6-1-2015

Availability of Higher Education and Educational Outcomes: Quantifying the Impacts of College Openings and Cohort Size

Teny Maghakian Shapiro
Santa Clara University, tshapiro@scu.edu

Availability of Higher Education and Educational Outcomes: Quantifying the Impacts of College Openings and Cohort Size

Abstract

Research has established the benefits of higher education and the importance of affordability, however less is known about how the availability of higher education affects educational attainment. By constructing a comprehensive dataset on college openings in the U.S. from 1969 to 1991, I show that exogenous variation in two-year and four-year college availability, caused by changed birth cohort sizes and local college openings, substantially affects educational attainment. New four-year colleges increase the likelihood of obtaining a Bachelor's degree, while new two-year colleges only affect Associate's degree attainment. Additionally, results show that students from larger cohorts are crowded out of four-year colleges. This crowd-out results in lower lifetime educational attainment by pushing students to two-year colleges.

JEL: I21, I23

Keywords: higher education, college access, demand for schooling, cohort size, college openings

May 20, 2013

1 Introduction

As the demand for technology goods, health care, and professional services rises, so does the demand for skilled labor. In 2009, the Council of Economic Advisers projected that the demand for workers

with at least an Associate's degree will grow twice as fast as demand for workers without such a degree. In spite of long term increases in the number of college graduates,[1] researchers anticipate that demand for workers with a college education will soon outpace supply by around 300,000 per year (Carnevale, Smith, and Strohl 2010). Many policymakers have expressed concern that if this demand is not met, millions of American workers be unemployed, and wage inequality in the U.S. will continue to grow (Greiner, Rubart, and Semmler 2004, Acemoglu 1999). As a result of this expected shortage, higher education reform has become a centerpiece of the United States' education policy. For example, in 2010, President Obama proposed the American Graduation Initiative (AGI) which would invest $12 billion into existing as well as new community colleges. The bill did not pass through Congress, however, as it was criticized for being too lofty in its expected impact on higher educational outcomes. The ambiguity of the correct prescription to fill the shortage of skilled labor highlights the need for a precise understanding of how policies can increase the number of Americans pursuing higher education.

While economic research on higher education has focused on the benefits of higher education and the effects of financial incentives on the decision to attend college, less is known about how the availability of college education affects educational attainment. The late 20th century was a time of great changes in higher education in the U.S.– new public colleges opened around the country, federal financial aid became more generous, and the college-aged population was fluctuating. This study provides an examination of how individuals respond to variation in the availability of higher educational opportunities by taking advantage of the great changes during this time period. This paper is the first to provide a unified empirical framework that can be used to understand how college availability affects educational attainment; jointly assessing individual components of college availability studied in the previous literature. Prior literature on this topic has examined the effect of cohort size (Card and Lemieux 2000, Bound and Turner 2011), access to two- and four-year colleges (Rouse 1998), as well as new college openings (Currie and Moretti 2003) on educational

[1]Source: U.S. Census Bureau. Educational Attainment in the United States: 2010

attainment. These studies have found that cohort size and school supply are both individually important for educational attainment. But because each study focuses on only one component of availability, they have been unable to establish a comprehensive understanding of the effects of college availability. This study combines insights from each of these previous studies with the goal of establishing a better overall understanding of the important relationships between the different components of availability and educational attainment. This stems from jointly quantifying how the higher education system accommodated changes in cohort size (that is, a change in the number of people seeking a seat in college), how new college openings affected the level of education obtained, and assessing the pathways between two- and four-year colleges.

Identification stems from isolating over-time variation in college openings and cohort size, as well as establishing the correct markets for two- and four-year schools. I assemble and use a comprehensive dataset on college openings in the United States, and exploit the timing of college openings and plausibly exogenous variation in cohort size. This allows me to determine how educational outcomes differed for those facing different amounts of college availability at college age. I focus this study on cohorts who were age 17 between 1969 and 1991 — a time period with great variation in both cohort size and college openings. The panel nature of my dataset allows for the inclusion of local area fixed effects. Thus, my identification strategy relies on the fact that, within a local area, certain cohorts receive distinct shocks in the availability of college.

This paper contributes to the existing literature in three ways. First, I consider the effects of cohort size and college openings separately to determine the differential effects of these two important components of college availability. Second, I assess the impact of availability on different levels of educational attainment, such as Associate's degree and Bachelor's degree attainment. This provides important insights on the pathways between two- and four-year colleges. As much of the investment in higher education is currently focused on two-year colleges, understanding the impact of two-year college openings on all levels of educational attainment is necessary for ensuring that current policies are preparing students for the jobs demanded in the labor market. Third, I establish

the relevant geographic markets for two- and four-year colleges. I argue that the state is the correct market for four-year colleges, while a more local market is appropriate for two-year colleges.

For this study, I created a dataset on the universe of all colleges in the United States that allows me to exploit the across area and over time variation in the number of schools available. The dataset, which includes information about the institution type (public, private, for-profit; four-year, two-year; vocational, academic), geographic location, and opening year, was created using data from the Integrated Postsecondary Education Data System (IPEDS) and Peterson's Nelnet LLC (Peterson's).[2] These data allow me to identify precisely the number of each type of college in any geographic area in a given year. I combine these data with the 1980, 1990, and 2000 Censuses which has individual-level data on educational attainment, assigning each individual information about the size of their cohort and number of colleges available to them at the age of 17 based on their location of residence.

Results establish that the number of four-year and two-year colleges available to an individual at age 17 has a positive effect on educational attainment. An additional four-year college increases the likelihood of obtaining a Bachelor's degree by .1 percentage points. That is, one additional person for every thousand will obtain a Bachelor's degree because a college opens in their state when they are 17. An additional two-year college increases the likelihood of obtaining an Associate's degree by .15 percentage points, but has no effect on the likelihood of obtaining a Bachelor's degree. This suggests that two-year colleges neither divert individuals away from four-year colleges nor induce individuals to eventually obtain a four-year degree. The estimated effect of cohort size provides additional insights. A two percent increase in state cohort size leads to a one percent *decrease* in the likelihood of obtaining a Bachelor's degree and a 1.5 percent *increase* in the likelihood of

[2]Peterson's collects data from its annual surveys, which are sent to thousands of accredited colleges and universities. All data is submitted by officials at each school. In addition, many of the institutions that submitted data were contacted directly by Peterson's research staff to verify unusual figures, resolve discrepancies, or obtain additional data. Their data are used by the U.S. government and researchers from various industries and are regarded as being accurate, objective, and comprehensive.

obtaining an Associate's degree.[3] The inability of four-year colleges to accommodate additional students suggests that they are operating at capacity. Two-year colleges are able to absorb some of the students who are crowded out of four-year schools; however, these students do not ultimately transition into four-year colleges, decreasing lifetime educational attainment.

My results have important implications for education policy. The four-year college system in the U.S. operates at capacity, thus, is unable to accommodate an increase in the demand for higher education. While community colleges will be able to accommodate some of the students crowded out of four-year colleges, many students will be left unable to obtain the post-secondary training necessary for the skill-intensive jobs available. While opening both two- and four-year colleges has a positive effect on overall educational attainment, policy makers should keep in mind the labor needs of the economy as they decide which type of college to invest in. Opening new four-year colleges will increase the number of people with a Bachelor's degree. Two-year colleges increase the likelihood of obtaining an Associate's degree, but they have no effect on Bachelor's degree attainment. These results contradict the argument that community colleges facilitate bachelor's degree attainment by serving as low-cost substitutes for the first two years of a four-year education (Jacobs 2011).

The remainder of this paper unfolds as follows: Section 2 discusses the previous literature in detail and gives a brief background on college expansion in the United states. Section 3 describes the data used in the analysis. Section 4 discusses the statistical methods I will employ. Main results and alternate specifications are presented in Sections 5 and 6, respectively. I discuss policy implications and conclude in Section 7.

[3]The mean annual cohort size change during my sample period is two percent.

2 Background

2.1 Previous Studies

A small set of studies indicate that the availability of educational opportunities is important in determining an individual's educational attainment. Card and Lemieux (2000) as well as Bound and Turner (2011) consider the effect of the size of one's cohort on collegiate attainment. They find that individuals from larger cohorts are less likely to obtain a Bachelor's degree. The authors attribute this finding to four-year colleges being capacity constrained and unable to accommodate the demand from large cohorts. This finding sheds light on the importance of demand-side factors in affecting college availability. However, these studies neither control for the supply of colleges — which was changing rapidly during their sample periods — nor do they assess the impact of cohort size on other levels of educational attainment, such as Associate's degree attainment.

Rouse (1998) uses variation in two- and four-year college systems across states to determine which type of college is more efficient at increasing the educational attainment of young adults. Identifying from differences in the number of two- and four-year colleges per 10,000 high school graduates across states, Rouse (1998) finds that increased access to two-year colleges increases the likelihood of attending a two-year college and increased access to four-year colleges increases the likelihood of attending a four-year college. By looking at different levels of educational attainment, Rouse (1998) is able to better understand the pathways between two- and four-year colleges. However, such a cross-sectional study may be biased if differences in college access across states is due to differences in preferences for education. Exploiting the over-time variation in college access has been difficult to do since there does not exist a comprehensive dataset on the number of colleges in a given geographic area in a given year. Currie and Moretti (2003) created their own such dataset to assess the impact of college access on maternal education. They find that the number of colleges per capita in one's county at age 17 have a positive effect on educational attainment. However, since the goal of their analysis was to determine the effect of maternal education on infant health,

the first stage of their study was isolated to a sample of white mothers and did not thoroughly examine the effect of college access on different levels of educational attainment.

