some (4)	Cert	AA	BA
Attempted Credits: Two-Year (Years)	1.681	0.402	1.755	2.494	0.590
	(0.049)	(0.031)	(0.081)	(0.086)	(0.031)
Attempted Credits: Four-Year (Years)	0.437	2.777	1.172	1.602	4.346
	(0.038)	(0.079)	(0.098)	(0.099)	(0.035)
Debt: Earnings (2011)[a]	4.540	4.300	1.396	1.896	2.176
	(2.321)	(3.003)	(0.339)	(0.442)	(0.310)
Debt: Earnings (2011)[b]	4.378	1.154	1.289	1.776	1.748
	(2.320)	(0.130)	(0.327)	(0.440)	(0.275)
First Year Postsecondary GPA	2.106	2.302	2.440	2.706	2.991
	(0.046)	(0.044)	(0.052)	(0.042)	(0.018)
Took Remedial Courses (1: Yes)	2.184	1.050	1.787	1.684	0.602
	(0.102)	(0.076)	(0.132)	(0.128)	(0.034)
Student Loan (1: Yes)	0.402	0.679	0.563	0.595	0.693
	(0.020)	(0.020)	(0.026)	(0.025)	(0.012)
Took Break From Education > 4 months	0.850	0.624	0.716	0.847	0.247
	(0.039)	(0.034)	(0.050)	(0.043)	(0.013)
Attended Full Time	0.480	0.786	0.633	0.733	0.912
	(0.020)	(0.018)	(0.026)	(0.022)	(0.006)
Major Choice					
Business	0.036	0.075	0.048	0.076	0.140
	(0.005)	(0.010)	(0.008)	(0.011)	(0.008)
Social Science	0.013	0.064	0.016	0.031	0.109
	(0.003)	(0.008)	(0.005)	(0.007)	(0.007)
Engineering	0.009	0.029	0.014	0.031	0.058
	(0.003)	(0.005)	(0.005)	(0.007)	(0.005)
Computer Science	0.012	0.016	0.004	0.016	0.019
	(0.003)	(0.004)	(0.003)	(0.006)	(0.003)
Humanities	0.029	0.076	0.030	0.077	0.149
	(0.005)	(0.008)	(0.006)	(0.012)	(0.007)
Education	0.032	0.048	0.016	0.035	0.075
	(0.005)	(0.007)	(0.005)	(0.007)	(0.006)
Health	0.041	0.050	0.117	0.153	0.075
	(0.005)	(0.007)	(0.012)	(0.015)	(0.005)
Life and Physical Sciences	0.014	0.040	0.020	0.021	0.064
	(0.003)	(0.006)	(0.005)	(0.005)	(0.005)
Math	0.001	0.002	0.000	0.001	0.011
	(0.001)	(0.001)	(0.000)	(0.001)	(0.002)
Vocational/Technical	0.033	0.020	0.051	0.090	0.027
	(0.005)	(0.005)	(0.009)	(0.012)	(0.004)
Missing Major	0.780	0.580	0.683	0.466	0.271
	(0.012)	(0.017)	(0.019)	(0.020)	(0.009)
N	1610	1310	810	800	3480

Standard errors in parentheses
a: Employment income: respondent + spouse
b: Employment income: respondent + spouse + non-employment income

job opportunities not requiring higher education are more likely to enter the labor force immediately.

The first two variables indicate full-time equivalent years of credits attempted in two and four-year institutions.[9] Total credits attempted in an institution are divided by 30 to obtain full-time equivalent units. The NCES normalized total credits to account for institutional differences in credit earning. For example, the average student who earns an associate degree has attempted 2.5 years in a two-year college and 1.6 years in a four-year college. Students with some two-year experience are more likely to take remedial courses, and the least likely to attend full-time. Particularly concerning are the 40% of students with some two-year college credits and 68% of students with some four-year college credits with student loan debt.

4 Methodology and Results

To consider the effect of educational decisions on wages and income, I consider three methods: ordinary least squares, two stage least squares, and propensity score matching. Ordinary least squares estimates assume that educational attainment is exogenous, which is unlikely, but provide a baseline for other models. On the other hand, two stage least squares is preferred if an ideal instrument is obtained, but identifying a variable that impacts education decisions but not income is difficult. Matching methods suffer from bias if there are unobservable factors that impact both the "treatment," or educational decision i, and "control," or educational decision j. However, with enough covariates, I hope to control for these. The dependent variables considered in each of the models below are the natural log of hourly wages and the respondent's income.

[9] I combine two and less than two-year colleges for the purposes of brevity.

4.1 Ordinary Least Squares

Many studies that consider labor market returns look to ordinary least squares estimates for the impact educational decisions have on outcomes. The covariates include individual and background characteristics: marital status, parental status, and race. Additionally, the ELS proctored math score in 2004 is included to control for ability, and labor market status in 2006 and on the job training are included to control for potential work experience. Urbanicity and census region dummies are included to control for differences in cost of living.[10] Finally, dummy variables for parental income between \$35,000 and \$75,000, and parental income over \$75,000 are included to control for background characteristics.[11] Further, all models are adjusted for stratification, weighted to account for population, and the standard errors are clustered at the school level.

As it is not likely that the returns to different levels of higher education are linear, I break down the "Years of Schooling" variable into three models which all consider five choices: Some Two-Year College, Some Four-Year College, Certificate, Associate Degree, and Bachelor's Degree. However, measuring "some" college can impact the results and implications. Thus, I consider three variations. The first is to separate those students with postsecondary experience but no degree into those who started at a two-year college and those who started at a four-year college. This model's advantage is that the educational choice variables are mutually exclusive, and is specified below in equation (1).

[10] I use census variables rather than state to increase comparability among results both within this paper and among other literature. While it has been shown that different areas see different returns to education, the sample sizes do not support such a degree of precision (Black et al. 2008).

[11] Replacing income variables with more descriptive dummies did not impact the models in a significant way, nor did replacing these variables with a measure of socio-economic status. Additionally, I included unemployment rates in 2004, but this had no impact on the models. I chose to not include student loans in these models as they may have confounding impacts.

$$Y_i = \gamma_1 Some(2) + \gamma_2 Some(4) + \gamma_3 Cert + \gamma_4 AA + \gamma_5 BA + \beta' X_i + \epsilon_i \tag{1}$$

Secondly, I include variables for the credits earned per year at a two or four-year institution rather than where the respondent began their postsecondary education. Total credits has the advantage of comparing individuals with similar invested years in higher education. Further, students who earn the same degree but take different amounts of credits may have different returns. Below, in equation (2), students who start at a two or four-year college but do not earn a postsecondary credential are replaced with the years of education attempted at a two or four-year college. These two sets of variables are correlated, but not identical; many students take coursework at both institutional levels. Further, in the second model, students who eventually earn a degree are also included in the credit earning variables so their interpretation includes all students who enroll in postsecondary education. This model is important because of the growing number of students who take longer than the defined time to complete a degree (eg. two years for an AA or four for a BA). Additional years might have a positive impact in that the student completes the degree that is best suited for them, or it could have negative implications in the labor market if employers see length to degree as a negative signal of ability.

$$Y_i = \gamma_1 FTE(2) + \gamma_2 FTE(4) + \gamma_3 Cert + \gamma_4 AA + \gamma_5 BA + \beta' X_i + \epsilon_i \tag{2}$$

Finally, I take the model above and restrict attempted credits to only students who do not complete a degree. This model has the benefit of being more similar to previous literature and mutually exclusive groups, although students who do not earn a degree could plausibly earn credits at both a two and four-year college.[12]

In each of the tables below, I see expected effects of marriage and children on women and men. Further, I see expected signs for math scores, job training, and family income, although the indicator for having worked in 2006, a proxy variable for whether a student had work experience, was not a significant determinant of wages or income. In the discussion below, however, I will focus on the returns to education.

[12]This is actually quite common in this data set, so most students with some credits have attempted credits in both two and four-year institutions.

Table 8: OLS Estimates (Dependent Variable ln(wage))

	Female			Male		
Started Two-Year	0.023			-0.014		
	(0.030)			(0.029)		
Started Four-Year	0.077**			-0.035		
	(0.031)			(0.034)		
Attempted Credits Two-Year		-0.012**			-0.028***	
		(0.006)			(0.008)	
Attempted Credits Four-Year		0.000			-0.026***	
		(0.005)			(0.006)	
Attempted Credits Two-Year (No Degree)			-0.007			-0.014
			(0.009)			(0.009)
Attempted Credits Four-Year (No Degree)			0.014**			-0.016*
			(0.007)			(0.008)
Certificate	0.139***	0.111***	0.111***	0.104**	0.129***	0.098***
	(0.029)	(0.020)	(0.022)	(0.041)	(0.037)	(0.037)
AA	0.204***	0.184***	0.177***	0.065*	0.139***	0.057*
	(0.032)	(0.023)	(0.026)	(0.037)	(0.030)	(0.031)
BA	0.281***	0.236***	0.253***	0.170***	0.257***	0.159***
	(0.028)	(0.022)	(0.025)	(0.031)	(0.025)	(0.025)
Married	0.086***	0.085***	0.086***	0.167***	0.163***	0.166***
	(0.016)	(0.016)	(0.016)	(0.019)	(0.018)	(0.019)
Children	-0.072***	-0.075***	-0.071***	-0.009	-0.017	-0.012
	(0.014)	(0.014)	(0.014)	(0.020)	(0.020)	(0.020)
Math Score (2004)	0.007***	0.007***	0.007***	0.004***	0.005***	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Worked in 2006	0.015	-0.013	0.001	0.013	-0.008	0.007
	(0.033)	(0.030)	(0.030)	(0.027)	(0.026)	(0.026)
Job Training	0.153***	0.154***	0.152***	0.157***	0.156***	0.156***
	(0.015)	(0.015)	(0.015)	(0.019)	(0.018)	(0.019)
Asian	0.068**	0.069**	0.067**	-0.007	0.007	-0.002
	(0.027)	(0.027)	(0.027)	(0.030)	(0.030)	(0.030)
African American	0.031	0.035*	0.033	-0.074***	-0.069**	-0.072***
	(0.021)	(0.021)	(0.021)	(0.027)	(0.027)	(0.027)
Hispanic	0.035	0.036	0.037	-0.059*	-0.061**	-0.059*
	(0.023)	(0.023)	(0.023)	(0.031)	(0.030)	(0.030)
Other	-0.017	-0.015	-0.013	-0.013	-0.014	-0.013
	(0.030)	(0.030)	(0.030)	(0.038)	(0.038)	(0.038)
Middle Income	0.045***	0.046***	0.045***	0.026	0.034*	0.029
	(0.016)	(0.016)	(0.016)	(0.021)	(0.020)	(0.021)
High Income	0.059***	0.060***	0.058***	0.056**	0.066***	0.060**
	(0.021)	(0.021)	(0.021)	(0.024)	(0.024)	(0.024)
Suburban	-0.006	-0.007	-0.006	0.034	0.035*	0.035
	(0.018)	(0.018)	(0.018)	(0.021)	(0.021)	(0.021)
Rural	-0.004	-0.004	-0.004	0.022	0.021	0.022
	(0.023)	(0.023)	(0.023)	(0.025)	(0.025)	(0.025)
Constant	2.196***	2.228***	2.221***	2.440***	2.416***	2.430***
	(0.061)	(0.058)	(0.059)	(0.071)	(0.071)	(0.071)
N	Observations: 4290			Observations: 3930		

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Census dummies included but not reported.

In Table 8, I look at the impact of education on wages for women and men for the above three models. For women, starting at a four-year college implies an 8% increase in wages, but men see no impact. Similarly, some four-year credits have a positive impact for women, but negative for men. When attempted credits are included for all students, women see negative impacts of two-year credits only, whereas men see negative impacts for both. This suggests that there are negative impacts to spending time in college without earning a degree, but only for men. On the other hand, earning an undergraduate certificate has positive impacts across the board on hourly wages. The effects range from 12-15% increases for women and 10-11% for men. Associate degrees have much stronger impacts for women, ranging from 19-23% versus only 6-9% improvements for men. Finally, bachelor's degrees have the greatest impacts, with 27-32% increases in hourly wages for women and about 17% increases in hourly wages for men. These results are similar to those presented in Marcotte et. al (2005), although I see less evidence of returns to some college than has previously been found.

Table 9: OLS Estimates (Dependent Variable ln(income))

	Female			Male		
Started Two-Year	0.132			0.008		
	(0.136)			(0.071)		
Started Four-Year	0.345**			-0.030		
	(0.139)			(0.075)		
Attempted Credits Two-Year		-0.048***			-0.059***	
		(0.018)			(0.014)	
Attempted Credits Four-Year		-0.018			-0.068***	
		(0.014)			(0.011)	
Attempted Credits Two-Year (No Degree)			-0.047			-0.024
			(0.040)			(0.020)
Attempted Credits Four-Year (No Degree)			0.029			-0.046***
			(0.022)			(0.017)
Certificate	0.294**	0.142*	0.105	0.136*	0.158***	0.087
	(0.137)	(0.076)	(0.089)	(0.080)	(0.061)	(0.065)
AA	0.272*	0.157**	0.079	0.116	0.243***	0.060
	(0.141)	(0.077)	(0.094)	(0.088)	(0.062)	(0.071)
BA	0.522***	0.350***	0.316***	0.219***	0.404***	0.150***
	(0.134)	(0.054)	(0.079)	(0.072)	(0.050)	(0.055)
Married	0.099**	0.099**	0.099**	0.405***	0.394***	0.402***
	(0.045)	(0.046)	(0.045)	(0.035)	(0.034)	(0.035)
Children	-0.157***	-0.176***	-0.163***	-0.018	-0.041	-0.029
	(0.052)	(0.052)	(0.053)	(0.045)	(0.044)	(0.045)
Math Score (2004)	0.019***	0.022***	0.020***	0.008***	0.010***	0.009***
	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
Worked in 2006	0.292***	0.152**	0.198**	0.063	-0.005	0.029
	(0.093)	(0.077)	(0.080)	(0.063)	(0.058)	(0.060)
Job Training	0.204***	0.208***	0.204***	0.233***	0.232***	0.232***
	(0.044)	(0.044)	(0.044)	(0.036)	(0.035)	(0.036)
Asian	-0.051	-0.046	-0.047	-0.032	-0.001	-0.018
	(0.080)	(0.079)	(0.078)	(0.061)	(0.061)	(0.061)
African American	-0.117	-0.094	-0.103	-0.147**	-0.127*	-0.139**
	(0.083)	(0.086)	(0.085)	(0.069)	(0.070)	(0.069)
Hispanic	-0.059	-0.060	-0.054	-0.096*	-0.104**	-0.094*
	(0.066)	(0.066)	(0.066)	(0.055)	(0.052)	(0.053)
Other	-0.113	-0.099	-0.102	-0.090	-0.096	-0.093
	(0.088)	(0.089)	(0.089)	(0.078)	(0.079)	(0.078)
Middle Income	0.078	0.080	0.075	0.042	0.059	0.050
	(0.054)	(0.055)	(0.055)	(0.046)	(0.046)	(0.046)
High Income	0.089	0.098*	0.090	0.081	0.102**	0.092*
	(0.057)	(0.058)	(0.058)	(0.051)	(0.051)	(0.051)
Suburban	0.020	0.013	0.016	0.090**	0.092**	0.089**
	(0.049)	(0.050)	(0.050)	(0.043)	(0.043)	(0.043)
Rural	0.104*	0.098*	0.101*	0.049	0.043	0.045
	(0.058)	(0.058)	(0.058)	(0.051)	(0.052)	(0.052)
Constant	8.496***	8.631***	8.640***	9.477***	9.449***	9.473***
	(0.219)	(0.190)	(0.191)	(0.143)	(0.141)	(0.138)
N	Observations: 4060			Observations: 3940		

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Census dummies included but not reported.

In Table 9, I consider the effect on income. As compared to wages in Table 8, I see similar impacts, although stronger in magnitude for some college attendance. The second model, which includes credits for all respondents, suggests that increasing time to graduation has negative implications in the labor market. However, the third model suggests that the positive returns to postsecondary attendance on income are only seen for bachelor's degree earners. Further, men who attempt but do not complete a bachelor's degree are worse off than high school graduates. Each year of attempted credits reduces earnings by 5%.

In Appendix A, I present similar results that additionally include controls for major choice in Tables A1 and A2. Including major choice dampens the attainment effects, but does not affect significance. For women, certificates increase wages by 8-11%, associate degrees by 14-17%, and bachelor's degrees by 26-29%. For men, certificates increase wages by 9-10%, associate degrees by 7%, and bachelor's degrees by 13-14%. Business degrees have additional positive labor market returns for both men and women. Health majors are associated with wage improvements for women, and Computer Science, Engineering, and Math majors indicate positive wage improvements for men. Interestingly, men see declines in wages if they major in Education or the Humanities, whereas women see no significant impact on majoring in the Humanities or Education. With respect to income, as in Table 9, I see negligible impacts on earning a certificate or associate degree in most cases, although earning a bachelor's degree still seems to generate positive returns regardless of major. In this case, Engineering, Health, and Vocational/Technical majors of both genders are rewarded in the labor market. Women see some positive returns to Education majors, though men do not. Men additionally see rewards to majoring in Computer Science, but not Math.

4.2 Two Stage Least Squares

The ordinary least squares estimates present evidence that there are returns to degrees, but negligible or even negative returns to some college attendance. Of course, the results above should be interpreted with caution. Educational decisions are generally thought to be endogenous, dependent upon motivation or persistence, as well as ability and parental background, so it is important to account for non-random educational decisions. While the models include measures for ability and background, there remains the selection bias; students do not randomly sort into two or four-year schools and they do not randomly choose whether to complete a degree or not. Formally, the direction of the bias on γ_k depends on the covariance between the error term and the education decision, as

$$\hat{\gamma_k} = \gamma_k + \frac{cov(educ, u)}{var(educ)}$$

Thus, if selection bias positively impacts education decisions, OLS estimates will be overestimated. Further, it is possible that selection bias has different impacts at different levels of education decisions.

To control for the selection bias, I consider a two stage least squares approach. Specifically, I use the optimal instrument method proposed by Newey (1990), and used in Adams et. al (2009), in which the instruments used are the predicted probabilities corresponding to the dummy variables indicating the education choice. This is a three step process; first I estimate a multinomial logit separately for men and women, including the covariates in the main equation as well as the excluded variables: parental education and student's own expectations. I do this instead of an ordered probit because I do not want to impose an order to degree attainment. Specifically for categories some two-year, some-four-year, and certificate, it is not obvious that one option is necessarily better than the other. Students may earn some four-year college credits and make connections that lead to better job market prospects than earning a certificate. On the other hand, earning a certificate may provide job training that it is more helpful in the labor market than a year of credits at either a two or four-year institution.

