Essays on Human Capital Investment

BY

THOMAS WILLIAM WALSTRUM
B.A., Wheaton College, 2005
M.A., University of Illinois at Chicago, 2015

THESIS

Submitted as partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Economics
in the Graduate College of the
University of Illinois at Chicago, 2015

Chicago, Illinois

Defense Committee:

Ben Ost, Chair and Advisor
Robert Kaestner
Joseph Persky
Steven Rivkin
William Testa, Federal Reserve Bank of Chicago

This thesis is dedicated to my wife Marisa, in gratitude for all the love and support that contributed to its completion.

ACKNOWLEDGEMENTS

I would like to thank my thesis committee, Ben Ost, Robert Kaestner, Joseph Persky, Steven Rivkin, and William Testa, for their readily available and thoughtful guidance. I am particularly grateful to Ben Ost, the committee chair and advisor, for all the time he spent helping to hone my thinking, results, and writing. I would also like to thank David Merriman for his role in mentoring me throughout the PhD program and Erin Toussaint for very helpful administrative assistance.

The research in Chapter 1 would not have been possible without the North Carolina Research Data Center, which made the data available and has always responded quickly to any requests. The research in Chapter 2 was made much easier thanks to the Minnesota Population Center, which has done excellent work in making the US Census and CPS microdata very accessible.

TABLE OF CONTENTS

CHAPTER PAGE

1. THE ROLE OF CAREER AND TECHNICAL EDUCATION IN HIGH SCHOOL GRADUATION 1
   - 1.1 Introduction 1
   - 1.2 Previous Research 4
     - 1.2.1 United States 5
     - 1.2.2 Europe 9
     - 1.2.3 Across countries 10
   - 1.3 Theory 11
   - 1.4 Evidence 14
     - 1.4.1 Data Description 14
     - 1.4.2 Summary Statistics 18
     - 1.4.3 Empirical Approach 20
     - 1.4.4 Main Results 23
     - 1.4.5 Robustness and Specification Checks 26
     - 1.4.6 Is CTE Supply Exogenous? 28
     - 1.4.7 Heterogeneity 31
   - 1.5 Discussion 31
   - 1.6 Cited Literature 34
   - 1.7 Figures and Tables 36
     - 1.7.1 Figures 36
     - 1.7.2 Tables 37

2. THE OPPORTUNITY COST OF A COLLEGE EDUCATION: HOW SHOCKS TO LOCAL LABOR DEMAND AFFECT ENROLLMENT AND THE GENDER GAP 52
   - 2.1 Introduction 52
   - 2.2 Theory and evidence in the literature 55
     - 2.2.1 On the cyclicality of enrollment 55
     - 2.2.2 On gender differences in college enrollment 58
   - 2.3 A model of the college enrollment decision 62
   - 2.4 Data and Empirical Strategy 67
     - 2.4.1 Data 67
     - 2.4.2 Empirical Strategy 71
   - 2.5 Results 73
     - 2.5.1 Main Results 73
     - 2.5.2 Accounting for inter-state migration 76
     - 2.5.3 Gender Heterogeneity 78
   - 2.6 Discussion 80
   - 2.7 Cited Literature 82
   - 2.8 Figures and Tables 85
     - 2.8.1 Figures 85
     - 2.8.2 Tables 89

3. VITA 99

LIST OF TABLES

TABLE PAGE

I. SUMMARY STATISTICS.......... 37

II. VARIABLE MEANS BY QUARTILE OF PROPORTION OF SCHOOL'S COURSES THAT ARE CTE.......... 38

III. STUDENT-LEVEL VARIABLE MEANS BY NUMBER OF CTE COURSES TAKEN IN NINTH GRADE.......... 39

IV. MEAN PROPORTION OF STUDENT'S CREDITS IN 9TH GRADE BY COURSE TYPE. 40

V. VARIABLE MEANS BY CTE CATEGORY.......... 41

VI. SUMMARY STATISTICS OF COHORT-LEVEL VARIABLES.......... 42

VII. LINEAR PROBABILITY MODEL OF HIGH SCHOOL GRADUATION IN FOUR YEARS - RELATIONSHIP WITH PROPORTION OF STUDENT'S TRANSCRIPT CTE IN 9TH GRADE.......... 43

VIII. LINEAR PROBABILITY MODEL OF HIGH SCHOOL GRADUATION IN FOUR YEARS - REDUCED FORM RELATIONSHIP WITH PROPORTION OF SCHOOL'S COURSES CTE.......... 44

IX. LINEAR PROBABILITY MODEL OF PROPORTION OF STUDENT'S TRANSCRIPT CTE IN 9TH GRADE - RELATIONSHIP WITH PROPORTION OF SCHOOL'S COURSES CTE.......... 45

X. LINEAR PROBABILITY MODEL OF HIGH SCHOOL GRADUATION IN FOUR YEARS - REDUCED FORM RELATIONSHIP WITH PROPORTION OF SCHOOL'S COURSES IN CTE CATEGORIES.......... 46

XI. ROBUSTNESS AND SPECIFICATION CHECKS.......... 47

XII. WLS MODEL OF PROPORTION OF SCHOOL'S COURSES CTE - CHECK FOR EXOGENEITY.......... 48

XIII. WLS MODEL OF PREDICTED COHORT GRADUATION RATE - CHECK FOR THE EXOGENEITY OF PROPORTION OF SCHOOL'S COURSES CTE.......... 49

XIV. LINEAR PROBABILITY MODEL OF HIGH SCHOOL GRADUATION IN FOUR YEARS - HETEROGENEITY IN THE REDUCED FORM RELATIONSHIP WITH PROPORTION OF SCHOOL'S COURSES CTE BY STUDENT-LEVEL CHARACTERISTICS.......... 50

XV. LINEAR PROBABILITY MODEL OF HIGH SCHOOL GRADUATION IN FOUR YEARS - HETEROGENEITY IN THE RELATIONSHIP WITH PROPORTION OF SCHOOL'S COURSES CTE BY SCHOOL-LEVEL CHARACTERISTICS.......... 51

XVI. SUMMARY STATISTICS.......... 89

XVII. SUMMARY STATISTICS FOR LOCAL LABOR DEMAND INDEX COMPONENTS.......... 90

LIST OF TABLES (continued)

<u>TABLE</u> <u>PAGE</u>

XVIII. CORRELATIONS BETWEEN LOCAL LABOR DEMAND PROXIES 91

XIX. LINEAR PROBABILITY MODEL OF COLLEGE ENROLLMENT WITH THE HIGH SCHOOL UNEMPLOYMENT RATE AS THE PROXY FOR LOCAL LABOR DEMAND 92

XX. LINEAR PROBABILITY MODEL OF COLLEGE ENROLLMENT WITH THE HIGH SCHOOL WAGE AS THE PROXY FOR LOCAL LABOR DEMAND 93

XXI. LINEAR PROBABILITY MODEL OF COLLEGE ENROLLMENT WITH THE HIGH SCHOOL LOW EDUCATION OCCUPATION WAGE AS THE PROXY FOR LOCAL LABOR DEMAND 94

XXII. LINEAR PROBABILITY MODEL OF COLLEGE ENROLLMENT WITH THE HIGH SCHOOL LOCAL LABOR DEMAND INDEX AS THE PROXY FOR LOCAL LABOR DEMAND 95

XXIII. MIGRATION-ADJUSTED LINEAR PROBABILITY MODEL OF ENROLLMENT WITH THE HIGH SCHOOL LOCAL LABOR DEMAND INDEX AS THE PROXY FOR LOCAL LABOR DEMAND 96

XXIV. LINEAR PROBABILITY MODEL OF COLLEGE ENROLLMENT – GENDER HETEROGENEITY BY DEMAND PROXY 97

XXV. LINEAR PROBABILITY MODEL OF COLLEGE ENROLLMENT – GENDER HETEROGENEITY BY GENDER-SPECIFIC DEMAND PROXY 98

LIST OF FIGURES

1. DISTRIBUTION OF THE PROPORTION OF A SCHOOL'S COURSES THAT ARE CTE 36

2. ENROLLMENT RATE FOR 18 TO 24 YEAR OLDS 85

3. THE CYCLICALITY OF THE ENROLLMENT RATE 86

4. ENROLLMENT RATES BY SEX FOR 18 TO 24 YEAR OLDS 87

5. ENROLLMENT GAP FOR AGES 18 TO 24 88

LIST OF ABBREVIATIONS

CTE	Career and Technical Education
ELS	Education Longitudinal Study
FAFSA	Free Application for Federal Student Aid
GE	General Education
NCERDC	North Carolina Research Data Center
NELS	National Education Longitudinal Study
NLS	National Longitudinal Study
NLSY	National Longitudinal Survey of Youth

SUMMARY

Given the large returns to education for the labor market and for other indicators of wellbeing, understanding the factors that determine educational attainment are an important concern of policymakers. In this thesis, I explore the contribution of two such factors: 1) the availability of career and technical education (CTE) in high school and 2) the strength of the local labor market in college.

In Chapter 1 of the thesis, I test the hypothesis that offering CTE courses encourages students to complete high school. I use data on public high school students from the North Carolina Education Research Data Center that allow me exploit within-school variation in the availability of CTE courses to obtain the causal effect of CTE availability on graduation. My study finds that a 10 percentage point increase in the proportion of a school's course offerings that are CTE (one standard deviation is 8 percentage points) raises the probability a student graduates by 0.63 percentage points. The average dropout rate in the data is about 10 percent, so such an increase reduces the dropout population by about 6 percent. I also find notable differences in the strength of the graduation response across student characteristics. The effect of increasing CTE supply is particularly strong for Hispanics, those with above average 8th grade math and reading test scores, and those with below average time spent on homework in 8th grade. The effect is also stronger when the local economy is stronger, suggesting that CTE helps keep people in school who would otherwise drop out and work.

In Chapter 2 of the thesis, I construct a local labor demand index that provides new causal evidence that college enrollment goes up in response to negative local labor demand shocks and down in response to positive ones. Unlike other proxies of local labor demand, such as the unemployment rate or wages, the index is exogenous to local labor supply shocks that may be correlated with other factors that determine the decision to enroll. In the end, I find little difference between estimates using demand proxies vulnerable to labor supply shocks and the demand index I construct. I also look for gender heterogeneity in the enrollment response to demand shocks, and find that men are more sensitive than women are. This indicates that the positive gap between women and men in enrollment grows during booms and shrinks during busts. I explore the possibility that the gender heterogeneity is the result of gender differences in labor demand, with mixed results.

1. THE ROLE OF CAREER AND TECHNICAL EDUCATION IN HIGH SCHOOL GRADUATION

1.1 Introduction

As the return to skill in the labor market has grown over the past few decades, the importance of good education policy to support skill formation has grown with it. One policy question that remains largely unresolved is about the relative merits of a general education (GE) curriculum versus a career and technical education (CTE) curriculum. GE courses teach broad skills useful in almost any occupation, whereas CTE courses teach detailed skills useful primarily in specific occupations. GE courses are designed to create workers who are adaptable to changing labor market conditions, while CTE courses are designed to create workers with specialized skills demanded by the current labor market. Thus GE is purported to have long run labor market benefits, while CTE is purported to have short run labor market benefits.

CTE may also encourage those on the verge of dropping out to stay in school. GE courses are typically more bookish and abstract, whereas CTE courses are typically more practical. For those with low skills or motivation who have a particularly difficult time in GE courses, the practicality of CTE courses may provide a more appealing way to stay in school and obtain marketable skills. Thus even though CTE may teach less flexible skills, it may still be better than the other alternative to GE: dropping out altogether.

This study tests the hypothesis that offering CTE courses encourages students to complete high school. Unlike previous US-based studies, my research design allows me to credibly address the empirical concern that schools tailor their course offerings to the long run characteristics of their student body. I use data from the North Carolina Education Research Data Center (NCERDC) to calculate the proportion of course offerings that are CTE in each public high school in North Carolina from 2005 to 2010. If this measure of CTE "supply" is uncorrelated with all the unobservable factors that determine high school graduation then my estimates of the effect of offering CTE courses on graduation are causal. Of course, it is reasonable to expect that there are unobservable differences across schools that affect both the supply of CTE and high school graduation. For example, social norms that encourage students to graduate and go to college may be much stronger in rich urban schools than in poor rural schools, leading administrators of poor rural schools to offer more CTE courses. To address this concern, I include school fixed effects in the empirical analysis so that I identify the effect of CTE supply on graduation using within-school variation. Specifically, I compare the graduation rates of different cohorts in the same school based on the level of CTE course offerings (across all grades) when they are in 9th grade (before most students are 16 and eligible to drop out). I also include cohort fixed effects to control for changes in statewide factors that could bias the estimates, such as macroeconomic conditions.

What school and cohort fixed effects cannot address is the possibility that administrators dynamically change a school's CTE supply in response to changes in incoming cohort quality. This is unlikely, however, because administrators probably have limited information about cohort quality until after the students enter their school and changing course offerings is a lengthy process. That said, I do find some evidence that some cohort characteristics are related to course offerings, though I also find that these cohort quality differences are small and unlikely to explain away the results. Specifically, I show that there is no difference in predicted graduation rates based on a school's CTE offerings.

My main results indicate that a 10 percentage point increase in the proportion of a school's course offerings that are CTE (one standard deviation is 8 percentage points) raises the probability a student graduates by 0.63 percentage points. The average graduation rate in the data is about 90 percent, so such an increase would reduce the dropout rate by a little over 6 percent. I also estimate the effect of offering particular CTE tracks on graduation and find that increasing the availability of most tracks has a positive effect (the agriculture, business and finance, and marketing tracks have the largest effects), but that offering technology and engineering courses actually has a negative effect.

The NCERDC provides data on a variety of student and school characteristics, which allow me to explore factors underlying the decision to take up CTE in 9th grade and the types of schools that offer CTE. Somewhat surprisingly, I find that the students most likely to take CTE courses in 9th grade are those with above average 8th grade math and reading test scores. However, conditional on test scores, CTE students were more likely to report spending little time in 8th grade doing homework. Thus the typical CTE student is someone who has moderate cognitive skills, but relatively less motivation for academic learning. Schools that offer higher proportions of CTE courses tend to be poorer, whiter, more rural, and have lower math and reading test scores.

I also look for heterogeneity in the effect of CTE on graduation and find notable differences in the strength of the response across student characteristics. The effect of increasing CTE supply is particularly strong for Hispanics and those with above average 8th grade math test scores. The effect is also stronger when the local economy is stronger, suggesting that CTE courses help keep people in school who would otherwise drop out and work. The effect is notably weaker for those who receive free lunch.

1.2 Previous Research

Studies of the impacts of CTE go back at least two decades and cover a variety of curriculums across the US and Europe. Most studies use student-level variation in CTE enrollment to examine outcomes in the labor market, such as wages and employment; human capital investment decisions, such as high school graduation and college enrollment; and standardized test scores. Many studies choose to include controls for enrollment in GE courses in their empirical specifications, so that the counterfactual is enrollment in non-GE, non-CTE courses such as study hall or physical education. This study is unique in that it focuses on school-level variation in CTE supply. I also do not include controls for GE course offerings, so that the counterfactual is the supply of all other courses, including GE.

Most studies focus on a single country, though a few studies use cross-country variation. European countries traditionally offer CTE earlier in students' academic careers and have more extensive CTE curriculums than the US does. Because of the large differences in CTE curriculums between Europe and the US, I review their literatures separately, and then discuss the few studies that explore differences across countries.

1.2.1 United States

Studies in the US using nationally representative data rely on two primary sources for information about CTE enrollment: 1) the NLSY:79 and NLSY:97, and 2) a series of longitudinal surveys undertaken by the US Department of Education, the NLS:72, High School & Beyond, NELS:88, and ELS:2002. These surveys provide very detailed information about a relatively small number of individuals, with sample sizes ranging from about 5,000 to 25,000. While the extent of individual-level information in these surveys provides useful controls for researchers studying CTE, the small sample size makes it difficult for researchers to obtain a plausible source of exogenous variation in CTE enrollment. This is important because, as all studies of CTE acknowledge, CTE enrollment is very likely to be endogenous. That said, studies that use datasets with very detailed information may be able to control for the most important factors that determine both CTE enrollment and the outcomes of interest, and at the very least provide baseline estimates.

Kang and Bishop (1989) conduct one of the first studies of CTE using data from the High School & Beyond survey. The study examines the relationship between labor market outcomes and the number of CTE courses taken by high school graduates who did not attend college. Importantly, the authors condition on the number of GE courses taken, so that counterfactual is taking non-GE, non-CTE courses. The study generally finds a positive relationship between CTE enrollment and wages, earnings, and employment. For example, the study finds that for men, 21 months after graduation, an additional CTE course is associated with an increase in wages of about 1.6 percent, an increase in earnings of about 3.7 percent, and an increase months employed of about 1.3 percent. It also finds that CTE and GE courses are complements, that is, that the return to CTE courses is greater for students who also take a large number of GE courses. Given the population being studied and the decision to include GE courses as controls, it is possible that the estimates are biased upward because higher-skilled students may be more likely to take CTE courses over non-CTE, non-GE courses.

There are a number of similar studies to Kang and Bishop (1989). Mane (1999) does the same analysis as Kang and Bishop (1989), but with the NELS:88 cohort, and obtains very similar results. Bishop and Mane (2004) also use the NELS:88 cohort, but they are able to look at longer-run outcomes. Again, the results are quite similar.

Acknowledging the concerns about endogeneity in CTE enrollment, Meer (2007) econometrically models the decision using a multinomial logit selection model and looks at labor market outcomes using the NELS:88. To make the model tractable, he divides course offerings into four tracks, college-prep, non-college-prep, technical, and business. Any person who does not complete a track (graduation requirements may be greater) is dropped from the sample. To identify the model, Meer treats a number of variables as exogenous: whether CTE classes are offered at a school, whether the school has full time CTE teachers, whether the school offers AP courses, whether more than 25% of the school's previous cohort went to a 2-year college, whether more than 25% of the school's previous cohort went to a 4-year college, and an individual's SES quartile (a composite variable). Meer (2007) finds that students who select into the technical and college-prep tracks receive higher returns than they would have if they had selected a different track. He also finds that students in the non-college-prep track would benefit from taking one of the CTE tracks and students in the business track would benefit from taking either the technical or college-prep track.

Studies using the NLSY are similar in flavor to those using the longitudinal datasets from the US Department of Education and generally obtain similar results. Neumark and Rothstein (2006) examine whether school-to-work programs (for example, cooperative education, CTE classes, and mentoring) predict college attendance and employment. Their econometric model conditions on observables, including school fixed effects, and they do not restrict the sample to those who graduated from high school. They find that most programs are positively associated with college attendance, while only cooperative education and internships are positively associated with employment. Cellini (2006) uses the NLSY:97 to assess the effectiveness of one particular school-to-work program, known as tech-prep, where students can obtain community college credit for taking technical courses in high school. While Neumark and Rothstein (2006) find that tech-prep programs do not encourage college enrollment, Cellini (2006) finds that after adding family fixed effects as controls, they have little effect on college enrollment, but raise the probability of completing 12th grade by about 9 percentage points. She also distinguishes between 2-year and 4-year college attendance and finds that tech-prep programs encourage 2-year college enrollment at the expense of 4-year college enrollment.