I combine insights from each of these previous studies with the goal of a providing a more comprehensive understanding of how cohort size and school supply each affect educational attainment. Unlike the previous literature, my regression includes both cohort size and the number of schools as separate covariates. This is in contrast to Card and Lemieux (2000) and Bound and Turner (2011) who look just at cohort size as well as Rouse (1998) and Currie and Moretti (2003) who include a colleges-per capita measure. My distinction with the two latter studies is important for a couple reasons. First, including the number of colleges and cohort size as separate variables allows for more flexibility and less bias in the parameter estimates. Second, it allows me to identify the over-time variation in the number of colleges —the impact of new college openings—as distinct from over-time variation in cohort size. There are likely important differences in the effect of variation in each on educational attainment.

As in Rouse (1998), I assess the impact of college availability on different *levels* of educational attainment. This allows me to assess the pathways between two- and four-year colleges. By extending my analysis to also include those with 12 years of education or less, I can also determine what types of students are drawn into higher education as a result of improvements in availability. I also establish the relevant geographic markets for two- and four-year colleges. Currie and Moretti (2003) and Rouse (1998) use different geographic markets — the county for the former and the state for the latter — in their analyses. I argue that the state is the correct geographic market for four-year colleges, while a more local area is appropriate for two-year colleges. Identifying correct markets reduces measurement error and improves identification in my analysis.

2.2 Defining the Geographic Market and Relevant Cohorts

In the analysis that follows, it is important to define the correct geographic market for each type of college as well as the relevant age groups that compete for seats in colleges within the market. I argue that relevant cohort is 17 to 20 year olds, and the relevant geographic market is the state for four-year public colleges and the county (Census Public Use Microdata Area(PUMA) for this analysis) for two-year public colleges.

Four-year colleges are largely under the control of state governments which control funding and outline plans and goals for higher education.[4] Only a small fraction of state higher education revenues come from the federal government and private endowments.[5] In turn, four-year colleges are most accessible to those living within the state. For example, in-state tuition is highly subsidized and admission standards are often lower for in-state students. Most students attending four-year colleges do not travel far from their homes, as the median distance between school and home for four-year college students is estimated to be between 30 and 94 miles.[6] Because identification in my analysis is based on students at the margin of collegiate attainment—students that Hoxby (1997) and Bound and Turner (2011) point out are unlikely to be apart of the highly integrated and national market for the premier universities—I can be less concerned about variation in the national market for higher education. It is also important to note that considering a smaller geographic area, such as county, to identify the effect of a four-year college opening may induce error, as individuals from different parts of the state move to the county for the purpose of attending the school or working at or near the school. Since two-year colleges, also referred to as community colleges, were created to increase the accessibility of higher education, they are built in locations where students can live

[4]See examples at http://www.sheeo.org/links/links_results.asp?regionID=53&issueID=17.

[5]Source: IPEDS Finance Tables

[6]Source: ESM Chaperone analysis of United States Department of Education National Center of Education Statistics

at home while attending college, further decreasing the cost of college attendance.[7] Accordingly, the average distance between school and home for those attending a public two-year college is 12 miles.[8] Because two-year colleges attract students from the local area, I assess the impact of two-year college access on educational attainment for the local market, defined at the PUMA.

I consider persons aged 17 to 20 as the relevant cohort of individuals competing for a seat in higher education. This age group was chosen because colleges likely have a fixed capacity for the entire school as opposed to a specific entering class, and those are the age groups that will be of college-going age when the individuals in our analysis are entering college. As most two-year schools have open-admissions policies, the relevant age group for two-year colleges is less clear.[9]

3 Data

3.1 The Dataset

This study uses a dataset that I created by combining multiple sources of data on institutions of higher education. The Integrated Postsecondary Education Data System (IPEDS) provides information about 7,150 institutions of higher education in the United States, that include institutional characteristics, such as address, degrees offered, and proprietorship. These characteristics allow me to identify whether a college is a two-year or four-year institution, and whether it is public, private, and for-profit. For years 1980 and 1984-2009, IPEDS also has data on annual enrollment at each of the institutions.

[7] Two-year colleges also allow part-time enrollment, low tuition (the average two-year college tuition is less than one-half that at public four-year colleges (Kane and Rouse 1999)), and an open admission policy– often not even requiring a high school diploma to enroll.

[8] ESM Chaperone analysis of United States Department of Education National Center of Education Statistics show that the average distance between school and home for those attending public two-year colleges is 12 miles.

[9] I also explored alternate definitions of the relevant cohort. Defining the cohort to be only 17 year olds or 17-18 year olds does not affect results.

IPEDS does not have information regarding the year in which each institution opened. Since institutions are not necessarily added to the IPEDS database the first year they are operating, inferring opening years from the IPEDS data is not a good strategy. Instead, I obtained a proprietary dataset of each institution's founding year from Peterson's and matched the data to the IPEDS database using each school's Office of Postsecondary Education Identification Number (a unique identification number given by the U.S. Department of Education). The merged dataset allows me to identify the number of each type of college in each zip code. For my analyses, I then aggregate the data to the appropriate geographic level. I focus on four-year and two-year public colleges, excluding vocational schools, graduate schools, and two-year colleges that do not offer associate's degrees. Focusing the analysis on public colleges strengthens identification for two reasons. First, public colleges account for a large proportion of collegiate enrollment; approximately 95% of students attending a two-year college and 70% of students attending a four-year college are enrolled at a public institution.[10] Second, the market for private colleges is a national one; just over half of all students who attend private colleges do so in another state. To the contrary, only 10% of students attend a public college outside their home state.[11]

To assess how changes in the availability of higher education affect individuals' educational attainment, I use data from the five percent sample of the 1980, 1990, and 2000 Censuses, which have individual-level data on educational attainment as well as demographic information. The quality of the educational attainment variables varied from census to census. The 1980 Census is the least detailed and offers information on the years of education obtained, but not the degree obtained. The 1990 and 2000 Censuses include information on degree obtained, which allows me to more precisely define educational attainment. For my main specification, I will use measures of educational attainment consistent across all Censuses, but will narrow the dataset to the 1990 and 2000 Censuses in robustness specifications. As I discuss further in Section 4, I define the market

[10]Source: U.S. Census. "Higher Education – Institutions and Enrollment 1980 to 2009"

[11]Source: National Center for Education Statistics

for four-year colleges to be the state and the market for two-year colleges to be a more local area. The smallest geographic area identified in the Census is the PUMA (Public Use Microdata Area). PUMAs generally follow the boundaries of county groups or single counties. If these areas exceed 200,000 residents, they are divided into as many PUMAS of 100,000+ residents as possible. PUMAs do not cross state lines. Since PUMAs vary from Census to Census, I use the Consistent PUMA — the most detailed geographic areas that can consistently be identified across samples from 1980 onward. Thus, PUMAs should be thought of as county groupings with a median population of approximately 250,000.

My analyses include individuals of all races. The sample includes individuals aged 25 to 47, since I only want to observe people after they have likely completed their schooling. The sample of the college opening data used is limited to 1969-1991. The cohort that was 17 in 1991 is the "youngest" cohort I am able to observe as adults since the 2000 Census is the most recent available. I assign each individual in the Census four measures of college availability from the year they were 17 years old: 1) the number of people in their cohort in their state, 2) the number of people in their cohort in their PUMA, 3) the number of four-year public colleges in their state, and 4) the number of two-year public colleges in their PUMA. I also assign PUMA-level economic characteristics from the Bureau of Economic Analysis for the year in which they were 17. Unfortunately, the Census does not specify where the individuals lived when they were 17 years old, thus, I assume that individuals live in the same PUMA at the time of survey as they did when they were 17. I account for potential movers by excluding everyone born outside the United States and those who lived in a different state five years before the survey from my sample. Additionally, if a cohort is surveyed in more than one Census, I only include data from their first survey year, when they are less likely to have migrated away from their home state/PUMA. The population data used in this study comes from the National Cancer Institute's Surveillance Epidemiology and End Results(SEER) population dataset. SEER population estimates are available at the county-level by age, race, and gender beginning in 1969 and are considered to be the most accurate intercensal estimates available.

With these data, I am able to measure the size of one's cohort at both the state and PUMA level in each year[12].

3.2 Descriptive Statistics

Before assessing the impact of college access on educational attainment, I begin by showing some stylized facts about college openings and cohort size. Understanding changes in the number of colleges and cohort size over the sample period, 1969-1991, is important not only for understanding the political economy of public higher education expansion in the U.S., but also for determining the source of identification in the later analyses. Table 3 shows summary statistics for the census data used in the analyses. Statistics are also shown separately for cohorts experiencing a college opening at age 17, by cohort size, and demographic group. Most notably, larger cohorts and cohorts experiencing an opening have more public four-year colleges, on average. Cohorts experiencing an opening are also larger than average.