After estimating the multinomial logit, I obtain predicted probabilities for each education choice, with high school graduation as the baseline case: $\hat{\gamma_{1M}}, ..., \hat{\gamma_{5M}}$ and $\hat{\gamma_{1F}}, ..., \hat{\gamma_{5F}}$. That is, for each respondent, I obtain the predicted probability they will earn some two-year college credits, some four-year college credits, a certificate, an associate degree, or a bachelor's degree.

I then regress γ_k, the true choice of the respondent, on each predicted probability, $\hat{\gamma_{1M}}, ..., \hat{\gamma_{5M}}$ or $\hat{\gamma_{1F}}, ..., \hat{\gamma_{5F}}$, and the covariates. Then, I obtain another set of predicted values, $\hat{\delta_{1M}}, ..., \hat{\delta_{5M}}$ and $\hat{\delta_{1F}}, ..., \hat{\delta_{5F}}$. Using the predicted values from the second stage in place of the dummy values, I estimate the final equation via ordinary least squares. This procedure avoids the "forbidden regression," and necessary standard error adjustment. I prefer this method to traditional two stage least squares in this case, as it allows for the educational decisions to be binary, it maintains consistency even if the first stage is misspecified, and the standard errors are correctly calculated. One caveat of the method I use is that it requires the use of the mutually exclusive choice set. That is, I first estimate the probability of choosing years of schooling equal to k, where k is: start at two-year (no degree), start at four-year (no degree), certificate, associate degree, or bachelor's degree.

The instruments used are parental education and students' expectations. These are highly correlated with eventual educational attainment. The maintained assumption is that these variables affect wages only through their impact on educational decisions.[13] [14] However, as this is an untestable assumption, the results below should

[13]The F-statistics, presented in Appendix B, are typically over 10, but some of the predictions are less conclusive than others. For example, predicting the likelihood of a male earning a certificate rather than another degree only has an F-statistic of 7.55.

[14]As alternative instruments, I used distance to the nearest two or four-year college. While the

be interpreted with caution.[15]

Table 10: Two Stage Least Squares Estimates

	Female		Male	
	$ln(Wage)$	$ln(Income)$	$ln(Wage)$	$ln(Income)$
Some Two-Year	-0.134	0.828*	-0.101	0.422
	(0.171)	(0.491)	(0.180)	(0.370)
Some Four-Year	-0.223	0.637	-0.117	0.038
	(0.196)	(0.653)	(0.173)	(0.340)
Certificate	0.269	1.101	0.398	-1.475*
	(0.274)	(0.951)	(0.378)	(0.819)
AA	0.567***	1.432***	-0.078	0.266
	(0.193)	(0.550)	(0.259)	(0.504)
BA	0.610***	1.492***	0.161	0.102
	(0.120)	(0.420)	(0.104)	(0.243)
Married	0.063***	0.085	0.167***	0.400***
	(0.020)	(0.053)	(0.021)	(0.038)
Children	-0.005	-0.087	-0.009	-0.032
	(0.021)	(0.078)	(0.027)	(0.056)
Math Score (2004)	0.002	0.012***	0.005***	0.006*
	(0.002)	(0.004)	(0.002)	(0.003)
Worked in 2006	0.072	0.619***	-0.002	0.034
	(0.046)	(0.154)	(0.046)	(0.103)
Job Training	0.098***	0.144**	0.154***	0.260***
	(0.020)	(0.065)	(0.024)	(0.048)
Asian	0.046	-0.085	-0.019	0.031
	(0.033)	(0.102)	(0.034)	(0.070)
African American	0.047*	-0.135	-0.072**	-0.144**
	(0.026)	(0.106)	(0.029)	(0.072)
Hispanic	0.060*	-0.067	-0.046	-0.153**
	(0.032)	(0.082)	(0.033)	(0.064)
Other	0.000	-0.101	-0.005	-0.136*
	(0.034)	(0.102)	(0.043)	(0.081)
Middle Income	0.020	0.030	0.030	0.032
	(0.020)	(0.058)	(0.021)	(0.048)
High Income	0.008	0.017	0.059**	0.081
	(0.028)	(0.072)	(0.025)	(0.056)
Suburban	0.000	0.017	0.032	0.109**
	(0.019)	(0.059)	(0.024)	(0.051)
Rural	0.009	0.099	0.012	0.085
	(0.026)	(0.070)	(0.029)	(0.060)
Constant	2.375***	8.174***	2.417***	9.656***
	(0.096)	(0.334)	(0.111)	(0.237)
N	4290	4060	3930	3940

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
IV: Parental education, student education expectations.
Census dummies included but not reported.

exclusion restriction assumption might be easier to argue, these instruments are not strong enough to warrant such an approach. I also tried using the number of vocational credits taken, the number of significant life events the respondent had experienced, and parental occupations. None of these alternatives have outperformed the current model.

[15] I cannot currently perform any over-identification tests as my model is just identified.

In Table 10, I consider the returns to education for both females and males. Females see significant returns to an associate degree; these estimates suggest such that they are extremely close to the returns to a bachelor's degree, relative to women who earn high school credentials only. This is true for both wages and income. There is some evidence that women see positive income effects to some two-year college attendance as well, relative to high school graduates. Men, on the other hand, only see negative impacts from earning a certificate on income, relative to men who earn high school credentials only. These results suggest that, once controlling for selection into college, men see little to no additional returns to their education in the labor market, and females see negligible effects for anything less than an associate degree. However, as the standard errors have grown significantly with the use of instrumental variables, this is also indicative of weak instruments.

4.3 Matching Methods

I additionally employ matching methods to identify the effect of education on income. In randomized control trials or natural experiments, assignment to control and treatment groups are random. In these cases, researchers are able to compare outcomes without controlling for additional covariates. However, in this scenario, random assignment is impossible, and so more rigorous forms of matching are required. Controlling for observable factors helps to remove preexisting differences between treatment and control groups. The benefits of using matching methods in this scenario are clear; identifying the causal impact of education on earnings is difficult without an exogenous variable that impacts educational decisions without affecting income. While the ordinary least squares estimates are useful for comparison, they are likely to be biased as variables that impact educational decisions also impact earnings.

I define the treatment, T_i, as the respondent's decision among no college, some two-year college, some four-year college, certificate, associate degree, or bachelor's degree. However, one of the main assumptions of matching is that the covariates are balanced between those that are treated and those that are not. Thus, as the entire sample contains a very heterogeneous group of respondents, I break the models down into binary decisions or treatments. That is, I consider several separate treatment models in which I utilize propensity score matching methods. Specifically, I consider the binary choices of a) no college/some two-year, b) no college/some four-year, c) some two-year/certificate or associate degree, d) some two-year/bachelor's degree, e) some four-year/certificate or associate degree, f) some four-year/bachelor's degree, and g) certificate or associate degree/bachelor's degree. This is a particular benefit to studying this problem via matching, as the various subgroups vary quite substantially. For example, respondents who earn bachelor's degrees differ significantly from those who earn only high school credentials.

As I am using observational data rather than a natural experiment, the relevant statistic is the average treatment effect on the treated (ATET). This estimate shows the effect of "treatment," which in my case is either attending college or earning a certificate or degree, on those who are treated.[16] In this case, Y_{1i} is the natural log of wages or income for respondent i if they have chosen the treatment (some college, certificate or associate degree, or bachelor's degree). Then, Y_{0i} represents the natural log of wages or income for those respondents who did not take the treatment (no college, some college, or certificate/associate degree). The average treatment effect on the treated is:

$$E[Y_{1i}|X, T = 1] - E[Y_{0i}|X, T = 1] \tag{3}$$

Since I only observe a respondents ultimate decision and not their counterfactual, I must employ a matching technique that requires additional assumptions

[16]This is not to say that the average treatment effects of of no interest; the ATE estimates the effect of educational decisions if everyone takes up the treatment.

(Rosenbaum and Rubin (1983), Wooldridge (2002))). Matching on several variables can be quite difficult, especially with one or more continuous covariates. This is the main advantage of propensity score matching, as this method reduces the match to one dimension, the propensity score. Using either a logit or probit model to predict the probability of treatment, each observation is matched to the closest observation in the alternate group by their propensity score. One drawback to matching methods is that there are many decisions the researcher must make. In this case, it begins with defining what "close" entails. In matching language, the caliper sets the maximum distance propensity scores may be from each other to constitute a match. The standard is to allow the caliper to be 0.1, although adjustments in the caliper can be made to test this baseline.[17]

To ensure that conditional on the covariates, treatment and control group assignment is "random," two assumptions are tested. First, the covariate means must be similar between the treatment and control groups after matching has occurred. This is often referred to as conditional mean independence, unconfoundedness, or balancing. In practice, only conditional mean independence is required in estimations of the ATET. That is, $E[Y_{0i}|X,T] = E[Y_{0i}|X]$ and $E[Y_{1i}|X,T] = E[Y_{1i}|X]$. Further, it is possible conditional independence is not necessary for estimating adjusted means (Frolich (2007)). Conditional upon a similar propensity score, respondents should have the same distribution of both observable and unobservable characteristics, whether or not they receive the treatment. In other words, something that impacts treatment does not impact potential outcomes, and something that impacts potential outcomes does not affect treatment status. To check the balancing as-

[17] Restricting the caliper to 0.01 removes some of the sample as these observations do not have good enough matches, but the results are similar and are presented in Appendix D. The one change is that the positive effect on income due to earning a certificate or associate degree instead of attending some two-year college loses significance, though the sign remains the same. Other estimates retain sign and significance, if not exact magnitudes as well. The number of dropped observations by subgroup are: none/some two: 34 (F), 2 (M); none/some four: 96 (F), 23 (M); some two/certificate/AA: 13 (F), 1 (M); some two/BA: 54 (F), 12 (M); some four/certificate/AA: 7 (F), 8 (M), some four/BA: 11 (F), 2 (M); certificate/AA/BA: 16 (F), 1 (M).

sumption, I use "pscore" in STATA. Additionally, tables with the conditional mean differences are available in Appendix C.[18] The user-written "pscore" tests for balance covariate by covariate, block by block. The number of blocks varies by case, and ranges from four to ten in this application. As matching on covariates is difficult enough, the included variables differ by subgroup to ensure the balancing property is satisfied in each case. That is, Table 11 uses the baseline model for matching with all subgroups, whereas Table 12 uses the variation in covariate inclusion. As shown below, the estimates remain stable, so the inclusion of, for example, math score or high school GPA does not seem to impact the results.[19]

Second, there must be sufficient overlap in the treatment and control populations such that each observation x_i can be matched with an observation in the alternative group. Given $\tilde{t}$, $Pr(T = \tilde{t}|x_i)$ must be strictly greater than zero and less than one. While this is easy to show with graphs of the propensity scores, I can also restrict the analysis to a region of common support. Evidence of the overlap assumption is in Appendix C in Figures C1 through C14. For example, there is strong evidence of overlap in Figures C5 and C6 in which respondents with some two-year experience are compared to those with a certificate or associate degree. Similarly Figures C7, C8, C9, and C10 compare respondents with some four-year experience to respondents with a bachelor's degree or certificate/associate degree respectively, and show strong overlap. The evidence of overlap is still fairly strong between students with no college and some two-year experience (Figures C1, C2) and those with a certificate/associate degree and a bachelor's degree (C13, C14). On the other hand, Figures C3 and C4 provide an example in which the overlap assumption may not hold. For both males and females, the propensity scores for some four-year experience are skewed towards zero and the propensity scores for no college experience are skewed slightly toward

[18]Different variations of the baseline model were tested based on the subgroup.

[19]While some have found that matching estimates are sensitive to included variables, I believe the variables chosen represent the data well. Alternative models suggest similar results and are presented in Appendix D (Agondini and Dynarski (2004) and Smith and Todd (2005).

one. A similar issue is seen in Figures C11 and C12, in which the propensity scores for earning a bachelor's degree are skewed toward zero. Here, the propensity scores for earning some two-year credits are slightly more evenly distributed, but there is no skew toward zero.

The results of the baseline model are presented below, in which the caliper was set to 0.1 and each observation was matched with one other observation.[20] Included covariates are: marital and parental status, whether the respondent worked in 2006 and had on the job training, math score (2004), race, urbanicity, and census region.

Table 11: Propensity Score Matching (ATET)

	Female		Male	
	$ln(Wage)$	$ln(Income)$	$ln(Wage)$	$ln(Income)$
Some Two-Year vs. No College	0.106***	0.041	-0.005	0.054
	(0.036)	(0.099)	(0.033)	(0.113)
N	1030	900	1240	1220
Some Four-Year vs. No College	0.099**	-0.002	-0.015	-0.180**
	(0.047)	(0.119)	(0.034)	(0.092)
N	900	810	1120	1100
Certificate/AA vs. Some Two-Year	0.125***	0.154**	0.126***	0.172***
	(0.022)	(0.076)	(0.033)	(0.066)
N	1610	1480	1330	1350
BA vs. Some Two-Year	0.259***	0.492***	0.255***	0.383***
	(0.050)	(0.138)	(0.057)	(0.079)
N	2490	2410	2200	2220
Certificate/AA vs. Some Four-Year	0.091***	-0.017	0.066*	0.147**
	(0.031)	(0.072)	(0.036)	(0.074)
N	1470	1390	1220	1230
BA vs. Some Four-Year	0.212***	0.285***	0.202***	0.288***
	(0.030)	(0.068)	(0.029)	(0.059)
N	2360	2320	2080	2110
BA vs. Certificate/AA	0.132***	0.382***	0.118***	0.204***
	(0.031)	(0.081)	(0.036)	(0.069)
N	2670	2620	2070	2100

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Variables Included: Married, Child, Work 2006, Job Training, Race,
Math Score, Urbanicity, Region

In Table 11, I consider the average treatment effect on the treated, or $E[Y_{1i} - Y_{0i}|T = 1]$. In Appendix A3, I estimate linear probability models for the same binary educational choices as a comparison. Relative to women who do not attend college,

[20] I could increase the number of matches used, but choosing a small number, namely one match, has the advantage of reducing finite sample bias that might occur with poor matches (Abadie and Imbens 2012).

women who attend two-year colleges without earning a degree see positive impacts on wages on the order of 11%, but not with respect to income. Neither estimate is significant with the linear probability model. Similarly, women see 10% gains in wages to four-year college attendance relative to women who do not attend college. This estimate is equivalent in the linear probability model, but the returns to income are 37%. This may be due to the number of observations which go unmatched in propensity score matching driving the OLS results. Wages and income for men who attend a two-year college are not significantly higher than men who did not go to college. Further, men see negative returns to income when starting at a four-year college and not earning a degree, relative to men who do not attend college. In the linear probability model, I see a 2% increase in male income between those who do not go to college and those who earn some two-year credits. Similar to the results for wages, I believe these differences demonstrate the importance of matching to create a comparable sample.

There are positive returns to earning a certificate or associate degree for both men and women, relative to those earning two-year college credits, of about 13% for wages and between 17-19% for income. Those who start at a two-year college and earn a bachelor's degree see very strong impacts; 30% increase in wages and between 47-64% increase in income. The linear probability model's results are similar, although the returns to male wages are smaller. Due in part to the lower probability of attaining a bachelor's degree when respondents start at a two-year college, I expect high returns to those that do earn a baccalaureate degree.

Comparing students who start at four-year colleges to those that earn a certificate or associate degree shows moderate returns to wages and some evidence of positive effects on income for men. These results persist in the linear probability model. This suggests that it may not be time in college, but what you do with that time that matters. Earning a degree has a higher weight than earning some college credits without the degree. Students who start at a four-year college and earn a bachelor's degree also see positive returns, although they are not quite as high as for those students who start at a two-year college. Again, these results are similar in the linear probability model. These smaller impacts are expected; students who start at a four-year college are much more likely to finish than those who start at a two-year college, as seen in Chapter 3.

Finally, comparing students who earn sub-baccalaureate degrees to those earning bachelor's degrees suggests that there is still a wage benefit to higher education; 13-14% increase in wages and 23-47% increase in income. While the magnitudes differ, the linear probability model produces results that are consistent with the propensity score estimates with the exception of male income, in which the linear probability model shows no effect. These results, while consistent with my results above, are in stark contrast to results in previous work which suggest positive impacts of any college experience, whether or not it results in a degree.

