Labor Market Returns to Community College: Evidence from Admissions Lotteries

Michel Grosz*

October 27, 2016

JOB MARKET PAPER

Abstract

Community colleges enroll a third of all postsecondary students and have great promise addressing recent increases in the demand for skilled workers. In this paper I estimate the labor market returns to a particularly large and important degree, the Associate's Degree in Nursing (ADN). I use student-level academic and earnings records across two decades for all community college students in California. I capitalize on random variation from admissions lotteries to produce causal estimates of the effect of the ADN on earnings and employment. Enrolling in the program increases earnings by 55% and the probability of working in the healthcare industry by 23 percentage points. I also use an individual fixed effects approach and show that there is substantial heterogeneity in earnings returns across nursing programs. The returns are higher in areas with more occupational opportunities for nurses, but there is little difference across measures of college quality. In light of concerns about nursing shortages, I estimate that the economic value of expanding an ADN program by one seat far outweighs the costs.

*Department of Economics, University of California, Davis, 1 Shields Ave., Davis, CA, 95616. Phone: (818) 926-1099, E-mail: mgrosz@ucdavis.edu, Website: http://www.michelgrosz.com

I am grateful to Ann Stevens, Michal Kurlaender, Scott Carrell, and Giovanni Peri for their helpful advice, guidance and comments. I also thank Tom Bailey, Marianne Bitler, Brian Duncan, Mark Hoekstra, Paco Martorell, Doug Miller, Erich Muehlegger, Dick Murnane, Marianne Page, Mitch Downey, Ryan Fuller, Lester Lusher, Stephanie Rennane, Vasco Yasenov, and seminar participants at UC Davis, Pomona College, the University of Nevada-Reno, the All California Labor Economics Conference, the Association for Education Finance and Policy, the Southern Economic Association conference, the Western Economic Association dissertation workshop, and the National Academy of Education Fall 2015 and Spring 2016 retreats. The California Community College Chancellor's Office generously provided data access, technical support and expertise. This work is supported by a National Academy of Education/Spencer Dissertation Fellowship.

1 Introduction

Community colleges have recently made a resurgence in debates about the future of public education. In 2015, for example, the Obama administration announced plans to make community college free for most students (Executive Office of the President, 2015).[1] There are a number of reasons for this increased attention. Community colleges are more accessible and affordable to students than four-year college, addressing issues of sagging postsecondary enrollment and attainment (Goldin and Katz, 2008; Cohen, Brawer and Lombardi, 2009). They also overwhelmingly enroll older, lower-income and first-generation students, making them drivers of upward socioeconomic mobility (Belfield and Bailey, 2011; Kane and Rouse, 1999). Career technical programs, which represent half of community college enrollments, are especially important in light of changing demand for skills in the labor force (Bailey et al., 2003; Acemoglu and Autor, 2011).[2] In recent years, policymakers have focused additional attention on expanding career technical programs.[3]

Career technical programs in health fields are of particular interest. As shown in Figure 1, the health workforce is booming, and employment rose even during the Great Recession. Employment grew the most for healthcare workers with less than a bachelor's degree, who predominantly receive their training from community colleges (Noy et al., 2008; Ross, Svajlenka and Williams, 2014; Lockard and Wolf, 2012). Health training programs are thus essential to provide workers the skills increasingly demanded in the labor market. Nevertheless, there is growing concern of shortages of skilled healthcare workers, and of training programs not expanding their capacity to meet demand (Buerhaus et al., 2013). Given these concerns, it is crucial to quantify the role of existing programs in affecting the earnings and employment of students. Such evidence is limited, and to my knowledge no study has yet used random variation to measure these effects.

In this paper I measure the labor market returns to an Associate's Degree in Nursing (ADN). I leverage the random lottery that assigns admission to a large program in California in order

[1] This plan echoed efforts underway in Tennessee, Oregon, and Chicago.

[2] The terms "vocational" and "career technical" are largely interchangeable terms for programs and coursework that train students for specific occupations. For the rest of this paper I exclusively use the term "career technical."

[3] A 2010 federal initiative contributed $2 billion to expanding community college career technical programs (Field, 2014), and there have been numerous state efforts as well.

to produce causal estimates of the program's effect. Surging demand for seats has forced many colleges to institute admissions policies to ration high-demand programs and courses, especially in health fields (Gurantz, 2015; Bohn, Reyes and Johnson, 2013; Bound and Turner, 2007). Because community colleges have a long history as open-door institutions, these policies often take the form of lotteries and waitlists, which do not depend on ability or merit (Grubb and Lazerson, 2009; Betts and McFarland, 1995).[4] This paper is the first to use variation from admissions lotteries to study an existing community college program, and one of very few in the context of higher education whatsoever (Ketel et al., 2016).

I rely on rich data that track community college students through their academic careers and into the labor market. I use detailed individual-level administrative records covering all students enrolled in California community colleges between 1992 and 2015, which include coursework, grades, credentials earned, demographics, and financial aid receipt. These data are also matched to individual quarterly earnings and industry of employment information from the state's unemployment insurance system. To these data I added information on the outcome of admissions lotteries to a large ADN program for cohorts since 2005. An important feature of this dataset is that I can track the academic and labor market trajectories of both admitted and rejected students. Moreover, I observe individual earnings and industry of employment before, during, and after students enroll in a college.

The lottery allows me to causally identify the effect of the program at one large college. In order to expand the analysis to all the programs across the state I also use an approach that is increasingly used in the community college literature. I estimate individual fixed effects models, which leverage student-level earnings variation across time before and after enrollment in an ADN program. Using this method I explore heterogenous earnings effects across students, colleges, and labor markets. This study is also the only one to my knowledge that produces estimates of the returns to a typical postsecondary program with both random and individual variation. Thus, I have an excellent opportunity to observe whether the two methods are similar, which I do by

[4] Open access is the norm for colleges across the country. In California, this policy was explicitly set in 1960 by the Master Plan for Higher Education, which delineated a hierarchy of higher education systems: the University of California, the California State University, and the California Community Colleges (Coons et al., 1960).

limiting the sample to only applicants at the lotteried program.

Using the results of the lottery as an instrument for enrollment, I find that the causal effect of enrolling in the lotteried program is a 55 percent increase in earnings within five to seven years of applying. This is a very large effect, especially given standard estimates of the returns to a year of school. While the program has a large earnings effect, the repeated lotteries delays students from entering into the labor market. I find that for each additional lottery application, which implies an additional semester of delay, the earnings effect decreases by five percent. This is not large enough to wash away the large benefits of the program, but is still substantial. The counterfactual students, who never win a lottery, are also much less likely to earn any type of community college credential or to transfer to a four-year college.

I do not find any impacts of the program on employment using the lottery estimates. This is perhaps due to a lack of statistical power, but more likely because the applicants to the program are already a select group; most have significant prior labor market experience. Instead, I find that students who enrolled in the program were 23 percentage points more likely to work in the healthcare sector after applying. While I cannot explicitly attribute the large earnings effects to this employment effect, it is at least suggestive evidence that students who complete the program enter careers as registered nurses.

Results when I use the individual fixed effects specification are broadly similar to those from the randomized lottery. However, I do find that modeling choices matter, especially the choice of comparison group students. I find that students who enroll in the program see their earnings jump by 54 percent, and those who graduate see earnings effects of 76 percent. In the statewide analysis, I find a great deal of heterogeneity across different nursing programs. This heterogeneity is primarily associated with local labor market opportunities in the health industry, and not necessarily program quality.

I use the results of my analysis to inform recent policy discussions. In particular, colleges may not be adequately expanding their nursing program capacity to meet student interest and the rising demand for healthcare. Using my most conservative estimates, I find that the private value of an additional student enrolling in an ADN program is approximately $150,000 over a 20 year career,

not taking into account any positive spillovers. On the other hand, a conservative estimate of the cost of expanding a program by one seat is approximately $20,000. Given state income tax rates, this expansion is also likely revenue neutral. Nevertheless, because colleges in California—and other states—are allocated funds based on overall enrollment, there is limited incentive for colleges to expand costly programs like nursing. Thus, an important policy implication of this study is that greater attention needs to be placed in developing strategies that make expansion more viable.

This paper makes several contributions to the education literature. First, I estimate the causal effect of an existing community college program on earnings and employment. The identifying variation comes from a random lottery, which is rare in studies of higher education. Second, I estimate two models of the returns to a program using the same sample, bridging the unique approach afforded me by the lottery and the observational estimates used by an increasing number of studies. Third, I show that there is substantial heterogeneity in earnings effects even within a single degree, and that this heterogeneity can be explained by regional economic opportunities and program characteristics. Fourth, I use the earnings results to suggest that the economic benefit of expanding nursing programs is far greater than the costs. This is especially important in light of concerns about the supply of registered nurses lagging behind demand for healthcare.