Kreisman and Stange (2015) also use the NLSY:97 to study the effects of CTE enrollment and at the time of their study, the sample had aged enough for them to examine wage outcomes. Like Kang and Bishop (1989), they do not group courses into tracks, but rather estimate the conditional-on-observables association between incremental increases in CTE enrollment and wages, with GE courses included as controls. They find a small wage return to taking advanced CTE courses, and find that while CTE enrollment reduces 4-year college enrollment, it doesn't reduce 4-year graduation rates.

Betts, Zau, McAdams, and Dotter (2014) use data from the San Diego public schools to estimate student fixed effect models of high school test scores, absences, and grade promotion, and find little evidence that CTE enrollment has an effect. They also estimate models of high school graduation and college enrollment where they instrument for CTE enrollment using a school's CTE supply. Thus their identifying variation is similar to this study's, though this study focuses on CTE supply as the primary explanatory variable. Because the dropout age is 18 in California, Betts, Zau, McAdams, and Dotter (2014) are only able to estimate the effect of CTE enrollment on the probability of 5-year graduation, conditional on having attended high school for four years. With such a short window for CTE to have an effect, it not surprising that the authors find no effect. In North Carolina, students are able to drop out at age 16, giving them much more control over the length of their high school career. The authors also find little evidence that taking CTE courses either encourages or discourages college enrollment.

1.2.2 Europe

As noted earlier, the typical curriculum covered in studies of CTE in European countries differs notably from the typical curriculum in the US. For example, Sweden, as discussed in Hall (2013), places students in CTE and GE tracks at age 16 and about 45 percent of students take the CTE track. The strong division between CTE and GE students in Europe lends itself to generating natural experiments when there are nationwide changes in the CTE track's curriculum. Such changes provide a credibly exogenous source of variation in CTE enrollment.

Three studies examine the labor market outcomes of students in the CTE track after the content or length of the track is changed. Malamud and Pop-Eleches (2010) study a change in Romania in 1973 where students in the CTE track were required to take two additional years of GE and two fewer years of CTE. The authors find that men affected by the change were less likely to work in manual or craft-related occupations, but had similar wages and labor force participation rates. Oosterbeek and Webbink (2007) exploit a change in 1975 in the Netherlands that required students in the CTE track to attend school for one additional year before graduating. They find that this exchange of one year of on-the-job training for one year of formal schooling had no effect on wages. Hall (2013) exploits a policy change in Sweden in the 1990s that lengthened the CTE track by one year and added more GE to it. She finds that the change actually increased the probability of unemployment for men who had low GPAs before entering the CTE track, likely because the change also increased dropout rates.

Another European study focuses solely on the apprenticeship track. Fersterer, Pischke, and Winter-Ebmer (2008) examine the wage outcomes of people in Austria who were apprentices at companies that happened to fail in the middle of the apprenticeship. They find that those who were a year closer to finishing their apprenticeships had wages that were 2.5 to 4 percent higher in their first post-apprenticeship job.

1.2.3 Across countries

Given the notable differences in CTE curriculum between the US and many European countries, some studies use cross-country variation to explore the relationship between CTE and educational and labor market outcomes. Bishop and Mane (2004) use data on twenty European countries and the US from the 2001 OECD report *Education at a Glance* and find a positive association between the percent of upperclassmen in CTE tracks and both high school graduation rates and college enrollment rates.

Using data from the International Adult Literacy Survey, which covers fifteen European countries plus the US, New Zealand, and Chile, Hanushek, Woessmann, and Zhang (2011) test the hypothesis that CTE has a smaller long run payoff because CTE creates workers who are less adaptable to structural changes in labor demand. They model the relationship between employment and CTE enrollment over the life cycle. Their model is causal under the assumption that selection into CTE based on skill has a level effect on the probability of employment but does not cause differences in employment probabilities over the life cycle. That is, skill does not have an effect on the age gradient of employment, so that the only reason employment rates differ between GE- and CTE-enrollees by age is because of their GE-CTE enrollment decision. The authors argue that the relationship is indeed causal and find that as people age, CTE takers have lower employment rates than GE takers.

1.3 Theory

In this section I present a theoretical framework of the decision to complete high school that highlights the trade-off between taking CTE and GE courses. In the framework, school administrators influence students' enrollment and graduation decisions through CTE and GE course availability. The framework builds on Becker's (1993) human capital model where the completion decision depends on whether the benefits of completion are greater than the costs. I treat the decision as a two-stage process. In the first stage, individuals decide on the optimal mix of CTE and GE courses to take. In the second stage, they decide whether to complete high school given their optimal CTE and GE course mix.

The first stage decision depends on three factors: 1) the relative returns Y to CTE (C) and GE (G) courses, which could be financial, health, or marital and are a function of skill[1] s, 2) the

[1] In this context, s is a summary measure of a student's endowment of cognitive and noncognitive skills. Thus even if students do not select into CTE courses based on cognitive skills, the psychic costs of GE enrollment may be decreasing in the noncognitive component of skill.

relative psychic costs P which are also a function of skill, and 3) the relative availability of CTE and GE courses at a student's school A which depends on the number of course tracks offered t. Notice that I assume t is exogenous, as it is not a function of s or any other variable in the model.

Thus in the first stage, students

$$\max\ f(Y,P)$$
$$subject\ to\ A.$$

where

$$Y = g(C,G,s)$$
$$P = g(C,G,s)$$
$$A = g(C,G,t_C,t_G).$$

If f is concave, there is an optimal mix of CTE and GE courses. In the constraint A, the "price" of taking a CTE or GE course would be inversely proportional to the number of tracks offered, t_C or t_G.

In the second stage, individuals decide whether to complete high school given their optimal GE and CTE course mix. This decision depends on the returns to GE and CTE courses and the psychic costs of taking them, but it also depends on the foregone earnings and leisure D individuals will have if they do not finish school. These are a function of skill as well. Thus individuals will complete high school if

$$h(Y, P, D) > 0,$$

where

$$D = g(s).$$

Note that the completion decision and the course mix decision both depend on skill. This will bias OLS estimates of the relationship between CTE enrollment and graduation, an issue widely discussed in the literature.

This study focuses on the effect of CTE supply on graduation and the framework indicates that school administrators face a tradeoff in choosing how much CTE to supply relative to GE. If, for example, the return to GE is increasing in skill, high skill students will have greater utility when more GE courses are offered, while low skill students will have greater utility when more CTE courses are offered. Thus any time administrators add CTE tracks, they are helping lower skilled students at the expense of higher skilled students. To maximize total utility across all students, school administrators should pick a GE-CTE course mix that best suits the average student. On the other hand, if school administrators wish to maximize high school graduation, they should pick a course mix that best suits students who are at the margin of graduation.

If administrators actually do offer CTE courses based on the students they serve, this would make CTE course offerings endogenous and bias my empirical estimates. Consistent with endogenous course offerings, in the summary statistics section I show that schools that offer many CTE courses have different average student characteristics than schools that offer few CTE courses. For this reason, I focus on within-school changes in CTE supply, which are exogenous under the assumption that principals do not set supply in response to a particular cohort's characteristics.

1.4 Evidence

1.4.1 Data Description

The data for my empirical analysis come from the North Carolina Education Research Data Center, which compiles a wide range of administrative data on the universe of public school students in North Carolina. Data are available at the local education agency (district) level, school level, and student level. I do my analysis primarily at the student level. Unfortunately, the data come from a number of separate administrative databases, so that it is necessary to match individuals across databases, and the probability of successfully matching is not random. Students who switch schools or who drop out are more difficult to track and tend to be lower performers. This problem could result in biased estimates if there is heterogeneity in the effect based on unobservable factors that also predict whether a student matches across databases. For this reason, it may not be valid to infer that this study's empirical estimates apply to the unmatched population of students.

To measure school-level CTE offerings, I calculate the credit-weighted proportion of a school's courses that are CTE across all grades. In the empirical analysis, I apply this measure of

CTE supply to cohorts when they are in 9th grade. While it may seem appropriate to measure CTE supply using only courses available to 9th graders, such a measure would discount the importance of more advanced CTE courses to the decisions of 9th graders. CTE courses are generally offered as a part of a track (such as business or IT), which includes a sequence of at least three courses. In 9th grade, students primarily take introductory CTE courses, but the availability of more advanced CTE courses surely matters for their decision to start on a CTE track.

North Carolina Public Schools classifies courses using a four-digit course code, but schools can create multiple course numbers under the umbrella of a single course code. For example, Algebra II and Honors Algebra II have the same course code, but different course numbers. To count courses, I aggregate up to the course code level, for two reasons. First, course codes are set by the state so that they are uniform across schools. Second, some schools create a new course number for each independent study course so that many course numbers only have one attendee. Such courses are more likely to be a function of demand than supply, and counting them could contaminate the measure of CTE supply.

When I calculate the proportion of courses a school offers, I weight by credits, meaning that a two-credit course code receives twice the weight that a one-credit course code does. For course codes where different course numbers have different credits levels (such as independent study course codes), I use the school’s average for that year.

The CTE availability data come from a course-level database that tracks student attendance by school and by course number. I count a course as offered only if there was positive attendance in the course. I cannot observe whether there were courses offered that no one decided to take. There is therefore some risk that my measure of course offerings captures course demand. Such mismeasurement would bias my results if there is sorting into CTE courses based on unobserved skill, which my theory suggests there is. I check for whether the measure of CTE availability is contaminated by demand shifts in Section 1.4.6 and find little evidence that it is.

There are 867 course codes in the database, 857 of which are not for disabled students and are not non-class codes (such as study hall or tutoring). Of the 857 codes I use, North Carolina Public Schools classifies 368 as CTE. However, not all of these courses offer training in marketable skills. I consider courses under the Career Development heading (such as Exploring Career Decisions) as CTE General Education and don't include them in my measure of CTE course offerings. I also don't include courses under the heading Family and Consumer Sciences. While many of the courses under this heading are clearly not career-oriented (such as Teen Living, Fundamentals of Foods, and Parent and Child Development), others (such as Culinary Arts and Hospitality and Clothing Design) are closer to the borderline of being career-oriented. My preferred estimates exclude these categories, but I present estimates that include them in the robustness section. I also count all computer science courses as CTE, even though the North Carolina Public Schools classifies some under the GE Computer Sciences heading and others under the CTE Technology heading.

To measure CTE enrollment in 9th grade, I calculate the proportion of a student's credits that are in CTE. I focus on 9th graders because most students cannot drop out until 10th grade (the dropout age is 16) and I cannot observe the mix of CTE and GE that dropouts would have taken if they had not dropped out. Thus, if I wanted to include course mix decisions for later grades, I would have to exclude those who dropped out after 9th grade. I use the proportion of a student's credits that are CTE instead of the raw count of CTE courses taken because not all courses have the same number of credits.

I use data from when students were in 8th grade as proxies for cognitive and noncognitive skills. In 8th grade students must take statewide standardized reading and math tests and I standardize their scores to be zero mean, standard deviation one. The database of 8th grade test scores also includes student-reported time spent on homework, time spent using a computer to do homework, and time spent reading for fun, all of which can be interpreted as proxies for noncognitive skills. I use principle component analysis to summarize the three time use measures, which I standardize to be mean 0, standard deviation 1. All three variables receive a positive loading. CTE course availability could affect measures of cognitive and noncognitive skills taken in 9th grade, meaning 9th grade measures are outcomes, not controls. Because the measures come from 8th grade, there is little reason to think that they were influenced by CTE course availability in 9th grade.

1.4.2 Summary Statistics

Table I shows the summary statistics for the variables used in the empirical analysis. The data cover cohorts that were in 9th grade in the years 2005 to 2010. I was able to successfully match 399,634 students from 586 public schools across all variables. The average match rate across schools was 82 percent. "School's average" variables are calculated across all years in the data using both matched and non-matched students. These variables are used only to check for heterogeneity, as they do not vary over time. They are unavailable for a small number of schools.

The average proportion of schools' courses that are CTE is 0.23 (Figure 1 shows the distribution). The sample is 27 percent black, which is about twice the rate for the US, and 7 percent Hispanic, which is about two-fifths of the rate for the US. The average 8th grade math score is 0.201 and the average 8th grade reading score is 0.188, indicating that the matched sample gets somewhat higher test scores than the universe of all students. Thirty-six percent of the sample receives free lunch, which is available to those with family incomes of less than 1.3 times the poverty line. Ninety-nine percent of students were rated proficient in English. The average time use index is close to zero, indicating that matched students were about average in how much time they spent in 8th grade doing homework and reading for fun.

Table II shows the school-level means from Table I broken down by quartile of the proportion of a school's courses that are CTE, where the quartile cutoffs are weighted by a school's student population so that a roughly equal number of observations are in each quartile.

Schools offering CTE have slightly lower graduation rates and are much more likely to be rural. Such schools were also whiter, poorer, had lower test scores, and lower time use index values. Table III shows the student-level means from Table I broken down by the number of CTE courses taken in 9th grade. Thirty-one percent of students took a CTE course in 9th grade, and of those, 86 percent took one course, 13 percent took two, and 1 percent took three or more. Males were more likely to take CTE courses, as were whites, and those not receiving free lunch. CTE takers tended to have higher math and reading test scores, but they also tended to have a lower time use index value. Thus the typical 9th grade CTE student had slightly above average cognitive skills, was less motivated to do academic work, and lived in a rural area.

The average distribution of 9th grade course enrollments across all course categories for the matched sample is shown in Table IV. On average, 13 percent of credits are in CTE courses. However, 55 percent of overall CTE credits are in the Career Development category, which I consider to be general education and do not count in my measures of CTE enrollment or availability. An additional 8 percent of 9th grade CTE credits are in the family and consumer sciences category, which covers skills necessary for a well-managed home, but not necessarily skills that are career-oriented. Thirty-seven percent of CTE enrollment is in what I consider to be career-oriented courses and the largest enrollments are in agriculture (20 percent), trades (16 percent), and health (15 percent). As stated earlier, for the empirical analysis that follows, my primary measure of CTE enrollment counts only courses that directly teach career-oriented skills.

Variable means by CTE category are shown in Table V. In general, men are more likely to take the career-oriented tracks, with the exception of the health track. Students who take IT and design tracks tend to have the highest math and reading scores, and students who take agriculture and trades courses tend to have the lowest time use index values.

1.4.3 Empirical Approach

My theory makes the prediction that students choose a CTE-GE course mix based on their skill level. This prediction has important implications for estimating the effect of CTE on high school graduation because it means that CTE enrollment is likely a function of skill. If lower skilled students tend to take CTE courses, OLS estimates of the effect of CTE enrollment on high school graduation will be biased downward. For this reason, I focus on CTE supply, which I argue below is exogenous in the empirical model under some reasonable assumptions. Exogeneity of CTE supply means that school administrators set supply based on factors that matter for graduation, such as the average skill level of the student body. For example, if administrators tend to offer more CTE courses in schools where the student body has below average skill, estimates of the effect CTE supply on graduation will be biased downward.

To understand why CTE supply is plausibly exogenous in the empirical model, it is important to distinguish between supply adjustments made based on the long run characteristics of a school's student body and adjustments made on a cohort-by-cohort basis. Adjustments made based on long run characteristics will largely be captured by the school fixed effects in the model. It is also possible that administrators make adjustments based on trends in student body quality, and I show in the robustness section that including linear school time trends in the model does not substantially change the results.

Because I am able to include school fixed effects in the empirical model, the primary requirement for CTE supply to be exogenous is that administrators do not make adjustments on a cohort-by-cohort basis. There are a number of theoretical reasons to believe they do not. First, school administrators are likely learn about cohort quality only after a cohort has entered their school, especially in larger districts where students can come from multiple middle schools. Second, CTE courses are part of a multiyear track where students typically take an introductory course in one year and take higher level courses in the years that follow. Thus adding or removing a track will affect multiple cohorts, not just one, making it difficult for administrators to make adjustments on the fly. Third, adding a new track likely requires more investment than simply hiring a teacher and purchasing textbooks. Many tracks require special equipment and classrooms for training. Decisions on such investments likely happen over multiple years, so it would not be possible to make them based on year-to-year differences in cohorts.

There are other reasons CTE supply can change that are unrelated to student body quality. For example, there may be changes in education policy at the national, state, or district level. Another possibility is that administrators are responding to long run structural changes in the economy that make some CTE courses obsolete and require the creation of new ones. The administrative response to factors such as these are likely to be random in timing, and therefore random across cohorts.

The first empirical model I estimate is an OLS model of the relationship between high school graduation and the proportion of a student's transcript that is CTE in 9th grade. The coefficient is likely to be biased downward because of selection on unobservable skill, but it provides an estimate that may be useful for comparison with estimates from other datasets. Next I use OLS to estimate the reduced form relationship between graduation and the proportion of a school's courses that are CTE (across all grades) when a student is in 9th grade. From the perspective of school administrators this is the empirical model of interest because it is the variable they decide on. The coefficient from this model is unbiased if my identification assumptions hold. Formally, the model is:

$$Grad4Yr_i = \alpha + \beta pCrsCTE_{sc} + \boldsymbol{\gamma}\boldsymbol{X_{ic}} + \pi_s + \varphi_c + u_i,$$

where $Grad4Yr_i$ is a binary variable for whether a student graduated in four years, $pCrsCTE_{sc}$ is the proportion of a school's courses that are CTE when cohort c is in 9th grade, $\boldsymbol{X}_{ic}$ is a vector of individual- and cohort-level covariates that include demographics, proxies for cognitive and noncognitive skills, and local employment growth, π_s is a school fixed effect, φ_c is a cohort fixed effect, and u_i represents all unobservable factors affecting high school graduation. Standard errors for all models are clustered at the school-by-year level.

Next I use OLS to estimate the relationship between the proportion of a student's transcript that is CTE in 9th grade and the proportion of a school's courses that are CTE. The coefficient from this model tells us how CTE enrollment responds to changes in CTE supply. The model also provides information on the student characteristics that predict CTE enrollment.

Finally, I explore which CTE tracks are driving the reduced form results, as the overall effect of CTE offerings on graduation is a weighted average of these individual effects. I rerun the reduced form model, replacing the overall proportion of a school's course offerings that are in CTE with the proportion of a school's course offerings that are in each track. I include the measures for all nine categories in one model so that the reported coefficients are those that are used in the weighted average.

1.4.4 Main Results

OLS estimates of the relationship between high school graduation and the proportion of a student's transcript that is CTE in 9th grade are show in Table VII. The coefficient moves around some initially as controls are added. For example, males are more likely to take CTE and less likely to graduate, so adding demographic controls that include a female dummy increases the coefficient. Also, students who take CTE typically have above average 8th grade math scores so after controlling for math, the coefficient decreases. After adding all the available controls, I estimate that taking one additional CTE course in 9th grade (typically one-eighth of a transcript) is associated with a 0.2 percentage point increase in the probability of graduation. Note again that if there is negative selection on unobservable skill, this estimate is biased downward.

The reduced form OLS estimates of the relationship between graduation and CTE supply are shown in Table VIII. If my identification assumption holds, the final estimate in column 7 indicates that a 10 percentage point increase (one standard deviation is 8) in the proportion of a school's courses that are in CTE results in a 0.63 percentage point increase in the probability of graduating. This translates into just over one additional graduate for a cohort of 200. Put another way, the average graduation rate in the data is about 90 percent, so a 10 percentage point increase in CTE supply would reduce the dropout population by about 6 percent.