Table 12A: Propensity Score Matching (ATET) Alternative Methods

	Female		Male	
	$ln(Wage)$	$ln(Income)$	$ln(Wage)$	$ln(Income)$
Some Two-Year vs. No College[a]				
Common Support	[.117, .984]		[.066, .990]	
Nearest Neighbor (Random Draw)	0.038	0.115	-0.06	-0.158
	(0.048)	(0.185)	(0.043)	(0.150)
Stratification	0.038	-0.031	-0.031	-0.088
	(0.026)	(0.093)	(0.031)	(0.079)
Kernel Matching	0.063**	0.043	-0.031	-0.074
	(0.027)	(0.116)	(0.031)	(0.085)
psmatch	0.064**	0.184	-0.018	-0.141
	(0.032)	(0.134)	(0.037)	(0.090)
Observations	1080		1260	
Some Four-Year vs. No College[b]				
Common Support	[.007, .999]		[.009, .998]	
Nearest Neighbor (Random Draw)	0.154*	0.004	-0.012	-0.135
	(0.090)	(0.336)	(0.059)	(0.131)
Stratification	0.117**	0.057	-0.045	-0.162**
	(0.047)	(0.084)	(0.042)	(0.074)
Kernel Matching	0.137***	0.188	-0.034	-0.149*
	(0.046)	(0.126)	(0.045)	(0.080)
psmatch	0.123***	0.074	0.002	-0.047
	(0.043)	(0.107)	(0.038)	(0.090)
Observations	980		1200	
Certificate/AA vs. Some Two-Year[c]				
Common Support	[.200, .793]		[.243, .602]	
Nearest Neighbor (Random Draw)	0.136***	-0.02	0.102***	0.093
	(0.027)	(0.080)	(0.032)	(0.069)
Stratification	0.144***	0.112**	0.103***	0.123***
	(0.019)	(0.057)	(0.023)	(0.047)
Kernel Matching	0.153***	0.147**	0.114***	0.141***
	(0.020)	(0.064)	(0.025)	(0.048)
psmatch	0.121***	0.153**	0.107***	0.119*
	(0.026)	(0.076)	(0.032)	(0.067)
Observations	1730		1470	
BA vs. Some Two-Year[d]				
Common Support	[.008, .999]		[.004, .999]	
Nearest Neighbor (Random Draw)	0.215***	0.337*	0.179***	0.28**
	(0.052)	(0.190)	(0.056)	(0.116)
Stratification	0.211***	0.332***	0.229***	0.289***
	(0.038)	(0.076)	(0.037)	(0.059)
Kernel Matching	0.228***	0.372***	0.202***	0.303***
	(0.031)	(0.076)	(0.033)	(0.070)
psmatch	0.184***	0.274**	0.222***	0.298***
	(0.052)	(0.124)	(0.064)	(0.078)
Observations	2480		2250	

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Variables Included in All: Married, Child, Work 2006, Job Training, Black, Hispanic, Other (d separates Asian and Other) Urbanicity, Region
a: Additional Variables: Math, HS GPA, Middle/High Income, Unemployment Rate, Per Capita Income
b: Additional Variables: Math, Income
c: Additional Variables: Math, Middle/High Income
d: Additional Variables: HS GPA, Middle/High Income, Unemployment Rate, Per Capita Income

Table 12B: Propensity Score Matching (ATET) Alternative Methods

	Female		Male	
	$ln(Wage)$	$ln(Income)$	$ln(Wage)$	$ln(Income)$
Certificate/AA vs. Some Four-Year[e]				
Common Support	[.147, .927]		[.105, .938]	
Nearest Neighbor (Random Draw)	0.092***	-0.072	0.095**	0.091
	(0.030)	(0.085)	(0.039)	(0.079)
Stratification	0.092***	-0.022	0.09***	0.115**
	(0.022)	(0.067)	(0.028)	(0.055)
Kernel Matching	0.091***	-0.029	0.092***	0.118**
	(0.022)	(0.064)	(0.028)	(0.055)
psmatch	0.110***	-0.021	0.068**	0.123
	(0.026)	(0.085)	(0.032)	(0.078)
Observations	1570		1340	
BA vs. Some Four-Year[f]				
Common Support	[.164, .960]		[.102, .974]	
Nearest Neighbor (Random Draw)	0.143***	0.176**	0.211***	0.269***
	(0.031)	(0.074)	(0.038)	(0.072)
Stratification	0.186	0.207	0.191***	0.261***
	-	-	(0.026)	(0.048)
Kernel Matching	0.198***	0.23***	0.225***	0.313***
	(0.024)	(0.052)	(0.029)	(0.050)
psmatch	0.147***	0.179***	0.208***	0.295***
	(0.030)	(0.066)	(0.035)	(0.057)
Observations	2370		2130	
BA vs. Certificate/AA[g]				
Common Support	[.055, .985]		[.039, .995]	
Nearest Neighbor (Random Draw)	0.098***	0.344***	0.082*	0.186**
	(0.035)	(0.09)	(0.049)	(0.091)
Stratification	0.095***	0.295***	0.101**	0.145**
	(0.027)	(0.068)	(0.048)	(0.066)
Kernel Matching	0.104***	0.315***	0.093**	0.164**
	(0.025)	(0.061)	(0.037)	(0.066)
psmatch	0.091***	0.259***	0.133**	0.174**
	(0.034)	(0.060)	(0.060)	(0.088)
Observations	2620		2110	

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Variables Included in All: Married, Child, Work 2006, Job Training, Black, Hispanic, Other (d separates Asian and Other) Urbanicity, Region
e: Additional Variables: Math, Middle/High Income
f: Additional Variables: Math, HS GPA, Middle/High Income, Unemployment Rate
g: Additional Variables: HS GPA, Race by HS GPA, Income, Unemployment Rate

I additionally present results using "pscore," which allows for propensity score estimation with different measures for how observations are matched. While appropriate for comparison, the results presented above, in which the standard errors account for the fact that the propensity score is estimated, are preferred. That said, I present estimates above in which I estimate the average treatment effect on the treated for the various groups presented above, matching by nearest neighbor, stratification, kernel, and "psmatch." Nearest neighbors are determined by the propensity score, and one neighbor is selected for each treated observation. In these models, the caliper is again set to 0.1. All ties are broken by random draw.

Rather than matching by the nearest neighbor, stratification matching restricts matches to others within the same bracket. The brackets for stratification matching are determined by the balancing property that is done prior to estimation. Optimal brackets are chosen based on the matching of the treatment and control groups. The number of brackets varies by subgroup and model.

The kernel matching technique matches each treatment observation to all control observations, with less weight attributed to matches that are further away. In this matching method, I bootstrap the standard errors.[21]

I finally estimate the "psmatch" model as in Table 11, which matches using the nearest neighbor method, but accounts for the fact that the propensity score is estimated. The other difference with the last estimate is that, while the researcher can restrict common support, the program enforces equivalent cut off points at the bottom and top threshold. For example, one can consider a common support of 0.1 to 0.9, but not 0.05 to 0.9. Currently, I do not restrict the support and include all possible observations when using this method. Nevertheless, I expect the first and last estimates to be similar due to their similarity in estimation.

The baseline variables that are included in all cases are: marital and parental status, whether the respondent worked in 2006 and had on the job training, race, urbanicity, and region (New England, South, Mid West, or West). Variations in other variables are also included in some but not all models, such as math score or high school GPA, unemployment rates, and Asian/Other students being separate or

[21] Currently, the bandwidth is set at 0.06 and the replications at 500.

combined.

In Table 12A, I again see a small, but significant return to wages for women earning some two-year college credits, and stronger returns to wages for some four-year college credits, relative to high school graduates. Men see negligible effects except for negative returns to income for earning some four-year credits relative to those who complete high school alone.

Compared to earning some two-year credits, women who earn a certificate or associate degree earn between 13-17% more in wages, and between 12-17% more in income. Women who instead earn a bachelor's degree see increases in wages and income of 20-26% and 32-45% respectively. Men who earn a certificate or associate degree rather than some two-year credits see increases to wages and income of between 11-12% and 13-15% respectively. Men who instead earn a bachelor's degree see wage increases of 20-26% and income increases of 32-35%.

In Table 12B, I consider the effect of earning some four-year college credits as compared to earning a certificate/associate degree or a bachelor's degree and finally, the effect of earning a certificate/associate degree versus a bachelor's degree. Women who earn a certificate or associate degree rather than those who earn some four-year college credits see wage improvements of 10-12% but do not see any significant improvements in income. Men see a 7-10% increase in wages from earning a certificate or associate degree, and some evidence of a 12-13% increase in income. For those who earn a bachelor's degree, the returns to wages for women are between 15-22% and the returns to income are between 19-26%. For males, the returns to wages are between 21-25% and the returns to income are between 30-37%. Interestingly, these alternative estimates produce slightly larger fluctuations in estimates of returns for some four-year credits compared to a bachelor's degree.

Finally, comparing respondents who earned a certificate or associate degree to those who earned a bachelor's degree again shows positive and significant returns.

Women who earn a bachelor's degree see returns of 10-11% in wages and 30-41% in income. Men see returns of 9-14% in wages and returns of between 16-20% in income.

While there are many choices to make with respect to matching methods, it is convincing to see so many methods produce similar results. While the magnitudes vary from model to model, the signs remain the same, and the impact increases with attainment. Further, the absence of an effect for many of the subgroups in which the treatment was some college versus no college is striking.

5 Conclusion

Returns to higher education are an important research topic for a number of reasons. Access to higher education is one way to promote equality; with more choices, students are better able to identify their optimal path after high school. While four-year colleges seem to get the most press as far as the "best" choice, there are many great career paths that only require an undergraduate certificate or associate degree and may be more optimal for some students. My results, taken together, provide some evidence that there are positive effects to some two or four-year college attendance for women, but not men. This lines up with some of the previous research, although most suggested positive impacts for even a year of postsecondary education, which is not what I see here. There is even some evidence that men might be harmed by earning credits without a degree. Additionally, an important contribution of this paper is the relative comparisons; students who complete bachelor's degrees are often quite different from those who complete high school, but high school graduates and those with some college attendance are much more similar. Estimating the impact for these smaller subgroups has the potential to inform policy in a more nuanced way. Some students might be better off with a certificate or associate degree rather than some four-year attendance.

Earning a certificate or associate degree has positive impacts for both genders, with much stronger associate degree impacts for women, but similar degree effects for men. Propensity score estimates suggest that both men and women see earnings benefits of between 13-19% if they earn a certificate or associate degree rather than some two-year college. The estimates are a bit smaller between the subgroup of respondents who earn some four-year college credits versus a certificate or associate degree, but generally still positive. As many of the students who attempted some four-year college but did not earn a degree come from high schools which seem to be "high" performing, it is possible that students are trying to do what they think is expected of them, instead of following what they might be better prepared for. From the propensity score estimates, there seems to be a small but positive effect to both men and women to earning a certificate or associate degree as opposed to some four-year college. This suggests that some students might be better off earning a sub-baccalaureate degree rather than trying to earn a bachelor's degree. That certificates have a positive impact is a promising result; most of the previous literature was unable to determine precise estimates. The returns to sub-baccalaureate degrees are similar, albeit a bit smaller in magnitude, to previous literature.

Bachelor's degrees maintain their large impacts for both men and women compared to students who earn some two-year or some four-year college credits. The increases are on the order of between 22% and 64%, although they are larger between two-year college attenders and baccalaureate attainers. I expect this to be due to the fact that there are more barriers for students who start at two-year colleges, and thus those who do make it through see larger improvements to their income. There are smaller, but still significant returns to earning a bachelor's degree over a certificate or associate degree. While not directly comparable to previous research, as the baseline case is not high school graduate, these results are similar to previous research which generally suggests a larger labor market impact with greater postsecondary credentials. Of the studies that consider propensity score matching methods, the comparison was no college to a bachelor's degree, which is not what I consider here. Further, Brand and Xie (2010) use data on respondents further along in their careers and Blundell et al. (2004) consider data on education decisions in Britain, both of which may have different implications as compared to American students early in their careers.

Combined, the results of this paper suggest a decline in the return to a sub-baccalaureate degree relative to a bachelor's degree. With recent attention to the necessity of a four-year degree, this is not entirely surprising. However, as there is only moderate evidence of positive returns to any college attendance, a departure from previous research, the push for all students to enter a four-year college may be misguided. There is evidence that students who earn a certificate or associate degree have higher earnings than those who earn only some college credits, at either a two or four-year college. Early in their career, it may be more beneficial for these students to be directed towards certificate or associate programs, rather than a four-year degree. Community colleges offer higher education opportunities to many who might not otherwise pursue it. However, if postsecondary attendance alone is not enough to increase wages, policies must be implemented to increase the incentive to complete a degree.

Appendices

Appendix A: OLS Results for Alternative Models

Table A1: OLS Estimates (Dependent Variable ln(wage))

	Female			Male		
Started Two-Year	0.009			-0.015		
	(0.030)			(0.029)		
Started Four-Year	0.054*			-0.034		
	(0.031)			(0.035)		
Attempted Credits Two-Year		-0.018***			-0.026***	
		(0.006)			(0.008)	
Attempted Credits Four-Year		-0.003			-0.028***	
		(0.005)			(0.006)	
Attempted Credits Two-Year (No Degree)			-0.013			-0.014
			(0.009)			(0.009)
Attempted Credits Four-Year (No Degree)			0.008			-0.017*
			(0.007)			(0.009)
Certificate	0.106***	0.095***	0.081***	0.092**	0.115***	0.084**
	(0.029)	(0.020)	(0.022)	(0.042)	(0.037)	(0.037)
AA	0.153***	0.155***	0.127***	0.050	0.117***	0.038
	(0.032)	(0.023)	(0.026)	(0.040)	(0.031)	(0.033)
BA	0.257***	0.232***	0.230***	0.152***	0.234***	0.134***
	(0.030)	(0.022)	(0.026)	(0.035)	(0.025)	(0.029)
Business	0.138***	0.148***	0.141***	0.110***	0.127***	0.116***
	(0.027)	(0.027)	(0.028)	(0.034)	(0.036)	(0.035)
Social Sciences	0.005	0.012	0.007	0.035	0.045	0.041
	(0.030)	(0.030)	(0.030)	(0.041)	(0.039)	(0.041)
Engineering	0.105	0.112*	0.099	0.210***	0.233***	0.220***
	(0.066)	(0.067)	(0.066)	(0.041)	(0.041)	(0.041)
Computer Science	0.087	0.091	0.092	0.193***	0.204***	0.199***
	(0.129)	(0.127)	(0.129)	(0.068)	(0.068)	(0.069)
Humanities	-0.029	-0.019	-0.027	-0.115***	-0.090**	-0.107***
	(0.028)	(0.028)	(0.028)	(0.036)	(0.036)	(0.037)
Education	-0.010	0.002	-0.009	-0.136***	-0.099**	-0.121**
	(0.029)	(0.029)	(0.029)	(0.050)	(0.050)	(0.050)
Health	0.199***	0.213***	0.202***	-0.057	-0.030	-0.047
	(0.021)	(0.022)	(0.022)	(0.039)	(0.040)	(0.040)
Life and Physical Sciences	-0.038	-0.030	-0.038	-0.013	0.008	-0.004
	(0.035)	(0.035)	(0.035)	(0.038)	(0.037)	(0.038)
Math	-0.059	-0.051	-0.057	0.273***	0.285***	0.279***
	(0.107)	(0.107)	(0.107)	(0.090)	(0.090)	(0.089)
Vocational/Technical	0.016	0.024	0.019	0.094**	0.103***	0.100**
	(0.052)	(0.052)	(0.053)	(0.040)	(0.039)	(0.039)
Constant	2.179***	2.202***	2.198***	2.479***	2.453***	2.469***
	(0.061)	(0.058)	(0.059)	(0.069)	(0.070)	(0.070)
N	Observations: 4290			Observations: 3930		

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Census dummies included but not reported.

Additional variables: married, children, worked in 2006, job training, math score, income, urbanicity.

Table A2: OLS Estimates (Dependent Variable ln(income))

	Female			Male		
Started Two-Year	0.092			-0.010		
	(0.137)			(0.071)		
Started Four-Year	0.276**			-0.054		
	(0.139)			(0.076)		
Attempted Credits Two-Year		-0.060***			-0.066***	
		(0.018)			(0.014)	
Attempted Credits Four-Year		-0.029**			-0.075***	
		(0.014)			(0.012)	
Attempted Credits Two-Year (No Degree)			-0.064			-0.036*
			(0.039)			(0.020)
Attempted Credits Four-Year (No Degree)			0.011			-0.056***
			(0.022)			(0.017)
Certificate	0.220	0.115	0.040	0.083	0.119**	0.032
	(0.137)	(0.077)	(0.088)	(0.080)	(0.060)	(0.065)
AA	0.163	0.102	-0.027	0.029	0.172***	-0.035
	(0.142)	(0.079)	(0.094)	(0.089)	(0.062)	(0.073)
BA	0.439***	0.322***	0.231***	0.140*	0.338***	0.055
	(0.135)	(0.052)	(0.080)	(0.075)	(0.051)	(0.059)
Business	0.310***	0.362***	0.339***	0.314***	0.361***	0.337***
	(0.070)	(0.070)	(0.071)	(0.053)	(0.052)	(0.053)
Social Sciences	0.078	0.122*	0.103	-0.067	-0.027	-0.039
	(0.065)	(0.065)	(0.066)	(0.106)	(0.106)	(0.107)
Engineering	0.273**	0.328**	0.281**	0.319***	0.381***	0.348***
	(0.138)	(0.139)	(0.142)	(0.067)	(0.064)	(0.066)
Computer Science	0.169	0.201	0.198	0.332***	0.368***	0.354***
	(0.169)	(0.166)	(0.168)	(0.113)	(0.109)	(0.112)
Humanities	-0.002	0.050	0.019	-0.049	0.020	-0.022
	(0.074)	(0.074)	(0.075)	(0.058)	(0.057)	(0.058)
Education	0.119*	0.185***	0.146**	-0.004	0.100	0.044
	(0.068)	(0.070)	(0.070)	(0.091)	(0.093)	(0.092)
Health	0.325***	0.378***	0.345***	0.134*	0.203***	0.160**
	(0.076)	(0.073)	(0.072)	(0.078)	(0.077)	(0.078)
Life and Physical Sciences	-0.023	0.035	-0.010	-0.038	0.025	-0.007
	(0.088)	(0.090)	(0.087)	(0.076)	(0.075)	(0.075)
Math	-0.358	-0.315	-0.345	0.073	0.119	0.094
	(0.379)	(0.387)	(0.380)	(0.162)	(0.152)	(0.157)
Vocational/Technical	0.332***	0.368***	0.354***	0.355***	0.380***	0.375***
	(0.116)	(0.116)	(0.118)	(0.062)	(0.060)	(0.061)
Constant	8.458***	8.555***	8.577***	9.483***	9.444***	9.472***
	(0.221)	(0.193)	(0.194)	(0.144)	(0.142)	(0.140)
N	Observations: 4060			Observations: 3940		

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Census dummies included but not reported.
Additional variables: married, children, worked in 2006, job training, math score, income, urbanicity.

Table A3: OLS Estimates: Binary Treatment

Dependent Variable ln(Wage)						
	Female			Male		
	Baseline	With Majors	N	Baseline	With Majors	N
Some Two-Year vs. No College	0.039	0.034	1030	-0.002	-0.008	1240
Some Four-Year vs. No College	0.092***	0.027	900	-0.021	-0.032	1120
Certificate/AA vs. Some Two-Year	0.147***	0.124***	1610	0.088***	0.079***	1330
BA vs. Some Two-Year	0.252***	0.244***	2490	0.162***	0.141***	2200
Certificate/AA vs. Some Four-Year	0.099***	0.080***	1470	0.087***	0.079***	1220
BA vs. Some Four-Year	0.191***	0.183***	2360	0.197***	0.176***	2080
BA vs. Certificate/AA	0.097***	0.126***	2670	0.065**	0.074**	2070
Dependent Variable ln(Income)						
	Female			Male		
	Baseline	With Majors	N	Baseline	With Majors	N
Some Two-Year vs. No College	0.098	0.075	900	0.020	-0.003	1220
Some Four-Year vs. No College	0.313*	0.214	810	-0.003	-0.064	1100
Certificate/AA vs. Some Two-Year	0.149**	0.107	1480	0.114**	0.076	1350
BA vs. Some Two-Year	0.388***	0.341***	2410	0.173***	0.120*	2220
Certificate/AA vs. Some Four-Year	-0.079	-0.133	1390	0.110*	0.078	1230
BA vs. Some Four-Year	0.182***	0.158***	2320	0.253***	0.186***	2110
BA vs. Certificate/AA	0.261***	0.283***	2620	0.091	0.104*	2100

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Variables: married, children, worked in 2006, job training, math score, income, urbanicity, census.