The rest of this paper proceeds as follows. Section 2 provides a review of the relevant literature and institutional background. Section 3 describes the data. An explanation of the instrumental variables methodology is in section 4, and main results are in section 5. In section 6 I present results using the student fixed effects approach. Section 7 includes a discussion of expansion costs, and section 8 concludes.

2 Background

2.1 The Returns to Community College

The key obstacle in estimating the effect of schooling on earnings is the issue of selection. Students who enter certain programs or earn credentials are systematically different than others in ways that can affect their labor market outcomes and that may also be unobservable to researchers. This

is a classic problem in the labor economics literature, and many econometric techniques have been developed in order to address it. However, the returns to schooling literature tends to focus on grade school, high school, and four-year college (eg. Angrist and Krueger, 1992; Ashenfelter and Krueger, 1994; Bedard, 2001). There is much less research on the returns to community college, perhaps because the multiple missions of community colleges make issues of selection particularly severe. Some students seek to transfer to a four-year college, others enter career technical programs, and many only aim to take a few courses or continuing education credits without earning a degree.[5]

Because of the many pathways community college students follow, it is crucial to be able to observe their individual courses of study, the content of the degrees and certificates they earn, and their earnings. Recent work finds substantial heterogeneity in labor market returns to four-year college major (Kirkeben, Leuven and Mogstad, 2014; Arcidiacono, 2004; Altonji, Blom and Meghir, 2012); these differences by field of study likely also exist in the community college setting. Thus, there is value in understanding the intricacies of returns specific programs. However, an obstacle to producing credible estimates of the returns to individual programs has been the availability of granular data and methods for causal identification.

Kane and Rouse (1995) find that the returns per credit hour at community college are roughly similar to those at four year colleges. Gill and Leigh (2003) find that vocational degrees yield a 46 percent return over just high school. These studies made great strides in this field, yet in using national survey datasets were limited in their ability to causally identify the returns or speak to issues of program heterogeneity.[6]

A newer line of research relies on state administrative datasets, pooling the academic and earnings records of many thousands of students as they swirl through community college systems. Jacobson, LaLonde and Sullivan (2005) examine retraining programs for displaced workers and find that older workers derive a lower benefit compared to younger workers. Using similar methods, recent work has measured the labor market returns to career technical awards by type and course of study (Stevens, Kurlaender and Grosz, 2015; Bahr et al., 2015; Liu, Belfield and Trimble, 2015;

[5]Many state administrative datasets have information on self-reported academic goals. Simply controlling for this measure is problematic, as it is notoriously unreliable (Zeidenberg, Scott and Belfield, 2015).
[6]See Belfield and Bailey (2011) for a full overview of this literature.

Jepsen, Troske and Coomes, 2014).[7] A key conclusion from this emerging literature is that the labor market returns to career technical programs vary immensely by subject and type of degree, and seem particularly large in health fields. What remains unclear is how well the models typically used in these analyses, which exploit within-individiual earnings changes, account for different types of bias. In this paper, in addition to employing a random lottery as the main source of identification, I am also able to replicate the methods used in these other studies.

Although experimental evaluations at community colleges are rare, there is a long history of such interventions in the related field of workforce development.[8] The Job Training Partnership Act of 1982, which had a randomized evaluation component, received a lot of attention from researchers, in particular because of its surprising negative findings (Doolittle and Traeger, 1990; Heckman et al., 2000; Heckman and Smith, 2000). Other prominent experimental studies of job training programs include the Job Corps (Schochet, Burghardt and McConnell, 2008; Flores-Lagunes, Gonzalez and Neumann, 2010), sectoral training programs (Maguire et al., 2010), and a program run through Food Stamps (Puma and Burstein, 1994). Much of the recent work in this field, though, has been in adapting non-experimental methods to evaluate more recent job training programs, such as the Workforce Investment Act and Trade Adjustment Assistance (Heinrich and Mueser, 2014; Heinrich et al., 2013; Hollenbeck and Huang, 2006; Hotz, Imbens and Mortimer, 2005). In addition to issues of external validity common to all experimental designs, few randomized interventions that study earnings are similar to existing career technical programs or are even set within community colleges, with some exceptions (Scrivener and Weiss, 2013; Visher et al., 2012).[9] These are special cases, though; they do not reflect standard community college programs that exist across the country, and their results would not necessarily hold if the programs were scaled up.

Two recent papers take another approach to randomized trials through audit studies, finding value in a two-year community college degree, especially relative to one from a for-profit institution

[7] Cellini and Turner (2016) use a similar method to compare the returns to community college and for-profit two-year colleges.

[8] See Barnow and Smith (2015) for a full overview of this literature.

[9] There are also large ongoing evaluations of training programs in cooperation with community colleges, though with no results yet. The first is the Health Professions Opportunity Grants (HPOG), which involves randomized admission into highly targeted and specialized health training programs (Anderson, Hall and Derrick-Mills, 2013), and another is the Accelerating Opportunities program (Anderson et al., 2016).

(Darolia et al., 2015; Deming et al., 2016). Using experimental or quasi-experimental variation in studying educational programs is more common at other levels of education, perhaps because higher education admissions are more likely based on merit.[10]

In sum, there are still significant gaps in understanding the labor market returns to community college programs, especially at the program level. Evidence from national surveys and state administrative datasets suggests large returns, especially in health, yet there is still concern about selection bias in these studies. On the other hand, in addition to issues of external validity, a drawback of the evidence from randomized control trials is that these interventions may not work when they scale, or may be too costly to implement at all.

2.2 The Labor Market for Registered Nurses

In this paper I focus on a program that awards an associate's degree in nursing (ADN), which is a requirement for work as a registered nurse (RN).[11] RNs are among the occupations with the highest employment in the national economy. There were more than 2.7 million RNs in 2012, with a median annual salary of $65,470 (Bureau of Labor Statistics, 2015).

The labor market for RNs has been dominated for years by fears of shortages (Kuehn, 2007). The nursing workforce, composed in large part of baby boomers, is aging and starting to retire, while expansions in the demand for healthcare require that new inflows of nurses exceed outflows (Buerhaus et al., 2013). However, there are signs that concerns of nursing shortages are perhaps overblown. Since 2005, nursing has seen a huge surge, especially among relatively younger new RNs, and perhaps due to the belief that the profession is "recession-proof" (Auerbach et al.,

[10] There is an established literature on the effect of charter schools on children, using variation from enrollment lotteries (Hanushek et al., 2007; Angrist, Pathak and Walters, 2013). At the postsecondary level, though, such variation is not as common. Hoekstra (2009) exploits the cutoffs in SAT scores and GPA for admission to a large flagship university. Leuven, Oosterbeek and de Wolf (2013) and Ketel et al. (2016) examine the the labor market returns to medical school in the Netherlands using a lottery stratified by student academic performance. I return to this paper in the methods section.

[11] RNs are the largest subset of nursing professionals. They are responsible for overseeing the total care of patients in hospitals, clinics, home health, and other settings such as schools and correctional facilities. They develop and manage plans for their patients' care, administer medicine and treatment, help perform tests, and are the point of contact with doctors and other healthcare professionals on the status of their patients' health. In most settings, an RN may manage other nursing professionals and, in turn, typically operate under the direction of a physician or a more advanced nursing professional such as a nurse practitioner (Bureau of Labor Statistics, 2015).

2013; Auerbach, Buerhaus and Staiger, 2011; Buerhaus, Auerbach and Staiger, 2009). The BLS projected that by 2022 there would be at least 500,000 more RNs, a 19 percent increase (Bureau of Labor Statistics, 2015). Part of the concern about nursing shortages has been whether educational providers could keep up with demand. There has been a huge growth in the number of new graduates since 2005, and for-profit private colleges have become a growing share (Buerhaus, Auerbach and Staiger, 2014).

As with most occupations in the healthcare sector, nursing is regulated by licensing boards and other regulatory institutions. The minimum requirement to become an RN is an ADN or a bachelor's degree in nursing (BSN) from a program approved by a state licensing board. Graduates of these programs must also pass a national licensing exam.[12] There is some debate regarding whether aspiring RN's should pursue a two-year ADN or a four-year BSN, both of which are sufficient qualification for certification.[13] An influential report by the Institute of Medicine (2011) showed that by 2008, an equal number of registered nurses had an ADN as a BSN, though many more had begun their career with the ADN and later earned the more advanced degree. The report also called for a greater focus on pivoting towards the four-year programs. There is little evidence, though, that BSNs do better in the labor market than ADNs, though some hospitals cite differential hiring practices across the degree types (Kovner et al., 2014). Auerbach, Buerhaus and Staiger (2015) show that while the unemployment rate for RNs with the two types of degree has indeed diverged since 2008, these rates have remained low, under 2 percent, even during the Great Recession. And, while there is an approximately $10,000 earnings difference between RNs with an ADN and a BSN, that difference has remained stable over time and is not causally identified.[14]

[12] Specifically, the National Council Licensure Examination for Registered Nurses (NCLEX-RN), administered nationally by the National Council of State Boards of Nursing. Students who do not pass the exam may retake as many times as they need in order to pass. In California, the first-time pass rates on the NCLEX-RN for all approved prelicensure programs has remained stable at approximately 88 percent over the past few years (California Board of Registered Nursing, 2015).