Figures 1 and 2 shows the time-series variation in the number of colleges and the population of 17-20 year olds from 1969-1991, respectively. Specifically, Figure 1 shows the number of two-year and four-year colleges during this time period. There are some notable differences between four-year and two-year college openings during this time. This may reflect the smaller size of two-year colleges or the need to have more schools as each school services a very specific geographic area. The fastest growth in both types of colleges occurred in the first part of the sample period. While four-year college growth stagnated beginning in 1978, growth of two-year colleges persisted throughout the 1980s. The population of 17-20 year olds peaks in the early 1980s, and declines fairly steadily afterward. The growth in population during the 1970s is reflective of the baby boom generation entering college-going age.

[12]In the Data Appendix, I detail how I aggregate county-level variables to the PUMA level.

The rise in the number of colleges during the 1970s (and also in the 1960s) was due to society's growing commitment to creating equality of opportunity in higher education (Rouse 1994, Thelin, Edwards, and Moyen 2011). The President's Commission on Higher Education in a Democracy and the Servicemen's Readjustment Act (The "G.I. Bill") ensured federal funds for universities. In addition to federal funding, growing states with enthusiastic governors and legislatures sought ways to accommodate an impending enrollment boom from the rising birth rate, increased migration, and deliberate extensions of college admissions (Rouse 1994, Thelin, Edwards, and Moyen 2011). Many states saw two-year colleges as a way to accommodate the enrollment demand; "community colleges could serve the dual purpose of allowing states to preserve the quality of the four-year institutions while not having to exclude individuals from higher education" (Rouse 1994).

To get more insight into the geographic variation in college openings, Figures 3 and 4 show the geographic distribution of the college openings from 1969-1991. States and PUMAs without any openings are in white. In Figure 3 the second lightest states are those with one opening and the darkest states are those with three or more openings. Twenty-two states had at least one four-year college opening during this time period.[13] Figure 4 shows the PUMAs in the U.S. with a two-year college opening. A drawback of using PUMAs in this analysis they cover a great deal of land area in less populated areas. In some states, such as New Mexico and Wyoming, the PUMA is the entire state. To account for this potential source of measurement error, in a robustness specification I limit my sample to PUMAs smaller than 7900 square miles and find that results are quantitatively similar.[14]

The validity of my research design requires exogeneity of cohort size and college openings within the appropriate geographic area. One may be concerned that college openings are correlated with demand for higher education and that it is this demand that is leading to differential trends in

[13]California and Texas had the most college openings. This is unsurprising, as they were both relatively young, growing, states with a clearly organized multi-tier public higher education system that directed the growth during this time period. For example, California's Master Plan of 1960 aimed at accommodating mass access to affordable higher education by channeling students into tiered institutions.

[14]7900 square miles is approximately the land area of a circular area with a radius of 50 miles.

educational attainment, not the college opening itself. To explore this potential source of bias, I compiled characteristics of the states and counties in 1960, years prior to the first opening in my sample. I begin by using these "pre" characteristics to predict whether or not a state (county) will experience a four-year (two-year) college opening at any time during my sample period. The dependent variable in the probit regression is an indicator variable for whether or not the state (county) had at least one four-year (two-year) college opening. The independent variables include log of the population, land area (in square miles), percent change in the population between 1950 and 1960, percent of the population that is living in an urban area, percent of land used for farming in 1964, birth rate in 1968, median education, unemployment rate, and median family income. 1 shows the results from these regressions. For the county-level regressions, I show results from models with and without state fixed-effects. I find that states with larger populations, more land area, higher percent of non-white population, and higher median family income are more likely to have a four-year college open between 1969 and 1991. Population growth between 1950 and 1960 and median education are negatively correlated with college openings. At the county-level, population, land area, unemployment rate, and median family income are positively correlated with two-year college openings during this time period.

Since my identification relies on over-time, rather than cross-sectional, variation in college availability, I next assess whether the "pre" characteristics can predict the timing of a state's (county's) college opening, conditional on having an opening during this time period. The dependent variable in this regression is the year the opening occurred and the independent variables are the same as above. 2 shows the results from these regressions. I find that none of the covariates individually explain the timing of four-year colleges within a state. However, these variables are jointly significant. For two-year college openings, I find that more populated counties open colleges earlier in the time period, while counties with a more urban population open colleges later. The covariates are not jointly significant at the county-level for either specification. While the state-level regressions show jointly statistically significant impacts of state characteristics on the timing of four-year college

openings, much of the variation in the timing of college openings remains unexplained. To account for possible trends within states that are correlated with college openings, I include a state linear time trend in my main specifications [15]

4 Methods: Determining the Effects of Access on Attainment

4.1 Estimation

The identification strategy used in this study exploits variation in both the number of schools and the size of the cohort over time within geographic area. Identification of the effects of school availability on educational attainment rests on how variation in these variables impacts individuals at the margin of receiving additional years of schooling. To do so, I estimate the following specification:

$$\begin{aligned} Education_{y-17pcrg} &= \alpha + \beta_1 NumPub4_{ys} + \beta_2 NumPub2_{yp} + \mu_1 \ln(Pop1720)_{ys} \\ &\quad + \mu_2 \ln(Pop1720)_{yp} + \delta X_{yp} + \gamma_r + \gamma_g + \gamma_{y-17c} + \gamma_p + \gamma_s * TREND + \epsilon_{y-17pcrg}, \end{aligned}$$

where $Education_{y-17pcrg}$ represents the educational attainment of a person observed in census year c, born in year $y-17$ that currently resides in PUMA p, which is in state s, is of race r and gender g. $NumPub4_{y+17,s}$ is the number of four-year public colleges in state s in year y, the year the individual is 17 years old. $NumPub2_{yp}$ is the number of two-year public colleges in PUMA p in year y. $\ln(Pop1720)_{ys}$ and $\ln(Pop1720)_{yp}$ are the population of 17 to 20 year-olds in year y in state s and PUMA p, respectively. X is a vector of economic variables included in the specification

[15] I include interactions of these 1960 pre-treatment state characteristics with time trends and find that my main findings are not impacted by the inclusion of these trends (Hoynes and Schanzenbach 2009).

in order to control for the conditions in the PUMA when the individual was 17 years old. It includes the employment to population ratio, per capita income, and the fraction of employment in construction and manufacturing. Unobservable differences in educational attainment that vary by age specific to a census year are controlled for with birth year-census year fixed effects, γ_{y-17c}. Unobservable differences in educational attainment between PUMAs are controlled for with PUMA fixed effects, γ_p. I also include race and gender fixed effects to control for unobservable differences in educational attainment across race and gender. Finally, I include a cohort trend specific to each state, which is intended to control for any unobservable variable specific to the state that would cause educational attainment to trend in a certain direction. As discussed above, this may pick up variation in the in-state demand for higher education that could be correlated with the decision to open new schools. Additionally, the trend captures variation in state financial aid policy (Cheslock and Hughes 2011, Doyle 2012).[16] Because I have included PUMA fixed effects, β_1 provides a measure of the effect of a change in the number of public four year schools in the state on the educational outcome of the person, while β_2 measures the effect of a change in the number of public two-year schools in the PUMA. The coefficients μ_1 and μ_2 provide a measure of the effect of a change in cohort size at the state and PUMA level, respectively. To account for arbitrary correlation in $\epsilon_{y-17pcrg}$ between people in the same state, I report standard errors clustered by state. The specification exploits variation the availability of higher education between cohorts within the same PUMA, holding fixed any differences attributable to demographic factors such as age, race, and gender. Thus, I am identifying differences in the educational outcome of persons within the same PUMA, who are of the same age and gender and were surveyed by the Census in the same year, but had different educational opportunities available at age 17.

To assess the impact of availability on higher educational outcomes, I consider multiple measures of educational attainment. These include a continuous variable of years of education completed, and indicator variables for whether they completed less than 12 years of schooling, obtained only a high

[16]Results are robust to including a PUMA specific cohort trend instead of a state trend.

school diploma, obtained an Associate's degree, or obtained a Bachelor's degree.[17] Assessing the impact of college availability on these different levels of educational attainment allows for important insights on the pathways between high school, community colleges, and four-year colleges.

5 Results

The first set of results are shown in Table 4. To show the importance of including control variables, I report results after progressively adding each type of fixed effect and state control. Estimates stabilize once PUMA, birth year-Census year, and race fixed effects are included in the specification. Column (5) is my preferred specification, as it includes a state-cohort trend which controls for other factors potentially correlated with both college openings and educational attainment, such as increasing demand for education.

Results indicate that the effect from a four-year college opening increases average educational attainment by approximately .012 years. The effect from a two-year college opening is smaller in magnitude, .004 years, and is not statistically different from zero. As in Bound and Turner (2011) and Card and Lemieux (2000), cohort size is negatively correlated with years of education. This effect, however, is statistically significant only for the PUMA cohort. Specifically, the estimate on state cohort implies that a two percent increase in the PUMA cohort size decreases average years of schooling by .18 years.[18] These estimates imply that even small changes in cohort size can have substantial effects on educational attainment that can't be compensated for with a single college opening.

To understand at what margin of educational attainment students are affected by changes in availability, I assess how college openings and cohort size affect different levels of educational

[17] As previously discussed, the 1980 Census does not specify degree attainment. Because of this, I assume that those with exactly 12 years of education have earned a high school diploma, exactly 14 years of education have earned an Associate's Degree, and exactly 16 years of education have earned a Bachelor's degree.