The stability of the coefficients in Table VIII as I add controls is one indicator of whether it is valid to assume that CTE supply is exogenous. If CTE supply is correlated with observable determinants of graduation, it is raises the concern that CTE supply is also correlated with the unobservable determinants of graduation, so that the estimates are biased. In general, the coefficients in Table VIII are stable, though they change with the addition of a couple of important controls. First, adding school and cohort fixed effects nearly doubles the coefficient. As discussed in the previous section, it is not surprising that schools choose CTE supply based on the long run characteristics of their student body, and it is not a violation of the exogeneity of CTE supply so long as I include school fixed effects. The other notable change in the coefficient comes when I include students' math scores. The coefficient goes up, which implies that cohorts with lower average math scores tend to be offered more CTE options (higher math scores strongly predict graduation). This finding does raise the concern that administrators are adjusting CTE supply on a cohort-by-cohort basis based on unobservable skill. I explore this concern in depth in Section 1.4.6.

The relationship between CTE enrollment and CTE supply is shown in Table IX. With all the controls included, I estimate that a 10 percentage point increase in CTE supply increases CTE enrollment by 0.45 percentage points. For a cohort size of 200, that translates into about 7 students taking 1 additional CTE course.

Some may be interested in knowing the implied effect of 9th grade CTE enrollment on graduation if we use CTE supply as an instrument for enrollment. For the IV estimate to be valid, we must assume that the entire effect of CTE course offerings operates through those who enroll in CTE courses in 9th grade. There are several reasons why this assumption may not be valid. First, CTE availability could affect CTE non-enrollees through peer effects. New CTE tracks will remove some students from GE courses, altering the learning environment in those GE courses, if, for example, CTE enrollees have a tendency to be disruptive. CTE availability may also lead friends of CTE enrollees to stay in school. Peer effects that work through these channels would bias instrumental variables estimates upward. Another reason that instrumental variables estimates may not be valid is that some may wait to enroll in CTE until 10th grade, even though their decision is based on CTE supply when they are in 9th grade. This would bias the first stage coefficient down and the second stage coefficient up. With these caveats in mind, using two stage least squares, I estimate that taking one additional CTE course in 9th grade (0.125 percent of the typical transcript) increases the probability of graduating by

18 percentage points. This magnitude is too high to be believable, which is why it is likely that CTE supply is acting on graduation through more channels than just a student's 9th grade CTE enrollment.

The reduced form estimates of the effect of individual CTE tracks on graduation are shown in Table X. Most tracks make positive contributions. Agriculture, business and finance, and marketing courses are the biggest contributors, though the benefits of offering agriculture courses are largely limited to rural schools. It is somewhat surprising that the results indicate that offering technology and engineering courses reduces graduation rates, when such courses offer some of the most marketable skills and are also typically precursors to further study in college. The estimates in Table X are generally stable after adding school and cohort fixed effects. As expected, the largest change to the coefficients from adding other controls comes from math, though the coefficient changes for only a few tracks (health and design).

1.4.5 Robustness and Specification Checks

In this section, I examine whether the conclusions drawn from the main results hold up if I use alternate measures of CTE supply and make modifications to the specification. I construct four alternate measures of CTE supply. The first alternate measure excludes CTE courses offered at community colleges that qualify for high school credit. The idea is that administrators do not have complete control over the availability of such courses, so they may not capture the true policy decision administrators face. The second alternate measure includes family and consumer sciences courses, some of which may teach career-oriented skills and therefore should be included in the measure of CTE supply. The third alternate measure is calculated at the course number level, not the course code level. For example, for this measure, Algebra I and Honors Algebra I are treated as separate course offerings, where my preferred measure treats them as only one course offering. The fourth alternate measure counts a course only if two or more students are enrolled in it.

I make four modifications to the specification. First, I add controls for the availability of family and consumer sciences and general education CTE courses. Without including them as controls, these courses are treated empirically as identical to GE courses, when it may be more appropriate to think of them as a third category of courses. Thus including the controls may more accurately reflect the tradeoff between offering CTE and GE courses. The second alternate specification excludes cohorts with fewer than 50 students. It may be that schools with small cohort sizes are fundamentally different than schools with bigger cohort sizes, especially in their capacity to offer CTE courses. The third alternate specification includes the square of the CTE supply measure, which is a check for nonlinearity. The fourth alternate specification includes school time trends, which control for long run trends in school and cohort quality.

The results from the robustness checks are shown in Table XI. Overall, the coefficient on the proportion of a school's courses that are in CTE is stable and ranges from 0.039 to 0.076. I do find evidence of nonlinearity in the effect of CTE supply. Interestingly, the estimates suggest that the effect is increasing in strength as CTE supply increases. For example, at the average level of CTE supply (0.23), the estimated effect is 0.045, while the estimated effect is 0.097 at one standard deviation above the average (0.31). These results contradict the theory that there are diminishing returns to CTE supply, though it may be that there are so few schools close to the point of diminishing returns that there is no support over the domain of diminishing returns.

1.4.6 Is CTE Supply Exogenous?

There are two primary reasons my measure of CTE supply could be endogenous. First, as discussed in Section 1.4.3, school administrators could change the supply of CTE courses on a cohort-by-cohort basis in response to cohort characteristics that also matter for graduation. If administrators do this, I should be able to predict CTE supply using observable characteristics. Second, schools may offer courses that no students take in a given year (likely higher level ones). Such courses would not show up in my measure of CTE supply, which means my measure could be a function of CTE demand. If this is the case, I should find that cohort characteristics that predict CTE enrollment also predict CTE supply.

To look for the patterns described above I estimate a weighted least squares model of the CTE supply using summary measures of cohort characteristics, where the weights are the size of the cohort. For most of the summary measures, I am able to use all the students in a cohort, not just the ones that match across NCERDC databases. The exceptions are for 8th grade test scores and the time use index. Because lower skilled students are more likely not to be matched, cohort averages of test scores and time use will be skewed and could result in biased estimates. For example, two cohorts may have the same average (higher) math score among matched students and the same average (lower) math score among unmatched students, but the first cohort may have more students who aren't matched. Thus while the first cohort's overall average math score is actually lower than the second cohort's is, their measured averages are the same. In this scenario, if school administrators are changing CTE supply based on average cohort math scores, I will not be able to detect it because there is no variation in the measured average scores. To address possible bias from this type of mismeasurement, I include the proportion of matched students in a cohort as an explanatory variable.

Summary statistics for all cohort summary measures used for the WLS model of CTE supply are shown in Table VI. Not surprisingly, the patterns are quite similar to the patterns in the student-level data. The results from the model are shown in Table XII. It is important to note that the magnitudes are all quite small, even though some are statistically significant. For example, a one standard deviation increase in the proportion of a cohort that is female (0.097) is associated with a 0.4 percentage point increase in the proportion of a school's courses that are CTE. And a one standard deviation increase in the average math score of a cohort (0.432) is associated with a 1.2 percentage point decrease in CTE supply, which is about 15 percent of the standard deviation of CTE supply. Thus if administrators are truly adjusting CTE supply on a cohort-by-cohort basis they are making relatively small adjustments. These results also suggest that there is little contamination from student demand in the measure. For example, people with higher math scores are more likely to enroll in CTE, but my measure of CTE supply has a slightly negative association with math scores. We should see these signs going in the same direction if the measure of CTE supply is picking up student demand.

To further test for the exogeneity of CTE supply, I look at its relationship with the graduation rate that would be predicted by a cohort's demographic characteristics. Because the signs on the demographic characteristics in the previous model are often in different directions, it is difficult to discern the direction (if any) of the overall association between a cohort's observable characteristics and CTE supply. This test allows us to do so. To generate the predicted cohort graduation rate, I take the linear projection of the actual cohort graduation rates on the cohort demographic variables from the previous model. I then estimate a WLS model of the predicted graduation rate using the CTE supply measure and school and cohort fixed effects.

Table XIII shows that CTE supply has little relationship with the predicted graduation rate. A 10 percentage point increase in CTE supply is associated with just a 0.0008 percentage point increase in the predicted graduation rate. One interpretation of this result is that while we would predict that cohorts that are offered more CTE options shouldn't have higher graduation rates, in reality they do. Specifically, we would expect a 10 percentage point increase in CTE supply to raise graduation rates by 0.0008 percentage points, when graduation rates actually go up by 0.63 percentage points. If administrators are adjusting CTE supply on a cohort-by-cohort basis, it does not appear that they are doing it in a way that will raise expected graduation rates.

Overall, the two models in this section provide evidence that CTE supply changes on a cohort-by-cohort basis are at most small and that it is unlikely that the measure of CTE supply is contaminated by CTE demand.

1.4.7 Heterogeneity

In this section I explore whether the reduced form effect of CTE supply on high school graduation varies by student and school characteristics. I estimate two models of graduation, one looking for heterogeneity by student-level variables and one looking for heterogeneity by school-level variables. Table XIV shows that there is heterogeneity by almost every student-level variable. For example, I estimate that a 10 percentage point increase in the supply of CTE increases the probability of graduating by 1.5 percentage points for Hispanics while the effect is 1.0 percentage points for whites. I also find that the effect is stronger for those with higher math scores and much smaller for students who receive free lunch.

Results for heterogeneity by school-level variables are shown in Table XV. With the exception of change in log local employment, these variables are fixed over time and are calculated for the same years as the student-level data, 2005-10. The school-level results are generally inconclusive, though I find that in times of strong local labor demand CTE is especially effective at keeping students in school. A one standard deviation increase in the growth rate of employment (0.045) increases the effect of CTE supply on graduation by 19 percent.

1.5 Discussion

This study provides the first causal evidence from US-based data that offering Career and Technical Education (CTE) helps raise high school graduation rates. I address the concern of nonrandom enrollment in CTE courses by employing a research design that compares outcomes for cohorts from the same school but that are offered different CTE options. Such variation is exogenous to the high school graduation decision so long as school administrators do not change CTE supply based on cohort characteristics. While I find some evidence that course offerings are correlated with cohort characteristics, I also find that the relationship is small and is unlikely to explain away my results.

In my preferred specification, I estimate that a 10 percentage point increase in the proportion of a school's courses that are in CTE when a student is in 9th grade raises graduation rates by 0.64 percentage points. For a cohort of 200 students, this means that about one additional student will graduate. Put another way, with an average dropout rate of about 10 percent in the data, the dropout population would decline by about 6 percent. When breaking CTE curriculum down by tracks, I find that the agriculture, business and finance, and marketing tracks make the largest contributions to increasing graduation rates, while the technology and engineering track actually makes a negative contribution.

While the stereotypical CTE student has below average academic skills, I find that 9th grade enrollees in CTE courses that teach marketable skills tend to have above average math and reading scores, but below average time spent on homework in 8th grade. This indicates that the typical CTE student has relatively good cognitive skills, but may lack the motivation to succeed in an academic environment.

When I look for heterogeneity in the effect of CTE availability, I find that the effect is stronger for students with higher 8th grade math test scores, but that the effect goes down significantly for students who receive free lunch. The effect is also stronger when the local economy is doing well, indicating that CTE courses help keep people in school who would otherwise drop out and work.

This study indicates that CTE can play a valuable role in the American education system alongside general education (GE). While GE may provide longer run benefits to many students because it teaches skills with broader applications in the labor market, CTE has clear benefits for a certain type of student: those with moderate cognitive skills, but who lack the motivation to apply themselves in an academic setting. For these students, CTE is not so much an alternative to GE, but an alternative to dropping out, and keeping them in school likely does help them in the long run. In the end, school administrators still face the difficult challenge of balancing the diverse needs of a student body. Devoting additional resources to CTE and the students who benefit from it likely means taking away resources from other academic programs that benefit other types of students. This tension is inherent to the education policy puzzle.

1.6 Cited Literature

Becker, Gary S. 1993. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education.* Chicago: University of Chicago Press.

Betts, Julian R., Andrew Zau, John McAdams, and Dallas Dotter. 2014. *Career and Technical Education in San Diego: A Statistical Analysis of Course Availability, Students' Course-Taking Patterns, and Relationships with High School and Postsecondary Outcomes.* San Diego, CA: University of California San Diego, San Diego Education Research Alliance.

Bishop, John H. and Ferran Mane. 2004. "The Impacts of Career-technical Education on High School and Labor Market Success." *Economics of Education Review* 23(2004): 381-402.

Cellini, Stephanie Riegg. 2006. "Smoothing the transition to college? The effect of Tech-Prep programs on educational attainment." *Economics of Education Review* 25(2006): 394-411.

Fersterer, Josef, Jörn-Steffen Pischke, and Rudolf Winter-Ebmer. 2008. "Returns to Apprenticeship Training in Austria: Evidence from Failed Firms." *Scandinavian Journal of Economics* 110(4), 773-753.

Hanushek, Eric, Ludger Woessmann, and Lei Zhang. 2011. "General Education, Vocational Education, and Labor-market Outcomes over the Life-cycle." *NBER Working Paper 17504.*

Hall, Caroline. 2013. "Does More General Education Reduce the Risk of Future Unemployment? Evidence from Labor Market Experiences During the Great Recession." *Institute for Evaluation of Labour Market and Education Policy Working Paper Series*: No. 2013:17.

Kang, Suk and John Bishop. 1989. "Vocational and Academic Education in High School: Complements or Substitutes?" *Economics of Education Review* 8(2): 133-148.

Kreisman, Daniel and Kevin Stange. 2015. "Vocational and Career Tech Education in American High Schools: Curriculum Choice and Labor Market Outcomes." *Unpublished manuscript.*

Malamud, Ofer and Cristian Pop-Eleches. 2010. "General Education Versus Vocational Training: Evidence from an Economy in Transition." *Review of Economics and Statistics* 92(1): 43-60.

Mane, Ferran. 1999. "Trends in the payoff to academic and occupation-specific skills: the short and medium run returns to academic and vocational high school courses for non-college-bound students." *Economics of Education Review* 18(1999): 417-437.

Meer, Jonathan. 2007. "Evidence on the Returns to Secondary Vocational Education." *Economics of Education Review* 26(2007): 559-573.

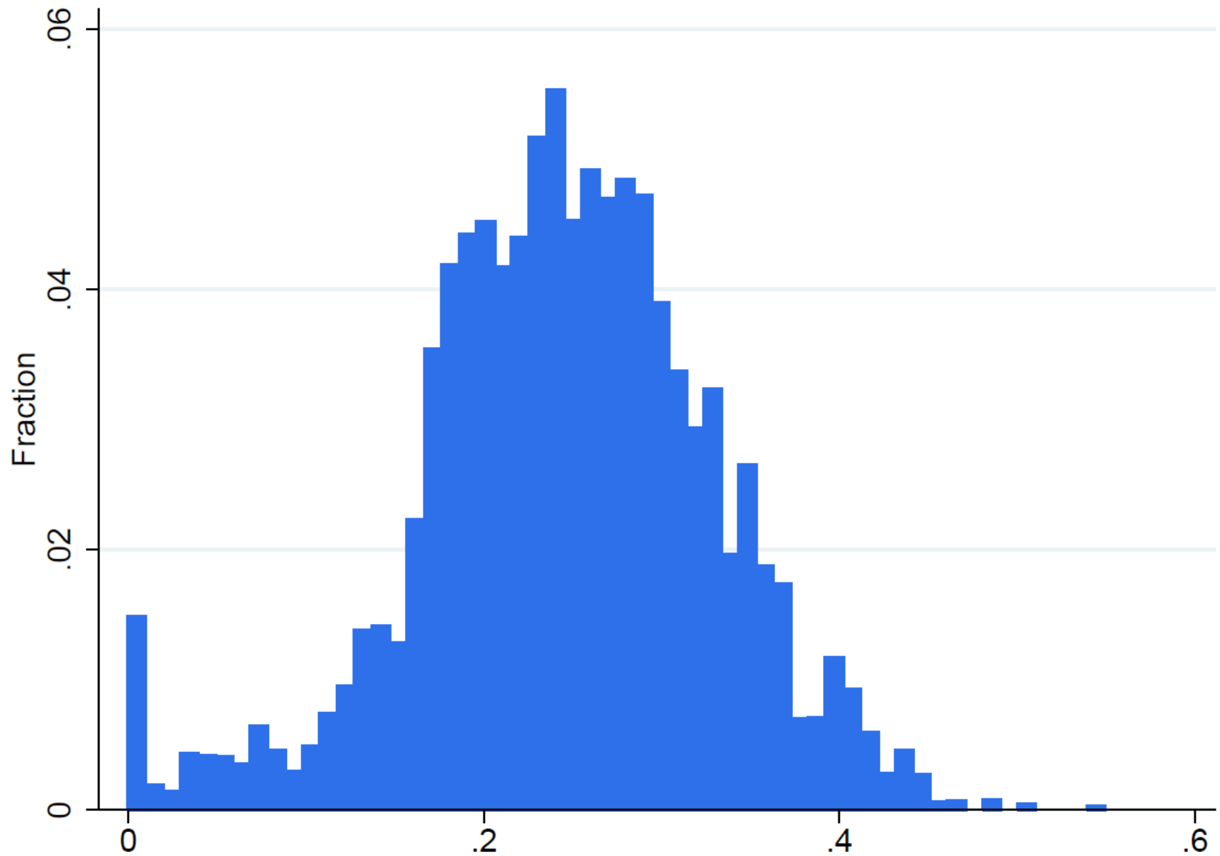
Neumark, David and Donna Rothstein. 2006. "School-to-Career Programs and Transitions to Employment and Higher Education." *Economics of Education Review* 25(2006): 374-393.

Oosterbeek, Hessel and Dinand Webbink. 2007. "Wage Effects of an Extra Year of Basic Vocational Education." *Economics of Education Review* 26(2007): 408-419.

1.7 Figures and Tables

1.7.1 Figures

Figure 1. Distribution of the Proportion of a School's Courses that are CTE



1.7.2 Tables

Table I. Summary Statistics

	Mean	Standard Deviation
Graduated in 4 Years	0.885	0.319
School's Average Graduation Rate	0.802	0.078
School's Average Cohort Size	320	136
Took CTE Course in 9th Grade	0.312	0.463
Proportion of Student's Transcript CTE in 9th Grade	0.047	0.077
Proportion of School's Courses CTE[a]	0.233	0.080
Female	0.521	0.500
Black, not Hispanic	0.266	0.442
Hispanic	0.069	0.253
Other Race, not Hispanic	0.059	0.236
Free Lunch	0.356	0.479
Limited English Proficient	0.014	0.118
Former Limited English Proficient	0.019	0.135
8th Grade Math Score	0.201	0.927
8th Grade Reading Score	0.188	0.899
8th Grade Time Use Index	-0.019	0.963
School's Average Proportion Female[b]	0.513	0.034
School's Average Proportion Black[b]	0.263	0.214
School's Average Proportion Hispanic[b]	0.079	0.060
School's Average Proportion Other Race, not Hispanic[b]	0.062	0.070
School's Average Proportion Free Lunch[b]	0.361	0.149
School's Average 8th Grade Math Score[b]	0.189	0.340
School's Average 8th Grade Reading Score[b]	0.177	0.289
School's Average 8th Grade Time Use Index[b]	-0.014	0.227
Change in Log Local Employment	-0.011	0.045
In City with Population greater than 250,000	0.178	0.383
In City with Population between 100,000 and 250,000	0.122	0.328
In City with Population between 25,000 and 100,000	0.062	0.241
In Town with Population between 2,500 and 25,000	0.129	0.335
In Rural Area	0.507	0.500
Number of Observations	399634	
[b]Number of Observations	390854	

Notes: 9th grade cohorts are from 2005 to 2010. The data cover "matched" students from 586 North Carolina public high schools who had data for all variables. The 8th grade time use index is the first principal component of reported time spent on homework per week, time spent reading for fun per week, and time spent using a computer for homework per month in 8th grade. Each variable is positively weighted in the first component. [a]Proportion of school's courses CTE is measured across all grades. [b]These variables are calculated using matched and nonmatched students and are not available for a small number of schools.