Appendix B: First Stage Results (IV)

Table B1: First Stage: Female

	Attainment	BA	Cert/AA	AA	Cert	Some	Some (2)	Some (4)
$\hat{p_2}$	1.273**	0.147	-0.069	-0.040	-0.028	0.803***	0.822***	-0.020
	(0.575)	(0.113)	(0.148)	(0.111)	(0.115)	(0.171)	(0.168)	(0.095)
$\hat{p_3}$	2.008***	0.018	-0.046	0.075	-0.121	0.970***	-0.041	1.011***
	(0.614)	(0.157)	(0.177)	(0.124)	(0.146)	(0.189)	(0.170)	(0.165)
$\hat{p_4}$	2.170***	-0.108	0.850***	0.012	0.839***	0.194	0.243	-0.049
	(0.743)	(0.138)	(0.223)	(0.148)	(0.184)	(0.236)	(0.221)	(0.139)
$\hat{p_5}$	4.185***	0.033	1.091***	1.084***	0.007	-0.108	0.122	-0.230
	(0.919)	(0.218)	(0.263)	(0.213)	(0.205)	(0.279)	(0.240)	(0.180)
$\hat{p_6}$	4.898***	1.027***	-0.058	-0.038	-0.020	0.008	0.040	-0.032
	(0.441)	(0.088)	(0.131)	(0.101)	(0.103)	(0.132)	(0.118)	(0.083)
Married	0.017	0.007	-0.004	-0.003	-0.000	-0.008	-0.015	0.006
	(0.071)	(0.019)	(0.018)	(0.014)	(0.014)	(0.022)	(0.019)	(0.016)
Children	-0.007	-0.003	0.002	-0.005	0.007	0.011	0.011	-0.001
	(0.074)	(0.016)	(0.025)	(0.019)	(0.020)	(0.026)	(0.024)	(0.018)
Math	0.000	0.000	0.000	0.000	-0.000	-0.001	-0.001	-0.000
	(0.005)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
Job Training	0.018	0.005	-0.001	0.002	-0.004	-0.009	-0.015	0.006
	(0.076)	(0.019)	(0.019)	(0.016)	(0.016)	(0.022)	(0.018)	(0.015)
Worked in 2006	-0.038	0.006	-0.016	0.005	-0.021	-0.003	0.019	-0.022
	(0.176)	(0.026)	(0.045)	(0.036)	(0.034)	(0.054)	(0.049)	(0.030)
Asian	0.058	0.011	0.007	0.000	0.006	-0.019	-0.021	0.001
	(0.105)	(0.033)	(0.031)	(0.020)	(0.025)	(0.032)	(0.028)	(0.022)
African American	0.029	0.004	0.006	-0.004	0.011	-0.007	-0.007	-0.001
	(0.090)	(0.022)	(0.028)	(0.021)	(0.023)	(0.030)	(0.026)	(0.024)
Hispanic	-0.050	-0.015	-0.000	0.004	-0.004	0.023	0.023	0.000
	(0.106)	(0.026)	(0.027)	(0.021)	(0.020)	(0.034)	(0.033)	(0.019)
Other	-0.013	0.001	-0.005	-0.009	0.004	0.002	-0.001	0.003
	(0.117)	(0.031)	(0.036)	(0.028)	(0.026)	(0.039)	(0.032)	(0.030)
Middle Income	-0.005	-0.002	-0.001	0.002	-0.003	0.003	-0.003	0.006
	(0.080)	(0.019)	(0.022)	(0.017)	(0.016)	(0.023)	(0.020)	(0.017)
High Income	0.016	0.001	0.003	0.007	-0.003	-0.008	-0.009	0.001
	(0.088)	(0.023)	(0.026)	(0.021)	(0.019)	(0.027)	(0.024)	(0.018)
Suburban	-0.020	-0.005	-0.004	-0.001	-0.003	0.011	0.004	0.007
	(0.077)	(0.020)	(0.019)	(0.015)	(0.015)	(0.024)	(0.021)	(0.015)
Rural	-0.029	-0.004	-0.008	-0.006	-0.002	0.013	0.002	0.011
	(0.105)	(0.025)	(0.025)	(0.020)	(0.018)	(0.033)	(0.026)	(0.023)
Constant	1.067***	-0.033	0.054	-0.013	0.067	0.053	0.023	0.030
	(0.300)	(0.063)	(0.086)	(0.068)	(0.064)	(0.093)	(0.087)	(0.058)
F-Statistic	52.33	70.58	20.40	12.88	11.65	14.85	16.25	10.90

Number of Observations: 4620
Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B2: First Stage: Male

	Attainment	BA	Cert/AA	AA	Cert	Some	Some (2)	Some (4)
$\hat{p_2}$	0.751	0.178	-0.064	-0.100	0.035	0.656***	1.157***	-0.501***
	(0.708)	(0.150)	(0.159)	(0.132)	(0.124)	(0.221)	(0.192)	(0.148)
$\hat{p_3}$	1.568***	-0.125	-0.001	0.117	-0.118	1.049***	0.018	1.031***
	(0.544)	(0.126)	(0.131)	(0.088)	(0.095)	(0.168)	(0.136)	(0.126)
$\hat{p_4}$	2.524*	-0.260	1.022***	-0.086	1.108***	0.169	-0.505	0.674*
	(1.455)	(0.323)	(0.390)	(0.260)	(0.315)	(0.440)	(0.371)	(0.344)
$\hat{p_5}$	4.568***	-0.094	1.279***	1.296***	-0.017	0.038	0.172	-0.133
	(1.019)	(0.260)	(0.307)	(0.237)	(0.225)	(0.332)	(0.300)	(0.236)
$\hat{p_6}$	5.136***	1.076***	-0.012	-0.062	0.050	-0.102	-0.057	-0.045
	(0.422)	(0.080)	(0.101)	(0.081)	(0.071)	(0.135)	(0.116)	(0.098)
Married	-0.016	-0.009	0.002	0.006	-0.004	0.008	-0.001	0.009
	(0.075)	(0.017)	(0.017)	(0.013)	(0.012)	(0.022)	(0.018)	(0.015)
Children	0.055	0.015	0.000	-0.010	0.010	-0.010	-0.007	-0.003
	(0.090)	(0.017)	(0.024)	(0.017)	(0.018)	(0.030)	(0.027)	(0.020)
Math	-0.003	-0.001	0.000	0.000	0.000	0.000	0.000	0.000
	(0.005)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
Job Training	-0.057	-0.003	-0.008	0.000	-0.008	-0.004	0.009	-0.013
	(0.078)	(0.019)	(0.019)	(0.013)	(0.014)	(0.024)	(0.018)	(0.019)
Worked in 2006	-0.014	-0.002	0.012	0.010	0.002	-0.027	-0.002	-0.025
	(0.175)	(0.028)	(0.040)	(0.029)	(0.028)	(0.051)	(0.047)	(0.035)
Asian	-0.005	-0.002	0.006	0.015	-0.009	-0.001	0.025	-0.026
	(0.118)	(0.033)	(0.027)	(0.020)	(0.022)	(0.033)	(0.024)	(0.028)
African American	0.023	-0.002	0.004	0.004	0.000	0.009	0.002	0.007
	(0.099)	(0.025)	(0.025)	(0.018)	(0.020)	(0.032)	(0.027)	(0.025)
Hispanic	0.025	-0.022	0.007	0.011	-0.004	0.041	-0.022	0.063**
	(0.122)	(0.028)	(0.028)	(0.023)	(0.019)	(0.039)	(0.035)	(0.029)
Other	-0.011	-0.008	-0.000	-0.003	0.003	0.008	-0.016	0.024
	(0.146)	(0.034)	(0.036)	(0.027)	(0.024)	(0.044)	(0.038)	(0.030)
Middle Income	-0.004	-0.004	-0.000	-0.001	0.000	0.006	-0.003	0.010
	(0.075)	(0.018)	(0.018)	(0.013)	(0.015)	(0.021)	(0.018)	(0.017)
High Income	-0.031	-0.006	-0.001	0.001	-0.002	0.003	0.003	-0.001
	(0.090)	(0.023)	(0.021)	(0.016)	(0.016)	(0.026)	(0.021)	(0.019)
Suburban	-0.019	-0.001	-0.007	-0.000	-0.007	0.003	0.000	0.003
	(0.077)	(0.020)	(0.019)	(0.014)	(0.014)	(0.023)	(0.021)	(0.020)
Rural	-0.022	-0.007	-0.002	0.011	-0.013	0.008	0.008	-0.000
	(0.095)	(0.026)	(0.024)	(0.018)	(0.018)	(0.029)	(0.023)	(0.026)
Constant	1.271***	0.046	-0.017	-0.004	-0.014	0.057	0.017	0.040
	(0.346)	(0.066)	(0.088)	(0.056)	(0.071)	(0.105)	(0.093)	(0.070)
F-Statistic	65.71	117.5	15.44	11.55	7.55	19.65	21.18	16.73

Number of Observations: 4310
Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Appendix C: Evidence of Conditional Independence and Overlap

Figure C1: None - Some Two-Year (Female)

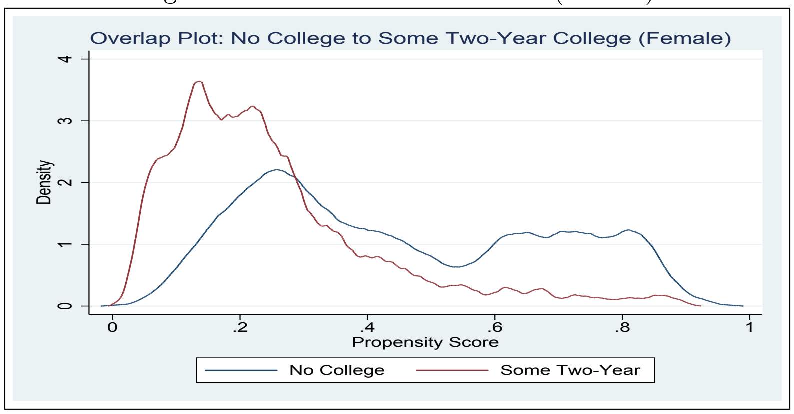


Figure C2: None - Some Two-Year (Male)

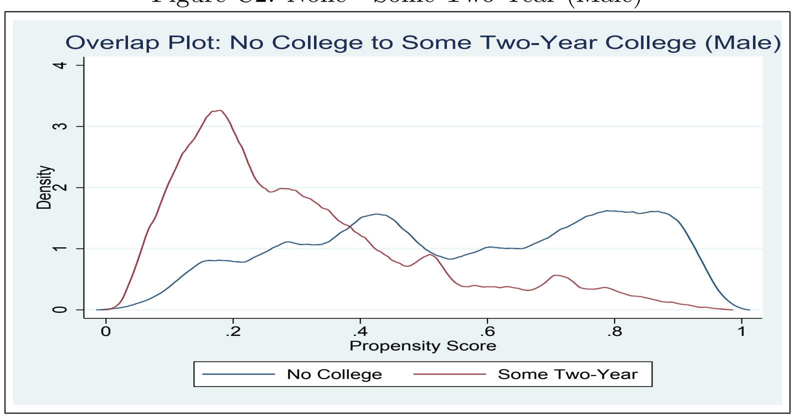


Figure C3: None - Some Four-Year (Female)

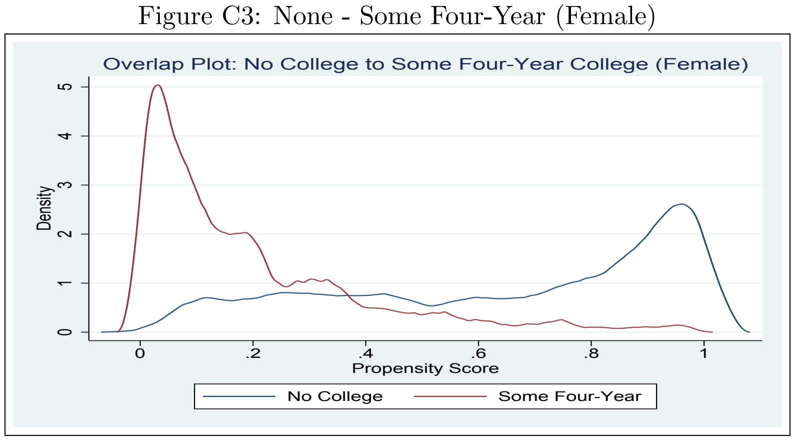


Figure C4: None - Some Four-Year (Male)

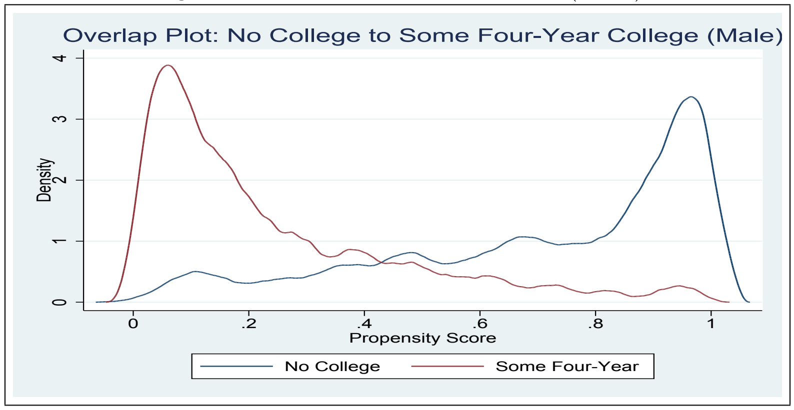


Figure C5: Some Two-Year - Certificate/AA (Female)

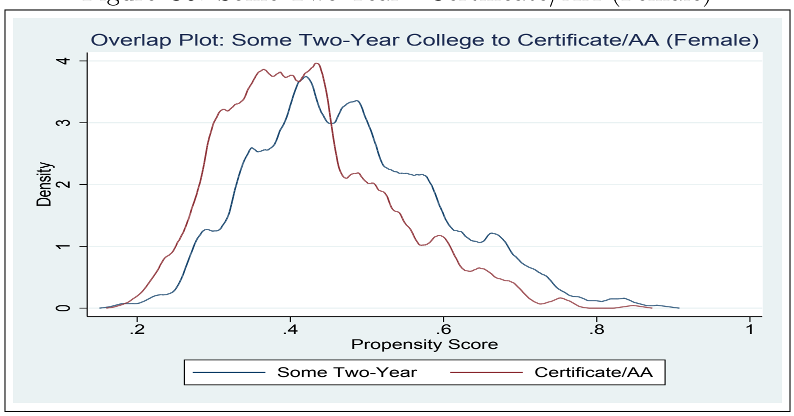


Figure C6: Some Two-Year - Certificate/AA (Male)

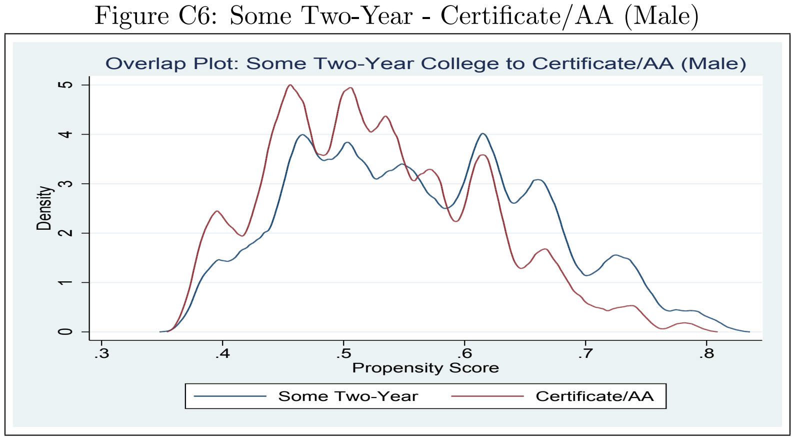


Figure C7: Some Four-Year - BA (Female)

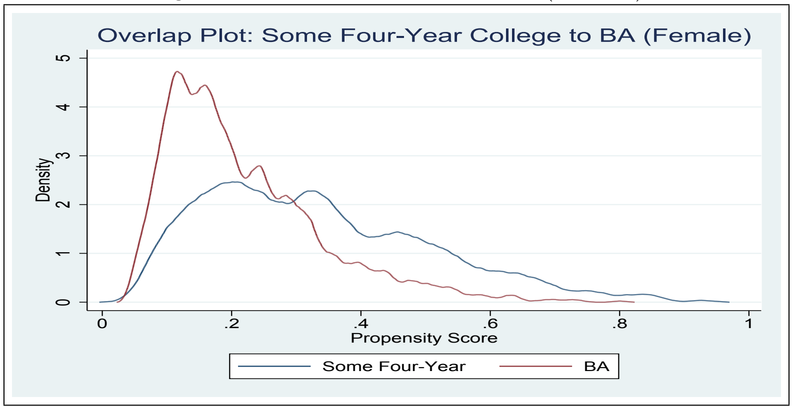


Figure C8: Some Four-Year - BA (Male)

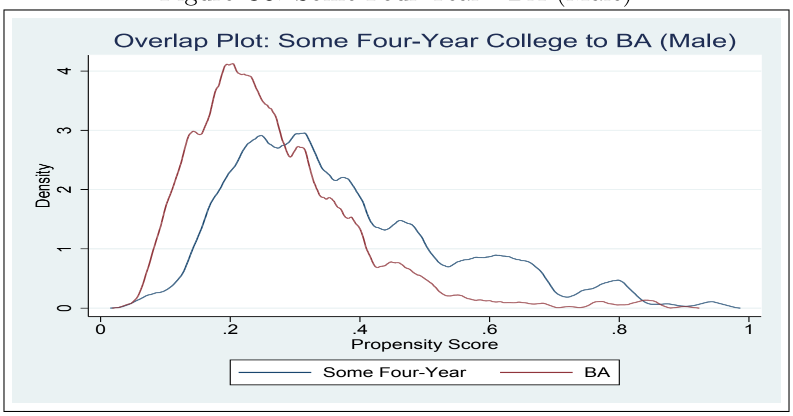


Figure C9: Some Four-Year - Certificate/AA (Female)

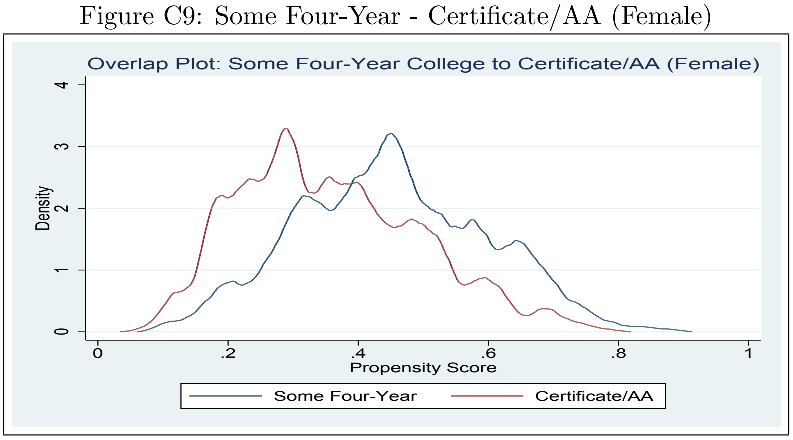


Figure C10: Some Four-Year - Certificate/AA (Male)