[18] I use two percent to describe changes in cohort size, since that is the mean yearly change in state cohort size during this time period.

attainment. Estimates are reported in Table 5. For ease of comparison with the previous table, column (1) reproduces estimates where the dependent variable is years of education. In Columns (2) - (5), the dependent variables are a indicator variable equal to one if the individual has less than 12 years of education, a high school diploma only, an Associate's degree only, and at least a Bachelor's degree, respectively. Results suggest that both cohort size and the number of colleges available affect educational attainment. Specifically, individuals experiencing a four-year college opening at the age of 17 are less likely to have a high school diploma only or less than 12 years of education. They are also more likely to obtain a Bachelor's degree. Approximately one extra person for every thousand will get a college degree because a college opens when they are 17. In a populated state such as California, this would result in approximately 550 additional individuals receiving a Bachelor's degree from a new college opening.

While the number of two-year colleges does not have a statistically significant effect on years of education, individuals with more two-year colleges available to them at the age of 17 are less likely to obtain less than 12 years of education, and more likely to obtain an Associate's degree. Approximately one and a half extra persons for every thousand will obtain an Associate's degree because a two-year college opens when they are 17. The number of two-year colleges does not have a statistically significant effect on obtaining a Bachelor's degree. This is important to note as researchers and policy makers have long been interested in determining the impact of community colleges on four-year college enrollment. While community colleges give those who might not have otherwise attended four-year college the ability to do so (referred to as the democratization effect), they may also attract students who might otherwise have attend a four-year college (the diversion effect) (Rouse 1995). These results suggest that while two-year colleges weren't diverting individuals out of four-year colleges during this time period, they were also not inducing individuals to eventually obtain a Bachelor's degree.

Cohort size affects different levels of educational attainment. Consistent with previous research, I find that individuals from large state cohorts are less likely to obtain a Bachelor's degree (Card and

Lemieux 2000, Bound and Turner 2011). A two percent increase in state cohort size decreases the probability of obtaining a Bachelor's degree by .18 percentage points, or approximately one percent at the mean. However, individuals from larger cohorts are *more* likely to obtain an Associate's degree or high school diploma only. Specifically, a two percent increase in state cohort size increases the probability of obtaining an Associate's degree by .10 percentage points (1.1 percent at the mean) and the probability of obtaining a high school diploma only by .16 percentage points (.5 percent at the mean). These results are very similar in magnitude to those of Bound and Turner (2011) who argue that large cohorts have lower college enrollment rates because of capacity constraints, leaving many young people crowded out of the four-year institutions.[19] PUMA cohort size does not have a statistically significant effect on the likelihood of obtaining these levels of education.

I find positive, but statistically insignificant, effects on state cohort size on the likelihood of obtaining an Associate's degree. Understanding the relationship between cohort size and two-year degree attainment is important for understanding the educational decisions of those who are potentially crowded out of four-year colleges. A positive and significant relationship between cohort size and Associate's degree attainment would refute previous theories that individuals from large cohorts are less likely to obtain a college education because of reduced returns to education, as two-year college attainment would also be negatively correlated with cohort size if this were the case (Murphy, Plant, and Welch 1988, Nothaft 1985, Stapleton and Young 1988, Welch 1979). This finding highlights the importance of understanding the effects of college availability on different levels of educational attainment, and not just bachelor's degree attainment.

6 Robustness Checks and Alternative Specifications

Because the 1980 Census provides less precise measures of educational attainment than the 1990 and 2000 Censuses, I next narrow my sample to exclude the 1980 Census. This allows me to

[19]Specifically, Bound and Turner (2011) find that a 10% increase in state specific cohort size decreases college completion rates by about 4%.

define educational attainment more precisely. Since I now observe the degree attainment of each individual, I no longer make any assumptions about the degree an individual received. Results are shown in Table 6 and are very similar to my main specification[20] The largest difference is that when I isolate my sample to the 1990 and 2000 Censuses, I find a positive and significant relationship between the number of four-year colleges in one's state and Associate's degree attainment. This may be a result of individuals pursuing higher education as a result of the college opening, but not successfully transferring to a four-year college.

As shown earlier in the paper, PUMAs vary greatly in land area as a result of low population. Some states, such as Wyoming and New Mexico, contain only one PUMA. To account for the large variance in PUMA land area, I narrow the sample to PUMAs that are 7900 square miles or smaller. This is approximately equal to a land area with a radius of 50 miles. Table 7 shows the estimates from this subsample. The main difference is that the effect of two-year college openings on Associate's degree attainment diminishes while the effect of PUMA cohort size on Associate's degree attainment increases.

In their analysis using college access as an instrument of maternal health, Currie and Moretti (2003) use a per capita measure of college access. I alter my specification to include the number of four-year colleges per 1,000 17-20 year olds in the state and the number of two-year colleges per 10,000 17-20 year olds in the PUMA instead of controlling for the number of colleges and cohort size separately. Estimates from this specification are shown in Table 8. I find that one additional four-year college per 1,000 17-20 year olds in a state increases years of education by 4.9 years and increases the likelihood of obtaining a Bachelor's degree by 78 percentage points. Unfortunately, Currie and Moretti (2003)'s measure of college availability is not as intuitive when considering the state to be the market for four-year colleges. One additional college per 1,000 17-20 year olds in the state would be equivalent to opening 47 more four-year colleges in the average

[20]When using the measures of educational attainment from my main analysis and isolating the sample to the 1990 and 2000 Censuses, results are very similar to Table 6. This dismisses concern that imprecision in the education measures are driving results.

state. The average state population of 17-20 year olds is roughly 470,000 and the average number of four-year colleges in a state is 16.7. Therefore, our estimate can roughly be thought of as approximately a 2 percentage point increase in the probability of obtaining a Bachelor's degree from an additional college opening. These estimates are not directly comparable to those of Currie and Moretti (2003) since their analysis of four-year colleges was at the county-level and their sample was isolated only to white mothers. However, they find that one new four-year college in one's own county per 1,000 18-22 year olds would increase the probability that these mothers had a college education by 19 percentage points. I also find that the number of four-year colleges per capita decreases the likelihood of obtaining only a high school diploma. There is also a negative, yet statistically insignificant, relationship between four-year colleges per capita and Associate's degree attainment. Additionally, a new two-year college per capita increases the likelihood of Associate's degree attainment by 5.1 percentage points. This equates to roughly a 2.2 percentage point increase in the likelihood of getting an Associate's degree from one more college opening in the average PUMA. These estimates are in line with those from my main specification which decomposes the effect of college openings and cohort size. Since I find relatively small effects from college openings and large effects from cohort size, the effect of this measure of college access is driven primarily by changes in population.

7 Discussion and Conclusion

The increasing demand for skilled labor requires a sound understanding of the factors that affect individuals' decisions to obtain higher education. This study provides an examination of the determinants of higher education by assessing how individuals responded to variation in the availability of higher educational opportunities during a time of great change in the higher education system in the United States. This study combines insights from previous studies with the goal of establishing a better overall comprehension of the important relationships between the different components

of availability and educational attainment. This stems from jointly quantifying how the higher education system accommodates changes in cohort size (that is, a change in the number of people seeking a seat in college), how new college openings affect the level of education obtained, and assessing whether there exist pathways between two- and four-year colleges. Identification stems from isolating over-time variation in school openings and cohort size, as well as establishing correct markets for two- and four year schools. I assemble and use a comprehensive dataset on college openings in the United States, and exploit the timing of college openings and plausibly exogenous variation in cohort size. This allows me to determine how educational outcomes differ for those facing different amounts of college availability at college age 17.

Results from this analysis add to the small existing body of literature on the determinants of higher educational attainment, and provides insight on what policy makers can do to increase the proportion of students that go on to higher education. Results suggest that the four-year college system during my sample period, 1969 -1991, was unable to accommodate students from larger cohorts, or, in other words, increases in the demand for four-year college education. Some of those crowed-out of four-year colleges attended two-year colleges, while others do not obtain any higher education. As the demand for higher education increases due to labor market forces, more and more students may be crowded out of four-year colleges. Expanding the capacity of existing colleges may be a first step in satisfying the unmet demand for a Bachelor's degree. However, since I am unable to use more recent college availability data in this study, an important next step in this research is understanding how the *current* higher education system reacts changes in the demand for college education. As community colleges are becoming increasingly overcrowded, they may be unable to accommodate the students crowded out of four-year colleges. Additionally, the growing for-profit and online college sectors may play a role in accommodating the excess demand for higher education in the current system.

I also find that opening new colleges induces those who would have otherwise not gone to college to obtain a degree. While opening new campuses is not financially feasible for many states, these

findings offer important insights into the unmet demand for higher education. I find that opening a new four-year college increases the number of Bachelor degree holders, suggesting that there is an unmet demand for college education. Similarly, opening a new two-year college increases the number of those with Associate's degrees. While the goal of many community colleges is to offer an inexpensive way for students to complete the first two years of college and then transition into four-year institutions (Rouse 1994), my results suggest that opening new two-year colleges does not lead to more individuals attaining bachelor's degrees. This may be addressed by streamlining the transfer process between two-year and four-year colleges. These results suggest that the decision to investment in each type of college should depend on the needs of the future labor market. That is, if the market demands workers with Bachelor's degrees, then the investments should be in four-year colleges. If the demand is for those with Associate's degrees, then investments should be in two-year colleges. Carnevale, Smith, and Strohl (2010) project that between 2007 and 2018, there will be a two percentage point increase in both the demand for workers with Associate's degrees and Bachelor's degrees. Thus, directing new investments only at two-year colleges, as proposed in the American Graduation Initiative, will not likely completely satisfy the needs of the labor market.