Table II. Variable Means by Quartile of Proportion of School's Courses that are CTE

	1st Quartile CTE Course Offerings	2nd Quartile CTE Course Offerings	3rd Quartile CTE Course Offerings	4th Quartile CTE Course Offerings
School's Average Graduation Rate	0.818	0.800	0.795	0.798
School's Average Cohort Size	326	356	294	305
Proportion of School's Courses CTE	0.131	0.211	0.258	0.332
School's Average Proportion Female[a]	0.524	0.513	0.509	0.505
School's Average Proportion Black[a]	0.325	0.284	0.244	0.200
School's Average Proportion Hispanic[a]	0.091	0.081	0.076	0.070
School's Average Proportion Other Race, not Hispanic[a]	0.072	0.064	0.050	0.063
School's Average Proportion Free Lunch[a]	0.349	0.351	0.368	0.374
School's Average 8th Grade Math Score[a]	0.210	0.210	0.172	0.168
School's Average 8th Grade Reading Score[a]	0.213	0.191	0.156	0.154
School's Average 8th Grade Time Use Index[a]	0.106	0.022	-0.072	-0.107
Change in Log Local Employment	-0.015	-0.006	-0.012	-0.010
In City with Population greater than 250,000	0.300	0.246	0.088	0.075
In City with Population between 100,000 and 250,000	0.185	0.160	0.093	0.052
In City with Population between 25,000 and 100,000	0.054	0.058	0.088	0.046
In Town with Population between 2,500 and 25,000	0.116	0.120	0.126	0.155
In Rural Area	0.342	0.414	0.604	0.671
Number of Observations	99592	100155	99843	99722
[a]Number of Observations	92643	99860	99067	98962

Note: See notes in Table I for an explanation of key variables. Quartile cutoffs are weighted by a school's student population so that a roughly equal number of observations are in each quartile.

Table III. Student-level Variable Means by Number of CTE Courses Taken in Ninth Grade

	Zero CTE Courses	One CTE Course	Two CTE Courses	Three or More CTE Courses
Graduated in 4 Years	0.883	0.888	0.891	0.878
Took CTE Course in 9th Grade	0.000	1.000	1.000	1.000
Proportion of Student's Transcript CTE in 9th Grade	0.000	0.133	0.257	0.377
Proportion of School's Courses CTE	0.227	0.244	0.263	0.284
Female	0.593	0.387	0.206	0.106
Black, not Hispanic	0.289	0.226	0.154	0.096
Hispanic	0.072	0.064	0.052	0.038
Other Race, not Hispanic	0.062	0.056	0.044	0.041
Free Lunch	0.369	0.334	0.292	0.280
Limited English Proficient	0.016	0.010	0.007	0.002
Former Limited English Proficient	0.019	0.018	0.013	0.014
8th Grade Math Score	0.170	0.254	0.362	0.362
8th Grade Reading Score	0.171	0.215	0.282	0.300
8th Grade Time Use Index	0.020	-0.091	-0.179	-0.253
Number of Observations	274877	107478	16394	885
Proportion of Observations	0.688	0.269	0.041	0.002

Note: See notes in Table I for an explanation of key variables.

Table IV. Mean Proportion of Student's Credits in 9th Grade by Course Type

	Mean
Career and Technical Education	**0.130**
General Education CTE	**0.547**
Career Development	1.000
Non-Career-Oriented CTE	**0.084**
Family and Consumer Sciences	1.000
Career-Oriented CTE	**0.369**
Business and Finance	0.128
Information Technology	0.021
Marketing	0.077
Agriculture	0.195
Health	0.152
Trades	0.159
Design	0.089
Technology and Engineering	0.143
Community College	0.024
General Education	**0.870**
Language Arts	0.187
Foreign Languages	0.039
Mathematics	0.204
Sciences	0.154
Social Studies	0.148
Arts	0.086
Physical Education	0.145
Miscellaneous	0.035
Community College	0.002

Note: Means of categories and courses under each heading sum to 1.

Table V. Variable Means by CTE Category

	Family and Consumer Sciences	Business and Finance	Information Technology	Marketing	Agriculture	Health	Trades	Design	Technology and Engineering	Community College
Graduated in 4 Years	0.893	0.898	0.919	0.905	0.877	0.929	0.844	0.918	0.873	0.909
School's Average Cohort Size	339	305	349	400	274	302	278	327	337	144
Took CTE Course in 9th Grade	0.182	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Proportion of Student's Transcript CTE in 9th Grade	0.025	0.162	0.176	0.162	0.171	0.151	0.177	0.174	0.168	0.161
Proportion of School's Courses CTE	0.249	0.231	0.239	0.233	0.289	0.268	0.282	0.264	0.254	0.154
Female	0.735	0.411	0.213	0.451	0.337	0.754	0.075	0.190	0.157	0.523
Black, not Hispanic	0.288	0.280	0.326	0.258	0.091	0.261	0.197	0.146	0.229	0.172
Hispanic	0.060	0.063	0.069	0.059	0.045	0.058	0.067	0.064	0.065	0.078
Other Race, not Hispanic	0.067	0.055	0.092	0.051	0.037	0.069	0.046	0.049	0.055	0.080
Free Lunch	0.377	0.339	0.310	0.279	0.300	0.349	0.369	0.252	0.315	0.380
Limited English Proficient	0.010	0.007	0.014	0.009	0.007	0.006	0.014	0.006	0.013	0.006
Former Limited English Proficient	0.015	0.018	0.023	0.019	0.010	0.018	0.018	0.017	0.019	0.028
8th Grade Math Score	0.186	0.298	0.510	0.408	0.169	0.316	0.123	0.577	0.308	0.402
8th Grade Reading Score	0.165	0.264	0.389	0.326	0.167	0.286	0.068	0.451	0.231	0.412
8th Grade Time Use Index	-0.020	-0.064	0.116	0.021	-0.220	0.027	-0.300	-0.055	-0.128	0.042
Change in Log Local Employment	-0.009	-0.012	-0.005	-0.008	-0.013	-0.011	-0.012	-0.011	-0.008	-0.021
In City with Population greater than 250,000	0.156	0.140	0.406	0.445	0.032	0.173	0.054	0.186	0.214	0.006
In City with Population between 100,000 and 250,000	0.137	0.088	0.167	0.062	0.048	0.075	0.072	0.117	0.133	0.102
In City with Population between 25,000 and 100,000	0.081	0.048	0.022	0.035	0.015	0.061	0.105	0.087	0.011	0.021
In Town with Population between 2,500 and 25,000	0.106	0.121	0.061	0.096	0.115	0.136	0.124	0.113	0.107	0.171
In Rural Area	0.518	0.602	0.344	0.362	0.789	0.553	0.644	0.498	0.533	0.700
Number of Observations	32354	20273	3058	11247	27616	22229	22381	12946	20794	3410

Note: See Table I for a description of key variables. Family and Consumer Sciences courses are not counted as CTE in the empirical analysis because the majority of them are not career-oriented.

Table VI. Summary Statistics of Cohort-level Variables

	Mean	Standard Deviation
Cohort Graduation Rate	0.781	0.154
Proportion of School's Courses CTE in 9th Grade	0.212	0.109
Cohort Proportion Female	0.514	0.097
Cohort Proportion Black, not Hispanic	0.290	0.251
Cohort Proportion Hispanic	0.078	0.077
Cohort Proportion Other Race, not Hispanic	0.060	0.082
Cohort Proportion Free Lunch	0.399	0.190
Cohort Proportion Limited English Proficient	0.024	0.034
Cohort Proportion Former Limited English Proficient	0.017	0.036
Cohort Average 8th Grade Math Score	0.076	0.432
Cohort Average 8th Grade Reading Score	0.089	0.393
Cohort Average 8th Grade Time Use Index	-0.055	0.292
Cohort Proportion with Matched Test Scores	0.819	0.119
Change in Log Local Employment	-0.012	0.046
Average Cohort Size	222	150
Number of Observations	2771	

Note: See Table I for a description of key variables.

Table VII. Linear Probability Model of High School Graduation in Four Years - Relationship with Proportion of Student's Transcript CTE in 9th Grade

	(1) Graduate in Four Years	(2) Graduate in Four Years	(3) Graduate in Four Years	(4) Graduate in Four Years	(5) Graduate in Four Years	(6) Graduate in Four Years	(7) Graduate in Four Years
Proportion of Student's Transcript CTE in 9th Grade	-0.008	-0.026***	0.032***	0.012	0.015*	0.017**	0.017**
	(0.011)	(0.009)	(0.009)	(0.008)	(0.008)	(0.008)	(0.008)
School Fixed Effects		X	X	X	X	X	X
Cohort Fixed Effects		X	X	X	X	X	X
8th Grade Math Score			X	X	X	X	X
8th Grade Reading Score				X	X	X	X
8th Grade Time Use Index					X	X	X
Sex and Race Fixed Effects						X	X
Free Lunch Fixed Effects						X	X
English Proficiency Fixed Effects						X	X
Change in Log Local Employment							X
N	399632	399632	399632	399632	399632	399632	399632
Adjusted R^2	0.000	0.041	0.061	0.120	0.124	0.124	0.124

Notes: Standard errors are clustered at the school by cohort level; * p<0.10, ** p<0.05, *** p<0.01. See notes in Table I for a description of key variables.

Table VIII. Linear Probability Model of High School Graduation in Four Years - Reduced Form Relationship with Proportion of School's Courses CTE

	(1) Graduate in Four Years	(2) Graduate in Four Years	(3) Graduate in Four Years	(4) Graduate in Four Years	(5) Graduate in Four Years	(6) Graduate in Four Years	(7) Graduate in Four Years
Proportion of School's Courses CTE[a]	0.022	0.038	0.067**	0.060**	0.060**	0.064**	0.063**
	(0.018)	(0.031)	(0.029)	(0.029)	(0.029)	(0.029)	(0.028)
School Fixed Effects		X	X	X	X	X	X
Cohort Fixed Effects		X	X	X	X	X	X
8th Grade Math Score			X	X	X	X	X
8th Grade Reading Score				X	X	X	X
8th Grade Time Use Index					X	X	X
Sex and Race Fixed Effects						X	X
Free Lunch Fixed Effects						X	X
English Proficiency Fixed Effects						X	X
Change in Log Local Employment							X
N	399632	399632	399632	399632	399632	399632	399632
Adjusted R^2	0.000	0.041	0.108	0.113	0.113	0.124	0.124

Notes: Standard errors are clustered at the school by cohort level; * p<0.10, ** p<0.05, *** p<0.01. See notes in Table I for a description of key variables. [a]This is the proportion of school's courses that are CTE (across all grades) when a cohort is in 9th grade.

Table IX. Linear Probability Model of Proportion of Student's Transcript CTE in 9th Grade - Relationship with Proportion of School's Courses CTE

	(1) Proportion of Transcript CTE	(2) Proportion of Transcript CTE	(3) Proportion of Transcript CTE	(4) Proportion of Transcript CTE	(5) Proportion of Transcript CTE	(6) Proportion of Transcript CTE	(7) Proportion of Transcript CTE
Proportion of School's Courses CTE[a]	0.119***	0.046***	0.047***	0.047***	0.048***	0.045***	0.045***
School Fixed Effects		X	X	X	X	X	X
Cohort Fixed Effects		X	X	X	X	X	X
8th Grade Math Score			0.004***	0.005***	0.005***	0.002***	0.002***
8th Grade Reading Score				-0.001***	-0.001**	-0.001***	-0.001***
8th Grade Time Use Index					-0.005***	-0.002***	-0.002***
Female						-0.034***	-0.034***
Black, not Hispanic						-0.011***	-0.011***
Hispanic						-0.009***	-0.009***
Other Race, not Hispanic						-0.007***	-0.007***
Free Lunch						-0.001***	-0.001***
Formerly Limited English Proficient						-0.011***	-0.011***
Limited English Proficient						-0.001	-0.001
Change in Log Local Employment							-0.007
N	399632	399632	399632	399632	399632	399632	399632
Adjusted R^2	0.016	0.140	0.142	0.142	0.146	0.196	0.196

Notes: Standard errors are clustered at the school by cohort level; * p<0.10, ** p<0.05, *** p<0.01. See notes in Table I for a description of key variables. [a]This is the proportion of school's courses that are CTE (across all grades) when a cohort is in 9th grade.

Table X. Linear Probability Model of High School Graduation in Four Years - Reduced Form Relationship with Proportion of School's Courses in CTE Categories

Proportion of School's Courses[a]:	(1) Graduate in Four Years	(2) Graduate in Four Years	(3) Graduate in Four Years	(4) Graduate in Four Years	(5) Graduate in Four Years	(6) Graduate in Four Years	(7) Graduate in Four Years
Business and Finance	-0.438***	0.120	0.141	0.136	0.134	0.143	0.157
	(0.122)	(0.135)	(0.128)	(0.127)	(0.127)	(0.127)	(0.127)
Information Technology	0.615***	0.072	0.116	0.110	0.110	0.104	0.102
	(0.092)	(0.106)	(0.100)	(0.099)	(0.099)	(0.098)	(0.099)
Marketing	0.088	0.223*	0.290**	0.276**	0.269**	0.277**	0.265**
	(0.099)	(0.135)	(0.131)	(0.129)	(0.129)	(0.128)	(0.128)
Agriculture	0.115**	0.283**	0.242*	0.243**	0.246**	0.232*	0.241**
	(0.056)	(0.129)	(0.124)	(0.122)	(0.122)	(0.121)	(0.121)
Health	0.203**	-0.123	-0.011	-0.030	-0.030	-0.026	-0.027
	(0.080)	(0.132)	(0.122)	(0.122)	(0.122)	(0.122)	(0.121)
Trades	-0.157***	0.037	0.053	0.050	0.050	0.054	0.052
	(0.036)	(0.046)	(0.043)	(0.042)	(0.042)	(0.042)	(0.042)
Design	0.184**	0.003	0.069	0.063	0.064	0.084	0.085
	(0.083)	(0.134)	(0.124)	(0.123)	(0.123)	(0.122)	(0.122)
Technology and Engineering	-0.110	-0.167	-0.173	-0.171	-0.169	-0.153	-0.154
	(0.111)	(0.171)	(0.156)	(0.155)	(0.155)	(0.155)	(0.154)
Community College Courses	0.024	0.074	0.090	0.078	0.076	0.081	0.081
	(0.051)	(0.058)	(0.056)	(0.054)	(0.054)	(0.055)	(0.054)
School Fixed Effects		X	X	X	X	X	X
Cohort Fixed Effects		X	X	X	X	X	X
8th Grade Math Score			X	X	X	X	X
8th Grade Reading Score				X	X	X	X
8th Grade Time Use Index					X	X	X
Sex and Race Fixed Effects						X	X
Free Lunch Fixed Effects						X	X
English Proficiency Fixed Effects						X	X
Change in Log Local Employment							X
N	0.024	0.074	0.090	0.078	0.076	0.081	0.081
Adjusted R^2	(0.051)	(0.058)	(0.056)	(0.054)	(0.054)	(0.055)	(0.054)

Notes: Standard errors are clustered at the school by cohort level; * p<0.10, ** p<0.05, *** p<0.01. See notes in table I for a description of the variables.
[a]This is the proportion of school's courses that are CTE (across all grades) when a cohort is in 9th grade.

Table XI. Robustness and Specification Checks

	(1) Graduate in Four Years	(2) Graduate in Four Years	(3) Graduate in Four Years	(4) Graduate in Four Years	(5) Graduate in Four Years	(6) Graduate in Four Years	(7) Graduate in Four Years	(8) Graduate in Four Years	(9) Graduate in Four Years
Proportion of School's Courses CTE[a]	0.063**	0.053*	0.040	0.039	0.046	0.063**	0.076**	-0.107	0.051
	(0.028)	(0.031)	(0.026)	(0.034)	(0.030)	(0.029)	(0.030)	(0.090)	(0.032)
Proportion of School's Courses CTE[a] Squared								0.326**	
								(0.166)	
All Controls in Table 8	X	X	X	X	X	X	X	X	X
N	399632	399632	399632	399632	399632	399632	382958	399632	399632
Adjusted R^2	0.124	0.124	0.124	0.124	0.124	0.124	0.118	0.124	0.128

Notes: Standard errors are clustered at the school by cohort level; * p<0.10, ** p<0.05, *** p<0.01. [a]This is the proportion of school's courses that are CTE (across all grades) when a cohort is in 9th grade.
Notes by specification:
(1): Baseline specification, same as column 7 in Table 8.
(2): CTE measure excludes community college CTE courses.
(3): CTE measure includes Family and Consumer Sciences Courses.
(4): CTE measure is calculated at the course number level, not the course code level. E.g., Algebra I and Honors Algebra I are counted as two courses, not one.
(5): CTE measure is based only on course codes with at least 2 students enrolled.
(6): Specification controls for the availability of Family and Consumer Sciences and General Education CTE courses.
(7): Specification excludes cohorts with fewer than 50 students.
(8): Specification includes Proportion of School's Credits CTE Squared.
(9): Specification includes school time trends.

Table XII. WLS Model of Proportion of School's Courses CTE - Check for Exogeneity

	(1) Proportion of School's Courses CTE[a]	(2) Proportion of School's Courses CTE[a]
School Fixed Effects	X	X
Cohort Fixed Effects	X	X
Cohort Average 8th Grade Math Scores		-0.03226***
		(0.00948)
Cohort Average 8th Grade Reading Scores		0.03773***
		(0.01285)
Cohort Average 8th Grade Time Use Index		-0.00998
		(0.00828)
Cohort Proportion with Matched Test Scores		-0.00011
		(0.01299)
Cohort Proportion Female		-0.04121*
		(0.02162)
Cohort Proportion Black, not Hispanic		-0.00046
		(0.03067)
Cohort Proportion Hispanic		-0.05114
		(0.05490)
Cohort Proportion Other Race, not Hispanic		0.00443
		(0.05382)
Cohort Proportion Free Lunch		-0.01278
		(0.02420)
Cohort Proportion Limited English Proficient		-0.01716
		(0.07482)
Cohort Proportion Formerly Limited English Proficient		-0.03346
		(0.05767)
Change in Log Local Employment		-0.02132
		(0.03327)
N	2771	2771
Adjusted R^2	0.85715	0.80804

Notes: Standard errors are clustered at the school by cohort level; * p<0.10, ** p<0.05, *** p<0.01. [a]This is the proportion of school's courses that are CTE (across all grades) when a cohort is in 9th grade. See Table I for a description of the key variables. The weighted least squares models use cohort size as the weight.