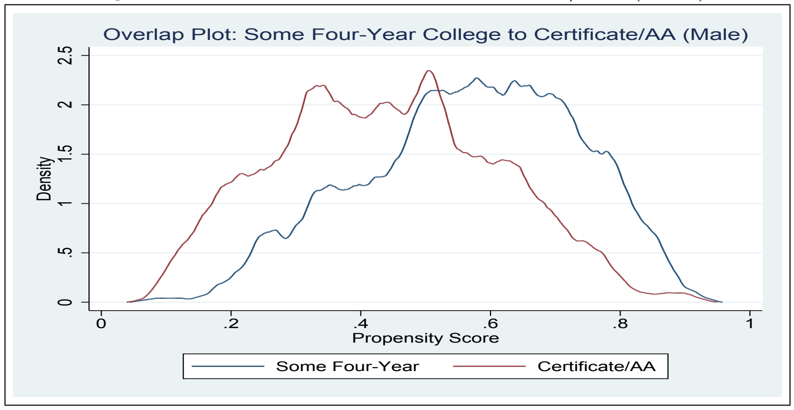


Figure C11: Some Two-Year - BA (Female)

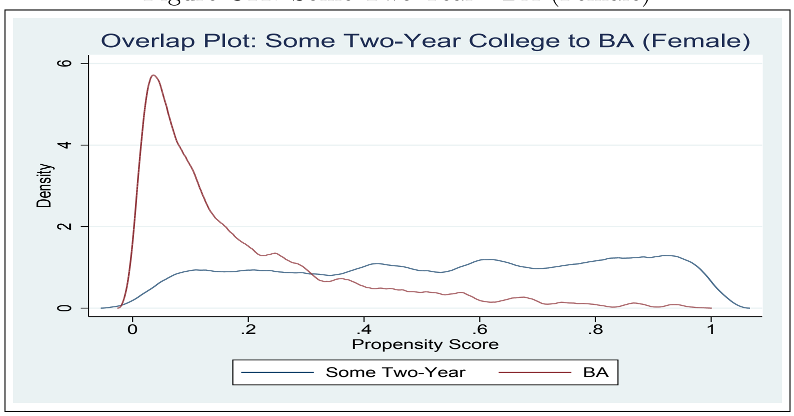


Figure C12: Some Two-Year - BA (Male)

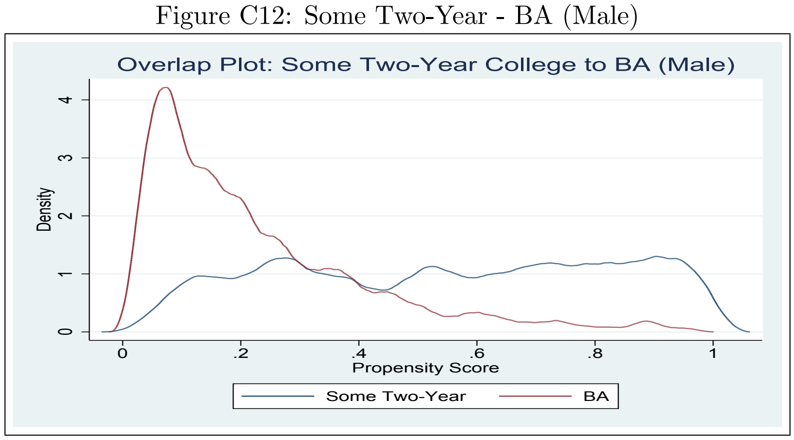


Figure C13: Certificate/AA - BA (Female)

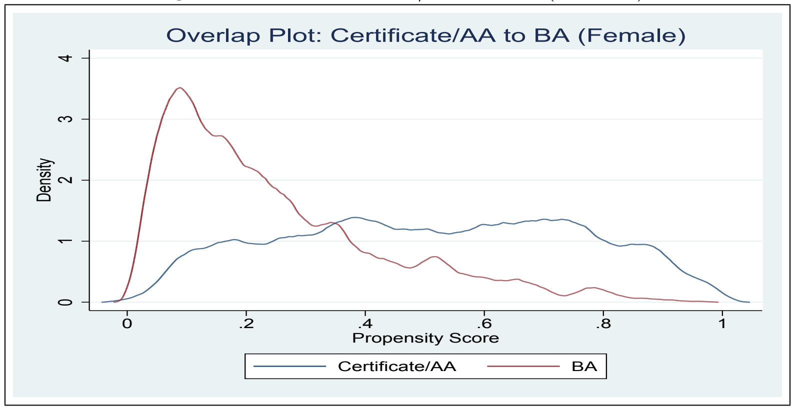


Figure C14: Certificate/AA - BA (Male)

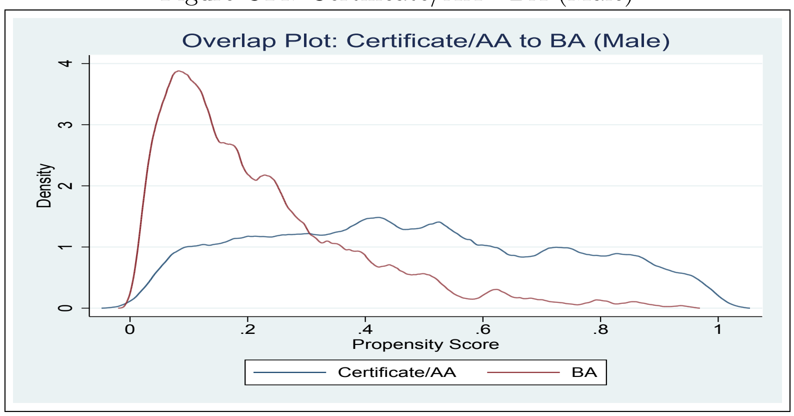


Figure C15: None - Some Two-Year (Female)

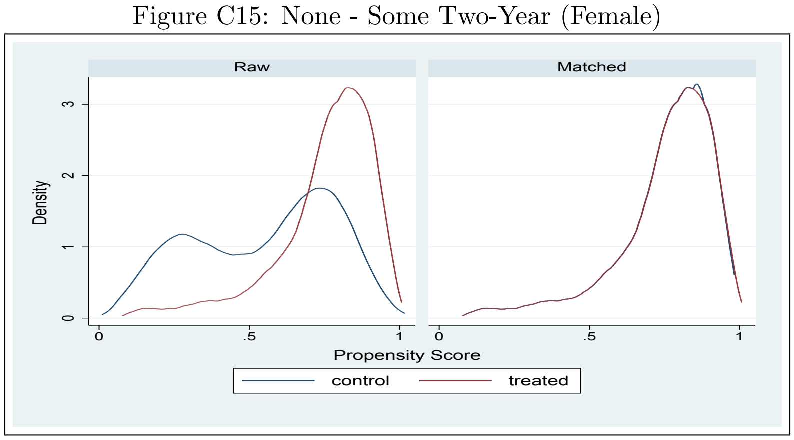


Figure C16: None - Some Two-Year (Male)

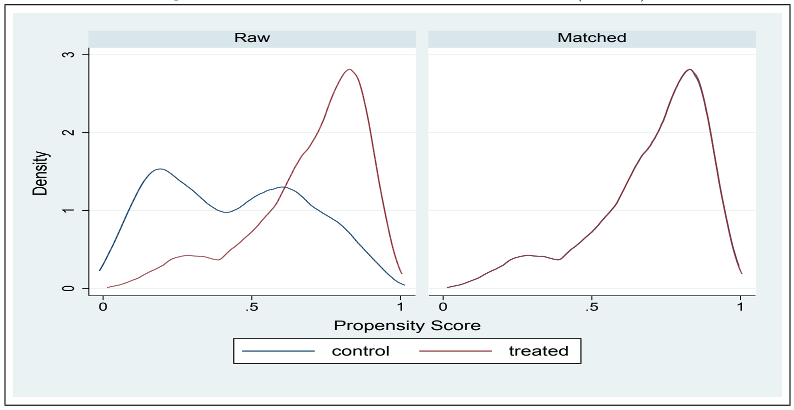


Figure C17: None - Some Four-Year (Female)

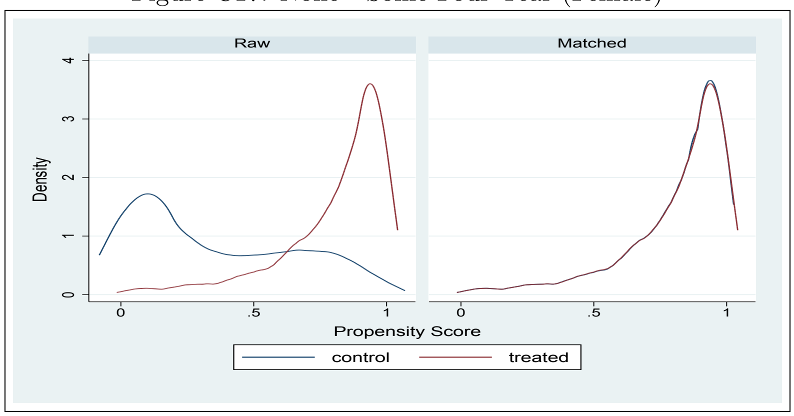


Figure C18: None - Some Four-Year (Male)

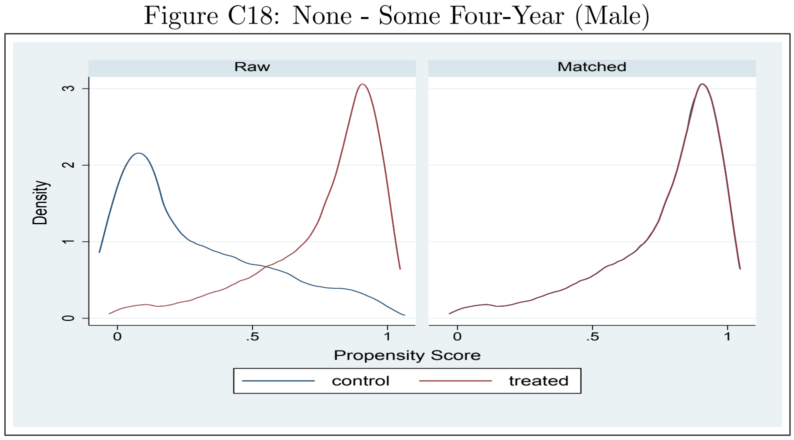


Figure C19: Some Two-Year - Certificate/AA (Female)

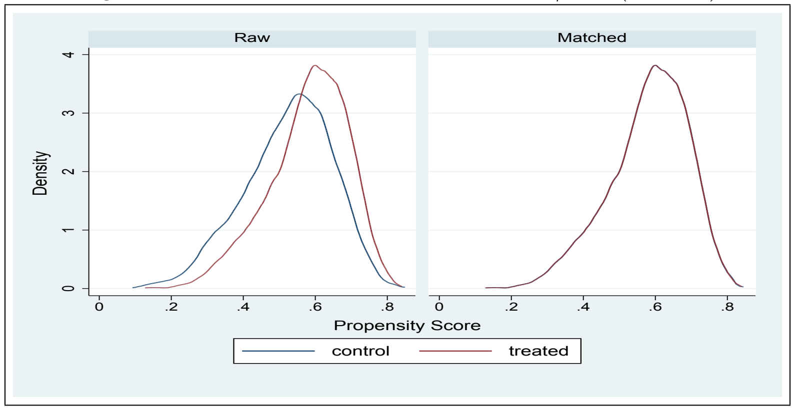


Figure C20: Some Two-Year - Certificate/AA (Male)

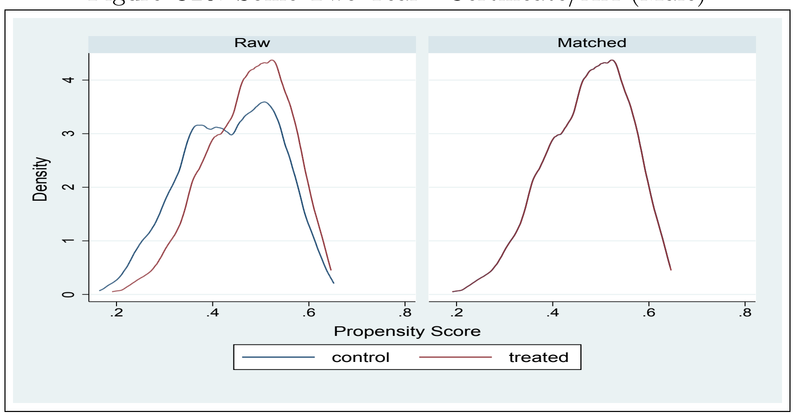


Figure C21: Some Four-Year - BA (Female)

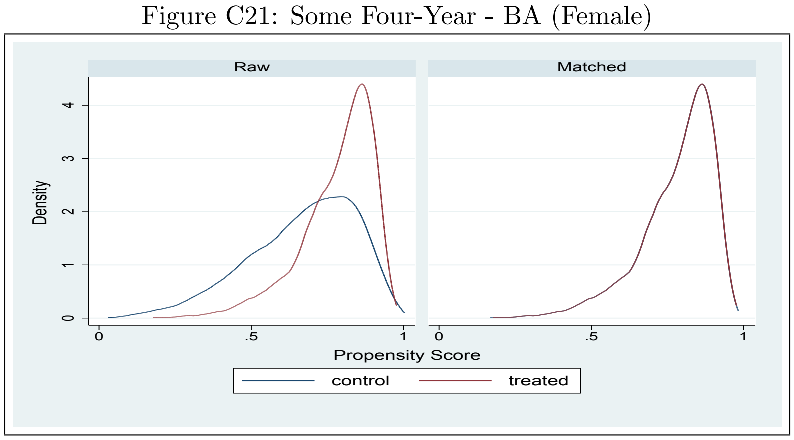


Figure C22: Some Four-Year - BA (Male)

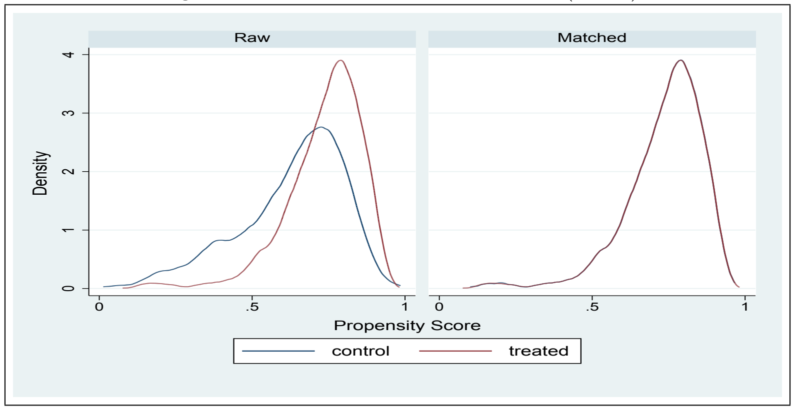


Figure C23: Some Four-Year - Certificate/AA (Female)

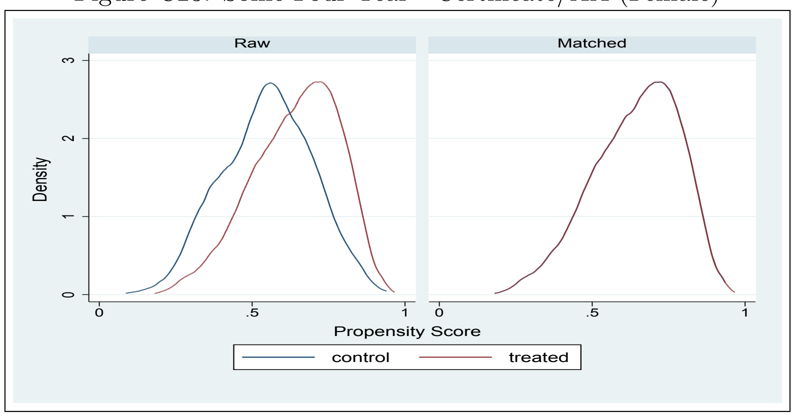


Figure C24: Some Four-Year - Certificate/AA (Male)

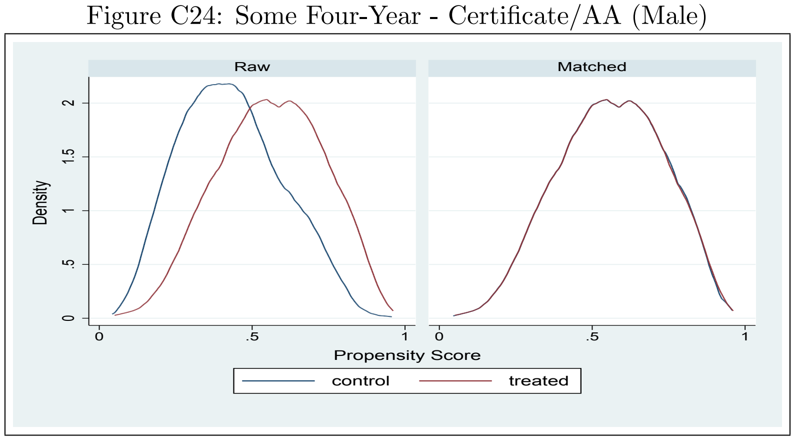


Figure C25: Some Two-Year - BA (Female)

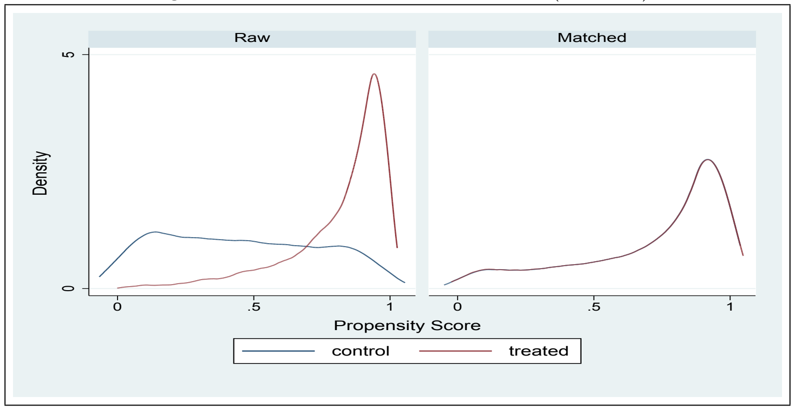


Figure C26: Some Two-Year - BA (Male)

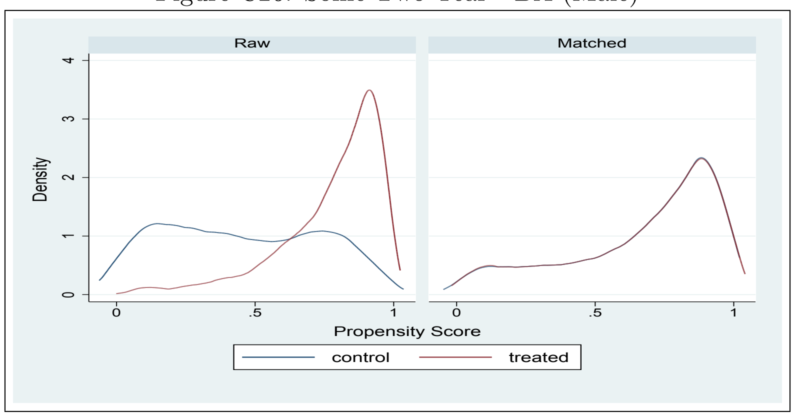


Figure C27: Certificate/AA - BA (Female)