[13] The content of the training differs slightly. BSNs tend to learn more about theory, public health and research, while ADN programs focus on clinical and practical skills.

[14] An additional difficulty is that many BSN-trained nurses start off with a ADN, while others study for the four-year degree first.

2.3 "Central College" and its ADN Lottery

My analysis is set in California, which has the largest system of community colleges in the country. The California Community Colleges consist of 113 campuses, with over 2.6 million students each year (California Community College Chancellor's Office, 2016*b*). By far the most popular career technical degree is in nursing: the state awarded 5,545 ADN's in 2013-2014.

Central College,[15] located in California's Central Valley, is among the largest community colleges in the state. Its course and program offerings are similar to other large colleges across the state, with a mix of career technical and academic course and program offerings. Apart from the nursing program, Central College's health division also offers associate's degrees in fields such as radiologic technology, respiratory care therapy, and dental hygiene. The ADN program is, however, by far the college's largest health program, and is also among the largest in the state. A new cohort of approximately 100 students begins each fall and spring semesters.[16] The ADN program is highly structured and takes four semesters to complete.[17]

Crucial to my identification strategy, Central College's ADN program has used a random lottery for more than 10 years. Of the 73 colleges in California that granted an ADN in 2014, 12 had admissions decided by a computerized lottery among eligible applicants, and an additional 12 had a lottery among students whose academic achievement surpassed a certain threshold.[18] In order to become eligible for the lottery, applicants must pass certain prerequisite courses.[19] Applicants apply on a paper form that they must pick up at Central College's administrative office,

[15] Anonymized for confidentiality reasons.

[16] Enrollment in some years exceeds 100 per semester if the college has extra capacity due to external grants. A summer cohort is also offered in some years.

[17] Students take a set schedule of courses in a pre-determined order, consistent with standards set by the state's Board of Nursing. Students must earn at least a "C" grade in each course in order to pass the program. Students who do not earn at least a "C" retake the class. Students have access to academic and career support in the form of counselors that serve the entire health sciences division, as well as free tutors for their classes. Beginning in the first semester of the program, students gain hands-on experience, working under the supervision of nurses in nearby hospitals and clinics with which Central College has affiliations. In practice, according to students and administrators at the college, many students ultimately find a job after graduation with one of their clinical placements during the program.

[18] Randomized lotteries are also common in other health fields. For example, four of the 24 radiologic technology programs in the state had a random lottery, as did three of the 15 dental hygiene programs.

[19] These include anatomy, physiology, chemistry, microbiology, and psychology, which are determined by the state's accrediting body and vary little across colleges. Students may fulfill their prerequisites at another college, though most applicants take their prerequisites at Central College.

but every other aspect of the process happens online.[20] One lottery is conducted each semester, using a computer random number generator. On average, 636 eligible applicants apply in each lottery.

Results of the lottery are posted online, according to a schedule made public on the college's website. If they are rejected, students may reapply to the next lottery, a semester later. Reapplication is easy; it does not involve filling out a new application form, but it does involve responding to the lottery result in a timely manner, usually within a week. Students who do not respond by the deadline have their application file marked as closed and, if they wish to reapply, must fill out a new application. Students who apply for a fifth consecutive time are entered into a "special lottery," with a higher change of admission.[21]

3 Data and Summary Statistics

I combine two sources of student-level administrative data for my analyses. The first consists of the lottery data from Central College's health programs. College administrators generously provided me with a list of all applicants and admitted students for each lottery since fall 2005. I drop students who applied despite being ineligible for admission, yielding a sample of 4,726 applicants.[22]

I match the Central College lottery data to detailed statewide data that track all California community college students in their academic careers and the labor market. I use administrative records from the California Community College Chancellor's Office (CCCCO) for students enrolled over a two-decade time span, between 1992 and 2015. The information contained in these data is extensive and detailed. I observe term-level coursework and grades for each student, academic

[20] Applicants submit a copy of their high school transcript, transcripts from all colleges attended, and catalog course descriptions for classes taken outside Central College. This is to verify eligibility and not used for the purposes of accepting or rejecting applicants to the program.

[21] Only consecutive applications count towards this special lottery. So, if a student was rejected and failed to reapply, any subsequent application is counted as that student's first application. Technically some students may apply more than five times in order to reach the special lottery, though in practice this is rare.

[22] Students often apply with incomplete prerequisite coursework. These students are included in the dataset I received from Central College. However, college administrators do not enter them into the application lottery, and thus I drop them altogether from the analysis.

outcomes such as the type and subject of each degree they earned or the four-year institution to which they transferred, financial aid information, and various demographic characteristics. The CCCCO matched these data to individual quarterly earnings and industry of employment information from the state's unemployment insurance system for 2000-2015.[23] The result is a dataset containing detailed information on a student's experience in the California Community College sytem as well as earnings and industry employment before, during, and after their schooling.

Because the lottery is administered at the college level, there is not a perfect match with the administrative data for the entire state system. Instead, I match students in the lottery dataset to students in the entire system's administrative dataset based on the first three letters of their first and last name, their birth date, and their gender, which are the personal identifiers common to both datasets. I am able to match 74 percent of applicants to the statewide data, yielding a sample of 3,506 students in the analysis dataset. This is a high match rate, considering that I must drop many students who have duplicate combinations of names and birth dates.[24]

In addition to the matched student-level dataset, I also learned institutional details from visitng Central College and speaking with students and faculty. I held semi-structured interviews with the dean of the health sciences division as well as the director of each of the individual health programs. In addition, I attended an orientation presentation for incoming ADN students and afterwards held a focus group to learn about their views on the lottery, their academic pathways, and their career goals. I also gathered information on prerequisite coursework, application requirements, admissions criteria, and graduation requirements directly from individual college websites and catalogs. This allows me to establish whether a student had fulfilled the necessary prerequisites to apply to the program, as well as whether they had enrolled in the program

Column 1 of Table 1 shows summary statistics for all applicants, using characteristics de-

[23] Approximately 93 percent of students in the college data are matched to earnings records. Students may be unobserved in the earnings records for several reasons apart from just a true lack of employment or earnings. The most likely other reasons for missing data including being self-employed over the period or having moved out of the state with no recorded earnings in California.

[24] Among all community college students, 15 percent have duplicate combinations of name and birthdate, since only the first three letters of the first and last name are reported in the data. I only drop 2 applicants due to duplicated name and birthdate. There is no statistical difference between matched and unmatched students in terms of lottery result and gender, which are the only two attributes I can observe in the lottery data.

termined before their first lottery. Applicants were predominantly female, a third were Hispanic and relatively few were African American. Applicants were 30 years old on average, which is not uncommon for community college students across the state, especially in career technical programs. Most students received some form of financial aid. Almost three quarters of students had a waiver to cover tuition, fees, and supplies.[25] Similarly, almost half of students received a Pell Grant, a national financial aid program intended for low-income students. A fifth of students received financial aid through the state-administered Cal Grant, and another seven percent had loans of other types.

Applicants had prior labor market attachment; 62 percent had been employed at least eight of the 16 quarters prior to applying. However, applicants were employed in low-paying jobs, with just an average of $4,717 per quarter. I categorize the employment industry variable into broad groupings and find that 41 percent of students had previously worked in the healthcare industry. Although I cannot observe occupation of employment, this is consistent with the idea that many applicants are nursing assistants, health aides, or licensed practical nurses who are looking to upgrade their skills, a trend that is also borne out in focus groups with recently admitted students. Applicants were also likely to have been employed in retail, food, and administrative service professions. Appendix Table A1 shows that, taking into account some demographic differences given its location in the rural Central Valley, the population of both Central College and its ADN program looks qualitatively similar to other colleges and ADN programs across the state.

Columns 2 and 3 of Table 1 show the validity of the randomized lottery by comparing mean characteristics across admitted and rejected applicants within each lottery occurence.[26] Since only students in their first through fourth lottery are among those chosen randomly, all fifth-lottery applications are excluded from the sample in this case. Overall, the two groups look balanced, and there is no evidence of systematic selection across the two groups. The last column of Table 1 shows the balance for just participants in their first lottery. Here, too, the lottery looks balanced.