References

Acemoglu, D. (1999): "Changes in Unemployment and Wage Inequality: An Alternative Theory and Some Evidence," *American Economic Review*, 89(6), 1259–1278.

Bound, J., and S. Turner (2011): "Cohort Crowding: How Resources Affect Collegiate Attainment," *Journal of Public Economics*, Forthcoming.

Card, D., and T. Lemieux (2000): *Dropout and Enrollment Trends in the Post-War Period: What Went Wrong in the 1970s?*University of Chicago Press.

Carnevale, A. P., N. Smith, and J. Strohl (2010): "Help Wanted: Projections of Jobs and Education Requirements through 2018," *Georgetown University Center on Education and the Workforce*.

Cheslock, J. J., and R. P. Hughes (2011): "Differences Across States in Higher Education Finance Policy," *Penn State Center for the Study of Higher Education Working Paper*, (5).

Currie, J., and E. Moretti (2003): "Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College openings," *Quarterly Journal of Eonomics*, VCXVII(4), 1495–1532.

Doyle, W. R. (2012): "The Politics of Public College Tuition and State Financial Aid," *The Journal of Higher Education*, Forthcoming.

Greiner, A., J. Rubart, and W. Semmler (2004): "Economic Growth, Skill-Biased Technical Change and Wage Inequality: A Model and Estimations for the U.S. and Europe," *Journal of Macroeconomics*, 26, 597–621.

Hoxby, C. (1997): "How the Changing Market Structure of U.S. Higher Education Explains College Tuition," *NBER Working Paper 6323*.

Hoynes, H. W., and D. W. Schanzenbach (2009): "Consumption Responses to In-Kind Transfers: Evidence from the Introduction of the Food Stamp Program," *American Economic Journal: Applied Economics*, 1(4), 109–139.

Jacobs, J. (2011): "A College Degree on the Cheap," *Commnity College Spotlight – The Hechinger Report*.

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

Murphy, K., M. Plant, and F. Welch (1988): *Economics of Changing Age Distribution in Developed Countries*chap. Cohort Size and Earnings in the USA. Clarendon Press: Oxford.

Nothaft, F. (1985): "The Effect of Cohort Size on Human Capital Investment and Earnings Growth," *Working Paper 21. Washington, DC: Board of Governors of the Federal Reserve System*.

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

Rouse, C. E. (1994): *Choices and Consequences: Contemporary Policy Issues in Education*chap. What to Do after High School: The Two-Year versus Four-Year College Enrollment Decision. Cornell University Press.

——— (1998): "Do Two-Year Colleges Increase Overall Educational Attainment? Evidence from the States," *Journal of Policy Analysis and Management*, 17(4), 595–620.

Stapleton, D. C., and D. Young (1988): "Educational Attainment and Cohort Size," *Journal of Labor Economics*, 6(3), 330–361.

Thelin, J. R., J. R. Edwards, and E. Moyen (2011): "Higher Education in the United States - Historical Development," *Education Encyclopedia - StateUniversity.com*.

Welch, F. (1979): "Effects of Cohort Size on Earnings: The Baby Boom Babies' Financial Bust," *The Journal of Political Economy*, 87(5), S65–S97.

Figure 1: The Number of Two- and Four-Year Public Colleges in the U.S.: 1969-1991

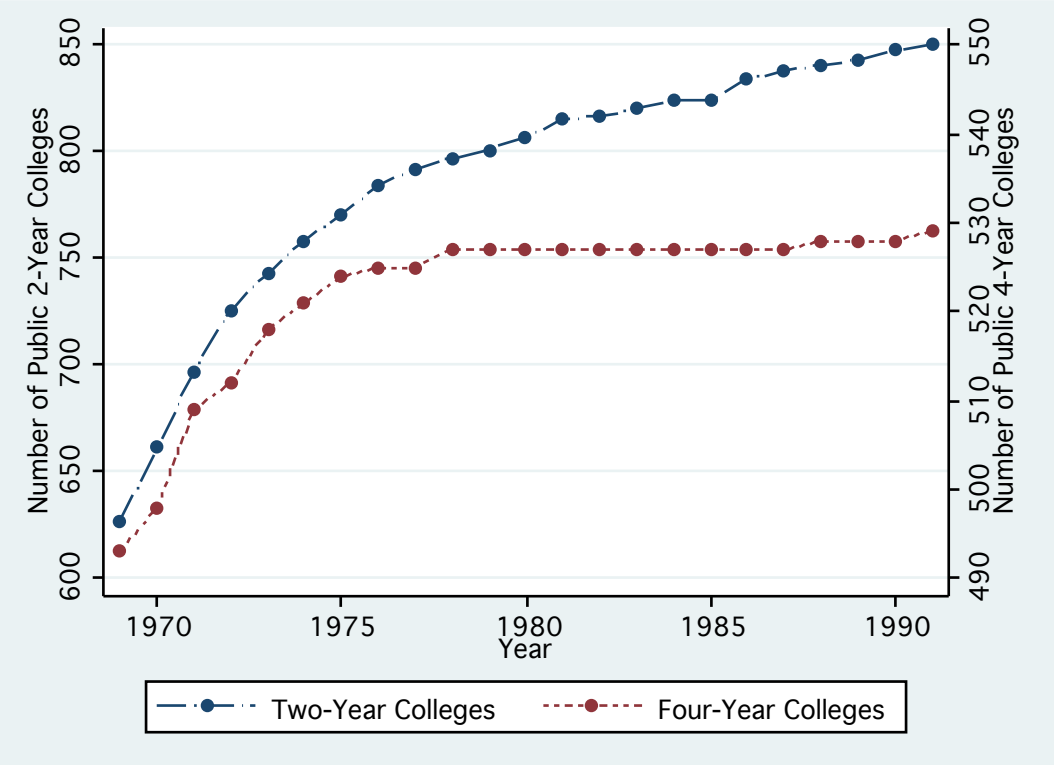
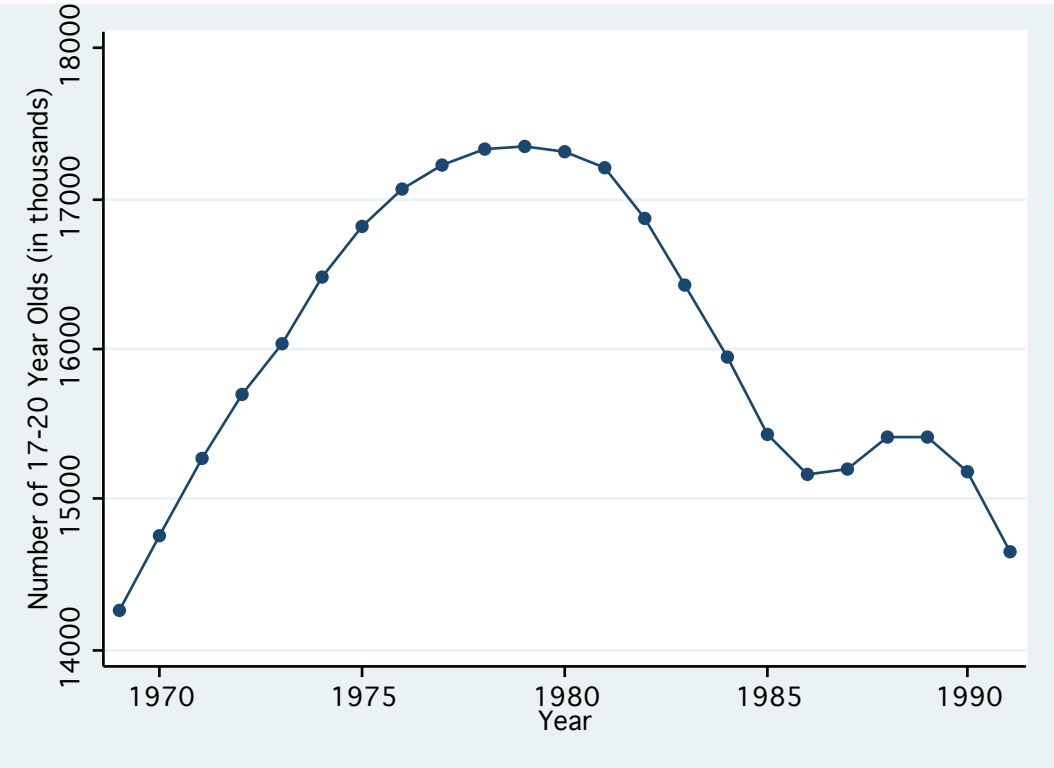