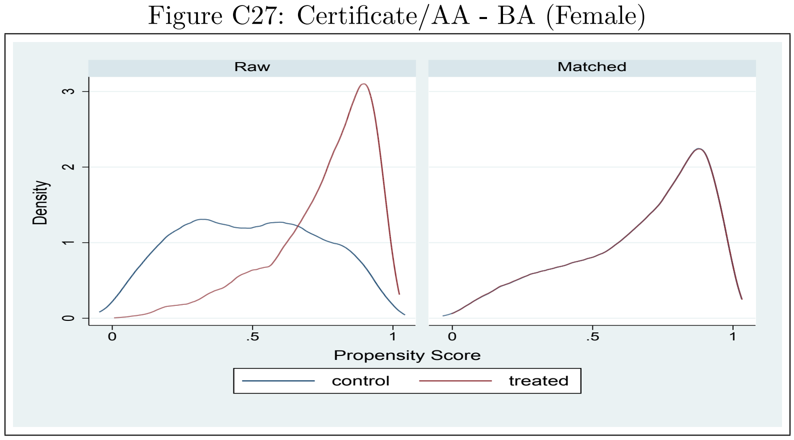
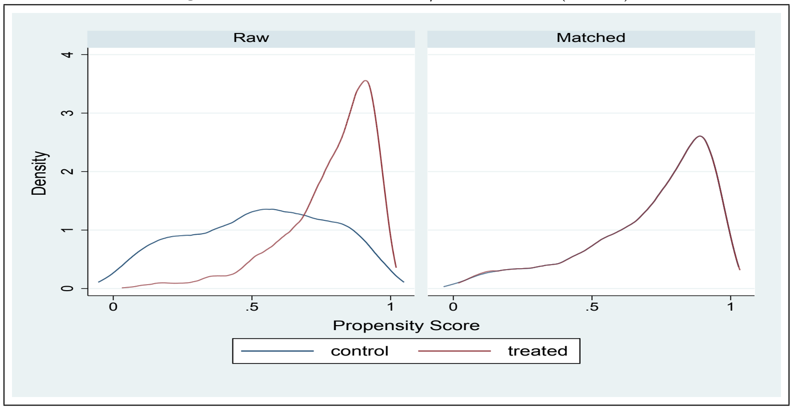


Figure C28: Certificate/AA - BA (Male)



Covariate Balance: Matched Means and Variances

Table C1: No College vs. Some Two-Year

	Female		Male	
	Std. Diff.	Variance Ratio	Std. Diff.	Variance Ratio
Married	0.12	1.14	-0.05	0.94
Children	-0.03	1.00	0.00	1.00
Math Score	0.03	1.02	0.10	0.99
Worked in 2006	0.00	0.99	0.00	1.00
Job Training	0.17	1.17	-0.10	0.95
African American	0.09	1.23	-0.04	0.92
Hispanic	0.16	1.29	-0.13	0.86
Other	-0.30	0.60	0.03	1.07
Background: Med Income	0.00	1.00	-0.09	0.99
Background: High Income	-0.04	0.94	0.12	1.20
Background: Suburban	0.27	1.07	-0.11	1.02
Background: Rural	-0.02	0.98	0.11	1.24
Census: Mid-Atlantic	0.14	1.56	0.03	1.07
Census: East North Central	-0.01	0.99	0.00	1.00
Census: West North Central	-0.15	0.69	0.00	0.99
Census: South Atlantic	-0.03	0.94	-0.02	0.96
Census: East South Central	0.02	1.05	0.08	1.31
Census: West South Central	0.07	1.21	0.03	1.09
Census: Mountain	0.03	1.13	0.04	1.21
Census: Pacific	-0.06	0.92	-0.06	0.94
	Number of Observations: 1150		Number of Observations: 1380	

Table C2: No College vs. Some Four-Year

	Female		Male	
	Std. Diff.	Variance Ratio	Std. Diff.	Variance Ratio
Married	-0.31	0.84	-0.11	0.87
Children	-0.13	0.91	-0.09	0.89
Math Score	-0.11	0.92	0.10	1.08
Worked in 2006	-0.03	0.83	-0.02	0.94
Job Training	0.40	1.59	-0.15	0.92
African American	0.08	1.16	-0.17	0.75
Hispanic	0.10	1.35	-0.01	0.98
Other	-0.12	0.81	0.08	1.19
Background: Med Income	-0.04	0.99	0.08	1.03
Background: High Income	0.09	1.11	0.02	1.02
Background: Suburban	0.26	1.15	0.13	1.04
Background: Rural	-0.44	0.65	0.08	1.19
Census: Mid-Atlantic	-0.35	0.53	0.05	1.10
Census: East North Central	0.13	1.22	0.04	1.07
Census: West North Central	0.13	1.89	0.06	1.28
Census: South Atlantic	0.22	1.40	-0.25	0.75
Census: East South Central	0.08	1.41	0.13	1.65
Census: West South Central	0.02	1.04	0.11	1.35
Census: Mountain	0.07	1.35	-0.05	0.83
Census: Pacific	-0.19	0.66	0.10	1.30
	Number of Observations: 960		Number of Observations: 1270	

Table C3: Some Two-Year vs. Certificate/AA

	Female		Male	
	Std. Diff.	Variance Ratio	Std. Diff.	Variance Ratio
Married	-0.04	0.98	-0.01	0.99
Children	0.03	1.01	-0.02	0.98
Math Score	0.01	0.95	0.02	1.12
Worked in 2006	-0.05	0.82	-0.01	0.97
Job Training	-0.01	1.00	0.00	1.00
African American	0.02	1.03	0.01	1.03
Hispanic	0.01	1.01	0.03	1.06
Other	0.08	1.20	-0.01	0.98
Background: Med Income	-0.04	0.99	0.02	1.00
Background: High Income	0.02	1.03	0.02	1.03
Background: Suburban	-0.03	1.00	0.07	1.00
Background: Rural	0.00	1.00	-0.04	0.94
Census: Mid-Atlantic	0.05	1.11	0.01	1.02
Census: East North Central	-0.01	0.98	-0.05	0.93
Census: West North Central	0.06	1.22	-0.02	0.96
Census: South Atlantic	-0.05	0.92	0.05	1.11
Census: East South Central	0.00	1.00	-0.03	0.88
Census: West South Central	0.01	1.03	-0.02	0.96
Census: Mountain	-0.02	0.93	0.00	1.00
Census: Pacific	0.03	1.06	0.03	1.07
	Number of Observations: 1720		Number of Observations: 1500	

Table C4: Some Two-Year vs. BA

	Female		Male	
	Std. Diff.	Variance Ratio	Std. Diff.	Variance Ratio
Married	0.04	1.04	0.15	1.24
Children	0.07	1.18	0.05	1.18
Math Score	-0.06	0.85	0.03	1.06
Worked in 2006	0.00	1.00	-0.05	0.60
Job Training	0.02	1.00	0.03	1.00
African American	0.07	1.28	-0.05	0.85
Hispanic	-0.02	0.95	-0.01	0.97
Other	-0.14	0.81	-0.05	0.92
Background: Med Income	0.04	1.02	0.07	1.04
Background: High Income	0.10	1.04	-0.07	0.99
Background: Suburban	-0.09	1.00	-0.20	1.02
Background: Rural	0.03	1.07	0.09	1.19
Census: Mid-Atlantic	0.02	1.05	-0.12	0.83
Census: East North Central	0.11	1.21	0.11	1.19
Census: West North Central	-0.03	0.92	0.03	1.11
Census: South Atlantic	0.02	1.03	-0.18	0.78
Census: East South Central	0.02	1.07	0.09	1.42
Census: West South Central	0.02	1.06	0.08	1.25
Census: Mountain	0.13	1.93	0.11	1.79
Census: Pacific	-0.25	0.67	0.03	1.07
	Number of Observations: 2660		Number of Observations: 2430	

Table C5: Some Four-Year vs. Certificate/AA

	Female		Male	
	Std. Diff.	Variance Ratio	Std. Diff.	Variance Ratio
Married	0.03	1.02	0.05	1.07
Children	-0.03	0.99	0.18	1.34
Math Score	0.00	1.03	-0.03	0.98
Worked in 2006	0.03	1.13	-0.14	0.69
Job Training	-0.03	0.99	0.01	1.00
African American	-0.01	0.98	-0.02	0.95
Hispanic	-0.02	0.96	0.07	1.16
Other	-0.03	0.95	-0.10	0.82
Background: Med Income	-0.05	0.99	0.01	1.00
Background: High Income	-0.04	0.95	0.05	1.06
Background: Suburban	-0.05	1.00	0.04	1.00
Background: Rural	0.03	1.04	-0.07	0.91
Census: Mid-Atlantic	-0.08	0.85	-0.05	0.90
Census: East North Central	-0.11	0.87	0.05	1.08
Census: West North Central	0.05	1.18	0.05	1.14
Census: South Atlantic	0.03	1.06	-0.06	0.91
Census: East South Central	0.06	1.21	0.04	1.16
Census: West South Central	0.05	1.12	-0.01	0.98
Census: Mountain	0.03	1.11	-0.05	0.84
Census: Pacific	0.02	1.05	0.02	1.04
	Number of Observations: 1530		Number of Observations: 1390	

Table C6: Some Four-Year vs. BA

	Female		Male	
	Std. Diff.	Variance Ratio	Std. Diff.	Variance Ratio
Married	0.02	1.02	-0.11	0.89
Children	0.04	1.09	-0.03	0.91
Math Score	0.05	1.06	0.03	1.10
Worked in 2006	0.05	2.66	-0.07	0.53
Job Training	0.00	1.00	0.02	1.00
African American	0.05	1.19	-0.04	0.89
Hispanic	-0.07	0.84	0.05	1.20
Other	-0.02	0.97	0.02	1.04
Background: Med Income	0.04	1.02	-0.04	0.98
Background: High Income	-0.02	1.00	0.00	1.00
Background: Suburban	-0.01	1.00	0.07	1.01
Background: Rural	-0.01	0.98	-0.06	0.90
Census: Mid-Atlantic	0.01	1.02	0.01	1.01
Census: East North Central	0.03	1.05	0.02	1.04
Census: West North Central	0.03	1.09	-0.02	0.94
Census: South Atlantic	0.04	1.08	0.01	1.02
Census: East South Central	0.01	1.03	0.05	1.20
Census: West South Central	-0.06	0.85	-0.04	0.91
Census: Mountain	-0.03	0.88	0.02	1.10
Census: Pacific	-0.03	0.94	0.00	1.00
	Number of Observations: 2470		Number of Observations: 2320	

Table C7: Certificate/AA vs. BA

	Female		Male	
	Std. Diff.	Variance Ratio	Std. Diff.	Variance Ratio
Married	0.02	1.02	0.08	1.11
Children	0.02	1.04	0.03	1.09
Math Score	-0.02	1.06	-0.01	1.01
Worked in 2006	0.01	1.14	0.00	1.00
Job Training	-0.05	1.00	-0.03	1.00
African American	0.10	1.44	-0.02	0.95
Hispanic	0.02	1.05	0.06	1.25
Other	-0.16	0.78	0.10	1.22
Background: Med Income	0.02	1.01	0.02	1.01
Background: High Income	-0.04	0.99	-0.03	0.99
Background: Suburban	-0.11	1.00	-0.04	1.00
Background: Rural	-0.02	0.96	0.05	1.09
Census: Mid-Atlantic	-0.02	0.96	0.02	1.04
Census: East North Central	-0.01	0.98	-0.04	0.94
Census: West North Central	-0.01	0.96	-0.07	0.82
Census: South Atlantic	0.11	1.22	-0.02	0.96
Census: East South Central	0.13	1.60	0.06	1.28
Census: West South Central	-0.04	0.88	0.04	1.13
Census: Mountain	-0.05	0.81	0.05	1.28
Census: Pacific	-0.11	0.82	-0.01	0.99
	Number of Observations: 2780		Number of Observations: 2310	

Appendix D: Alternative Propensity Score Models

Table D1: Propensity Score Matching (ATE)

	Female		Male	
	$ln(wage)$	$ln(Income)$	$ln(wage)$	$ln(Income)$
Some Two-Year vs. No College	0.084***	0.010	0.009	-0.001
	(0.030)	(0.095)	(0.034)	(0.095)
N	1030	900	1240	1220
Some Four-Year vs. No College	0.088**	0.114	-0.019	-0.232
	(0.038)	(0.102)	(0.037)	(0.158)
N	900	810	1120	1100
Certificate/AA vs. Some Two-Year	0.114***	0.101	0.115***	0.168***
	(0.020)	(0.071)	(0.029)	(0.057)
N	1610	1480	1330	1350
BA vs. Some Two-Year	0.222***	0.446***	0.239***	0.396***
	(0.039)	(0.110)	(0.041)	(0.064)
N	2490	2410	2200	2220
Certificate/AA vs. Some Four-Year	0.076***	-0.059	0.113***	0.122**
	(0.026)	(0.066)	(0.031)	(0.059)
N	1470	1390	1220	1230
BA vs. Some Four-Year	0.202***	0.262***	0.203***	0.279***
	(0.026)	(0.062)	(0.027)	(0.054)
N	2360	2320	2080	2110
BA vs. Certificate/AA	0.102***	0.343***	0.108***	0.175***
	(0.025)	(0.066)	(0.031)	(0.060)
N	2670	2620	2070	2100

Variables: Married, Children, Race, Income, Job Training, Worked in 2006, Urbanicity, Census Region
Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table D2: Propensity Score Matching (ATE) - Alternative Model

	Female		Male	
	$ln(wage)$	$ln(Income)$	$ln(wage)$	$ln(Income)$
Some Two-Year vs. No College	0.030	-0.070	0.029	-0.054
	(0.030)	(0.115)	(0.040)	(0.088)
N	1030	900	1240	1220
Some Four-Year vs. No College	0.081**	0.230*	-0.003	-0.095
	(0.038)	(0.138)	(0.032)	(0.136)
N	900	810	1120	1100
Certificate/AA vs. Some Two-Year	0.143***	0.104	0.127***	0.147**
	(0.022)	(0.072)	(0.027)	(0.060)
N	1610	1480	1330	1350
BA vs. Some Two-Year	0.261***	0.431***	0.197***	0.447***
	(0.057)	(0.142)	(0.033)	(0.069)
N	2490	2410	2200	2220
Certificate/AA vs. Some Four-Year	0.100***	-0.051	0.108***	0.117**
	(0.024)	(0.068)	(0.031)	(0.059)
N	1470	1390	1220	1230
BA vs. Some Four-Year	0.185***	0.172***	0.195***	0.268***
	(0.024)	(0.056)	(0.029)	(0.061)
N	2360	2320	2080	2110
BA vs. Certificate/AA	0.119***	0.364***	0.089***	0.153**
	(0.024)	(0.069)	(0.029)	(0.062)
N	2670	2620	2070	2100

Variables: Married, Children, Race, SES, Job Training
Variables: Worked (2006), Urbanicity, Unemployment Rate, Census
Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table D3: Propensity Score Matching (ATET) - Alternative Model

	Female		Male	
	$ln(wage)$	$ln(Income)$	$ln(wage)$	$ln(Income)$
Some Two-Year vs. No College	0.029	-0.103	0.001	-0.070
	(0.036)	(0.135)	(0.050)	(0.113)
N	1030	900	1240	1220
Some Four-Year vs. No College	0.092**	0.130	-0.006	-0.015
	(0.039)	(0.182)	(0.041)	(0.105)
N	900	810	1120	1100
Certificate/AA vs. Some Two-Year	0.150***	0.099	0.124***	0.196***
	(0.026)	(0.077)	(0.032)	(0.065)
N	1610	1480	1330	1350
BA vs. Some Two-Year	0.285***	0.481***	0.222***	0.531***
	(0.076)	(0.184)	(0.045)	(0.086)
N	2490	2410	2200	2220
Certificate/AA vs. Some Four-Year	0.124***	-0.039	0.103***	0.119*
	(0.025)	(0.073)	(0.036)	(0.070)
N	1470	1390	1220	1230
BA vs. Some Four-Year	0.199***	0.176***	0.187***	0.280***
	(0.026)	(0.057)	(0.033)	(0.073)
N	2360	2320	2080	2110
BA vs. Certificate/AA	0.146***	0.405***	0.099***	0.177**
	(0.029)	(0.081)	(0.034)	(0.072)
N	2670	2620	2070	2100

Variables: Married, Children, Race, SES, Job Training
Variables: Worked (2006), Urbanicity, Unemployment Rate, Census
Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table D4: Propensity Score Matching (ATE) - Caliper 0.01

	Female		Male	
	$ln(wage)$	$ln(Income)$	$ln(wage)$	$ln(Income)$
Some Two-Year vs. No College	0.042	0.067	0.009	-0.001
	(0.028)	(0.118)	(0.034)	(0.095)
N	1000	870	1240	1220
Some Four-Year vs. No College	0.089***	0.236*	-0.019	-0.232
	(0.034)	(0.121)	(0.037)	(0.158)
N	780	710	1120	1100
Certificate/AA vs. Some Two-Year	0.136***	0.093	0.115***	0.168***
	(0.022)	(0.074)	(0.029)	(0.057)
N	1590	1470	1330	1350
BA vs. Some Two-Year	0.251***	0.449***	0.239***	0.396***
	(0.040)	(0.107)	(0.041)	(0.064)
N	2430	2370	2200	2220
Certificate/AA vs. Some Four-Year	0.067***	-0.077	0.113***	0.122**
	(0.025)	(0.071)	(0.031)	(0.059)
N	1460	1370	1220	1230
BA vs. Some Four-Year	0.187***	0.238***	0.203***	0.279***
	(0.030)	(0.057)	(0.027)	(0.054)
N	2350	2310	2080	2110
BA vs. Certificate/AA	0.101***	0.324***	0.108***	0.175***
	(0.026)	(0.068)	(0.031)	(0.060)
N	2660	2610	2070	2100

Variables: Married, Children, Race, Income, Job Training, Worked in 2006, Urbanicity, Census Region
Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table D5: Propensity Score Matching (ATET) - Caliper 0.01

	Female		Male	
	$ln(wage)$	$ln(Income)$	$ln(wage)$	$ln(Income)$
Some Two-Year vs. No College	0.055*	0.089	-0.005	0.054
	(0.032)	(0.130)	(0.033)	(0.113)
N	1000	870	1240	1220
Some Four-Year vs. No College	0.116***	0.151	-0.015	-0.180**
	(0.040)	(0.154)	(0.034)	(0.092)
N	780	710	1120	1100
Certificate/AA vs. Some Two-Year	0.128***	0.102	0.126***	0.172***
	(0.026)	(0.081)	(0.033)	(0.066)
N	1590	1470	1330	1350
BA vs. Some Two-Year	0.276***	0.474***	0.255***	0.383***
	(0.051)	(0.134)	(0.057)	(0.079)
N	2430	2370	2200	2220
Certificate/AA vs. Some Four-Year	0.065**	-0.055	0.066*	0.147**
	(0.029)	(0.076)	(0.036)	(0.074)
N	1460	1370	1220	1230
BA vs. Some Four-Year	0.183***	0.234***	0.202***	0.288***
	(0.034)	(0.063)	(0.029)	(0.059)
N	2350	2310	2080	2110
BA vs. Certificate/AA	0.121***	0.362***	0.118***	0.204***
	(0.032)	(0.082)	(0.036)	(0.069)
N	2660	2610	2070	2100

Variables: Married, Children, Race, Income, Job Training, Worked in 2006, Urbanicity, Census Region
Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

References

Abadie, A., and G. W. Imbens. 2012. "Matching on the estimated propensity score." *Harvard University and National Bureau of Economic Research.* http://www.hks.harvard.edu/fs/aabadie/pscore.pdf.