[25] Specifically, this is a Board of Governor's (BOG) fee waiver, a program exclusive to the California Community Colleges that covers all tuition, fees, and sometimes books and supplies. The main BOG fee waiver has eligibility based on 150% of the federal poverty line.

[26] Specifically, I regress each variable of interest on a dummy variable for whether the student was accepted, as well as cohort fixed effects.

Appendix A.2 provides further evidence of the randomization of the lottery.

Table 2 shows dynamics of the lottery process. Approximately 10 percent of applicants were accepted their first time. Of those who were not accepted, almost all—82 percent—reapplied to be considered in the next lottery. Reapplication rates are relatively consistent across each lottery. Most applicants are accepted in their fifth lottery, though not all. Overall, 41 percent of applicants are ever admitted

4 Methods

I am interested in the effect of the ADN program on subsequent labor market outcomes. I begin by considering the following estimating equation:

$$y_{ic} = \beta_0 + \beta_1 D_{ic} + X_i\beta + \mu_c + \varepsilon_{ic} \tag{1}$$

For student i who applied to the program as part of cohort c, y_{ic} is an outcome such as earnings or employment at some period and D_{ic} is a dummy variable taking a value of one if the student received the treatment. Even controlling for observable student characteristics with the matrix X_i and cohort fixed effects μ_c the treatment is correlated with the error term, thus biasing estimates of β_1. I resolve this bias by exploiting the random variation produced by the admissions lotteries.

I consider two different treatments D_{ic}. The first is enrollment in the program, which is the treatment most directly manipulated by the lottery. It is also a policy-relevant treatment, since it is the lever that college administrators and policymakers can explicitly move by expanding programs. I also consider earning an ADN as another treatment, since I am interested in the effect of degree attainment on earnings. There are a few caveats to using ADN attainment, however. Approximately half of students who start the program earn the degree, which decreases the first stage power. Moreover, students who start the program but do not finish are partially treated, since they likely accumulated human capital despite not graduating.[27]

[27] This also raises some issues for interpretation of who are compliers and non-compliers. If completion is the treatment,

If students were only allowed to apply once, estimating the effect of the treatment on earnings would be straightforward. Admission through the lottery process is both uncorrelated with the error term and also a strong predictor of enrollment, making it a valid instrument. Although a benefit of examining Central College's ADN program is that it is the product of a natural experiment and not one run by researchers, a challenge is that the institutional details of the admissions procedure necessitate a departure from this relatively simple estimation strategy. The key complicating factor is that students may reapply if they are not admitted.

I first adapt the strategy used in the case of a similar repeated lottery for medical school in the Netherlands. Like Leuven, Oosterbeek and de Wolf (2013) and Ketel et al. (2016), I note that the first lottery a student enters is a valid instrument through the following first stage:

$$D_{ic} = \delta_0 + \delta_1 Admit_{ic} + \mu_c + \epsilon_{ic} \tag{2}$$

The coefficient δ_1 in Equation (2) reflects the difference in enrollment among winning and losing compliers of the first lottery. Non-complying lottery losers include two important and distinct groups. First, there is an empirically small group of students who gain admission through channels outside the lottery process. A more prominent group of non-complying lottery losers consists of students who lose their first lottery but win a subsequent lottery and enroll. They are non-compliers of the first lottery, but perhaps compliers of later lotteries. The local average treatment effect of this approach, though, implies a control group consisting of compliers. When a student's first lottery is the instrument, this control group is composed of applicants who were either pushed to never reapply, or who reapplied but were never admitted. Thus, it excludes a large number of students who do reapply and eventually win admission. In addition to affecting the interpretation of the estimates, this feature of the lottery also lowers the power and coefficient of the first stage.[28]

One way to resolve this issue is to notice that each lottery a student applies to is a valid

then some compliers are applicants who may have lost a lottery but still enrolled in the program by winning later lotteries and even potentially took some of the program's courses.

[28] This discussion focuses on enrollment as the treatment. When using program completion as the treatment, an additional set of complying losers include students who win later lotteries, enroll, but do not complete.

instrument for enrollment, not just the first. For example, among all students in their second lottery attempt, admission is random and also a valid instrument for enrollment. Thus, I have four potentially valid instruments for enrollment: the result of each of the four lotteries a student could apply to. I estimate the following first stage equation:

$$D_{icg} = \alpha_0 + \sum_{g=1}^{4} \alpha_g Admit_{icg} + \mu_c + \nu_g + \eta_{cg} + \varepsilon_{icg} \tag{3}$$

where there are four individual reduced form effects for each lottery application g. I include lottery instance fixed effects ν_g and lottery term fixed effects μ_c as before, as well as the interaction term η_{cg}, in order to separately identify the effect of each individual lottery pool. This is equivalent to estimating Equation (2) for each individual lottery.

To increase power and precision I also estimate a modified version of Equation (3), where I only estimate a single effect of lottery admission. This is similar to a method used by Gelber, Isen and Kessler (2016) when studying a treatment that affects outcomes over time, and where treatment in one period affects the probability of treatment in a subsequent period.[29] The first stage equation is:

$$D_{icg} = \alpha_0 + \alpha_1 Admit_{icg} + \mu_c + \nu_g + \varepsilon_{icg} \tag{4}$$

Here, α_1 is essentially the average of the four individual estimates from Equation (3). One concern is that there is selection in who reapplies among the set of lottery losers. This would make the local average treatment effect of each lottery different. However, Table A2 shows that observable characteristics do not predict reapplication among lottery losers. Moreover, the determinants of enrolling in the program conditional on being admitted are not markedly different across different lotteries. This is likely because the cost of reapplying, which only involves clicking a button on a computer screen, is relatively low.

Since all applicants apply for a first time but not necessarily in subsequent lotteries, the

[29] In the context studied by Gelber, Isen and Kessler (2016), a summer youth employment program, treatment was positively correlated with reapplication: program participants were more likely to reapply. In the Central College context, however, lottery losers are more likely to reapply, and lottery winners do not reapply.

first lottery effect is overly represented when using the stacked data specification. I weight the regressions by $w_i = \frac{1}{max_i(k)}$ where k takes values one through four, which downweights students who applied many times and who would otherwise be over-represented.

An important feature of the lottery process is that it not only affects whether a student enrolls in the program, but also possibly delays students from fully participating in the labor market. Time spent applying to lotteries is an endogenous variable, a result of an individual's choices conditional on previous lottery outcomes. However, the result of a lottery is a valid instrument for time spent applying to lotteries. Because the specification from Equation (3) has four separate instruments, I can simultaneously predict enrollment in the program and time spent applying to lotteries. For this reason, this specification yields my preferred estimates.

For most of the analysis I focus on mean quarterly earnings five to seven years following the lottery. Ideally I would examine long-term earnings, but five to seven years is as far out as the data allow. In most specifications I report results in log earnings, though I show the main result in levels in a robustness exercise.

5 Results

5.1 Academic and Labor Market Outcomes

Table 3 shows summary statistics of academic and labor market outcomes. The first column consists of students who ever won a lottery and also enrolled in the program. The second column consist of students who won a lottery but did not enroll. The third column consists of students who never won a lottery.

Approximately half of the students who were offered admission actually enrolled, and half the students who start the program complete it. The low completion rate is still higher than those of other non-nursing community college programs (Bound, Lovenheim and Turner, 2010). In addition, four percent of students who rejected the offer of admission and 15 percent of students who were never offered admission to Central College's program also earned an ADN. This is likely because they applied to multiple programs in addition to Central College. However, students who

never won a lottery, or who did win a lottery but did not accept the offer, had low completion rates in general. Only eight percent received any type of degree from community college, and less than five percent transferred to a four year college.[30]

Panel A of Table 3 also shows the long time to completion among students who did complete the degree. The program is supposed to take fewer than four semesters, or less than two years. However, students who earned the ADN took approximately three years between their first application and completing the degree. This is because of the wait times associated with applying and reapplying to the lottery. Panel B of the table shows labor market outcomes, including mean quarterly earnings, employment, and industry of employment five to seven years after the first lottery. Mean quarterly earnings for lottery winners who enrolled (Column 1) were close to the BLS estimate of $65,000 in annual earnings for registered nurses nationally. Employment rates were high for all groups. A large share of students subsequently worked in the health industry.

5.2 First Stage Results

Table 4 displays first stage results using the three possible sets of instruments for several potential outcomes. Panel A shows estimates of Equation (2), where only the first lottery is used. Winning the first lottery also increases the probability of enrolling in the Central College ADN program—the endogenous regressor of interest—by 24 percentage points, with a large F statistic (31.8). There is a smaller effect on earning an ADN, of 8 percentage points, with the F statistic of 5.4.