Figure 2: The 17-20 Year Old Population in the U.S.: 1969-1991



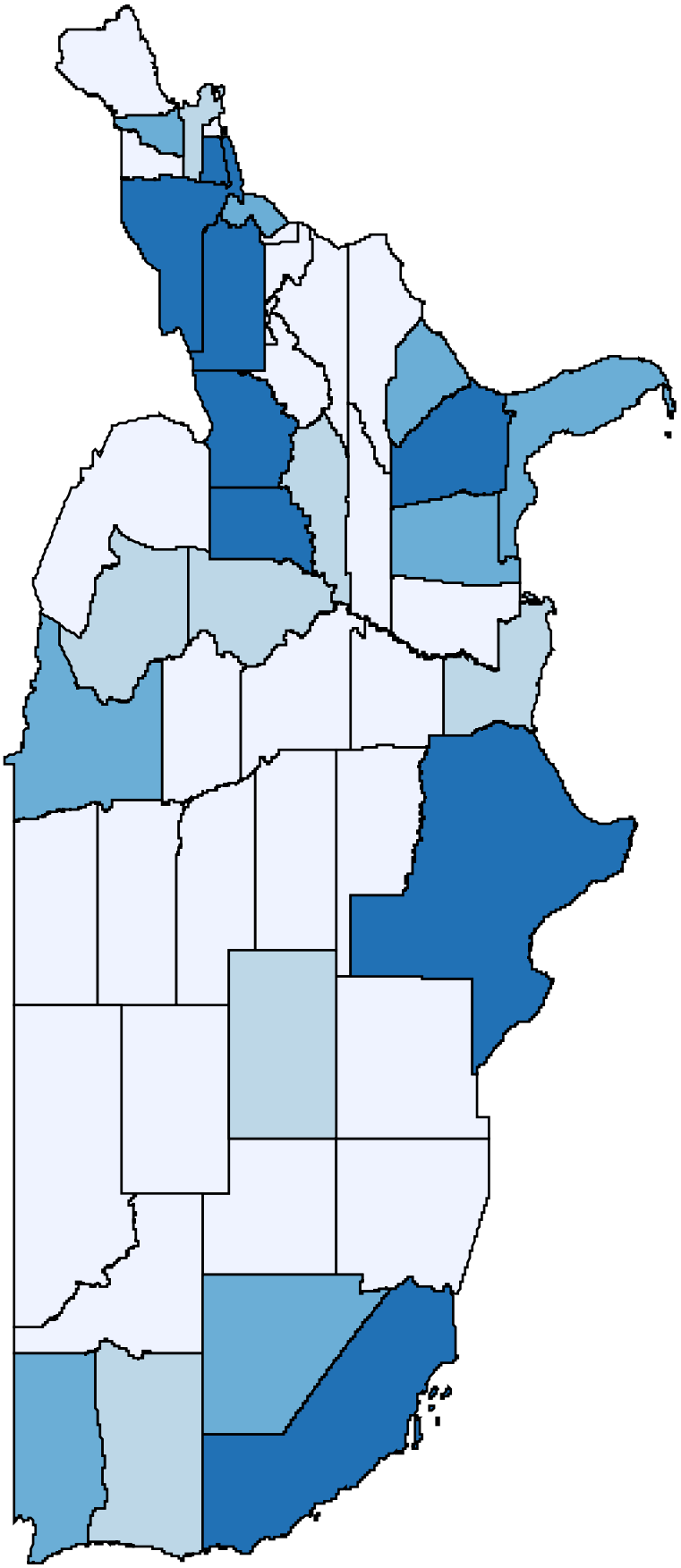


Figure 3: Four-Year College Openings in the U.S.: 1969-1991

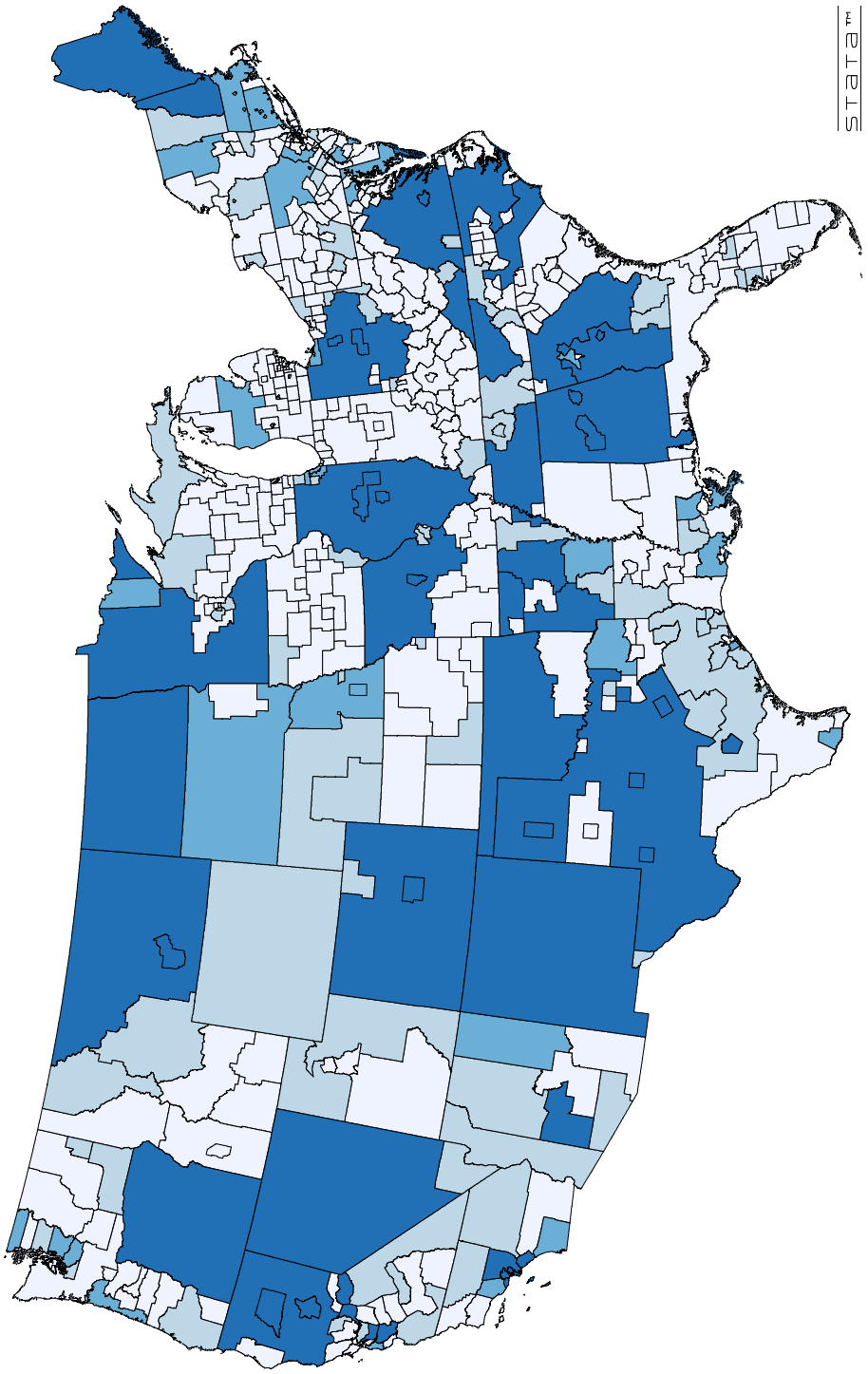


Figure 4: Two-Year College Openings in the U.S.: 1969-1991

Table 1: Determinants of State and County Level College Openings using 1960 Census Data

	4-Year Colleges	2-Year Colleges	2-Year Colleges
	(1)	(2)	(3)
Log Population	0.496*	0.328***	0.429***
	(0.281)	(0.072)	(0.089)
Land Area	0.0123**	0.0945***	0.0389
(in square miles)	(0.006)	(0.026)	(0.042)
Percent Change in	-0.0499**	0.00214	0.00205
Population: 1950-1960	(0.020)	(0.002)	(0.002)
Percent Population	-0.0416	0.00241	0.00084
Urban	(0.030)	(0.003)	(0.003)
Percent Farm Land 1964	-0.0114	0.00308	0.00309
	(0.013)	(0.002)	(0.003)
Population Percent	0.0374*	0.00656	0.00671
NonWhite	(0.020)	(0.004)	(0.005)
Birth Rate in 1968	0.246	-0.0046	-0.00293
	(0.188)	(0.013)	(0.013)
Median Education	-0.730*	-0.055	-0.0973
	(0.416)	(0.054)	(0.062)
Unemployment Rate	-0.393	0.0692***	0.0863***
	(0.277)	(0.019)	(0.020)
Median Family Income	2.079***	0.133	0.192**
(in thousands)	(0.607)	(0.098)	(0.098)
State Fixed Effects	--	No	Yes
Observations	51	2860	2579
Chi^2	30.84	191.77	148.5

Notes: The data are at the state level in column (1) and the county level in columns (2) and (3). The dependent variable is an indicator variable for whether a college opened in that state (county) between 1969 and 1991. The control variables come from the City and County Book for 1960 and Census state statistics for 1960.

Table 2: Determinants of the Timing of State and County Level College Openings using 1960 Census Data

	4-Year Colleges	2-Year Colleges	2-Year Colleges
	(1)	(2)	(3)
Log Population	-1.476	-1.938***	-0.725
	(1.402)	(0.520)	(0.482)
Land Area	-0.00181	0.177	0.0854
(in square miles)	(0.016)	(0.181)	(0.116)
Percent Change in	0.0142	-0.00085	0.00883
Population: 1950-1960	(0.105)	(0.011)	(0.018)
Percent Population	0.169	0.0713**	0.0111
Urban	(0.111)	(0.027)	(0.026)
Percent Farm Land 1964	-0.00504	-0.00819	0.00589
	(0.025)	(0.017)	(0.019)
Population Percent	-0.0596	-0.0429	-0.0399
NonWhite	(0.092)	(0.032)	(0.033)
Birth Rate in 1968	1.057	0.126	-0.0638
	(1.190)	(0.128)	(0.151)
Median Education	-0.938	0.113	-0.169
	(1.272)	(0.272)	(0.534)
Unemployment Rate	0.214	-0.0108	-0.128
	(0.520)	(0.159)	(0.254)
Median Family Income	0.0858	-0.972	-0.416
(in thousands)	(1.190)	(0.800)	(0.672)
F-Statistic	18.57	2.58	1.34
State Fixed Effects	--	No	Yes
Observations	19	231	231
R^2	0.65	0.14	0.46

Notes: The data are at the state level in column (1) and the county level in columns (2) and (3). The dependent variable is the year the college opened. The sample is for the years 1969 and 1991. The control variables come from the City and County Book for 1960 and Census state statistics for 1960.