Abadie, A. and G. W. Imbens. 2016. "Matching on the Estimated Propensity Score," *Econometrica,* 84(2):781-807.

Abdulkadiroglu, A., J. D. Angrist, S. M. Dynarski, T J. Kane, and P. A. Pathak. 2011. "Accountability and Flexibility in Public Schools: Evidence from Boston's Charters and Pilots." *The Quarterly Journal of Economics* 126(2): 699-748.

Adams, R., Almeida, H., and D. Ferreira. 2009. "Understanding the relationship between founder-CEOs and firm performance." *Journal of Empirical Finance,* 16 (1): 136-150.

Agodini, R. and M. Dynarski. 2004. "Are experiments the only option? A look at dropout prevention programs." *Review of Economics and Statistics.* 86(1): 180-194.

Alba, R. D., and D. E. Lavin. 1981. "Community colleges and tracking in higher education." *Sociology of Education.* 54(4): 223-237.

Alfonso, M. 2006. "The Impact of Community College Attendance on Baccalaureate Attainment." *Research in Higher Education.* 47(8):873-903.

Alon, S., T. Domina, and M. Tienda. 2010. "Stymied Mobility or Temporary Lull? The Puzzle of Lagging Hispanic College Degree Attainment." *Social Forces.* 88(4): 1807-1832.

Anderson, K. L. 1981. "Post-high school experiences and college attrition." *Sociology of Education.* 54(1): 1-15.

Anderson, K.L. 1984. "Institutional Differences in College Effects." *ERIC 256-204.* Boca Raton, FL: Florida Atlantic University.

Angrist, J. D. and J-S Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton, NJ. Princeton University Press.

Angrist, J. 2013. "The Perils of Peer Effects." *NBER Working Paper 19774.*

Arellano, M. and S. Bond. 1991. "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations." *Review of Economic Studies, Wiley Blackwell* 58(2): 277-297.

Bailey, M.J. and S.M. Dynarski. 2011. "Gains and Gaps: Changing Inequality in U.S. College Entry and Completion." *NBER Working Paper No. 17633.*

Bailey, T., Kienzl, G. S., and D.E. Marcotte. 2004. "The returns to a sub-baccalaureate education: The effects of schooling, credentials, and program of study on economic outcomes." New York, NY: Teachers College, Community College Research Center, Columbia University.

Bailey, T.R., and Averianova, I. 1998. "Multiple Missions of Community Colleges. Conflicting or Complementary?" *Community College Research Center,* Teachers College, Columbia University. New York.

Bardallis, D. 1999. "And Now a Word from Our Sponsors - Your Local Public Schools." Midland, MI: Mackinac Center for Public Policy.

Becker, G. S. 1962. "Investment in human capital: A theoretical analysis." *Journal of Political Economy,* 70(5.2): 949.

Belfield, C. R., and T. Bailey. 2011. "The benefits of attending community college: A review of the evidence." *Community College Review.* 39(1):46-68.

Black, D., Kolesnikova, N., and L. Taylor. 2009. "Earnings Functions When Wages and Prices Vary by Location." *Journal of Labor Economics,* 27 (01): 21-47.

Black, D., and Smith, J. 2004. "How robust is the evidence on the effects on college quality? Evidence from matching." *Journal of Econometrics,* 121: 99-124.

Blundell, R., L. Dearden and B. Sianesi. 2005. "Evaluating the Effect of Education on Earnings: Models, Methods and Results from the National Child Development Survey." *Journal of the Royal Statistical Society. Series A (Statistics in Society),* 168 (3): 473-512.

Blundell, R. and M.C. Dias. 2009. "Alternative Approaches to Evaluation in Empirical Microeconomics." *The Journal of Human Resources,* 44 (3): 565-640.

Bohte, J. 2004. "Examining the Impact of Charter Schools on Performance in Traditional Public Schools." *Policy Studies Journal* 32(4): 501 - 520.

Boozer, M. and C. Rouse. 2001. "Intraschool Variation in Class Size: Patterns and Implications," *Journal of Urban Economics,* 50(1): 163-189.

The Boston Foundation.*Informing the Debate: Comparing Boston's Charter, Pilot, and Traditional Schools.* https://folio.iupui.edu/bitstream/handle/10244/726/InformingTheDebate_Final.pdf?sequence=2 January 2009.

Bosworth, B. 2010. "Certificates count: An analysis of sub-baccalaureate certificates." Washington, DC: Complete College America and Future Works.

Bound, John, Michael Lovenheim and Sarah Turner. 2010. "Why Have College Completion Rates Declined? An Analysis of Changing Student Preparation and Collegiate Resources." *American Economic Journal: Applied Economics*, 2(3): 129-57.

Bowen, W., Chingos, M., and McPherson, M. 2009. *Crossing the finish line.* Princeton, NJ: Princeton University Press.

Bozick, Robert. 2009. "Job Opportunities, Economic Resources, and the Postsecondary Destinations of American Youth." *Demography,* 46(3): 493-512.

Bozick, R. and S. DeLuca. 2005. "Better Late than Never? Delayed enrollment in the high school to college transition." *Social Forces,* 84: 527-550.

Brand, J.E., and Y. Xie. 2010. "Who Benefits Most from College? Evidence for Negative Selection in Heterogeneous Economic Returns to Higher Education." *American Sociological Review,* 75(2): 273-302.

Breneman, D., and Nelson, S. 1981. "Financing Community Colleges: An Economic Perspective." *Brookings Institution,* Washington D.C.

Bresnahan, T. F., and P. C. Reiss. 1991. “Entry and Competition in Concentrated Markets,” *The Journal of Political Economy.* 99(5): 977-1009.

Budde, Ray. 1988. “Education by Charter: Restructuring School Districts. Key to Long-Term Continuing Improvement in American Education.” *Publication Sales, Regional Laboratory for Educational Improvement of the Northeast and Islands.*

Bureau of Labor Statistics Table 7, 2013. “’Employment by summary education and training assignment, 2012 and projected 2022,” *BLS on the Internet at* *http://www.bls.gov/news.release/archives/ecopro_12192013.pdf* (visited on February 16, 2016).

Bureau of Labor Statistics Table 8, 2013. “Occupations with the largest projected number of job openings due to growth and replacement needs, 2012 and projected 2022,” *BLS on the Internet at* *http://www.bls.gov/news.release/archives/ecopro_12192013.pdf* (visited on February 16, 2016).

Bureau of Labor Statistics Economic News Release 2015. “’Employment Projections: 2014-24 Summary,” *BLS on the Internet at* *http://www.bls.gov/news.release/ecopro.nr0.htm* (visited on February 16, 2016).

Bureau of Labor Statistics, U.S. Department of Labor, The Economics Daily, Unemployment rates by educational attainment in April 2015 on the Internet at `http://www.bls.gov/opub/ted/2015/unemployment-rates-by-educational-attainment-in-april-2015.htm` (visited November 11, 2015).

Bureau of Labor Statistics, U.S. Department of Labor, The Economics Daily, More education still means more pay in 2014 on the Internet at `http://www.bls.gov/opub/ted/2015/more-education\-still-means-more-pay-in-2014.htm` (visited November 11, 2015).

Caliendo, M. and Kopeinig, S. 2005. “Some practical guidance for the implementation of propensity score matching.” IZA Discussion paper No. 1588.

CAPE: Council for American Private Education. 2012. `http://www.capenet.org/facts.html.` December 12, 2013.

Card, David. 1999. “The causal effect of education on earnings.” In *The handbook of labor economics, vol. 3A*, ed. Orley C. Ashenfelter and David Card, 180163. Amsterdam: Elsevier.

Card, David. 2001. “Estimating the return to schooling: Progress on some persistent econometric problems.” *Econometrica* 69: 11271160.

Carnevale, A.P., Smith, N., and J. Strohl. “Recovery: Job Growth and Education Requirements Through 2020.” *Georgetown Public Policy Institute: Center on Education and the Workforce.* on the Internet at `https://cew.georgetown.edu/wp-content/uploads/2014/11/Recovery2020.ES_.Web_.pdf` (visited November 11, 2015).

Cejda, B.D., Kaylor, A. J., and Rewery, K. L. 1998. “Transfer shock in an academic discipline: The relationship between students’ majors and their academic performance.” *Community College Review,* 26(3): 1-13.

Center for Education Reform. 2008. “Charter School Funding: Follow the Money.”

Center for Research on Education Outcomes (CREDO). 2009. *Multiple Choice: Charter School*

Performance in 16 States. http://credo.stanford.edu/reports/MULTIPLE_CHOICE_CREDO.pdf.

Center for Research on Education Outcomes (CREDO). 2013. *Charter School Performance in Massachusetts.* http://credo.stanford.edu/documents/MAReportFinal_000.pdf.

Christie, R. L., and Hutcheson, P. 2003. "Net effects of institutional type on baccalaureate degree attainment of "traditional" students." *Community College Review.* 31(2): 1-20.

Chubb, J. and T. M. Moe. 1990."Politics, Markets, and America's Schools." *Washington, DC: The Bookings Institution Press.*

Cohen, A., and Brawer, F.B. 2003. *The American Community College* Fourth Edition. San Francisco: Jossey: Bass.

Cutler, D.M., and A. Lleras-Muney. 2006. "Education and Health: Evaluating Theories and Evidence." *NBER Working Paper No. 12352.*

Dadgar, M., and M.J Trimble. 2015. "Labor Market Returns to Sub-Baccalaureate Credentials: Does a Community College Degree or Certificate Pay?" *Educational Evaluation and Policy Analysis,* 37(4): 399-418.

Dehejia, R. H., and Wahba, S. 2002. "Propensity score matching methods for non-experimental causal studies." *Review of Economics and Statistics,* 84: 151-161.

Deming, D. J., J. S. Hastings, T. J. Kane, and D. O. Staiger. 2014. "School Choice, School Quality, and Postsecondary Attainment." *The American Economic Review,* 104 (3): 991-1013.

Department of Education *Massachusetts Department of Education* http://www.doe.mass.edu/ September 11, 2013.

Department of Education *Report on Charter School Waitlists. July 2013.* http://www.doe.mass.edu/charter/reports/2013Waitlist.pdf.

Dougherty, K. J. 1992. "Community Colleges and Baccalaureate Attainment." *The Journal of Higher Education.* 63 (2): 188-214.

Dowd, A. C. and T. Coury. 2006. "The Effect of Loans on the Persistence and Attainment of Community College Students." *Research in Higher Education,* 47(1): 33-62.

Downes, T. A., and S. M. Greenstein. 1996. "Understanding the Supply Decisions of Nonprofits: Modeling the Location of Private Schools." *The RAND Journal of Economics.* 27(2): 365- 390.

Doyle, W. R. 2009. "The Effect of Community College Enrollment on Bachelor's Degree Completion." *Economics of Education Review,* 28(2): 199-206.

Dynarski, Susan. 2005. "Building the Stock of College-Educated Labor." *NBER Working Paper No. 11604.*

Freedman, D. A. and J.S. Sekhon. 2010. "Endogeneity in Probit Response Models." *Political Analysis.* 18(2): 138-150.

Freeman, K. 2005. *African Americans and college choice: The influence of family and school.* Albany, NY: State University of New York Press.

Friedman, M. 1962. *Capitalism and Freedom.* Chicago, IL. University of Chicago Press.

Fry, R. 2004. "Latino Youth Finishing College: The Role of Selective Pathways." *Washington, DC: Pew Hispanic Center.*

Ganderton, P.T., and Santos, R. 1995. "Hispanic college attendance and completion: Evidence from the High School and Beyond Surveys." *Economics of Education Review.* 14(1): 35-46.

Gill, A. M. and D. E. Leigh. 1997. "Labor market returns to community colleges: Evidence for returning adults." *Journal of Human Resources,* 32(2): 334-353.

Gill, A. M. and D. E. Leigh. 2003. "Do the Returns to Community Colleges Differ between Academic and Vocational Programs?" *Journal of Human Resources.* 38(1): 134-155.

Gonzalez, A., and M.J. Hilmer. 2006. "The role of two-year colleges in the improving situation of Hispanic postsecondary education." *Economics of Education Review,* 25(3): 249-257.

Goodman, J., Hurwitz, M., and J. Smith. 2015. "College Access, Initial College Choice and Degree Completion." *NBER Working Paper No. 20996.*

Gottfried, M. A. 2010. "Evaluating the Relationship Between Student Attendance and Achievement in Urban Elementary and Middle Schools: An Instrumental Variables Approach." *American Educational Research Journal,* 47(2): 434-465.

Greene, J. P. and G. Forster. 2002. "Rising to the Challenge: The Effect of School Choice on Public Schools in Milwaukee and San Antonio" *Mimeo, The Manhattan Institute.*

Greene, W. H. 2003. *Econometric Analysis.* Upper Saddle River, NJ. Pearson Education.

Grubb, W. N. 1993a. "The varied economic returns to postsecondary education: New evidence from the class of 1972." *Journal of Human Resources,* 28: 365382.

Grubb, W. N. 1993b. "The long-run effects of proprietary schools on wages and earnings: Implications for federal policy." *Educational Evaluation and Policy Analysis,* 15(1): 17-33.

Grubb, W. N. 1995. "Postsecondary Education and the Sub-Baccalaureate Labor Market: Corrections and Extensions." *Economics of Education Review.* 14(3): 285-299.

Grubb, W. N. 1997. "The returns to education in the sub-baccalaureate labor market, 1984-1990." *Economics of Education Review,* 16: 231-245.

Grubb, W. N. 2002a. "Learning and earning in the middle, Part I: National studies of pre-baccalaureate education." *Economics of Education Review,* 21: 299-321.

Grubb, W. N. 2002b. "Learning and earning in the middle, part II: State and local studies of pre-baccalaureate education." *Economics of Education Review,* 21: 401414.

Hanushek, E.A. 1986. "The Economics of Schooling: Production and Efficiency in Public Schools." *Journal of Economic Literature.* 24(3): 1141-1177.

Hanushek, E.A. 1997. "Assessing the Effects of School Resources on Student Performance: An Update." *Educational Evaluation and Policy Analysis.* 19(2): 141-164.

Hanushek, E., G. Schwerdt, L. Woessmann, and L. Zhang. forthcoming. "General education, vocational education, and labor-market outcomes over the life-cycle." *Journal of Human Resources.* (Winter 2017).

Harvey, J. A., and H. Busher. 1996. "Marketing Schools and Consumer Choice." *International Journal of Educational Management* 10(4): 2632.

Heckman, J. J., Ichimura, H., and Todd, P. 1997. "Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme." *Review of Economic Studies,* 64: 605-654.

Hoachlandar, G., Sikora, A. C., Horn, L., and Carroll, C.D. 2003. "Community College Students: Goals, Academic Preparation, and Outcomes (No. NCES 2003-164)." *National Center for Education Statistics,* U.S. Department of Education, Washington, D.C.

Holmes, G. M., J. DeSimon, and N. G. Rupp. 2004. "Does School Choice Increase School Quality?" *National Bureau of Economic Research.*

Hoxby, C. M. 1994. "Do Private Schools Provide Competition for Public Schools?" *NBER.*

Hoxby, C. M. 1998. "The Effects of Class Size and Composition on Student Achievement: New Evidence from Natural Population Variation." *NBER Working Paper 6869.*

Hoxby, C. M. 2002. *How School Choice Affects the Achievement of Public School Students* in Paul Hill, ed., Choice with Equity. Stanford, CA. Hoover Press.

Hoxby, C. M. 2004. "School Choice and School Competition: Evidence from the United States". *Swedish Economic Policy Review*, 10(2).

Hoxby, C. M. 2003. *A Nation at Risk, Then and Now: What has Changed and What has Not* in P. Peterson, ed. Our Schools and Our Future. Stanford, CA. Hoover Institution Press.

Hoxby, C. M. 2001. "Rising Tide: New Evidence on Competition and the Public Schools, *Education Next* 1(4), 69-74.

Hoxby, C. M. 2000. "Does Competition Among Public Schools Benefit Students and Taxpayers?" *American Economic Review* 90(5): 1209-1238.

Hoxby, C. M. 1998. "Analyzing School Choice Reforms Using America's Traditional Forms of School Choice," in Bryan Hassel and Paul Peterson, eds., Learning from School Choice. Washington, D.C. Brookings Institution Press.

Hoxby, C. M. 1998. "What Do America's 'Traditional' Forms of School Choice Teach Us About School Choice Reforms? *Federal Reserve Bank of New York Economic Policy Review* 4(1).

Hoxby, C. M. 2003. *School choice and school productivity (or, could school choice be a tide that lifts all boats?)*, in C. Hoxby (ed.), *The Economics of School Choice, University of Chicago Press*, Chicago.

Hoxby, C. M., and J. Rockoff. 2005. "The Impact of Charter Schools on Student Achievement," *NBER Working Paper.*

Jacobson, L., LaLonde, R., and D. Sullivan. 2005. "Estimating the returns to community college schooling for displaced workers. *Journal of Econometrics,* 125: 271-304.

Jacobson, L., and Mokher, C. 2009. "Pathways to boosting the earnings of low-income students by increasing their educational attainment." Washington, DC: Hudson Institute Center for Employment Policy.

Jennings, J. L., D. Deming, C. Jencks, M. Lopuch and B. E. Schueler. 2015. "Do Differences in School Quality Matter More Than We Thought? New Evidence on Educational Opportunity in the Twenty-first Century." *Sociology of Education,* 88(1): 56-82.

Jepson, C., Troske, K., and Coomes, P. 2014. "The Labor-Market Returns to Community College Degrees, Diplomas, and Certificates." *Journal of Labor Economics.* 32(1): 95-121.