Table XIII. WLS Model of Predicted Cohort Graduation Rate - Check for the Exogeneity of Proportion of School's Courses CTE

	Predicted Cohort Graduation Rate	Predicted Cohort Graduation Rate
Proportion of School's Courses CTE[a]		0.00008
		(0.02303)
School Fixed Effects	X	X
Cohort Fixed Effects	X	X
N	2771	2771
Adjusted R^2	0.81533	0.81524

Notes: Standard errors are clustered at the school by cohort level; * p<0.10, ** p<0.05, *** p<0.01. [a]This is the proportion of school's courses that are CTE (across all grades) when a cohort is in 9th grade. The predicted cohort graduation rate is the linear projection of graduation on the control variables in Table 12, excluding school and cohort fixed effects.

Table XIV. Linear Probability Model of High School Graduation in Four Years - Heterogeneity in the Reduced Form Relationship with Proportion of School's Courses CTE by Student-level Characteristics

	(1) Graduate in Four Years	(2) Graduate in Four Years	(3) Graduate in Four Years	(4) Graduate in Four Years	(5) Graduate in Four Years	(6) Graduate in Four Years	(7) Graduate in Four Years	(8) Graduate in Four Years
Proportion of School's Courses CTE[a]	0.063**	0.069**	0.085***	0.059**	0.061**	0.064**	0.102***	0.098***
	(0.028)	(0.030)	(0.029)	(0.029)	(0.029)	(0.028)	(0.028)	(0.031)
School's Courses CTE X Female		-0.011						-0.005
		(0.013)						(0.013)
School's Courses CTE X Black			-0.062***					-0.014
			(0.021)					(0.023)
School's Courses CTE X Hispanic			0.009					0.055*
			(0.029)					(0.031)
School's Courses CTE X Other Race			-0.021					0.004
			(0.027)					(0.028)
School's Courses CTE X 8th Grade Math				0.040***				0.036***
				(0.013)				(0.014)
School's Courses CTE X 8th Grade Reading					0.025**			-0.009
					(0.012)			(0.011)
School's Courses CTE X 8th Grade Time Use Index						0.004		-0.002
						(0.007)		(0.006)
School's Courses CTE X Free Lunch							-0.090***	-0.081***
							(0.018)	(0.018)
School Fixed Effects	X	X	X	X	X	X	X	X
Cohort Fixed Effects	X	X	X	X	X	X	X	X
Sex and Race Fixed Effects	X	X	X	X	X	X	X	X
Free Lunch Fixed Effects	X	X	X	X	X	X	X	X
English Proficiency Fixed Effects	X	X	X	X	X	X	X	X
8th Grade Math and Reading Scores	X	X	X	X	X	X	X	X
8th Grade Time Use Index	X	X	X	X	X	X	X	X
Change in Log Local Employment	X	X	X	X	X	X	X	X
N	399632	399632	399632	399632	399632	399632	399632	399632
Adjusted R^2	0.124	0.124	0.124	0.124	0.124	0.124	0.124	0.124

Notes: Standard errors are clustered at the school by cohort level; * p<0.10, ** p<0.05, *** p<0.01. See Table 1 for a description of key variables. [a]This is the proportion of school's courses that are CTE (across all grades) when a cohort is in 9th grade.

Table XV. Linear Probability Model of High School Graduation in Four Years - Heterogeneity in the Relationship with Proportion of School's Courses CTE by School-level Characteristics

	(1) Graduate in Four Years	(2) Graduate in Four Years	(3) Graduate in Four Years	(4) Graduate in Four Years	(5) Graduate in Four Years	(6) Graduate in Four Years
Proportion of School's Courses CTE[a]	0.070**	0.067**	0.069**	0.088	0.103**	0.108
	(0.029)	(0.027)	(0.029)	(0.066)	(0.041)	(0.074)
School's Courses CTE X School's Average Proportion Free Lunch		0.046				0.063
		(0.162)				(0.202)
School's Courses CTE X Change in Log Local Employment			0.450**			0.444*
			(0.226)			(0.229)
School's Courses CTE X 2nd Quartile of School's Average Graduation Rate				-0.027		-0.025
				(0.087)		(0.086)
School's Courses CTE X 3rd Quartile of School's Average Graduation Rate				0.052		0.043
				(0.088)		(0.091)
School's Courses CTE X 4th Quartile of School's Average Graduation Rate				-0.077		-0.073
				(0.078)		(0.092)
School's Courses CTE X In City with Population greater than 250,000					-0.063	-0.039
					(0.066)	(0.068)
School's Courses CTE X In City with Population between 100,000 and 250,000					0.057	0.076
					(0.091)	(0.093)
School's Courses CTE X In City with Population between 25,000 and 100,000					0.039	0.012
					(0.160)	(0.162)
School's Courses CTE X In Town with Population between 2,500 and 25,000					-0.144	-0.143
					(0.094)	(0.096)
School Fixed Effects	X	X	X	X	X	X
Cohort Fixed Effects	X	X	X	X	X	X
Sex and Race Fixed Effects	X	X	X	X	X	X
Free Lunch Fixed Effects	X	X	X	X	X	X
English Proficiency Fixed Effects	X	X	X	X	X	X
8th Grade Math and Reading Scores	X	X	X	X	X	X
8th Grade Time Use Index	X	X	X	X	X	X
Change in Log Local Employment	X	X	X	X	X	X
N	390852	390852	390852	390852	390852	390852
Adjusted R^2	0.120	0.120	0.120	0.120	0.120	0.120

Notes: Standard errors are clustered at the school by cohort level; * p<0.10, ** p<0.05, *** p<0.01. [a]This is the proportion of school's courses that are CTE (across all grades) when a cohort is in 9th grade. See Table 1 for a description of key variables.

2. THE OPPORTUNITY COST OF A COLLEGE EDUCATION: HOW SHOCKS TO LOCAL LABOR DEMAND AFFECT ENROLLMENT AND THE GENDER GAP

2.1 Introduction

In a standard human capital model, the opportunity cost of college enrollment is a small consideration relative to the lifetime benefits of a college degree. Yet, there is evidence that foregone earnings may be quite important to potential enrollees who lack information about the payoff to college (Bettinger, Long, Oreopoulos, and Sanbonmatsu (2012)), face credit constraints that force them to put a high value on current income (Lochner and Monge-Naranjo (2011)), or are trying to time their enrollment to the business cycle (Dellas and Sakellaris (2003)). In these cases, shocks to the opportunity cost of enrollment can have wide-ranging implications for enrollment decisions and, therefore, public policy.

In this study, I present new evidence that changes to the opportunity cost of enrollment resulting from local labor market demand shocks provides a causal explanation for why we observe a decrease in enrollment rates during booms and an increase during recessions. To isolate shocks to foregone earnings, I construct a "shift-share" labor demand index (see, for example, Bartik (1991)) for workers with a high school diploma but no college experience. For a given state in a given year, the index is a weighted average of *national* industry employment growth, where the weight is the proportion of workers in the industry with a high school diploma times the historical proportion of employment in the industry in the state. My findings using this index corroborate those of previous studies, which have relied primarily on local unemployment rates as a proxy for cyclical shocks to the opportunity cost of enrollment (for example, Dellas and Sakellaris (2003)). The use of local unemployment rates is potentially problematic because they are calculated using the labor force participation rate, which is a measure of labor supply. If there are labor supply shocks that are caused by changes in other factors that determine the demand for college education, estimates of the relationship between opportunity cost and enrollment will be biased. For example, an increase in financial aid offerings will increase the demand for college, draw people out of the low education labor force (a labor supply shock), and lower the unemployment rate. If financial aid increases (not caused by local labor market shocks) are positively correlated with strong demand for high school labor, estimates of the relationship between opportunity cost and enrollment will be biased downward. Because the index I construct does not use contemporaneous local labor market information, it is not vulnerable to such shocks.[2]

My primary data sources are the decennial Census and the March CPS. I construct the demand index using the 1980 Census for a state's baseline industrial composition and the 1989 to 2013 March CPS for national industry employment growth rates. I use 25 to 64 year olds with twelve years of schooling for both of the index's components. The enrollment data are for 18 to 24 year olds and come from the March CPS. Splitting the sample by age eliminates the possibility of a mechanical relationship between enrollment and the index, where increases in

[2] One approach to alleviating the concern about changes in the demand for college affecting unemployment rates is to use rates for older workers, which is what I do in the empirical work below. However, this solution is not entirely sufficient as these unemployment rates will provide unbiased results only if the supply of young low skill workers has no effect on the unemployment rate of older low skill workers. While young and old workers are probably not perfect substitutes, they are likely to be at least somewhat substitutable.

employment would necessarily result in decreases in enrollment.[3] I estimate that a one standard deviation decrease in the index causes the college enrollment rate to increase by 0.011 percentage points. This countercyclical relationship is robust to a variety of specification checks and tests, including adjustments for inter-state migration. The results I obtain using the index are quite similar to the results I obtain using other proxies for shocks to foregone earnings, including the high school unemployment rate, the high school wage, and the wage of high school occupations. My estimates corroborate both the direction and the magnitude of estimates using proxies that are vulnerable to labor supply shocks, suggesting that such bias is of little concern.

Next, I look for heterogeneity by gender in the enrollment response to local labor demand shocks and find evidence that men are more sensitive to the shocks than women are. The result highlights a phenomenon known as the college gender gap, which is that women currently enroll in college at higher rates than men do. Thus because men appear to be more sensitive to shocks to foregone earnings, the gender gap tends to shrink during recessions and grow during booms. One possible source of the heterogeneity is gender differences labor market demand. For example, workers in some industries and occupations are predominantly male or female. A demand shock to a male industry or occupation, such as construction, should matter more for the male enrollment rate than for the female enrollment rate. To explore whether gender differences in labor demand provide an explanation for gender heterogeneity, I

[3] This concern would be especially important if I were using local employment growth rates and not national ones.

construct gender-specific versions of the four proxies I use in the earlier analysis. I calculate male and female high school unemployment rates and wages; I categorize low education occupations as male-segregated, female-segregated, or gender-integrated and calculate their wages separately; and for the demand index, I calculate the historical industry shares separately for men and women. Replacing the original local labor demand proxies in the model of college enrollment with their gender-specific versions gives mixed results. In the end, it is difficult to say how much gender differences in labor demand contribute to gender heterogeneity and whether other factors, such as gender differences in psychic costs, may matter as well.

2.2 Theory and evidence in the literature

2.2.1 On the cyclicality of enrollment

As discussed earlier, it is not readily clear from a standard human capital model that enrollment should be responsive to the cyclical component of labor demand because foregone earnings should be relatively small compared to the return to attending college. The literature provides a number of reasons for why foregone earnings could matter to enrollment decisions.

First, some people may have a very high discount rate that makes the present value of the payoff to college relatively small, and the present value of current income relatively large (Becker (1993)).

Second, the distribution of college payoffs is wide and some may be uncertain about where in the distribution they will fall (Charles and Luoh (2003)). If people are risk averse, the uncertainty associated with the payoff to and the psychic costs of college may significantly reduce the discounted present value of the return to college, raising the relative importance of certain income right now. Stange (2012) finds evidence that college students do face uncertainty about the costs and benefits of attending. He shows that while in college, students learn both about their ability to complete it and the labor market opportunities associated with it. He estimates that the average high school graduate would be willing to pay about $15,000 (in 1992 dollars) to have the option to drop out of college over being required to complete it once they started.

Third, people may not know what the cost of attending is or how to get financial aid. Bettinger, Long, Oreopoulos, and Sanbonmatsu (2012) conducted an experiment where they randomly offered low income tax filers assistance with completing the Free Application for Federal Student Aid (FAFSA) for themselves or their children. The FAFSA is the first step a student takes in getting student loans for attending college. They also provided financial aid estimates and tuition information for local colleges. The authors found that high school seniors whose parents received assistance were 8 percentage points more likely to have completed two years of college. Thus access to information about college and its costs can make a large difference in enrollment decisions, especially for certain segments of the population.

Fourth, credit constraints could limit some individuals' ability to borrow from the future payoff to college, forcing them to finance tuition and current consumption from current income. Lochner and Monge-Naranjo (2011) demonstrate that credit constraints play an important role in college enrollment decisions, even with the availability of government student loans. They document that in recent years in the US, after conditioning on ability, college attendance is increasing in income. However, in the early 1980s, this was not the case. They show that changes in access to sufficient credit to cover tuition can explain the change. Thus, particularly for students from low income families, limits on their ability to finance tuition will make current income quite important for their enrollment decisions.

Finally, some may time their enrollment based on the business cycle (Dellas and Sakellaris (2003)). That is, during an economic boom, they may decide to take advantage of relatively high wages and wait to enroll until a downturn hits.

High discount rates, uncertainty or lack of information about returns and costs, and timing enrollment to match the business cycle all suggest that enrollment should be countercyclical. Credit constraints, on the other hand, suggest that enrollment should be procyclical because household incomes are higher during booms and there are more resources available to finance tuition.

Figure 2 shows the evolution of the college enrollment rate of 18 to 24 year olds since 1986. There is a clear upward trend, which I trace using a quadratic function, following Dellas and Sakellaris (2003). There is also evidence of cyclicality about the trend. Enrollment tends to be below trend leading up to a recession, jumps above trend after a recession, and slowly moves back below trend before the next recession hits. Figure 3 shows the relationship between the unemployment rate (the most common measure of labor market cyclicality in the literature) and the enrollment rate when both are detrended using a quadratic trend. As Figure 2 suggests, the two clearly move together. They have a correlation coefficient of 0.56.

Most studies empirically examining the cyclicality of enrollment agree with the visual evidence in Figures 2 and 3 (Barrow and Davis (2012), Betts and McFarland (1995), Black and Sufi (2002), Clark (2011), Dellas and Sakellaris (2003), and Rivkin (1995)), though some find only a weak or even no relationship (Card and Lemieux (2001) and Kane (1994)). Like this study, most rely on regional variation in unemployment rates for identification and many are able to include regional fixed effects because their data cover multiple years. Studies with a similar empirical setup to this one (Barrow and Davis (2012), Betts and McFarland (1995), Clark (2011), and Dellas and Sakellaris (2003)) estimate that a one percentage point increase in the unemployment rate raises enrollment by between 0.11 and 0.80 percentage points.

To date no study examining the enrollment response to cyclical labor demand has used a demand proxy that is not vulnerable to bias from labor supply shocks. Thus this study is the first study to provide causal evidence that college enrollment is countercyclical.

2.2.2 <u>On gender differences in college enrollment</u>

This study's finding that male enrollment is more responsive to labor demand shocks than female enrollment is points to a growing literature that seeks to understand gender differences in college enrollment more broadly. While women now go to college at higher rates than men, this has not always been the case. For birth year cohorts born before the early 1950s, men attended at higher rates than women did, especially cohorts born between 1920 and 1950. Figure 4 shows that in 1986, enrollment rates for 18 to 24 year olds were nearly equal for men and women, but since then, female enrollment growth has outpaced male enrollment growth. The difference in enrollment rates has steadily grown since 1986, as shown in Figure 5.

Explaining the reversal of the college gender gap in favor of women has been the focus of much of the literature to date. Goldin, Katz, and Kuziemko (2006, hereafter GKK) attribute the upward trend in the gender gap to two main factors: increasing women's labor force participation and declining discrimination against women. GKK provide evidence that more women entered the labor force as 1) social norms changed, 2) access to contraception reduced the number of childbearing years, and 3) rising demand for high-education workers increased the return to college and the opportunity cost of remaining out of the labor force. This trend led girls from younger generations to expect to spend most of their adult life in the labor force just as men do, further raising the return to college. Use data that asks students directly about future college attendance expectations, Fortin, Oreopoulos, and Phipps (2013) find that changes in such expectations can help explain why girls are attaining higher grades in high school than they used to.

GKK also speculate that rising women's labor force participation has helped to reduce discrimination against working women. They further speculate that discrimination may have decreased more in high education jobs, because higher educated people may have more progressive attitudes toward working women.

Much of the rest of the literature focuses on whether differences between women and men in the psychic costs of enrollment can explain both the trend and current level of the enrollment gap (for example, Aucejo (2014), Jacob (2002), and Becker, Hubbard, and Murphy (2010)). Women tend to possess higher levels of noncognitive skills,[4] which should make it easier to navigate an academic setting. Cornwell, Mustard, and Van Parys (2013) find evidence of this as early as in elementary school, where they show that differences in non-cognitive skills can help explain the gender gap in grades. Becker, Hubbard, and Murphy (2010, BHM hereafter) point to similar research in arguing that the distribution of noncognitive skills is wider for men than for women, so that more men are in the tails of the overall distribution. To explain the trend in the gender gap, BHM argue that increasing demand for college-educated workers has led to greater enrollment, so that those who are still choosing not to enroll are increasingly coming from the left tail of the noncognitive skill distribution. Because more men are in the tail, they enroll at lower rates.

There is also a growing literature that explains the level and trend in the gap by arguing that returns in the marriage market have grown for women relative to men (Chiappori, Iyigun,

[4] Noncognitive skills are "soft" skills such as motivation, self-control, mental focus, organizational skills, and collaborative skills. This contrasts with cognitive, or "hard" skills, such as problem solving, critical thinking, systematic thinking, adaptability, and creativity.

and Weiss (2009), Chiappori, Salanié, and Weiss (2010), Peña (2006), and Becker, Hubbard, and Murphy (2010)). Chiappori, Iyigun, and Weiss (2009) argue that when home production technology was less developed and social norms were different, women could get a return to education in the labor market or the marriage market, but not both. Today, women can get both returns, so that their return to college has risen relative to men. However, BHM argue that the empirical evidence is still too weak to support this story. In particular, they note that while low and high educated men used to marry at similar rates, high educated men are now more likely to be married, suggesting that the marriage market payoff to college has grown for both men and women.

While the rising labor force participation rate of women can explain the trend in the gender gap, it cannot explain why women now attend college at higher rates than men, especially because the overall labor force participation rate of men is still higher. Noncognitive skill differences can explain why more women go to college than men, but differences in labor market returns may play an important role as well. The literature exploring this possibility is conflicted. Jacob (2002) and Dougherty (2005) find that the labor market return to college is higher for women. Jacob (2002) finds that returns along with other labor market variables can explain about half of the gender gap. Charles and Luoh (2003) argue that the distribution of college returns is wider for men than women so that greater uncertainty reduces the expected utility of college for men. On the other hand, BHM (2010) and Hubbard (2011) argue that measured gender differences in the return in the earlier literature are the result of insufficient upward adjustments to topcoded wages in the CPS (which disproportionally affects men). After making more reasonable topcoding assumptions, Hubbard (2011) finds that there is little difference presently or historically. However, Hubbard (2011) focuses on differences in average returns. Given what Charles and Luoh (2003) find, it is possible that while average returns are close among men and women, more men are in the left tail of the distribution than women. Thus gender differences in labor market returns could matter even if average returns are the same.