Table 3: Puma-Level Census Summary Statistics

	All	Cohorts with a 4-year college opening	Cohorts with a 2-year college opening	Smallest Cohorts	Largest Cohorts
	(1)	(2)	(3)	(4)	(5)
Age	33.9	35.91	35.91	34.03	33.85
	(6.36)	(7.19)	(7.08)	(6.47)	(6.32)
Black	0.11	0.11	0.09	0.08	0.12
	(0.31)	(0.31)	(0.29)	(0.27)	(0.32)
Years of Education	13.40	13.43	13.37	13.32	13.43
	(0.84)	(0.90)	(0.74)	(0.98)	(0.07)
Less than 12 Years of Education	0.067	0.070	0.069	0.067	0.066
	(0.064)	(0.066)	(0.053)	(0.082)	(0.057)
Exactly 12 Years of Education	0.331	0.324	0.334	0.346	0.326
	(0.123)	(0.120)	(0.105)	(0.156)	(0.111)
Exactly 14 Years of Education	0.091	0.088	0.092	0.092	0.090
	(0.055)	(0.051)	(0.035)	(0.070)	(0.048)
16+ Years of Education	0.257	0.269	0.252	0.241	0.261
	(0.139)	(0.142)	(0.114)	(0.162)	(0.130)
Number of Public 4-Year Colleges in State	16.69	21.04	15.89	15.43	17.09
	(9.71)	(9.17)	(8.28)	(9.66)	(9.69)
Number of Public 2-Year Colleges	4.12	3.95	9.73	0.739	5.195
	(5.657)	(5.620)	(6.910)	(0.986)	(6.056)
Log 17-20 Year Old Population in State	10.76	13.39	13.11	12.92	13.11
	(1.070)	(0.689)	(0.901)	(0.835)	(0.845)
Log 17 Year Old Population in PUMA	9.38	9.49	10.31	7.92	9.80
	(1.071)	(0.875)	(0.875)	(0.319)	(0.803)
Fraction of Employment in Construction	0.052	0.051	0.049	0.053	0.051
	(0.102)	(0.147)	(0.010)	(0.020)	(0.014)
Fraction of Employment in Manufacturing	0.189	0.201	0.197	0.196	0.187
	(0.087)	(0.088)	(0.082)	(0.108)	(0.080)
Income Per Capita (in 1999 dollars)	11510	10894	10556	11427	11586
	(2788)	(2640)	(2208)	(2937)	(2739)
Employment to Population Ratio	0.472	0.446	0.455	0.440	0.481
	(0.116)	(0.080)	(0.095)	(0.111)	(0.116)
Observations	76744663	5755491	6932634	19637352	61613052

The mean of each variable is reported with standard errors in parentheses. The statistics are weighted by the number of individuals in each cell. The smallest cohorts are cohorts in the bottom 50th percentile of population size. The largest cohorts are those in the top 50th percentile of population size. The PUMA-level variables are from the year when the individual was 17 years old. Years of Education is a continuous variable of years of education (equivalent to grades in the U.S. school system). The other education dependent variables are indicator variables equal to one if the indidivual has that level of education. Data on individual educational attainment are from the 1980, 1990, and 2000 Censuses. Population data are from the National Cancer Institute's SEER population database. Data on the number of schools were created using data from the National Center for Education Statistics and Peterson's. County-level economic variables are from the Bureau of Economic Analysis. Variables are aggregated from the county-level using a county to PUMA crosswalk.

Table 4: Effect of College Availability on Years of Education

Dependent Variable is Years of Education	(1)	(2)	(3)	(4)	(5)
Number of Four-Year Colleges	0.0044	0.00825***	0.00819**	0.00835***	0.0116***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)
Log State Cohort	-0.562***	-0.392***	-0.392***	-0.377***	-0.246
	(0.143)	(0.119)	(0.119)	(0.118)	(0.163)
Number of Two-Year Colleges	0.00192	0.00106	0.00103	0.00211	0.00423
	(0.005)	(0.005)	(0.005)	(0.006)	(0.005)
Log PUMA Cohort	-0.0934	-0.109*	-0.110*	-0.0899*	-0.0919*
	(0.061)	(0.056)	(0.056)	(0.053)	(0.055)
Observations	79,614,311	79,614,311	79,614,311	79,472,538	79,472,538
R^2	0.534	0.692	0.697	0.697	0.699
PUMA Fixed Effects	Yes	Yes	Yes	Yes	Yes
Year Birth*Census Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
Race Fixed Effects	No	Yes	Yes	Yes	Yes
Gender Fixed Effects	No	No	Yes	Yes	Yes
PUMA Controls	No	No	No	Yes	Yes
State Linear Time Trend	No	No	No	No	Yes

The dependent variable in each specification is years of education. Fixed effects are progressively added in each column. A PUMA is a geographic area defined by the census and can be thought of as a county group. PUMA controls include the employment to population ratio, the fraction of employment in manufacturing and construction, and mean income per capita. The cohort is the population of 17-20 year olds. Standard errors are clustered by state. All regressions are weighted by the number of individuals in each cell.

* Significant at the 0.10 level, ** Significant at the 0.05 level, *** Significant at the 0.01 level.

Table 5: Effect of College Availability on Various Educational Outcomes

	(1)	(2)	(3)	(4)	(5)
	Years of Education	Less than 12 Years of Education	High School Diploma Only	Associate's Degree Only	Bachelor's Degree
Number of Four-Year Colleges	0.0116***	-0.000625***	-0.000845*	0.000339	0.000997**
	(0.0020)	(0.0002)	(0.0005)	(0.0003)	(0.0004)
Log State Cohort	-0.246	-0.004	0.0847**	0.051	-0.0896**
	(0.1630)	(0.0151)	(0.0383)	(0.0391)	(0.0357)
Number of Two-Year Colleges	0.004230	-0.000453*	0.000001	0.00137***	-0.000376
	(0.0051)	(0.0002)	(0.0006)	(0.0004)	(0.0008)
Log PUMA Cohort	-0.0919*	-0.00018	-0.00260	0.00429	-0.01500
	(0.0547)	(0.0028)	(0.0111)	(0.0034)	(0.0101)
Observations	79472538	79494185	79494185	79494185	79494185
R^2	0.6990	0.3960	0.5090	0.2090	0.6680

The dependent variable in each specification is specified in the column title. Years of Education is a continuous variable of educational attainment. The other education dependent variables are indicator variables equal to one if the indidivual has that level of education. The unit of observation is the PUMA-year of birth-race-gender-census year. A PUMA is a geographic area defined by the census and can be thought of as a county group. Those regressions include PUMA, year birth * census year, race, and gender fixed effects as well as PUMA controls and a state linear time trend. PUMA controls include the employment to population ratio, the fraction of employment in manufacturing and construction, and mean income per capita. The cohort is the population of 17-20 year olds. Standard errors are clustered by state. All regressions are weighted by the number of individuals in each cell.

* Significant at the 0.10 level, ** Significant at the 0.05 level, *** Significant at the 0.01 level.

Table 6: Effect of College Availability on Various Educational Outcomes: 1990 & 2000 Census Only

	(1)	(2)	(3)	(4)	(5)
	Years of Education	Less than 12 Years of Education	High School Diploma Only	Associate's Degree Only	Bachelor's Degree
Number of Four-Year Colleges	0.0107***	-0.000563**	-0.00110*	0.000981***	0.000882*
	(0.0027)	(0.0003)	(0.0006)	(0.0003)	(0.0005)
Log State Cohort	-0.093	-0.027	0.0899**	0.051	-0.0561**
	(0.1450)	(0.0273)	(0.0363)	(0.0324)	(0.0274)
Number of Two-Year Colleges	0.003570	0.000066	-0.000258	0.00128***	-0.000458
	(0.0042)	(0.0004)	(0.0007)	(0.0004)	(0.0007)
Log PUMA Cohort	-0.164***	0.00406	0.01200	0.00066	-0.0287***
	(0.0593)	(0.0035)	(0.0114)	(0.0035)	(0.0110)
Observations	73299898	73320785	73320785	73320785	73320785
R^2	0.70	0.44	0.51	0.22	0.68

The dependent variable in each specification is specified in the column title. This sample is limited to the 1990 and 2000 Censuses. Years of Education is a continuous variable of educational attainment. The other education dependent variables are indicator variables equal to one if the indidivual has that level of education. The unit of observation is the PUMA-year of birth-race-gender-census year. A PUMA is a geographic area defined by the census and can be thought of as a county group. Those regressions include PUMA, year birth * census year, race, and gender fixed effects as well as PUMA controls and a state linear time trend. PUMA controls include the employment to population ratio, the fraction of employment in manufacturing and construction, and mean income per capita. The cohort is the population of 17-20 year olds. Standard errors are clustered by state. All regressions are weighted by the number of individuals in each cell.