The lottery not only manipulates whether a student enrolls in the program, but when. Effects on earnings might be affected by delay in entering the labor market and the loss of earnings growth from experience. Winning a lottery reduced the number of lotteries entered, by about 1.6 lotteries, but had a negligible effect on the time spent enrolled in community college.[31] This is because losing a lottery results in an increase in the time spent enrolled through inducing students to reapply; on the other hand, winning a lottery also increases time spent in college since the ADN

[30] Some students who were never admitted also enrolled, perhaps gaining admission through other means outside the lottery.

[31] I measure this as the number of years between the quarter of first enrollment in any community college course to the last observed quarter of enrollment in any community college course.

program takes two years to complete.

Panel B of Table 4 shows first stage estimates of Equation (3). This is, again, equivalent to individually estimating these effects in separate regressions. In fact, the first row of panel B is identical to panel A. For the most part, the first stage effects follow an intuitive pattern across all the potential lotteries. Column 1 shows that winning any lottery has a strong effect on ever being admitted, but this effect decreases in later lotteries because of the potential to enter a fifth application. Winning later lotteries seems to increase the likelihood of enrolling. This makes sense, since there are fewer always-takers implied by later lotteries to drive down the first stage. The fourth lottery has little effect on enrollment, likely because of the relatively small group of students admitted and also because so many of the losers reapply. The final panel of Table 4 shows the first stage estimate of Equation (4). This is essentially a weighted average of the effects from panel B. The results are consistent with the results in panel A.[32] However, the first stage power is stronger when pooling the lotteries in panel C. Another important implication of the table is that the lotteries may not be a strong enough instrument for ADN completion, consistent with the reasons outlined in section 4.

5.3 Results on Earnings

Figure 2 shows estimates of the reduced form effect of winning a lottery on earnings over time. The figure plots the first-lottery estimate as well as the pooled lottery estimate that includes information from all four lotteries a student can enter. Not surprisingly, there is no statistically significant effect in the years prior to the lottery occurring. There is a slight increase in the coefficient—albeit not a statistically significant one—in the first year. This is because the program takes four consecutive semesters to complete, sometimes including a summer, so completers who exit on time see large earnings returns between their first and second year following the lottery. The main effects of the program can start to be seen, however, in the second full year. The point estimate is approximately 0.10 log points, suggesting an 11 percent difference in earnings between lottery winners and lottery

[32] Appendix Table A3 shows comparable estimates, this time including covariates. The results are almost identical, which is not surprising given the instrument is a random lottery.

losers.[33] The effect flattens out in the fifth year, stabilizing at approximately 0.17 log points, or 19 percent. In order to estimate this stable, relatively long-term effect, for the rest of the paper I limit the outcome to mean earnings five to seven years after the lottery. Appendix Table A4 shows estimates of the reduced form using all three sets of possible instruments, and shows that the results are robust to inclusion of demographic, academic, labor market, and financial aid covariates.[34]

An important implication of Figure 2 is that using the first lottery only or using all four possible lotteries yields almost identical estimates. The latter approach, though, leverages more variation and has a smaller confidence interval, so this is what I show for the rest of the paper.

Table 5 displays the main instrumental variables estimates of the effect of enrolling in the program. I include four instruments, one for each lottery outcome. Column 1 only controls for cohort and lottery instance, suggesting that enrollment in the program leads to an earnings effect of 0.42 log points, or 52 percent. Inclusion of covariates in the next four columns does not significantly affect the estimate. To account for the effect of the lottery on delaying students' entry into full-time employment, in the sixth column I also instrument for the number of lotteries applied.[35] The effect of an additional lottery on earnings is negative, albeit not statistically significant. Including the number of lotteries applied also lowers the estimate of the effect of starting the program. These estimates suggest that applying to an additional lottery—and thus delaying entry into the labor market for an additional six months— decreases earnings by four percent. This drop is consistent with the literature on early career wage growth: earnings in the first few years tend to grow the fastest, with estimates ranging between six and 12 percent annually (Adda et al., 2013; Buchinsky et al., 2010; Munasinghe, Reif and Henriques, 2008; Altonji and Williams, 2005; Dustmann and

[33] In the log earnings specification, the percent earnings increase can be calculated as $e^{\beta} - 1$, where β is the estimated coefficient on log earnings.

[34] Demographic controls include race, gender, and age at application. Academic controls include the number of units attempted in community college up to the term of application, as well as GPA. Labor market covariates include mean quarterly earnings in the two years prior to application, as well as a dummy for the whether the student had ever worked in the health industry. Financial aid controls include receipt of a fee waiver, Pell grants and any type of financial aid loan. All the controls are interacted with the lottery instance.

[35] As shown previously, the first stage effect of admission on additional time spent enrolled in community college is not strong, because the lottery lengthens the process for both students who enroll in the two year program as well as losers who reapply. However, winning a lottery is highly predictive of applying to fewer lotteries.

Meghir, 2005).

The last column of Table 5 shows estimates of the effect of receiving an ADN, as opposed to merely enrolling. This is effectively a rescaling of the reduced form by the smaller first stage of winning the lottery on completion. The results suggest that finishing an ADN, either at Central College or at another community college in the state, increases earnings by 91 percent. Again, caution should be used to interpret these coefficients: the substantial attrition in the program means that many compliers—in particular the subset who won a subsequent lottery, enrolled, but did not finish—are partially treated by virtue of having taken a few courses in nursing. Still, the results are consistent with the program having very large earnings effects.

Figure 3 shows the estimated return for enrollment and completion over time. In both panels, there is no effect prior to the lottery, and the effect stabilizes within four years of the lottery.

5.4 Results on Employment

Taken together, the results so far suggest large impacts of an ADN on earnings. The large earnings effects may come from increased working hours or wages if the program provides students the skills and network connections to obtain stable employment in any occupation.[36] On the other hand, the degree may steer graduates into high-paying jobs by conferring upon them the necessary credentials to enter registered nursing. The data I use do not contain information on occupation, wages, or work hours, so it is not possible to explicitly parse through these arguments. I draw some suggestive evidence, however, from detailed information on industry of employment.

Column 1 of Table 6 shows estimates of the effect of enrolling in the program on having any employment five to seven years after the lottery.[37] While the estimate is positive, it is small and not statistically significant. Similarly, Column 2 shows the effect on the number of quarters employed, and again the results are not statistically significant. Column 3 of Table 6, however, shows that there is a 22 percentage point effect of enrolling in the program on being employed in the health

[36] In fact, nationwide, based on data in the 2014 ACS, workers in the healthcare sector with an associate's or bachelor's degree were 20 percentage points more likely to work full-time than healthcare workers without these credentials. Only 22 percent of workers employed as registered nurses worked part-time.

[37] Employment is defined as having nonzero earnings.

industry.[38] This is a large effect, especially since so many of the applicants had worked in the health industry prior to applying to the program.

These findings suggest that the large effects on earnings comes from earnings conditional on employment. That is, the program seems to lead to significant occupational sorting. This is at least suggestive evidence that the program drives participants to more lucrative occupations, as opposed to increasing their likelihood of employment or improving their hours.

5.5 Robustness Checks

Table 7 shows a series of robustness checks. First, Column 1 shows the main results without weighting according to the number of applications a student entered, which effectively upweights the participants of later lotteries. The estimated coefficient, 0.398, is similar to the main estimate of 0.446.

Column 2 excludes all students who applied for a fifth time. Students who are admitted on their fifth attempt and enroll are always-takers for all the lotteries. The result is slightly larger than the main estimate, but not by a statistically or economically meaningful amount. In a similar exercise, Column 3 excludes students who gained admission through some method other than the lottery itself. Although I do not know the specific reason for why these students were admitted, in conversations with program administrators I learned that these students are often military veterans or students with a special arrangement from a local hospital.[39] In the main analyses I code these students as not gaining admission. Excluding them from the analysis altogether does not significantly affect the estimates.

In the fourth column I limit the sample to students who had non-zero earnings prior to first application. The estimated coefficient is slightly larger than the main estimate, but not statistically different. One potential concern is that the cause of the large returns may be from students

[38] In the UI earnings data I can observe industry of employment, not occupation. I use the two-digit NAICS code 62, which indicates Health Care and Social Assistance. This category encompasses smaller categories such as Offices of Physicians; Outpatient Care Centers; Medical and Diagnostic Laboratories; Hospitals; and Nursing and Residential Care Facilities. Because registered nurses can and do work in all these settings, I do not disaggregate the industry codes beyond the two-digit code.

[39] I had these conversations over a period of two days as part of a qualitative study on the goals and pathways of students in career technical programs.

transferring to four-year colleges, making the ADN itself just an intermediary step. Column 5 excludes students who transferred to a four-year institution, and still reveals a large and unchanged coefficient.