This paper contributes to the literature on the college gender gap by providing evidence that because men are more sensitive than women are to cyclical labor demand, the gender gap tends to shrink during recessions and expand during booms. This paper also explores whether gender differences in labor demand can help explain the cyclicality of the gender gap, with inconclusive results.

2.3 A model of the college enrollment decision

I now present a theoretical model that serves to highlight the role of opportunity cost in the enrollment decision and to help clarify the factors that determine the strength of the relationship between changes in foregone earnings and enrollment rates. It is a human capital model based on the pioneering work of Becker (1993) and most closely related to the model in Jacob (2002).

To begin, assume that all individuals have graduated from high school and must decide whether to attend college. College attendance takes one period and comes with a guarantee of graduation. The costs of attending college are 1) direct costs d, which are tuition and fees net of financial aid, 2) psychic costs p, which are a negative function of noncognitive skills, and 3) opportunity costs or foregone high school earnings Y_{h0}, which depend on the prevailing wage. The benefits of attending college are purely monetary and depend on the difference between college and high school earnings. Define college earnings in period t as

$$Y_{ct}, t = 1, 2, \dots, T$$

and high school earnings in period t as

$$Y_{ht}, t = 1, 2, \dots, T.$$

Then the labor market return to attending college in period t is

$$\Pi_{ct} = Y_{ct} - Y_{ht}, t = 1, 2, \dots, T.$$

If we assume that credit markets are complete,[5] individuals know their return with certainty, and the interest rate is r, the present a value of the payoff to college is

$$\sum_{t=1}^{T} \frac{\Pi_{ct}}{(1+r)^t}.$$

Individuals will attend college if the benefits for periods $t = 1, 2, \dots T$ exceed the costs from the current period $t = 0$:

$$\sum_{t=1}^{T} \frac{\Pi_{ct}}{(1+r)^t} > Y_{h0} + d + p.$$

[5] As discussed in the literature review, this assumption may have significant implications for the importance of foregone earnings to the college enrollment decision. Complete capital markets mean that people can borrow from their full lifetime wealth to smooth consumption. If instead, people are credit constrained and must rely on current income to finance their education, current income can loom large and may actually have a positive effect on the probability of going to college. The effect of foregone earnings on college attendance then becomes an empirical question.

The college enrollment rate C for the current period, then, depends on the distribution of the net benefits of enrollment Z within the population:

$$C = \Pr(Z > 0) = Pr\left(\sum_{t=1}^{T} \frac{\Pi_{ct}}{(1+r)^t} - Y_{h0} - d - p > 0\right).$$

I am interested in how the enrollment rate changes in response to changes in a demand index I, which acts through shifting foregone earnings Y_{h0}. My empirical estimates will be of the derivative β, where

$$\beta = \frac{dC}{dZ}\frac{\partial Z}{\partial Y_{h0}}\frac{\partial Y_{h0}}{\partial I}.$$

There are a couple of important implications of this derivative. First, β is a constant only if all of the sub-derivatives are constant. By definition $\frac{\partial Z}{\partial Y_{h0}} = -1$. However, it is possible that the foregone earnings of some segments of the population are more responsive to changes in a demand index than others are, so that $\frac{\partial Y_{h0}}{\partial I}$ or $\frac{\partial log(Y_{h0})}{\partial I}$ are unlikely to be constants. Moreover, $\frac{dC}{dZ}$ will be a constant only if net benefits are uniformly distributed in the population. That is, for the relationship between the enrollment rate and a labor demand index to be linear, the population distributions of payoffs, foregone earnings, tuition, and psychic costs must all be uniform. Thus it is likely that my empirical estimates of β are linear approximations of a nonlinear function.

Second, because the distribution of net benefits (and therefore $\frac{dC}{dZ}$) depends on the population distribution of the four factors that determine enrollment, changes in the shape of their distributions or changes in their means (if they aren't uniformly distributed) will change the proportion of the population that is at the margin of enrollment, changing OLS estimates of β. For example, suppose that over time, skill biased technical change increases the average return to college, shifting the distribution of net benefits to the right. If net benefits are normally distributed and the initial enrollment cutoff is to the left of the mean, there will now be a smaller proportion of the population at the margin of enrollment, reducing estimates of β over time. A reduction in the average psychic or tuition costs of college would have a similar effect on β.

Now suppose that the payoff to college, foregone earnings, direct costs, and psychic costs can vary by gender so that enrollment rates are gender-specific:

$$C^g = \Pr(Z^g > 0) = Pr\left(\sum_{t=1}^{T} \frac{\Pi_{ct}^g}{(1+r)^t} - Y_{h0}^g - d^g - p^g > 0\right).$$

The college gender gap can arise from gender differences in any of the four terms. For example, if women have lower average psychic costs, they will have a higher enrollment rate.

To understand possible sources of gender heterogeneity in the response to changes in a local labor demand index, consider the derivative β^g:

$$\beta^g = \frac{dC^g}{dZ^g} \frac{\partial Z^g}{\partial Y_{h0}^g} \frac{\partial Y_{h0}^g}{\partial I^g}.$$

Gender heterogeneity can arise from differences in the distribution of net benefits $\left(\frac{dC^g}{dZ^g}\right)$ or differences in the response of foregone earnings to a demand index $\left(\frac{\partial Y_{h0}^g}{\partial I^g}\right)$. As noted earlier, the literature has largely focused on how gender differences in psychic costs affect net benefits and it possible for such differences to fully explain gender heterogeneity. For example, assume that net benefits are normally distributed with identical variances for both genders, but that the mean of the male distribution is less than the mean of the female distribution because men face greater average psychic costs than women do. If the margin of enrollment is in the left tail of both the male and female net benefits distributions, then more men will be at the margin than women will be, so that shifts in foregone earnings that move the margin will affect more men than women.

But while differences in psychic costs are sufficient to explain gender heterogeneity, it is also possible for gender-specific labor demand to explain the heterogeneity through either the foregone earnings or college returns channels. For example, men may tend to work in industries where employment demand is more cyclical, so that their foregone earnings are more cyclical. It could also be that male-specific high school labor demand is stronger than female-specific high school demand, creating a greater return to college for women. To explore these possibilities I model gender-specific enrollment using gender-specific labor demand proxies. The theoretical model above suggests that men should respond to male-specific demand proxies and women should respond to female-specific demand proxies. I find mixed evidence of this, highlighting the role that other factors, such as psychic costs are likely to play.

2.4 Data and Empirical Strategy

2.4.1 Data

I use the 1989 to 2013 March CPS as provided by IPUMS (King et al, 2010) for enrollment, demographic, and employment data. The CPS asks individuals age 16 to 24 whether they are enrolled in high school or college. Because so few people enroll in college before age 18, I limit the sample to those ages 18 to 24 with positive survey weights, yielding 391,772 observations.

I also use the March CPS to calculate state-level unemployment rates and average wages for those ages 25 to 64 and with twelve years of education. I focus on this population because it should best reflect the opportunity cost of attending college while limiting the influence of shifts in college demand on their measurement. I calculate wages using only full-time-full-year workers, because the 1994 redesign of the CPS had a meaningful impact on the measurement of employment and wages for women who work part time. Full-time-full-year workers must have reported usual hours of 35 or higher and have worked more than 50 weeks in the last year. I adjust topcoded wages using the scaling factors in Armour, Burkhauser, and Larrimore (2014), which go through 2007. For 2008 and later I use the average of the scaling factors from 2007 and earlier, which is about 2.4.

I classify occupations by their typical education level and gender and calculate the average wages for their typical worker at the state level. I classify an occupation as low education if from 1989 to 2013 more than 50% of its workers ages 25 to 64 had never attended college and classify it as high education otherwise. I classify an occupation as gender-segregated if more than 80% of its workers were of a single gender (this follows Blau, Brummund, and Liu (2012)). When I calculate the wages, I use only those whose education level is the typical education level. For example, for low education male occupation wages, I use only those with twelve years of schooling.

It is important to note that the CPS occupation classification scheme changes over time as new occupations develop and old occupations become obsolete. Every time the coding scheme changes, the CPS provides a crosswalk that maps old occupation codes to new ones. Thus it is possible to have a time-consistent occupational coding scheme by combining more specific occupations, where codes change, into more general occupations that are consistent over time. I use a time-consistent occupational classification scheme provided by IPUMS that is based on the 1990 Census Bureau classifications. IPUMS created 389 occupational categories from an original 514 in the 1990 Census Bureau scheme that are consistent from 1968 to present.

To proxy for the cost of attending college, I use data on statewide average costs from the Department of Education's Digest of Education Statistics. Data for average total costs (including room and board) for in-state public institutions are available starting in 1989. There are a handful of states where the data are not available for some of the years after 1989. I linearly interpolate and extrapolate the missing data for every state except the District of Columbia, which has no data.

2.4.1.1 Local Labor Demand Index

I construct the local labor demand index based on an approach used widely in the labor literature, commonly referred to as a shift-share model or the Bartik instrument.[6] The index predicts state-level employment growth by interacting a state's historical industry employment shares with national-level industry growth rates. I modify the index by weighting industry growth rates by education shares. This allows me to focus on demand for workers with twelve years of schooling and no college experience, which again, is the relevant population for representing the opportunity cost of enrollment. I use the 1980 US Census for historical industry shares and the 1989 to 2013 March CPS for national industry trends.[7] Formally, the demand index is

$$D_{ste} = \sum_{i} G_{it} \cdot \frac{E_{ite}}{E_{it}} \cdot \frac{E_{is0}}{E_{s0}},$$

where G_{it} is the national growth rate of industry i at time t, $\frac{E_{ite}}{E_{it}}$ is the employment share of education group e in industry i at time t at the national level, and $\frac{E_{is0}}{E_{s0}}$ is the employment share of industry i in state s in 1980. I calculate all components of the index using full-time-full-year workers ages 25 to 64. For ease of interpretation, I scale the index to have a standard deviation

[6] For examples, see Freeman (1980), Bartik (1991), Katz and Murphy (1992), Blanchard et al (1992), Charles, Hurst, and Notowidigdo (2013), and Schaller (2013).

[7] I base my industry categories on Katz and Murphy (1992) and Schaller (2013). They are: 1) agriculture, forestry, and fishing, 2) mining, 3) construction, 4) basic manufacturing (primary metals, fabricated metals, machinery, electrical equipment, automobile, other transportation equipment (excluding aircraft), tobacco, paper, printing, rubber, and miscellaneous manufacturing), 5) low-tech manufacturing (lumber, furniture, stone, clay, glass, food, textiles, apparel, and leather, 6) high-tech manufacturing (aircraft, instruments, chemicals, and petroleum), 7) transportation, 8) telecommunications, 9) utilities, 10) wholesale trade, 11) retail trade, 12) finance, insurance, and real estate, 13) business and repair services, 14) personal services, 15) entertainment and recreation services, 16) professional and related services, and 17) public administration.

of one. To calculate gender-specific versions of the index, I replace the historical industry employment shares in 1980 with gender-specific employment shares. That is, for the female version, I calculate the state industry employment shares in 1980 for females only, and for the male version, I calculate the state industry employment shares in 1980 for males only.

2.4.1.2 <u>Summary Statistics</u>

Summary statistics for the data used for the empirical models of college enrollment are shown in Table XVI. Thirty-four percent of the sample is enrolled in college. The average high school unemployment rate is 6.7 percent overall, with an average of 5.8 percent for women and an average of 7.4 percent for men. The average value of the demand index is 0.17 and it is slightly positive because it is based on employment growth rates.

Table XVII shows summary statistics for the components of the high school demand index. The first four columns are the shares of industry employment by education group $\frac{E_{ite}}{E_{it}}$ for the US for the entire sample period of 1989-2013. Those with twelve years of schooling but no college experience make up the largest share in many industries, though their share is especially large for construction, low-tech manufacturing, transportation, and retail trade. The fifth column is US industry shares as a percent of total employment $\frac{E_{is0}}{E_{s0}}$ in 1980. The largest industries in terms of employment are basic manufacturing, professional services, and retail trade. Combining the second and fifth columns gives the average weights for industry growth rates. The weights are scaled to add to 100 in the table, though that is not the case when the index is actually calculated. The final column gives the average yearly industry growth rates from 1989 to 2013. These rates reflect the well-known long run decline in basic and low-tech manufacturing and the long run growth of service industries.

While the high school unemployment rate, high school wage, low education occupation wage, and demand index are all reasonable proxies for local labor demand shocks, they do not necessarily move together. Table XVIII shows that some measures are much more correlated than others. The high school unemployment rate and demand index have a correlation coefficient of -0.49 and the high school wage and low education occupation wage have a correlation coefficient of 0.89. However, the wage measures are not closely correlated with the unemployment rate or demand index. Thus the wage measures appear to capture a different component of labor demand than the unemployment rate and demand index do.

2.4.2 <u>Empirical Strategy</u>

The primary empirical question this paper addresses is whether there is bias in estimating the relationship between college enrollment and local labor demand shocks when using standard measures of cyclical labor demand such as the unemployment rate or wages. Such measures may be contaminated by labor supply shocks that are caused by unobservable determinants of college enrollment. For example, an increase in financial aid offerings will increase the demand for college, draw people out of the low education labor force, raise wages, and lower the unemployment rate. If financial aid increases (not caused by labor market shocks) are positively correlated with strong demand for high school labor, then estimates of the relationship between labor demand shocks and enrollment will be biased downward.

The labor demand index that I construct is not vulnerable to local labor supply shocks because it uses state-level data as a baseline only and uses national-level data contemporaneously. In the context of the data used for this study, this means that state-level labor supply shocks in 1989 and later would have to be correlated with the industrial composition of states in 1980 to cause biased estimates. Returning to the financial aid example, using the demand index will result in biased estimates if a state's industrial composition in 1980 is a good predictor of whether there was a shock to financial aid in 1989. While it is impossible to prove that this is not the case, it is highly unlikely.

Because the local labor demand index is unique in not being vulnerable to local labor supply shocks, it may be tempting to use it to instrument one of the other demand proxies. But the very existence of multiple other demand proxies indicates that the instrumental variables estimates would be invalid because the demand index must act on enrollment solely through the endogenous regressor. For example, it would be invalid to instrument the unemployment rate using the demand index because the demand index is also likely acting on enrollment decisions through wages.

To estimate the relationship between local labor demand shocks and college enrollment I use a linear probability model. The model is

$$Enroll_i = \alpha + \beta ForegoneEarnings_{st} + \gamma Tuition_{st} + \delta \boldsymbol{X}_i + \theta_s + \lambda_t + \varepsilon_i.$$

where $Enroll_i$ is an indicator for whether an individual is enrolled and $ForegoneEarnings_{st}$ is one of the proxies for shocks to foregone earnings, either the high school unemployment rate, the high school wage, the low education occupation wage, or the high school demand index. $Tuition_{st}$ is the average total cost (tuition plus room and board) of attending an in-state college, which is intended to proxy for shifts in the supply of college. $\boldsymbol{X}_i$ is a vector of individual-level demographic controls**,** which contains gender, race, age, and metropolitan area status. I include year fixed effects λ_t to nonparametrically capture the upward trend in enrollment. I include state fixed effects θ_s to control for any long run state-specific factors that may affect enrollment, such as culture, industrial structure, and policy differences. I cluster standard errors at the state-by-year level.

2.5 Results

2.5.1 Main Results

I present results from the linear probability models of college enrollment for each of the four demand proxies in Tables XIX through XXII, which show how the coefficients evolve as I add additional controls. Table XIX shows the relationship between enrollment and the high school unemployment rate. Column 1 gives estimates of the relationship without any controls, and suggests that a 1 percentage point increase in the unemployment rate is associated with a 0.76 percentage point increase in the enrollment rate. The large difference in the coefficients between columns 1 and 2 indicates that unemployment and enrollment likely share a time trend and that states with higher long run unemployment rates also have higher enrollment rates. After adding state and year fixed effects, the coefficient on the high school unemployment rate is relatively stable. In column 3, the addition of demographic controls makes the model more precise and raises the estimate slightly. Adding the cost of college in column 4 raises the estimate slightly because the unemployment rate and college costs are positively correlated and higher college costs are negatively correlated with enrollment. The estimate in column 4, with all the covariates included in the model, suggests that a 1 percentage point increase in the high school unemployment rate raises the enrollment rate by 0.34 percentage points.

Results from the model of enrollment using the high school wage as the local labor demand proxy are shown in Table XX. In this model, the inclusion of state and year fixed effects is also important as states with higher wages tend to have higher college enrollment. The estimates are once again stable as demographic and tuition controls are added, and with all the controls included, I estimate that a 0.1 log point increase in the high school wage is associated with a 1.2 percentage point decrease in college enrollment. The low education occupation wage follows the same pattern as the high school wage. The results in Table XXI indicate that a 0.1 log point increase in the low education occupation wage is also associated with a 1.2 percentage point decrease in enrollment. The similarity in estimates is not surprising, as the only difference between the measures' universes is that the low education occupation wage excludes individuals who have twelve years of education but work in high education occupations.

The stability of the unemployment and wage estimates after adding state and year fixed effects indicates that conditional on the state and year fixed effects, the additional controls (which do increase the adjusted R-squared of the models) are not correlated with the demand proxies. This is a somewhat reassuring signal that the unobservable factors that determine enrollment may also be uncorrelated with local unemployment and wages. That said, stability is only a signal: it does not prove that there are not correlated unobservables. Because the demand index is not vulnerable to correlated unobservables, it provides new information that the other demand proxies cannot.

Table XXII shows the results of the model with the demand index as the local labor demand proxy. This time the estimates are stable across all specifications, including those without state and year fixed effects. In column 1, I estimate that a one standard deviation increase in the index causes the enrollment rate to decrease by 0.012 percentage points, and when I include all the controls in column 4, the estimate is 0.011 percentage points.

Because the demand proxies have different scales (with the exception of the wage proxies), it is difficult to compare their magnitudes. To make the proxies more comparable, I rescale the unemployment rate and wages to have a standard deviation of one. Columns 5 through 7 of Table XII show the results from the enrollment model with the rescaled proxies.

Interestingly, the magnitudes from all the proxies are quite close. While this suggests that there is little bias in the results from the unemployment rate and wage models, it is weakly suggestive, as the wages and the employment measures are barely correlated in the raw data, so that it is clear that they are capturing very different components of labor demand. Moreover, because the demand proxies are sample-based measures, their standard deviations are functions of both the population standard deviation and sampling error. That said, this finding, together with the stability of the estimates for all the proxies after including state and year fixed effects provides some reassurance that in a model of college enrollment, all the measures are reasonable proxies for labor demand.

2.5.2 Accounting for inter-state migration

One concern for the validity of my estimates is that people may migrate across state borders to go to college or in response to changes in local labor market conditions. To the extent that this is the case, my estimates are subject to measurement error. For example, if there is a large negative shock to the demand for high school workers in Michigan but not elsewhere, some of the people who are induced to go to college may choose to enroll in Indiana. If inter-state migration were prohibited, the enrollees would be appropriately counted in Michigan's enrollment rate. Instead, the enrollment rate increases in Indiana when its demand for high school workers did not change and the enrollment rate in Michigan does not go up enough. While the demand index is not vulnerable to migration because it uses historical and national data, mismeasurement of enrollment rates as described above will bias my estimates toward zero.[8]

While such bias is possible, there are two good theoretical reasons to suggest that migration bias is of little concern. First, migrating is expensive. For marginal college enrollees, migrating is particularly expensive because in-state tuition is subsidized. In addition, many marginal students live at home while attending college to save on room and board. Second, the average inter-state net migration rate for 18 to 24 year olds is small (around 0.1 percent per year), suggesting that even if there is bias, it is likely to be small as well.