Jepsen, C. and S. Rivkin. 2002. "What is the Tradeoff Between Smaller Classes and Teacher Quality?" *NBER Working Paper No. 9205.*

Kalogrides, D. 2007. "The effects of community college attendance on baccalaureate attainment (Working Paper)." Davis: University of California, Davis.

Kane, T. J., and C. E. Rouse. 1995. "Labor-market returns to two- and four-year colleges." *American Economic Review,* 85: 600-614.

Kane, T. J., and C. E. Rouse. 1999. "The Community College: Educating Students at the Margin between College and Work." *Journal of Economic Perspectives*, 13(1): 63-84.

Kerckhoff, A. C., and L. Bell. 1998. "Hidden capital: Vocational credentials and attainment in the United States." *Sociology of Education,* 71: 152174.

Kienzl, G.S., Alfonso M., and T. Melguizo. 2007. "The Effect of Local Labor Market Conditions in the 1990s on the Likelihood of Community College Students' Persistence and Attainment." *Research in Higher Education,* 48(7): 751-774.

Kim, D. 2004. "The Effect of Loans on Students' Degree Attainment: Differences by Student and Institutional Characteristics." *Harvard Educational Review.* 77(1): 64-100.

Kinsler, J. and R. Pavan. 2011. "Family Income and Higher Education Choices: The Importance of Accounting for College Quality." *Journal of Human Capital,* 5 (4): 453-477.

Klein, R. and F. Vella. 2010. "Estimating a class of triangular simultaneous equations models without exclusion restrictions," *Journal of Econometrics.* 154(2): 154-164.

Klein, R., F. Vella and L. Farré. 2009. "Does increasing parents' schooling raise the schooling of the next generation? Evidence based on conditional second moments," *Ivie Working Paper.*

Konstantopoulous, S. 2006. "Trends of School Effects on Student Achievement: Evidence from NLS:72, HSB:82, and NELS:92." *Teachers College Record.* 108: 2550-2581.

Krueger, A. B. 1997. "Experimental Estimates of Education Production Functions." *NBER Working Paper No. 6051.*

Lanaan, F.S. 2007. "Studying Transfer Students: Part II: Dimensions of transfer students' adjustment." *Community College Journal of Research and Practice,* 31(1): 37-59.

Lee, V. E., Mackie-Lewis, C., and Marks, H. M. 1993. "Persistence to the baccalaureate degree for students who transfer from community college." *American Journal of Education.* 102(1): 80-114.

Leigh, D.E. and A.M. Gill. 2003. "Do community colleges really divert students from earning bachelor's degrees?" *Economics of Education Review.* 22(1): 23-30.

Leigh, D. E., and A.M. Gill. 1997. "Labor market returns to community colleges: Evidence for returning adults." *Journal of Human Resources,* 32: 334353.

Leuven, E., and Sianesi, B. 2004. "PSMATCH2: Stata module to perform full mahalanobis and propensity score matching, common support graphing and covariate imbalance test?" Retrieved from *http://ideas.repec.Org/c/boc/bocode/s432001.html* April 24, 2016.

Levin, H.M. 2005. "The Social Costs of Inadequate Education." *Teachers College Symposium on Educational Equity.* http://www.schoolfunding.info/resource_center/research/2005symposium.pdf.

Liu, V.Y.T., Belfield, C.R., and M.J. Trimble. 2015. "The Medium-Term Labor Market Returns to Community College Awards: Evidence From North Carolina." *Economics of Education Review,* 44: 42-55.

Lochner, L., and E. Moretti. 2004. "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports." *American Economic Review,* 94(1): 155-189.

Long, B. T. and M. Kurlaender. 2009. "Do Community Colleges provide a Viable Pathway to a Baccalaureate Degree?" *Educational Evaluation and Policy Analysis* 31(1): 30-53.

Lubienski, C. 2005. "Public Schools in Marketized Environments: Shifting Incentives and Unintended Consequences of Competition-Based Educational Reforms." *American Journal of Education* 11(4): 464-486.

Massachusetts Budget and Policy Center. 2010. *Demystifying the Chapter 70 Formula: How the Massachusetts Education Funding.* http://www.massbudget.org/report_window.php?loc=Facts_10_22_10.html. May 8, 2016.

Massachusetts Department of Education. 2015. *Chapter 70 Program.* http://www.doe.mass.edu/finance/chapter70/. May 8, 2016.

Massachusetts Department of Education. 2013a. *Understanding District Aid for Commonwealth Charter School Tuition.* http://www.doe.mass.edu/charter/finance/tuition/Reimbursements.html. May 8, 2016.

Massachusetts Department of Education. 2013b. *2013-2014 Charter School Pre-Enrollment.* http://www.doe.mass.edu/charter/enrollment/FY14.html. May 8, 2016.

Massachusetts Department of Education 2013c. *Report on Charter School Waitlists.* http://www.doe.mass.edu/charter/enrollment/fy2014Waitlist.pdf. May 8, 2016.

MassGIS Data. 2015. *Public School Districts.* http://www.mass.gov/anf/research-and-tech/it-serv-and-support/application-serv/office-of-geographic-information-massgis/datalayers/schooldistricts.html. May 25, 2016.

MCPSA. Massachusetts Charter Public School Association. http://www.masscharterschools.org/. May 8, 2016.

Marcotte, D. E., T. Bailey, C. Borkoski, and G. S. Kienzl. 2005. "The Returns of a Community College Education: Evidence from the National Education Longitudinal Survey." *Educational Eval-*

uation and Policy Analysis, 27(2): 157-175.

Massachusetts Charter Public School Association `http://www.masscharterschools.org`. February 4, 2013.

Mead, S. and A. J. Rotherham. 2007. *A Sum Greater Than the Parts: What States Can Teach Each Other About Charter Schooling.* Washington, D.C: Education Sector.

Medsker, L. L. 1960. *The junior college: Progress and prospect.* New York: McGraw-Hill.

Meier, K. J., R. D. Wrinkle, and J.L. Polinard. 2000. "Bureaucracy and Organizational Performance: Causality Arguments about Public Schools." *American Journal of Political Science,* 44: 590-602.

Melguizo, T., Kienzl, G. S., and Alfonso, M. 2011. "Comparing the Educational Attainment of Community College Transfer Students and four-year College Rising Juniors Using Propensity Score Methods." *The Journal of Higher Education.* 82(3): 265-291.

Merseth, K. K.; with K. Cooper, J. Roberts, M.C. Tieken, J. Valant, C. Wynne, and forward by S. P. Reville. 2009. *Inside Urban Charter Schools Promising Practices and Strategies in Five High-Performing Schools.* Harvard Education Press.

Mincer, J. 1974. "Schooling, Experience and Earnings." *New York: National Bureau of Economic Research*

Mintrom, M. 2000. "Leveraging Local Innovation: The Case of Michigan's Charter Schools," East Lansing: Michigan State University.

Moffitt, R. 1992. "Incentive effects of the U.S. welfare system: a review" *Journal of Economic Literature.* 30(1): 1-61.

Monk-Turner, E. 1994. "Economic returns to community and four-year college education." *Journal of Socio-Economics,* 23: 441447.

National Alliance for Public Charter Schools. *Measuring up to the Model: A Tool for Comparing State Charter School Laws.* `http://www.publiccharters.org/law/ViewState.aspx?state=MA`. September 11, 2013.

National Alliance for Public Charter Schools. 2013. *Public Charter Schools Dashboard: A Comprehensive Data Resource from the National Alliance for Public Charter Schools.* `http://dashboard.publiccharters.org/dashboard/students/page/age/year/2013` September, 11 2013

National Alliance for Public Charter Schools. 2012. *Public Charter Schools Dashboard: A Comprehensive Data Resource from the National Alliance for Public Charter Schools.* `http://dashboard.publiccharters.org/dashboard/students/page/race/year/2010` August 24, 2014.

National Education Association. *Rankings of the States 2011 and Estimates of School Statistics 2012.* `http://www.nea.org/assets/docs/NEA_Rankings_And_Estimates_FINAL_20120209.pdf`. October 4, 2013.

NCES. 2014. Table 330.10: "Average undergraduate tuition and fees and room and board rates charged for full-time students in degree-granting postsecondary institutions, by level and control of institution: 1963-64 through 2013-14" *NCES* on the Internet at `http://nces.ed.gov/programs/`

digest/d14/tables/dt14_330.10.asp (visited November 11, 2015)

NCES. 2014. Table 331.30: "Average amount of grant and scholarship aid and average net price for first-time, full-time students receiving Title IV aid, and percentage distribution of students, by control and level of institution and income level: 2009-10, 2010-11, and 2011-12" *NCES* on the Internet at https://nces.ed.gov/programs/digest/d13/tables/dt13_331.30.asp (visited November 11, 2015)

NCES. 2014. Table 326.10: "Graduation rates of first-time, full-time bachelor's degree-seeking students at 4-year postsecondary institutions, by race/ethnicity, time to completion, sex, and control of institution: Selected cohort entry years, 1996 through 2006." *NCES* on the Internet at http://nces.ed.gov/programs/digest/d13/tables/dt13_326.10.asp (visited on November 11, 2015.)

NCES. 2013a. Table 302.30: "Percentage of recent high school completers enrolled in 2-year and 4-year colleges, by income level: 1975 through 2012" *NCES.* on the Internet at http://nces.ed.gov/programs/digest/d13/tables/dt13_302.30.asp (visited on November 11, 2015.)

NCES. 2013b. Table 321.20. "Associate's degrees conferred by postsecondary institutions, by race/ethnicity and sex of student: Selected years, 1976-77 through 2012-13." *NCES on the Internet at http://nces.ed.gov/programs/digest/d14/tables/dt14_318.10.asp* (visited on February 16, 2016).

NCES. 2013c. Table 320.20. "Certificates below the Associate's degree level conferred by postsecondary institutions, by race/ethnicity and sex of student: 1998-99 through 2012-13." *NCES on the Internet at http://nces.ed.gov/programs/digest/d14/tables/dt14_321.20.asp* (visited on February 16, 2016).

NCES. 2013d. Table 318.10. "Degrees conferred by postsecondary institutions, by level of degree and sex of student: Selected years, 1869-70 through 2024-25." *NCES on the Internet at http://nces.ed.gov/programs/digest/d14/tables/dt14_318.10.asp* (visited on February 16, 2016).

NCES. 2012a. "Education Longitudinal Study of 2002 Base-year to Third Follow-up Data File Documentation." http://nces.ed.gov/pubs2014/2014364.pdf. May 9, 2016.

NCES. 2012b. Table 306.20: "Total fall enrollment in degree-granting postsecondary institutions, by level and control of institution and race/ethnicity of student: Selected years, 1976 through 2012." *NCES* on the Internet at http://nces.ed.gov/programs/digest/d13/tables/dt13_306.20.asp (visited on November 11, 2015.)

NCES. 2012c. Table 334.10: "Expenditures of public degree-granting postsecondary institutions, by purpose of expenditure and level of institution: 2005-06 through 2011-12" *NCES* on the Internet at http://nces.ed.gov/programs/digest/d13/tables/dt13_334.10.asp (visited on November 11, 2015.)

NCES. 2012d. Table 502.30. "Median annual earnings of full-time year-round workers 25 to 34 years old and full-time year-round workers as a percentage of the labor force, by sex, race/ethnicity, and educational attainment: Selected years, 1995 through 2012." *NCES on the Internet at http://nces.ed.gov/programs/digest/d13/tables/dt13_502.30.asp* (visited on February 16, 2016).

NCES. 2011. Table 326.40: "Percentage distribution of first-time postsecondary students starting at 2- and 4-year institutions during the 2003-04 academic year, by highest degree attained, enrollment status, and selected characteristics: Spring 2009." *NCES* on the Internet at http:

//nces.ed.gov/programs/digest/d13/tables/dt13_326.40.asp (visited on November 11, 2015.)

NCES. 2010. Table 9. "Percentage distribution of private school students, by racial/ethnic background and selected characteristics: United States, 200910" *NCES* on the Internet at http://nces.ed.gov/surveys/pss/tables/table_2009_09.asp (visited on May 8, 2016).

Newey, W. K. 1990. "Efficient Instrumental Variables Estimation of Nonlinear Models. *Econometrica.* 58 (4): 809837.

Nunez, A.M. and D. Kim. 2012. "Building a Multicontextual Model of Latino College Enrollment: Student, School, and State-Level Effects." *The Review of Higher Education.* 35(2): 237-263.

Nunley, C. R., and D. W. Breneman. 1988. "Defining and Measuring Quality in Community College Education." In *Colleges of Choice,* edited by Judith Eaton, pp. 62-92. New York: Macmillan.

Oreopoulos, P. and U. Petronijevic. 2013. "Making College Worth It: A Review of the Returns to Higher Education." *The Future of Children,* 23(1): 41-65.

Perna, L. W. 2007. "The sources of racial-ethnic group differences in college enrollment: A critical examination." *New directions for institutional research.* 2007(133): 51-66.

Perna, L.W. 2000. "Differences in the Decision to Attend College among African Americans, Hispanics, and Whites." *The Journal of Higher Education.* 71(2): 117-141.

Powers, J. M. 2008. *Charter Schools.* Encyclopedia of the Social and Cultural Foundations of Education. SAGE Publications. December 5, 2011.

Psacharopoulos, G. 2006. "The Value of Investment in Education: Theory, Evidence, and Policy." *Journal of Education Finance.* 32(2): 113-136.

Reynolds, L. 2012. "Where to attend? Estimates of the effects of beginning at a two-year college." *Economics of Education,* 31(4): 345-362.

Reynolds, C. L., and Desjardins, S. L. 2009. "The use of matching methods in higher education: Answering whether attendance at a two-year institution results in differences in educational attainment." In *J. C. Smart (Ed.), Higher Education: Handbook of Theory and Practice,* 12: 47-104. New York: Springer.

Rivkin, S. G., E. A. Hanushek, and J. F. Kain. 2005. "Teachers, Schools, and Academic Achievement." *Econometrica,* 73(2): 417-458.

Roderick, M., Coca, V., and J. Nagaoka. 2011. "Potholes on the Road to College: High School Effects in Shaping Urban Students' Participation in College Application, Four–Year College Enrollment, and College Match," *Sociology of Education.* 84(3): 178-211.

Rosenbaum, R R., and Rubin, D. B. 1983. "The central role of the propensity score in observational studies for causal effects." *Biometrika,* 70: 41-55.

Rothstein, J. 2005. "Does Competition Among Public School Benefit Students and Taxpayers? A Comment on Hoxby (2000)." NBER Working Paper 11215.

Rouse, C. E. 1995. "Democratization or Diversion? The Effect of Community Colleges on Educational Attainment." *Journal of Business and Economic Statistics,* 13(2): 217-224.

Sacerdote, B. 2001. "Peer Effects with Random Assignment: Results for Dartmouth Roommates." *The Quarterly Journal of Economics,* 116 (2): 681-704.

Sandy, J., Gonzalez, A., and M.J. Hilmer. 2006. "Alternative paths to college completion: Effect of attending a 2-year school on the probability of completing a 4-year degree." *Economics of Education Review.* 25(5): 463-471.

Schneider B., Martinez S., and A. Owens. 2006. *Barriers to Educational Opportunities for Hispanics in the U.S.* In: Tienda M Hispanics and the Future of America. Washington, DC: National Academies Press; pp. 179-227.

Shanker, A.; President, American Federation of Teachers. 1988. *National Press Club Speech.* https://www.reuther.wayne.edu/files/64.43.pdf. Washington, D.C. March 31, 1988.

Sirin, S. R. 2005. "Socioeconomic Status and Academic Achievement: A Meta-Analytic Review of Research." *Review of Educational Research,* 75(3): 417-453.

Smith, J. and K. Stange. 2015. "A New Measure of College Quality to Study the Effects of College Sector and Peers on Degree Attainment." *NBER Working Paper 21605.*

Smith, J. and Tood, P. 2005. "Does matching address Lalonde's critique of nonexperimental estimators?" *Journal of Econometrics,* 125(12): 305-353.

Sugarman, S. D. 2002. *Charter School Funding Issues.* Education Policy Analysis Archives 10(34). 2002.

Thomas B. Fordham Institute. 2005. *Charter School Funding: Inequity's Next Frontier.*

Townsend, B. 1995. "Community College transfer students: A case of survival." *The Review of Higher Education.* 18(2): 175-193.

Trostel, P. 2010. "The Fiscal Impacts of College Attainment." *Research of Higher Education.* 51(3): 220-247.

U.S. Census Table S1501. 2014. "Educational Attainment: 2010-2014 American Community Survey 5-Year Estimates." *U.S. Census on the Internet at http://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_14_5YR_S1501&src=pt* (visited on February 16, 2016).

U.S. Census 2014. "Educational Attainment in the United States: 2014 - Detailed Tables" *U.S. Census* on the Internet at http://www.census.gov/hhes/socdemo/education/data/cps/2014/tables.html (visited on November 11, 2015.)

U.S. Census. Population and Housing Occupancy Status: 2010 Census Redistricting Data (Public Law 94-171) Summary File. http://factfinder2.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=DEC_10_PL_GCTPL2.ST16&prodType=table. 3/26/14

U.S. Census Table GCT-PL2. 2010. "Population and Housing Occupancy Status: 2010 - State - County Subdivision; 2010 Census Redistricting Data (Public Law 94-171) Summary File." http://factfinder2.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=DEC_10_PL_GCTPL2.ST16&prodType=table. May 8, 2016.

United States Department of Education. 1983. "A Nation at Risk: The Imperative for Educational Reform." *The National Commission on Excellence in Education* http://datacenter.spps.org/

uploads/SOTW_A_Nation_at_Risk_1983.pdf. April 1983. October 4, 2013.

United States Education Dashboard. 2015. “The President’s 2020 College Attainment Goal.” http://dashboard.ed.gov/dashboard.aspx?i=d4&id=25&wt=40. May 25, 2016.

Velez, W. 1985. “Finishing College: The Effects of College Type.” *Sociology of Education.* 58(3): 191-200.

Wells, R. S. and C. M. Lynch. 2012. “Delayed College Entry and the Socioeconomic Gap: Examining the Roles of Student Plans, Family Income, Parental Education, and Parental Occupation.” *The Journal of Higher Education,* 83(5): 671-697.

Whitaker, D. G., and Pascarella, E. T. 1994. “Two-year college attendance and socioeconomic attainment: Some additional evidence.” *Journal of Higher Education.* 65(2): 194-210.

Wood, J. L. and R. C. Williams. 2013. “Persistence Factors for Black Males in the Community College: An Examination of Background, Academic, Social, and Environmental Variables.” *Spectrum.* 1(2): 1-28.

Wooldbridge, J. 2002. *Econometric Analysis of Cross Section and Panel Data.* MIT Press, Cambridge, MA.

Zimmer, R. and R. Buddin. 2009.“Is Charter School Competition in California Improving the Performance of Traditional Public Schools?” *Public Administration Review,* 69(5): 831-845.