So far, the robustness exercises have used the same methods, but changed the sample. Next, in Columns 6 and 7, I consider two different methods of estimating the effect of the program on earnings. First, I consider a conventional panel, which uses a single instrument based on whether a student ever won a lottery. The timing of the outcomes depend on the last lottery a student entered, either because they were admitted or because they chose not to reapply. This is conceptually simpler than the stacked panel analysis I have used thus far, although it does have some concerns about endogenous selection. In particular, the timing of a student's last lottery is not controlled for. Nevertheless, as shown in Column 6, these concerns about selection seem to be slight, however, since the estimates are not too different from the main result.

In Column 7 I show results when I adapt a method used by Cellini, Ferreira and Rothstein (2010) and Gelber, Isen and Kessler (2016), which takes into account the dynamic nature of treatment in this case. Because students may reapply if they lose a lottery, the treatment effect identified by the stacked lottery specifications does not account for how reapplication following a loss affects earnings. The "one-step" estimator proposed by Cellini, Ferreira and Rothstein (2010) takes these into account. I describe this estimator in more detail in appendix A.3. The coefficient of 0.461 is quite similar to the main estimate, suggesting that reapplication has a limited effect on earnings apart from the effect of admission itself.

Column 8 of Table 7 shows estimates of the main effect in quarterly earnings levels. These estimates are not as precise as the log earnings estimates because the presence of some very high earners increases the dispersion. However, taking the $3,600 per quarter point estimate at face value, the results are similar in scale to the log earnings estimates.[40] In the final column I focus only on female students, and find similar results.

[40] To be precise, the main coefficient of 0.461 suggests a quarterly earnings effect of 59 percent. Mean earnings for the sample prior to application are $4,717, and thus the implied levels effect is $2,760, which is similar to the coefficient I find when I estimate the equation in levels.

6 Individual Fixed Effects

6.1 Method and Main Results

The results shown so far suggest a very large effect of the program on earnings. With the same sample of Central College applicants, in this section I produce estimates using a method increasingly used in the literature. Jacobson, LaLonde and Sullivan (2005) note that individual fixed effects models, which identify effects from within-student changes in earnings, are valid estimates when students have considerable—and observable— pre-enrollment labor market experience. Recent work has applied this method at the community college level to estimate the returns to different career technical degree programs, credentials, and coursework using state administrative datasets.[41]

Because individual fixed effects models do not account for time-varying differences in ability, there are lingering concerns about whether they produce causal estimates. Thus, I am in a unique position to investigate this issue by comparing the results to those from the Central College lottery. This a similar approach as LaLonde (1986), who compares observational and experimental evaluations of training programs. Another reason to estimate individual fixed effects models is that they have tighter confidence intervals, which allows me to explore heterogenous effects.

I estimate a model of the form:

$$y_{it} = \alpha_i + \gamma D_{it} + \Phi Z_{it} + \mu_t + u_{it} \tag{5}$$

For student i in quarter t, D_{it} takes a value of one after enrolling in an ADN program or, in another specification, finishing the program. The matrix Z_{it} consists of time-varying individual characteristics, including dummies for age and whether the student was taking any courses that quarter. The individual fixed effect, α_i, accounts for time-invariant characteristics, so that the effect of enrollment or finishing on earnings, γ, is identified off within-individual changes in earnings. I also include calendar year and quarter fixed effects, μ_t.

[41] For example, in Arkansas (Belfield, 2015); California (Stevens, Kurlaender and Grosz, 2015; Bahr, 2014); Kentucky (Jepsen, Troske and Coomes, 2014); Michigan (Bahr et al., 2015); North Carolina (Liu, Belfield and Trimble, 2015; Xu and Trimble, 2015); Washington (Dadgar and Trimble, 2015) and two unnamed states (Liu and Belfield, 2014).

Equation (5) can be estimated with just students who enrolled or finished the program, but it is also useful to have a comparison group to account for the counterfactual earnings trajectories of students who never receive the treatment D_{it}. In particular, this comparison group serves to identify the year and age effects, as well as other covariates. Ideally, this comparison group should represent a group of students with similar patterns of earnings prior to enrollment. Recent studies have differed in how to construct a comparison group whose earnings trajectory would present a natural counterfactual to that of graduates. Some have compared graduates and non-graduates among cohorts of new students (Liu, Belfield and Trimble, 2015; Bahr et al., 2015), or categorize students by their stated academic goals (Jepsen, Troske and Coomes, 2014) or course-taking behavior (Stevens, Kurlaender and Grosz, 2015).[42] Applicants to the program make a natural comparison group. Although they are highly selected among typical community colleges, their earnings trajectories are the most similar to students who earn an ADN, especially since admission is based on a random lottery.

The first column of Table 8 shows my preferred IV estimates of the effect of enrolling in and finishing an ADN program. The second column shows the corresponding OLS estimate. The selection bias seems to be substantial and also negative. The third column shows estimates of Equation (5) using the same sample of applicants, with a coefficient of 0.43 for enrollment and 0.57 for completion. Both coefficients in the first column are statistically indistinguishable from their counterparts in the third column. This similarity has an important implication. The instrumental variables estimates are not susceptible to issues of omitted variables bias. On the other hand, any time-varying shocks that affect both the earnings and academic performance of students render the individual fixed effects estimates invalid, and are usually assumed to be small (Jepsen, Troske and Coomes, 2014). The results in Table 8 show the two estimates are relatively similar, and thus lend additional credence to the growing recent literature producing such estimates from large observational datasets.

In order to further link these results to that literature and to later explore heterogeneity across colleges, I expand the sample to include all ADN programs in the state. This necessitates a slight

[42] Although they do not use individual fixed effects, Zeidenberg, Scott and Belfield (2015) focus on this issue of the degree intentions of new students by creating an algorithm to identify the program a student was likely interested in.

change of the comparison group since I cannot identify program applicants at any college except Central College. Instead of comparing statewide ADN earners to applicants, I compare them to all students who started a program.[43]

The fourth column of Table 8 shows estimates of the return to an ADN at Central College with the sample being any student who enrolled in the program. The coefficient, 0.46, is smaller than in Column 3, which suggests that much of the effect of the program comes from the degree as opposed to enrolling.[44] In prior work, where the control group consisted of likely degree seekers in California, I found a log earnings result of 0.695 (0.01) (Stevens, Kurlaender and Grosz, 2015), which is higher than the estimates from Table 8. Bahr et al. (2015), who use an even broader sample consisting of first-time students, find an estimate of 1.030 (0.049) for women and 0.862 (0.136) for men. Taken together, there is a pattern of declining estimates as the comparison group grows more similar to the treatment group. Of course, this observation should be taken with the caveat that these studies also involve different policy contexts across states and time periods.

In the final two columns of Table 8 I expand the sample to include students who started an ADN program at any of the state's colleges. This improves precision dramatically by increasing the number of observations. The OLS estimate is, again, quite small. However, the fixed effects coefficient of 0.50 is similar though slightly larger to that in Column 5, suggesting that Central College students saw an earnings effect slightly lower than the state average. In the next analyses I explore heterogeneity across individuals and across the state's many ADN programs in more detail. Appendix A.4 describes various robustness exercises for these specifications.

6.2 Heterogeneity by Student and College Characteristics

The instrumental variables sample from the Central College lottery is too small to adequately explore heterogenous treatment effects. Because the individual fixed effects estimates are quite

[43] The set of all students who start an ADN program is still a reasonable comparison group for students who completed a degree. Unfortunately, I only have information on applicants for one college: Central College. In expanding the analysis to other colleges the best comparison group for completers is other enrollees. I also cannot estimate an enrollment effect with this sample since the comparison group also enrolled.

[44] For example, if the impact were entirely on enrolling and not on the degree itself, the coefficient in column 3 would be zero. The fact that it is lower but still quite close to the estimate in Column 2 suggests that much of the effect depends on actually earning the degree.

similar, I leverage the added precision of this approach to explore heterogeneity.

I am particularly interested in differences across programs producing the same degree. Recent evidence has demonstrated that there is a great deal of heterogeneity in outcomes at the community college level, even accounting for student characteristics (Stange, 2012; Clotfelter et al., 2013; Cunha and Miller, 2014; Kurlaender, Carrell and Jackson, 2016). Figure 4 plots estimates of the return to an ADN from Equation (5), calculated for each individual college, arranged in ascending order of the estimated coefficient. The vast majority of estimates are large and positive, yet there is a considerable range of estimates: weighted in terms of the number of graduates, the difference between the college at the 25th percentile and the 75th percentile of the estimated coefficient is 0.30, or 35 percent.[45] This range is surprising, given that all ADN programs offer a similar curriculum, have similar prerequisites, are overseen by a state board, and are more likely to draw students from the whole state than other programs.