While inter-state migration bias may be of little concern, the CPS has migration information that allows me to investigate empirically whether inter-state migration potentially biases my results. In every year except 1995, the CPS reports what state individuals lived in in the previous year. I adjust the CPS data by putting all the people who have moved across state lines over the last year back in their state of origin. The theory behind the adjustment is that an unbiased measure of the local enrollment rate should also consider the work and enrollment decisions of recent migrants. By returning people to where they lived a year ago, we can see whether there are differences between counting recent immigrants and counting recent emigrants. Table XIII shows the same progression of results as Table XXII but with the

---

[8] This is not classical measurement error in which measurement of a variable is randomly too high or too low with mean zero mismeasurement. In the scenario I describe, the outcome variable is either not negative enough or not positive enough, so that is it is always closer to zero than it should be.

migration adjustment. The number of observations is lower because I must exclude data from 1995. The results are nearly identical to those in Table XXII, providing some empirical evidence that bias from of inter-state migration is of little concern.

2.5.3 Gender Heterogeneity

In this section I explore whether there is gender heterogeneity in the effect of local labor demand shocks on college enrollment. I estimate the same set of models as before, but split the sample by gender. Table XXIV shows the results for each of the demand proxies. Every measure indicates that male enrollment is more responsive than female enrollment, with the exception of high school wages, which indicate that the responses are about equal. For example, I estimate that a one standard deviation increase in the demand index lowers the female enrollment rate by 0.6 percentage points, while it lowers the male enrollment rate by 1.4 percentage points. While none of the differences in coefficients between genders for any of the proxies are statistically significant, taken as a whole, the results are suggestive that men are more sensitive to local labor demand shocks than women are.

Why might this be the case? As I discussed in the theory section, gender differences in psychic costs are sufficient to explain the heterogeneity, but it is also possible for gender-specific labor demand to explain it, through either the foregone earnings or college returns channels. For example, because there are industries and occupations where workers are predominantly male or female, a demand shock to such industries or occupations should matter for the enrollment decisions of one gender more than for the other. To explore this possibility I model male and female enrollment separately using gender-specific labor demand proxies. The theoretical model I developed suggests that men should respond more to male-specific demand proxies and women should respond more to female-specific demand proxies.

Table XXV shows the results from gender-specific models of college enrollment where the demand proxies are also gender specific. The results are difficult to interpret. When using the male and female high school unemployment rates as the demand proxies, I find that men are more responsive to both rates but that they respond most strongly to the male unemployment rate. Women are unexpectedly unresponsive to the female rate, but quite strongly responsive to the male rate. When using the gender-specific high school wages as the demand proxies, the results are essentially equal for both genders for both the male and female wage. The occupation wage results are the most consistent with what the theory predicts. Here men are more responsive to the male occupation wage, women are more responsive to female occupation wage, and men are only somewhat more responsive to the integrated occupation wage. Finally, the gender-specific demand indexes are inconclusive, likely because they are so highly correlated.

Unfortunately, it is difficult to make any conclusions from these results. They do not rule out that gender differences in labor demand could contribute to the gender heterogeneity, but they also do not strongly confirm it. Gender differences in psychic costs could help explain why men appear to be more sensitive than women to the unemployment measures, but it could also be that women are more likely to move from employment into homemaking, as opposed to moving from employment into college enrollment.

2.6 Discussion

This study presents new evidence that local labor market demand shocks cause a decrease in enrollment rates during booms and an increase during recessions. Unlike previous studies of the relationship, I provide evidence using a demand index that is not vulnerable to contamination from labor supply shocks caused by changes in the demand for college. I estimate that a one standard deviation increase in the index causes the college enrollment rate to decrease by 0.011 percentage points. The results I obtain using the index are quite similar to the results I obtain using three other proxies for shocks to foregone earnings: the high school unemployment rate, the high school wage, and the low education occupation wage. My estimates, then, corroborate both the direction and the magnitude of estimates that use proxies that are vulnerable to labor supply shocks, suggesting that such bias is of little concern.

The study also presents evidence on heterogeneity by gender in the enrollment response to demand shocks and finds that men are more sensitive to the shocks than women are. The result has implications for the college gender gap. Because men are more sensitive to shocks to foregone earnings, this gap tends to shrink during recessions and grow during booms. I explore the possibility that the gender heterogeneity is the result of gender differences in labor demand by constructing gender-specific versions of the four proxies used in the earlier analysis. In the end, it is difficult to say how much gender differences in labor demand contribute to gender heterogeneity and whether other factors, such as gender differences in psychic costs, may matter as well.

This study provides additional evidence that foregone earnings play an important role in the college enrollment decision, even though the lifetime benefits of attending college can be substantial. If foregone earnings are important because potential enrollees lack information about the potential benefits and costs of college or because they are credit constrained, then there is room for public policy to do more to meet these needs.

2.7 Cited Literature

Armour, Philip, Richard V. Burkhauser, Jeff Larrimore. 2014. "Using the Pareto Distribution to Improve Estimates of Topcoded Earnings." *NBER Working Paper 19846.*

Aucejo, Esteban M. 2014. "Explaining Cross-Racial Differences in the Education Gender Gap." *Unpublished manuscript.*

Barrow, Lisa and Jonathan Davis. 2012. "The Upside of Down: Postsecondary Enrollment in the Great Recession." *Economic Perspectives* 36(4Q): 117-129.

Bartik, Timothy J. 1991. *Who Benefits from State and Local Economic Development Policies?* Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Becker, Gary S. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. University of Chicago Press: 1993.

Becker, Gary S., William H. J. Hubbard, Kevin M. Murphy. 2010. "Explaining the Worldwide Boom in Higher Education of Women." *Journal of Human Capital* 4(3): 203-241.

Bettinger, Eric P., Bridget Terry Long, Philip Oreopoulos, and Lisa Sanbonmatsu. 2012. "The Role of Application Assistance and Information in College Decisions: Results from the H&R Block FAFSA Experiment." *Quarterly Journal of Economics*: 1205-1242.

Betts, Julian R. and Laurel L. McFarland. 1995. "Safe Port in a Storm: The Impact of Labor Market Conditions on Community College Enrollments." *The Journal of Human Resources* 30(4): 741-765.

Black, Daniel A., Terra G. McKinnish, and Seth G. Sanders. 2005. "Tight Labor Markets and the Demand for Education: Evidence from the Coal Boom and Bust." *Industrial and Labor Relations Review* 59(1), 3-16.

Black, Sandra E. and Amir Sufi. 2002. "Who Goes to College? Differential Enrollment by Race and Family Background." *NBER Working Paper 9310*.

Blanchard, Olivier J. and Lawrence F. Katz. 1992 "Regional Evolutions." *Brookings Papers on Economic Activity* 23(1): 1-76.

Blau, Francine D., Peter Brummund, and Albert Yung-Hsu Liu. 2012. "Trends in Occupational Segregation by Gender 1970-2009: Adjusting for the Impact of Changes in the Occupational Coding System." *NBER Working Paper 17993.*

Boffy-Ramirez, Earnest, Benjamin Hansen, and Hani Mansour. 2013. "The Effect of Business Cycles on Educational Attainment." *Unpublished Manuscript.*

Card, David and Thomas Lemieux. "Dropout and Enrollment Trends in the Postwar Period: What Went Wrong in the 1970s?" In *Risky Behavior Among Youths: An Economic Analysis*, edited by Jonathan Gruber, 439-482. National Bureau of Economic Research.

Charles, Kerwin Kofi, Erik Hurst, and Matthew J. Notowidigdo. 2013. "Manufacturing Busts, Housing Booms, and Declining Employment: A Structural Explanation." *Unpublished manuscript.*

Charles, Kerwin Kofi and Ming-Ching Luoh. 2003. "Gender Differences in Completed Schooling." *The Review of Economics and Statistics* 85(3): 559-577.

Chiappori, Pierre-André, Murat Iyigun, and Yoram Weiss. 2009. "Investment in Schooling and the Marriage Market." *American Economic Review* 99(5): 1689-1713.

Chiappori, Pierre-André, Bernard Salanié, and Yoram Weiss. 2011. "Partner Choice and the Marital College Premium." *Columbia University Department of Economics Discussion Paper Series* No. 1011-04.

Clark, Damon. 2011. "Do Recessions Keep Students in School? The Impact of Youth Unemployment on Enrolment in Post-compulsory Education in England." *Economica* 78: 523-545.

Cornwell, Christopher, David Mustard, and Jessica Van Parys. 2013. "Noncognitive Skills and the Gender Disparities in Test Scores and Teacher Assessments: Evidence from Primary

Dellas, Harris and Plutarchos Sakellaris. 2003. "On the Cyclicality of Schooling: Theory and Evidence." *Oxford Economic Papers* 55: 148-172.

Dougherty, Christopher. 2005. "Why Are the Returns to Schooling Higher for Women than for Men?" *Journal of Human Resources* 40(4): 969-988.

Freeman, Richard B. 1980. "An Empirical Analysis of the Fixed Coefficient Manpower Requirements Model, 1960-1970." *Journal of Human Resources* 15(2): 176-199.

Fortin, Nicole M., Philip Oreopoulos, and Shelley Phipps. 2013. "Leaving Boys Behind: Gender Disparities in High Academic Achievement." *NBER Working Paper 19331*.

Goldin, Claudia, Lawrence F. Katz, and Ilyana Kuziemko. 2006. "The Homecoming of American College Women: The Reversal of the College Gender Gap." *Journal of Economic Perspectives* 20(4): 133-156.

Hubbard, William H. J. 2011. "The Phantom Gender Difference in the College Wage Premium." *Journal of Human Resources* 46(3): 568-586.

Jacob, Brian A. 2002. "Where the boys aren't: non-cognitive skills, returns to school and the gender gap in higher education." *Economics of Education Review* 21: 589-598.

Kane, Thomas J. 1994. "College Entry by Blacks since 1970: The Role of College Costs, Family Background, and the Returns to Education." *Journal of Political Economy* 102(5) 878-911.

Katz, Lawrence F. and Kevin M. Murphy. 1992. "Changes in relative wages, 1963-1987: Supply and demand factors." *Quarterly Journal of Economics* 107(1) 35-78.

King, Miriam, Steven Ruggles, J. Trent Alexander, Sarah Flood, Katie Genadek, Matthew B. Schroeder, Brandon Trampe, and Rebecca Vick. 2010. "Integrated Public Use Microdata Series, Current Population Survey: Version 3.0." [Machine-readable database]. Minneapolis: University of Minnesota.

Lochner, Lance J. and Alexander Monge-Naranjo. 2011. "The Nature of Credit Constraints and Human Capital." *American Economic Journal* 101: 2487-2529.

Peña, Ximena. 2006. "Assortative Matching and the Education Gap." *Unpublished Manuscript.*

Rivkin, Steven. 1995. "Black/White Differences in Schooling and Employment." *Journal of Human Resources* 30(4): 826-852.

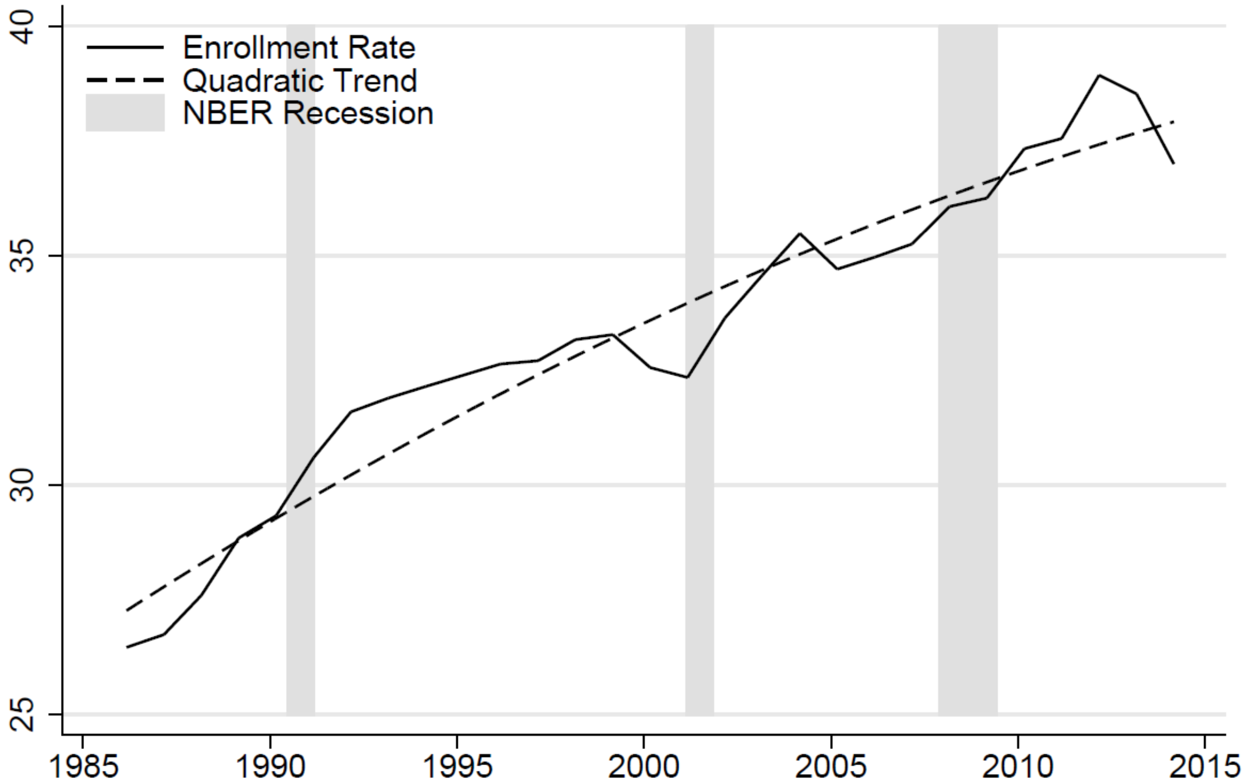
Schaller, Jessamyn. 2015. "Booms, Busts, and Fertility: Testing the Becker Model Using Gender-specific Labor Demand." *Unpublished Manuscript*.

Stange, Kevin M. 2012. "An Empirical Investigation of the Option Value of College Enrollment." *American Economic Journal: Applied Economics* 4(1): 49-84.