* Significant at the 0.10 level, ** Significant at the 0.05 level, *** Significant at the 0.01 level.

Table 7: Effect of College Availability on Various Educational Outcomes: Small PUMAs Only

	(1) Years of Education	(2) Less than 12 Years of Education	(3) High School Diploma Only	(4) Associate's Degree Only	(5) Bachelor's Degree
Number of Four-Year Colleges	0.0128***	-0.000732***	-0.00089	0.000427	0.00132**
	(0.0025)	(0.0002)	(0.0007)	(0.0004)	(0.0006)
Log State Cohort	-0.216	-0.006	0.0877**	0.075	-0.104**
	(0.2020)	(0.0167)	(0.0442)	(0.0465)	(0.0471)
Number of Two-Year Colleges	0.003290	-0.00176**	-0.000094	0.00012	-0.000918
	(0.0142)	(0.0007)	(0.0019)	(0.0007)	(0.0016)
Log PUMA Cohort	(0.091)	-0.00055	-0.00452	0.00782*	-0.0197*
	(0.0587)	(0.0026)	(0.0114)	(0.0042)	(0.0117)
Observations	53540199	53558279	53558279	53558279	53558279
R^2	0.6830	0.3500	0.4930	0.1850	0.6600

The dependent variable in each specification is specified in the column title. This sample is limited to PUMAS with land area smaller than 7,900 square miles. Years of Education is a continuous variable of educational attainment. The other education dependent variables are indicator variables equal to one if the indidivual has that level of education. The unit of observation is the PUMA-year of birth-race-gender-census year. A PUMA is a geographic area defined by the census and can be thought of as a county group. Those regressions include PUMA, year birth * census year, race, and gender fixed effects as well as PUMA controls and a state linear time trend. PUMA controls include the employment to population ratio, the fraction of employment in manufacturing and construction, and mean income per capita. The cohort is the population of 17-20 year olds. Standard errors are clustered by state. All regressions are weighted by the number of individuals in each cell.

* Significant at the 0.10 level, ** Significant at the 0.05 level, *** Significant at the 0.01 level.

Table 8: Effect of the Number of Colleges Per Capita on Various Educational Outcomes

	(1)	(2)	(3)	(4)	(5)
	Years of Education	Less than 12 Years of Education	High School Diploma Only	Associate's Degree Only	Bachelor's Degree
Number of Four-Year Colleges Per Capita	4.881***	-0.159	-1.029***	-0.196	0.780***
	(1.657)	(0.106)	(0.284)	(0.225)	(0.271)
Number of Two-Year Colleges Per Capita	0.326*	-0.0249**	-0.038	0.0506**	0.009
	(0.187)	(0.012)	(0.029)	(0.023)	(0.030)
Observations	79472538	79494185	79494185	79494185	79494185
R^2	0.70	0.40	0.51	0.21	0.67

The dependent variable in each specification is specified in the column title. The per capita measures of college availability are the number of four-year colleges in one's state at age 17 divided by the number of 17-20 year olds in one's state (in thousands) and the number of two-year colleges in one's PUMA at age 17 divided by the number of 17-20 year olds in one's PUMA (in thousands). Years of Education is a continuous variable of educational attainment. The other education dependent variables are indicator variables equal to one if the indidivual has that level of education. The unit of observation is the PUMA-year of birth-race-gender-census year. A PUMA is a geographic area defined by the census and can be thought of as a county group. Those regressions include PUMA, year birth * census year, race, and gender fixed effects as well as PUMA controls and a state linear time trend. PUMA controls include the employment to population ratio, the fraction of employment in manufacturing and construction, and mean income per capita. Standard errors are clustered by state. All regressions are weighted by the number of individuals in each cell.

* Significant at the 0.10 level, ** Significant at the 0.05 level, *** Significant at the 0.01 level.

Appendix

State-Level Analyses

The main analyses in this paper are at the PUMA-level. That is, I assign individuals the characteristics (cohort size, the number of two-year colleges, and local macroeconomic variables) of the PUMA they currently reside in for the year they were 17 years old. This assumes that individuals live in the same PUMA at the time of survey as they did when they were 17 years old. I relax this assumption and do an analogous analysis at the state-level and assign each person the characteristics of the *state* they currently live in for the year they were 17 years old. Thus, I estimate the following equation:

$$\begin{aligned} Education_{yscrg} &= \alpha + \beta_1 NumPub4_{y+17,s} + \beta_2 NumPub2_{y+17,s} + \mu_1 \ln(Pop1720)_{y+17,s} \\ &\quad + \delta X_{y+17,s} + \gamma_r + \gamma_g + \gamma_{yc} + \gamma_s + \gamma_s * TREND + \epsilon_{yscrg}, \end{aligned}$$

Estimates are shown in Table 9, and are similar to results from the main analysis. A key difference is that at the state-level, the number of two-year colleges does not have a significant effect on educational attainment. The effect of two-year colleges on obtaining an Associate's degree is negative (contrary to previous results), but only marginally significant and very small in magnitude.

Data Appendix

For the handful of institutions in IPEDS, but not in the Peterson's data, I identified founding dates by visiting the schools' websites and reviewing the school's history page. Since institutions do not begin instruction in the year they are founded (in most cases), I infer the opening year to

be two years after the founding year. I compared the opening and founding dates of approximately 50 institutions, and found two years to be the difference for all but a few of the schools. IPEDS identifies each institution to be public, private (not-for-profit), or for-profit. I exclude for-profit institutions from this study[21]. Next, I classify each institution as either a four-year or two-year school. While IPEDS designates each school as a two- or four-year institution, I further require that schools classified as four-year grant primarily baccalaureate degrees or above[22]. I require two-year schools to grant at least associate degrees. Finally, I classified each school as academic or vocational based on both their Carnegie Classification and school name[23]. For this study, I narrow my focus to two types of institutions: four-year academic public and two-year academic public.

While the variables I use in my analyses are initially at the county-level, I aggregate them to the PUMA level. Using a county to PUMA crosswalk[24], I identify the PUMA each county is assigned to. For PUMAs comprised of multiple counties, the PUMA-level variable is the sum (or mean, when appropriate) of the county-level variables. "Summing" variables include population, total employment, the number of colleges, enrollment at colleges, and land area. "Meaning" variables include fraction of employment in construction and manufacturing and per capita income. [25] When a county is divided into multiple PUMAs, the PUMA-level variable is derived by dividing county-level "summing" variables by the number of divisions of the county. For "meaning" variables, the county-level value is assigned to the PUMA.[26]

[21]I exclude for-profits for a few reasons. First, enrollment in for-profit colleges is very small relative to the other institutions. Additionally, many for-profit schools offer distance-learning programs which makes proximity to such a school less important. Finally, Peterson's did not have founding dates for many of the for-profit schools and schools often did not provide information about when specific branches of the school opened.

[22]Schools that primarily grant graduate degrees are not considered to be four-year schools.

[23]By identifying keywords from the institutions' names, I determined the type of vocation taught by the school. Beauty, art, nursing, theology, and technical schools as well as career colleges can often be identified by their names. For example, Florida Career College or Youngstown College of Massotherapy

[24]Source: Lisa Neidert; University of Michigan.

[25]For example, suppose PUMA A is made up of three counties, x, y, and z. If the total population in counties x, y, and z are 10, 11, and 12, respectively, the total population for PUMA A will be 33. However, for certain variables, it is appropriate to use the mean instead of the sum. For example, if per capita income in counties x, y, and z are 100, 150, and 350, respectively, the per capita income for PUMA A would be 200.

[26]For example, Los Angeles County is divided into six PUMAs. If the county has six public four-year colleges, each PUMA will be assigned one public four-year college. If the fraction of employment in construction is .20 for Los Angeles, then it will be .20 for all six of the L.A. County PUMAs.

Table 9: State-Level Analysis

	(1)	(2)	(3)	(4)	(5)
	Years of Education	Less than 12 Years of Education	High School Diploma Only	Associate's Degree Only	Bachelor's Degree
Number of Four-Year Colleges	0.0166	-0.00352***	-0.0038	-0.00359**	0.0009
	(0.0169)	(0.0010)	(0.0030)	(0.0016)	(0.0043)
Log State Cohort	-0.2270	-0.0384*	0.1020	0.0312	-0.111*
	(0.2720)	(0.0226)	(0.0754)	(0.0329)	(0.0662)
Number of Two-Year Colleges	0.0036	0.000548***	-0.0011	-0.0003	0.0006
	0.00290)	(0.0002)	(0.0011)	(0.0005)	(0.0006)
Observations	26681065	26681930	26681930	26681930	26681930
R^2	0.85	0.71	0.73	0.52	0.82

The dependent variable in each specification is specified in the column title. Years of Education is a continuous variable of educational attainment. The other education dependent variables are indicator variables equal to one if the indidivual has that level of education. The unit of observation is the state-year of birth-race-gender-census year. Those regressions include state, year birth * census year, race, and gender fixed effects as well as state controls and a state linear time trend. State controls include the unemployment rate, the fraction of employment in manufacturing, military, and construction, mean earnings per job, and mean income per capita. The cohort is the population of 17-20 year olds. Standard errors are clustered by state. All regressions are weighted by the number of individuals in each cell.

* Significant at the 0.10 level, ** Significant at the 0.05 level, *** Significant at the 0.01 level.