Program-level heterogeneity has a few possible explanations. First, there may be differences in program quality. Quality is difficult to measure, especially since so many program inputs, such as curriculum and faculty-student ratios, are determined by the state board of nursing. I examine program pass rates, program size, and licensing exam pass rates.[46]

Program-level heterogeneity may also be due to local economic conditions. As a measure of the density of job opportunities, I collected data on the capacity of hospitals in the state, from which I created measures of hospital beds and long-term care beds per capita.[47] As a rough estimate of demand for healthcare, I categorized counties by median share of the population age 60 and over in the 2000 Census. Another important aspect of local labor markets that might affect the returns to an ADN are the employment prospects of workers who typically enter nursing programs, as well as the employment and earnings of other nurses. I created measures of the employment level and earnings of nurse assistants, medical assistants, and registered nurses relative to overall

[45] Weighting instead by the number of students who start the program yields an identical estimate.

[46] All potential nurses must take the NCLEX-RN. Program-level pass rates for first-time test takers are published online for the past five years. There is surprisingly little variation: most colleges had pass rates above 80%, and students may retake the exam multiple times.

[47] These data are available from the California Office of Statewide Health Planning and Development, and include information on all licensed facilities in the state. Facility data only exist at one point in time, so the measures I use here do not take into account any facility closures or openings during the time period.

employment and earnings.

Finally, a key and policy-relevant institutional characteristic is admission type. A common refrain when talking to ADN program administrators is that the lottery system does not allow them to admit the most qualified students. To investigate this further, I categorized the admissions process of each of the programs that granted an ADN by reading about its policies in the course catalog and student handbook, available on program and college websites.[48] There were 12 programs whose policies featured a lottery among all eligible applicants, like at Central College. There were an additional 12 programs at which admission had some form of randomized component, as well as some type of screening.[49] There were 40 "competitive" programs, at which admission decisions were purely based on the applicant's qualifications and there was no randomization. In an additional nine programs admission was based on a waitlist or first-come-first-served.[50] Appendix Table A6 shows that there are no substantial differences in the characteristics of incoming students by type of admission.

There may also be differences in earnings returns across different types of students. Jacobson, LaLonde and Sullivan (2005) show that older students, for example, see smaller returns to retraining programs.

To investigate which characteristics are most correlated with earnings returns, I interact the main variable of interest with an indicator for the particular group:

$$y_{it} = \alpha_i + \gamma_1 D_{it} + \gamma_2 (D_{it} * 1(X_i = x)) + \Phi Z_{it} + \mu_t + u_{it} \tag{6}$$

The term $1(X_i = x)$ is a dummy variable indicating whether a student is a member of a certain group. Table 9 displays estimates of the coefficient on the interaction terms γ_2 for institutional and individual level characteristics. The interaction terms are defined as programs above or below the

[48] Programs may change their admissions requirements across years, but I can only observe the policy for the years that college catalogs are available. I use the most recent year available. In recent years, more colleges in the state have moved towards admissions based on multiple-screening criteria (Moltz, 2010)

[49] For example, at some of these programs students scored points based on their grades and prior work experience, and then a lottery was instituted among the highest point scorers.

[50] In the waitlist setup, students add their names to the list whenever they complete their application requirements, and incoming cohorts are admitted from the top of the list. In the first-come-first-served setup, applications are only accepted in a narrow time window each semester, and spots are filled in the order the applications arrive.

median.[51] Appendix Table A7 shows the results of each interaction in a separate regression.

Table 9 suggests that program-level heterogeneity is primarily associated with labor market opportunities. I find no correlation between college quality and estimated returns; if anything, students graduating from colleges with higher NCLEX pass rates saw lower earnings returns. On the other hand, students in counties with a higher density of hospital beds had earnings returns that were approximately 8-9% higher than other students. Similarly, students in counties where medical assistants made relatively lower wages saw higher returns, as did students in counties with higher nursing wages. Column 2 shows that the relative wage of medical assistants to registered nurses is positively correlated with earnings returns. In terms of program characteristics, admissions type is also important. Predictably, students in lotteried programs saw lower returns than students in competitive and wait-list programs, providing more support for the idea of introducing muiltiple-criteria screening admissions processes.

I also find that older and minority students see relatively smaller earnings gains than other students. In addition, Figure 5 shows a strong downward gradient in the earnings effect across students' pre-enrollment annual earnings. This suggests that ADN programs can be significnat sources of upward mobility, especially for students at with the largest potential for earnings gains. A downward gradient is also apparent when estimating the effects in levels. Figure A1 shows that the earnings returns are relatively consistent across cohorts.

7 Costs and Program Expansion

With ADN programs oversubscribed and a growing demand for healthcare workers, a crucial question is to determine whether it is cost-effective to increase capacity to ADN programs. In this section I calculate the private benefit to students implied by the estimates discussed previously, and then compare these to a rough calculation of the cost of a marginal expansion of a nursing program.

Table 10 shows calculations of the present discounted value of enrolling and completing an

[51] Because so many colleges are in Los Angeles County, it is not possible to explore heterogeneity across more than two quantiles

ADN based on my preferred estimates. I assume that the counterfactual earnings for students not enrolled in the program are $5,757 per quarter, which is the mean earnings for students when they apply to the Central College program. To provide conservative estimates I assume that students forego all earnings while enrolled and experience no earnings gain over time. In order to be able to compare this number to the cost estimates, I include supplies and uniforms in this calculation, but not tuition and fees.[52] I assume a 20 year career and a 3% real interest rate. Column 2 shows a PDV of $141,986 for enrolling and $210,913 for completing the Central College program. In column 3 I allow for a modest 2% earnings growth per year, and in column 4 I assume students only forego half their earnings while studying. All told, the private value of enrolling in the program is approximately $150,000 and the value of completing the degree is approximately $220,000. There are obviously also social benefits to additional nurses. Dall et al. (2009) and Needleman et al. (2006) estimate that avoided adverse health outcomes and cost savings from an additional nurse are approximately $40,000 to $57,000.

How do the benefits compare to the costs of program expansions?[53] Estimates of the cost of marginally increasing the capacity of a nursing program are difficult to find.[54] To calculate a back-of-the-envelope estimate, I break down the costs into operating costs and capital costs. I obtained per-student operating costs from two nursing programs in the state, from which I estimate that each student in a nursing program costs $7,600 to $9,200 per year to educate. A similar estimate comes from a California legislature initiative over the past few years that granted expansion funds to nursing programs. These range between $6,500 and $9,100 per new full-time equivalent student (California Community College Chancellor's Office, 2015, 2016*a*). Rounding up, a conservative estimate is that operating costs are $10,000 per year per student.

Much less information is available about capital, infrastructure, and equipment costs. Health-care programs are expensive and require specialized machinery and teaching equipment, and

[52] The Central College ADN program estimates a cost to students of $5,710 over two years in supplies, immunization and other records, and books. This does not include approximately $2,100 in tuition and fees at $46 per credit hour.

[53] This discussion is not meant to be exhaustive. In particular, I am assuming that there are no general equilibrium effects of expanding capacity, which would likely decrease the estimated return.

[54] There is no study to my knowledge that explicitly estimates these costs. Information for the costs of new buildings tend to include costs of expanding other programs in addition to nursing.

instruction often occurs in dedicated facilities. In the past five years two community colleges completed construction for new nursing buildings. Both projects cost approximately $12 million for the addition of 50 new students each year. If a new nursing building is replaced after 30 years, this amounts to $8,000 per new student in up-front expansion costs. In addition, I estimated that teaching equipment, such as practice maniquins, cost $1,000 per additional student if replaced every five years.[55] Thus, the infrastructure and capital spending associated with adding an additional student is approximately $9,000. In sum, a conservative estimate of the total cost of increasing capacity to a nursing program by one student is $19,000. While large, this is a fraction of the smallest private benefit implied by the earnings returns and much smaller when incorporating the likely positive spillovers. Moreover, since California's state income marginal tax rate for workers earning in a similar range as the students is between 6 and 9%, the expansion would almost pay for itself given the private benefit, making expansion likely revenue neutral (Franchise Tax Board, 2015).

Despite evidence that nursing programs are overwhelmingly cost-effective, however, there are a number of reasons why colleges do not increase their capacity. In California, as in many other states, the incentive structure of college expenses and revenue is not aligned for expansion. Regardless of the program, colleges receive a set per-pupil allocation. At $4,900 in the most recent year, this per-pupil allocation is approximately half the cost of regular operating expenses for a nursing program. Colleges tend to recoup the costs of expensive programs by increasing enrollment in less costly programs or through external grants. Moreover, expanding a program involves upfront costs in addition to regular operating costs, such as buying new simulation and teaching equipment or hiring new faculty to meet state-mandated ratios (Kuehn, 2007). Instead, capacity expansions have overwhelmingly occured at the state level, if at all. Thus, it is perhaps not surprising that despite large earnings returns and concerns about nursing shortages, demand for seats in nursing programs still outpaces supply.