2.8 Figures and Tables

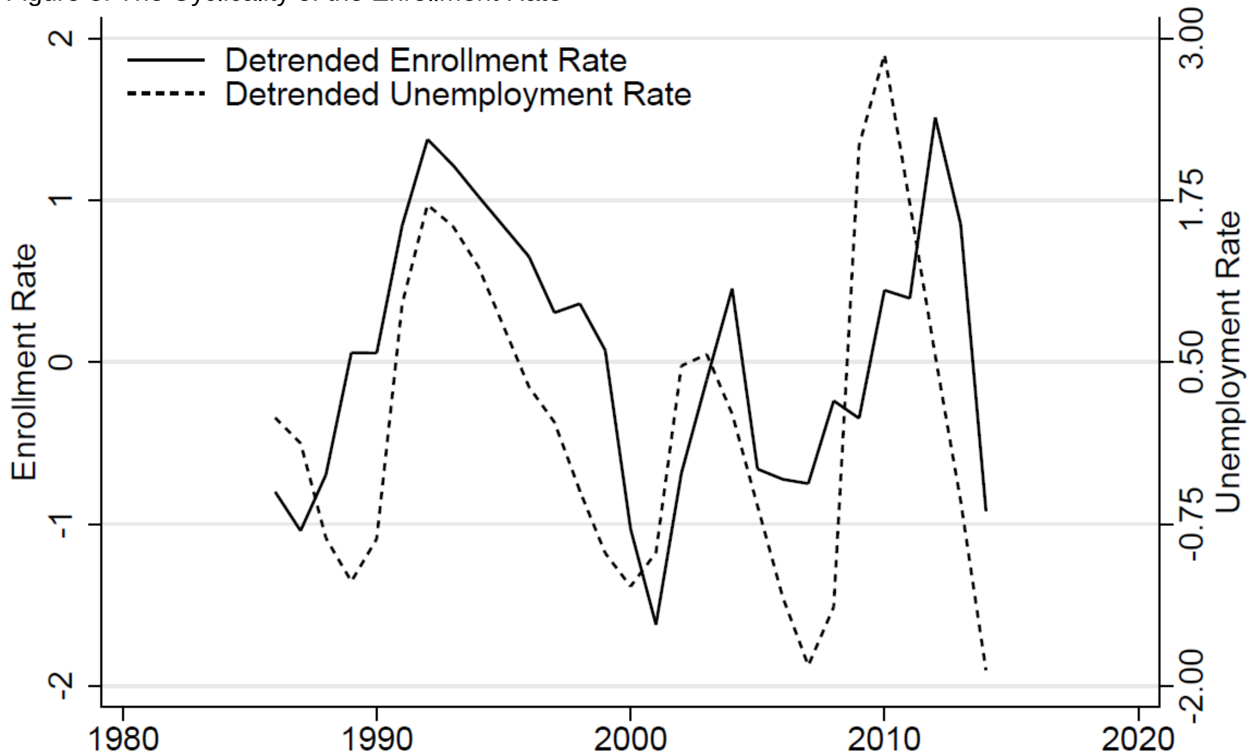
2.8.1 Figures

Figure 2. Enrollment Rate for 18 to 24 Year Olds



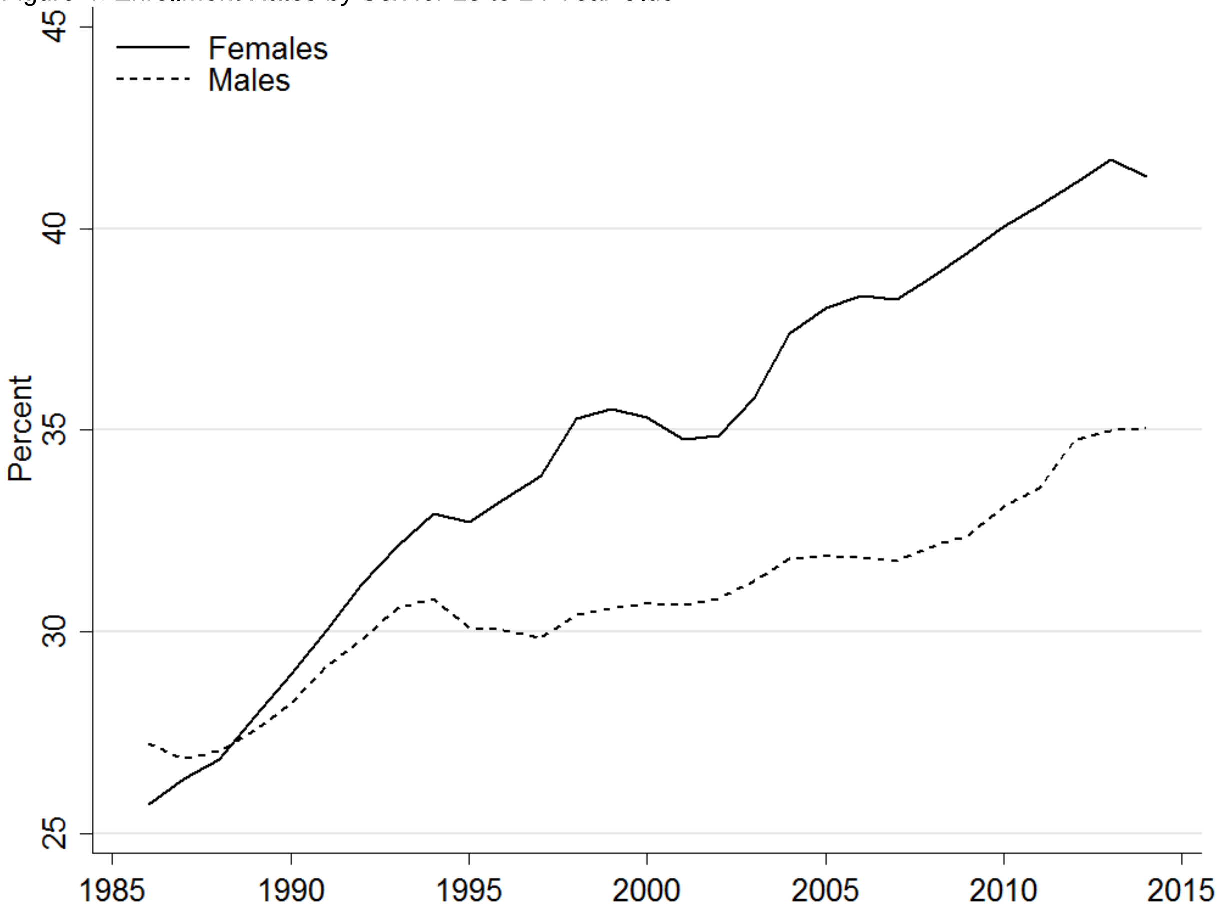
Source: March CPS

Figure 3. The Cyclicality of the Enrollment Rate



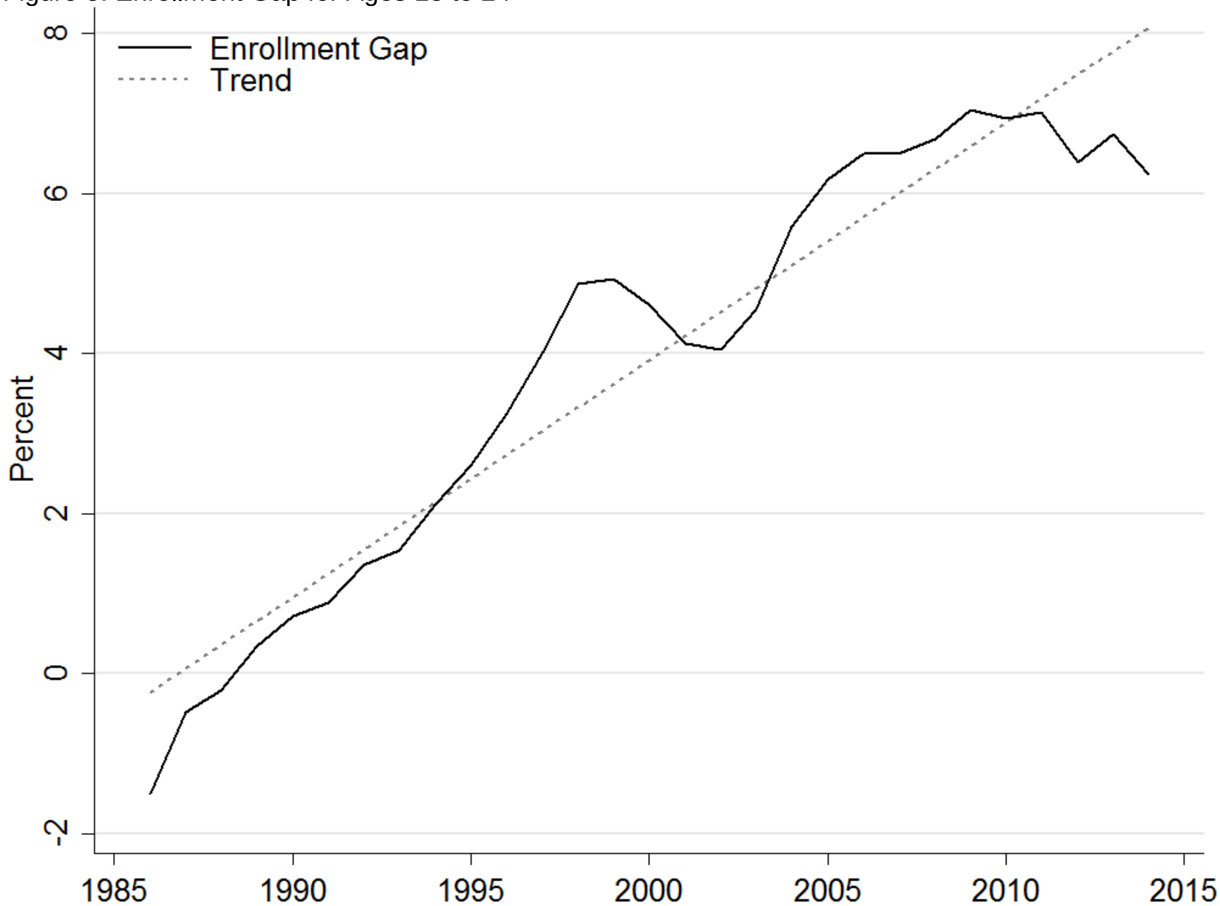
Source: March CPS

Figure 4. Enrollment Rates by Sex for 18 to 24 Year Olds



Source: March CPS

Figure 5. Enrollment Gap for Ages 18 to 24



Source: March CPS
Note: The enrollment gap is a three-year moving-average.

2.8.2 Tables

Table XVI. Summary Statistics

	Mean	Standard Deviation
Enrolled in College	0.340	0.474
Age	21.014	2.004
Male	0.501	0.500
White, Not Hispanic	0.636	0.481
Black, Not Hispanic	0.139	0.346
Hispanic	0.167	0.373
Other Race, Not Hispanic	0.059	0.235
Log Average Cost at Public College	9.500	0.271
Reside in Metropolitan Area	0.824	0.380
High School Unemployment Rate	0.067	0.030
Female High School Unemployment Rate	0.058	0.026
Male High School Unemployment Rate	0.074	0.037
Log High School Wage	2.815	0.085
Log Female High School Wage	2.657	0.097
Log Male High School Wage	2.929	0.089
Log High School Low Education Occupation Wage	2.742	0.094
Log High School Low Education Female Occupation Wage	2.419	0.148
Log High School Low Education Integrated Occupation Wage	2.611	0.111
Log High School Low Education Male Occupation Wage	2.908	0.106
High School Local Labor Demand Index	0.174	0.995
Female High School Local Labor Demand Index	-0.243	0.889
Male High School Local Labor Demand Index	-0.217	0.953
N	391772	

Notes: The sample is for those ages 18 to 24. Unemployment rates, wages, and the demand indexes are calculated at the state level using those ages 25 to 64 with 12 years of education. Wages are for full- time-full-year workers and are adjusted using the CPI to be in 2013 dollars. Topcoded wages are adjusted using the scaling factors in Armour, Burkhauser, and Larrimore (2014). For years when their scaling factors are unavailable, the average scaling factor for available years is used. Low education occupations are those where 50% or more of workers never attended college. Gender-segregated occupations are those where 80% or more of workers are of the same gender. See Table XVII for a description of how the demand indexes are calculated.
Sources: 1989 to 2013 March CPS and 1980 US Census.

Table XVII. Summary Statistics for Local Labor Demand Index Components

	Share of Industry Employment 1989-2013				Share of Total Employment 1980	High School Average Weight on Growth	Average Growth 1989-2013
	Dropout	High School	Some College	Bachelor's			
Agriculture, Forestry, and Fishing	23	39	21	17	3	3	0.4
Mining	11	42	25	22	1	2	1.5
Construction	16	45	25	13	5	7	1.7
Basic Manufacturing	10	40	26	25	17	20	-1.5
Low-tech Manufacturing	22	45	19	14	6	9	-1.7
High-tech Manufacturing	5	30	26	39	4	4	0.6
Transportation	8	43	32	17	6	7	1.3
Telecommunications	1	27	34	38	2	2	0.3
Utilities	6	40	32	23	2	2	0.0
Wholesale Trade	7	35	28	29	5	5	-0.6
Retail Trade	11	42	28	20	12	14	1.5
Finance, Insurance, and Real Estate	2	25	30	43	7	5	1.2
Business Services	9	30	27	34	3	3	2.9
Personal Services	16	41	27	16	2	2	1.0
Entertainment and Recreation	7	28	29	36	1	1	4.1
Professional Services	3	19	24	54	17	9	2.1
Public Administration	1	25	35	39	7	5	1.0
Overall	8	32	27	33	Sum = 100	Sum = 100	1.0

Notes: The demand index for a given state and year is a weighted average of national industry growth rates for the year, where the weights are the proportion of high school workers in the industry for the year at the national level times the proportion of overall employment in a state in each industry in 1980. Gender-specific versions of the index use gender-specific overall employment in a state in each industry in 1980. All statistics are for full-time-full-year workers age 25 to 64. High school average weight on growth is the high school share of industry employment from 1989-2013 times the share of total employment in 1980 scaled to add to 100. The division of manufacturing types is based on Katz and Murphy (1992). Basic manufacturing is primary metals, fabricated metals, machinery, electrical equipment, automobile, other transportation equipment (excluding aircraft), tobacco, paper, printing, rubber, and miscellaneous manufacturing. Low-tech manufacturing is lumber, furniture, stone, clay, glass, food, textiles, apparel, and leather. High-tech manufacturing is aircraft, instruments, chemicals, and petroleum.

Table XVIII. Correlations between Local Labor Demand Proxies

	High School Unemployment	High School Wage	Low Ed Occ Wage	Demand Index
High School Unemployment Rate	1.0000			
Log High School Wage	0.0825	1.0000		
Log Low Education Occupation Wage	-0.0015	0.8929	1.0000	
Demand Index for High School Workers	-0.4874	-0.0069	0.0152	1.0000

Note: See Tables XVI and XVII for a description of the variables.

Table XIX. Linear Probability Model of College Enrollment with the High School Unemployment Rate as the Proxy for Local Labor Demand

	(1) Enroll	(2) Enroll	(3) Enroll	(4) Enroll
High School Unemployment Rate	0.761***	0.316***	0.320***	0.335***
	(0.088)	(0.087)	(0.075)	(0.077)
State and Year Fixed Effects		X	X	X
Gender, Race, Age, and Urban Dummies			X	X
Average Total Cost In-state College				X
N	391772	391772	391772	391772
Adjusted R^2	0.002	0.011	0.088	0.088

Note: Standard errors are clustered at the state-by-year level; * p<0.10, ** p<0.05, *** p<0.01. See Tables XVI and XVII for a description of the variables.

Table XX. Linear Probability Model of College Enrollment with the High School Wage as the Proxy for Local Labor Demand

	(1) Enroll	(2) Enroll	(3) Enroll	(4) Enroll
Log High School Wage	0.203***	-0.104***	-0.124***	-0.124***
	(0.026)	(0.031)	(0.030)	(0.030)
State and Year Fixed Effects		X	X	X
Gender, Race, Age, and Urban Dummies			X	X
Average Total Cost In-state College				X
N	391772	391772	391772	391772
Adjusted R^2	0.001	0.010	0.088	0.088

Note: Standard errors are clustered at the state-by-year level; * p<0.10, ** p<0.05, *** p<0.01. See Tables XVI and XVII for a description of the variables.

Table XXI. Linear Probability Model of College Enrollment with the High School Low Education Occupation Wage as the Proxy for Local Labor Demand

	(1) Enroll	(2) Enroll	(3) Enroll	(4) Enroll
Log Low Education Occupation Wage	0.079***	-0.104***	-0.116***	-0.115***
	(0.024)	(0.023)	(0.022)	(0.022)
State and Year Fixed Effects		X	X	X
Gender, Race, Age, and Urban Dummies			X	X
Average Total Cost In-state College				X
N	391772	391772	391772	391772
Adjusted R^2	0.000	0.011	0.088	0.088

Note: Standard errors are clustered at the state-by-year level; * p<0.10, ** p<0.05, *** p<0.01. See Tables XVI and XVII for a description of the variables.

Table XXII. Linear Probability Model of College Enrollment with the High School Local Labor Demand Index as the Proxy for Local Labor Demand

	(1) Enroll	(2) Enroll	(3) Enroll	(4) Enroll	(5) Enroll	(6) Enroll	(7) Enroll
High School Local Labor Demand Index	-0.012*** (0.002)	-0.011* (0.006)	-0.010* (0.005)	-0.011* (0.005)			
High School Unemployment Rate (Rescaled)					0.010*** (0.002)		
Log High School Wage (Rescaled)						-0.010*** (0.003)	
Log High School Low Education Occupation Wage (Rescaled)							-0.011*** (0.002)
State and Year Fixed Effects		X	X	X	X	X	X
Race, Age, and Urban Dummies			X	X	X	X	X
Average Total Cost In-state College				X	X	X	X
N	391772	391772	391772	391772	391772	391772	391772
Adjusted R^2	0.001	0.010	0.088	0.088	0.088	0.088	0.088

Note: Standard errors are clustered at the state-by-year level; * p<0.10, ** p<0.05, *** p<0.01. See Tables XVI and XVII for a description of the variables. The high school unemployment rate, high school wage, and high school low education occupations wage are rescaled to have a standard deviation of 1.

Table XXIII. Migration-adjusted Linear Probability Model of Enrollment with the High School Local Labor Demand Index as the Proxy for Local Labor Demand

	(1) Enroll	(2) Enroll	(3) Enroll	(4) Enroll
Demand Index for High School Workers	-0.011***	-0.013**	-0.011**	-0.011**
	(0.002)	(0.006)	(0.006)	(0.006)
State and Year Fixed Effects		X	X	X
Race, Age, and Urban Dummies			X	X
Average Total Cost In-state College				X
N	378508	378508	378508	378508
Adjusted R^2	0.001	0.010	0.088	0.088

Note: Standard errors are clustered at the state-by-year level; * p<0.10, ** p<0.05, *** p<0.01. See Tables XVI and XVII for a description of the variables. The high school unemployment rate, high school wage, and high school low education occupations wage are rescaled to have a standard deviation of 1.

Table XXIV. Linear Probability Model of College Enrollment – Gender Heterogeneity by Demand Proxy

	(1) Female Enroll	(2) Male Enroll	(3) Female Enroll	(4) Male Enroll	(5) Female Enroll	(6) Male Enroll	(7) Female Enroll	(8) Male Enroll
High School Unemployment Rate	0.266***	0.412***						
	(0.088)	(0.096)						
Log High School Wage			-0.127***	-0.121***				
			(0.034)	(0.040)				
Log Low Education Occupation High School Wage					-0.107***	-0.125***		
					(0.026)	(0.030)		
Demand Index for High School Workers							-0.006	-0.014**
							(0.007)	(0.007)
State and Year Fixed Effects	X	X	X	X	X	X	X	X
Race, Age, and Urban Dummies	X	X	X	X	X	X	X	X
Average Total Cost In-state College	X	X	X	X	X	X	X	X
N	198765	193007	198765	193007	198765	193007	198765	193007
Adjusted R^2	0.091	0.083	0.091	0.083	0.091	0.083	0.091	0.083

Note: Standard errors are clustered at the state-by-year level; * p<0.10, ** p<0.05, *** p<0.01. See Tables XVI and XVII for a description of the variables. None of the differences in coefficients between genders is statistically significant at the p<0.10 for any of the proxies.

Table XXV. Linear Probability Model of College Enrollment – Gender Heterogeneity by Gender-specific Demand Proxy

	(1) Female Enroll	(2) Male Enroll	(3) Female Enroll	(4) Male Enroll	(5) Female Enroll	(6) Male Enroll	(7) Female Enroll	(8) Male Enroll
Female High School Unemployment Rate	0.040	0.175*						
	(0.079)	(0.094)						
Male High School Unemployment Rate	0.194***	0.230***						
	(0.070)	(0.071)						
Log Female High School Wage			-0.067**	-0.061**				
			(0.028)	(0.030)				
Log Male High School Wage			-0.066**	-0.059*				
			(0.029)	(0.033)				
Log Low Edu Female Occupation HS Wage					-0.031***	-0.020*		
					(0.011)	(0.012)		
Log Low Edu Integrated Occupation HS Wage					-0.038**	-0.046**		
					(0.018)	(0.021)		
Log Low Edu Male Occupation HS Wage					-0.036*	-0.058**		
					(0.021)	(0.023)		
Female High School Demand Index							0.010	0.001
							(0.011)	(0.012)
Male High School Demand Index							-0.013	-0.005
							(0.009)	(0.010)
State and Year Fixed Effects	X	X	X	X	X	X	X	X
Race, Age, and Urban Dummies	X	X	X	X	X	X	X	X
Average Total Cost In-state College	X	X	X	X	X	X	X	X
N	198765	193007	198765	193007	198765	193007	198765	193007
Adjusted R^2	0.091	0.083	0.091	0.083	0.091	0.083	0.091	0.083

Note: Standard errors are clustered at the state-by-year level; * p<0.10, ** p<0.05, *** p<0.01. See Tables XVI and XVII for a description of the variables.

3. Vita

Name:	Thomas William Walstrum
Education:	B.A., Economics, Wheaton College, Wheaton, Illinois, 2005
	M.A., Economics, University of Illinois at Chicago, Chicago, Illinois, 2015
	Ph.D., Economics University of Illinois at Chicago, Chicago, Illinois, 2015 Fields: Labor, Urban and Regional
Employment:	Federal Reserve Bank of Chicago, Chicago, Illinois Senior Associate Economist, Regional Analysis Group, 2010-2015 Business Economist, Regional Analysis Group, 2015-present
Publications:	Walstrum, Thomas. 2014. "Crowding Out – Servicing Illinois' Pension Debt. 2011. *Tax Facts*, 64(5), 1-7. Taxpayers' Federation of Illinois.
	Brave, Scott and Thomas Walstrum. 2014. "Estimating Marginal Treatment Effects Using Parametric and Semiparametric Methods." 2014. *Stata Journal*, 14(1), 191-271.
	Brave, Scott and Thomas Walstrum. 2014. "Using the Federal Reserve's Beige Book to Track Economic Activity." *Chicago Fed Letter*, No. 328.
Other Writing:	Lead author, Chicago Fed Beige Book. 2013-present.
Research Assistance:	National Community Investment Fund, Chicago, Illinois, 2010
	Federal Reserve Bank of Chicago, Chicago, Illinois, 2010-2015
	Taxpayers' Federation of Illinois, Chicago, Illinois, 2011
Teaching:	Principles of Macroeconomics, 2013