[55] One college provided me a list of its inventory, and I searched for the prices of the listed prices online.

8 Conclusion

In this paper I estimate the effect of an associate's degree in nursing on earnings and employment. I leverage the random assignment of admissions to a large community college program, thus providing one of the first estimates of the returns to a postsecondary degree using variation resembling an experiment. By taking advantage of a rich dataset describing the academic and labor market experiences of millions of students, I also estimate models using methods that are more common in the literature.

I find that enrolling in a nursing program results in a 55 percent increase in quarterly earnings. Accounting for tuition and other costs, I estimate that the value of earning an ADN is approximately $200,000, which far outweighs conservative estimates of the costs of expanding programs. Despite the large economic benefit to nursing programs, there are limited incentives to community college administrators to expand enrollments. This is important in light of recent discussions and debates in the policy arena: my results suggest that state and federal efforts to increase address nursing shortages by expanding training programs are likely cost effective. Another potential solution might be adopting differential pricing or funding of programs, which are policies that have received recent attention (Stange, 2012; Smith, 2016; Long, 2016).

This paper also contributes to a more general discussion of accessibility in higher education. Calls of "college-for-all" and free community college have in the past few years reached the highest levels of policy.[56] The results of this paper suggest some caution when taking action on these initiatives. Apart from the low completion rates in the program, a product of the lottery process is that the majority of applicants are not offered admission to their program of choice. These applicants, who have completed a battery of intensive prerequisite coursework, are among the highest-performing community college students in the state. Thus, any policy that intends to increase access to more students should also take care in improving access to programs for students who have already embarked on academic pathways within the community colleges.

Some relevant questions remain outside the scope of this paper. A key issue, especially

[56] As mentioned previously, a number of states have free community college programs. Free four-year college, either for all or a majority of students, has been a main item of discussion in the presidential campaigns of Hillary Clinton and Bernie Sanders, and the subject of a proposed policy by the Obama administration.

as demand for entry to nursing programs outpaces supply, is how to efficiently allocate these spots. I show that students in lottery programs see smaller earnings returns than students in other programs, but these results are suggestive. More research is needed in studying the implications of how colleges manage the scarce resources available to run high-demand programs.

Another open question is what portion of the large earnings effects I estimate come from restricting access to seats in programs. More generally, it is important to understand how community college career technical programs affect local labor supply. Recent work has explored inefficiencies from occupational licensing and credentialism (Kleiner, 2016; Gittleman, Klee and Kleiner, 2015), but less is known about the effects of restricted educational supply on the local economy.

References

Acemoglu, Daron, and David Autor. 2011. "Skills, Tasks and Technologies: Implications for Employment and Earnings." , ed. O. Ashenfelter and D. Card Vol. 4 of *Handbook of Labor Economics*, Chapter 12, 1043–1171. Elsevier.

Adda, Jerome, Christian Dustmann, Costas Meghir, and Jean-Marc Robin. 2013. "Career Progression, Economic Downturns, and Skills." National Bureau of Economic Research Working Paper 18832.

Altonji, Joseph G., and Nicolas Williams. 2005. "Do Wages Rise with Job Seniority? A Reassessment." *Industrial & Labor Relations Review*, 58(3): 370–397.

Altonji, Joseph G., Erica Blom, and Costas Meghir. 2012. "Heterogeneity in Human Capital Investments: High School Curriculum, College Major, and Careers." *Annual Review of Economics*, 4(1): 185–223.

Anderson, Theresa, Jamie Hall, and Teresa Derrick-Mills. 2013. "Introduction to the Health Profession Opportunity Grants (HPOG) Program and First Year Implementation and Outcomes." The Urban Institute Practice Brief.

Anderson, Theresa, Lauren Eyster, Robert I. Lerman, Maureen Conway, Ranita Jain, and Marcela Montes. 2016. "Implementation of Accelerating Opportunity: Final Implementation Findings with Lessons from the Field." The Urban Institute Practice Brief.

Angrist, Joshua D., and Alan B. Krueger. 1992. "Estimating the Payoff to Schooling Using the Vietnam-Era Draft Lottery." National Bureau of Economic Research, Inc NBER Working Papers 4067.

Angrist, Joshua D., Parag A. Pathak, and Christopher R. Walters. 2013. "Explaining Charter School Effectiveness." *American Economic Journal: Applied Economics*, 5(4): 1–27.

Arcidiacono, Peter. 2004. "Ability sorting and the returns to college major." *Journal of Econometrics*, 121(12): 343 – 375. Higher education (Annals issue).

Ashenfelter, Orley, and Alan B Krueger. 1994. "Estimates of the Economic Returns to Schooling from a New Sample of Twins." *American Economic Review*, 84(5): 1157–73.

Auerbach, David I., Douglas O. Staiger, Ulrike Muench, and Peter I. Buerhaus. 2013. "The Nursing Workforce in an Era of Health Care Reform." *New England Journal of Medicine*, 368(16): 1470–1472. PMID: 23594001.

Auerbach, David I, Peter I Buerhaus, and Douglas O Staiger. 2011. "Registered nurse supply grows faster than projected amid surge in new entrants ages 23–26." *Health Affairs*, 30(12): 2286–2292.

Auerbach, David I, Peter I Buerhaus, and Douglas O Staiger. 2015. "Do associate degree registered nurses fare differently in the nurse labor market compared to baccalaureate-prepared RNs?" *Nursing Economics*, 33(1): 8.

Bahr, Peter Riley. 2014. "The labor market return in earnings to community college credits and credentials in California." *Ann Arbor, Michigan: Center for the Study of Higher and Postsecondary Education, University of Michigan. Working Paper*.

Bahr, Peter Riley, Susan Dynarski, Brian Jacob, Daniel Kreisman, Alfredo Sosa, and Mark Wiederspan. 2015. "Labor Market Returns to Community College Awards: Evidence From Michigan." CAPSEE Working Papers.

Bailey, Thomas R, D Timothy Leinbach, Marc Scott, Mariana Alfonso, Gregory S Kienzl, Benjamin Kennedy, and D Marcotte. 2003. "The characteristics of occupational sub-baccalaureate students entering the new millennium." *National Assessment of Vocational Education, US Department of Education, Washington, DC*.

Barnow, Burt S., and Jeffrey Smith. 2015. "Employment and Training Programs." In *Economics of Means-Tested Transfer Programs, volume 2. NBER Chapters*. National Bureau of Economic Research, Inc.

Bedard, Kelly. 2001. "Human Capital versus Signaling Models: University Access and High School Dropouts." *Journal of Political Economy*, 109(4): 749–775.

Belfield, Clive. 2015. "Weathering the Great Recession with Human Capital? Evidence on Labor Market Returns to Education from Arkansas. A CAPSEE Working Paper." *Center for Analysis of Postsecondary Education and Employment*.

Belfield, Clive R., and Thomas Bailey. 2011. "The Benefits of Attending Community College: A Review of the Evidence." *Community College Review*, 39(1): 46–68.

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

Bohn, Sarah, Belinda Reyes, and Hans Johnson. 2013. "The impact of budget cuts on Californias community colleges." *Public Policy Institute of California*.

Bound, John, and Sarah Turner. 2007. "Cohort crowding: How resources affect collegiate attainment." *Journal of Public Economics*, 91(5): 877–899.

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

Buchinsky, Moshe, Denis Fougere, Francis Kramarz, and Rusty Tchernis. 2010. "Interfirm mobility, wages and the returns to seniority and experience in the United States." *The Review of Economic Studies*, 77(3): 972–1001.

Buerhaus, Peter I., David I. Auerbach, and Douglas O. Staiger. 2009. "The Recent Surge In Nurse Employment: Causes And Implications." *Health Affairs*, 28(4): w657–w668.

Buerhaus, Peter I, David I Auerbach, and Douglas O Staiger. 2014. "The rapid growth of graduates from associate, baccalaureate, and graduate programs in nursing." *Nursing Economics*, 32(6): 290.

Buerhaus, Peter I, David I Auerbach, Douglas O Staiger, and Ulrike Muench. 2013. "Projections of the long-term growth of the registered nurse workforce: a regional analysis." *Nursing Economics*, 31(1): 13–17.

Bureau of Labor Statistics. 2015. " US Department of Labor." *2014-2015 Occupational Outlook Handbook*.

California Board of Registered Nursing. 2015. "NCLEX Pass Rates." http://www.rn.ca.gov/schools/passrates.shtml [Accessed: 2015-05-22].

California Community College Chancellor's Office. 2015. "Nursing Educational Programs."

California Community College Chancellor's Office. 2016*a*. "Nursing Educational Programs."

California Community College Chancellor's Office. 2016*b*. "Student success scorecard: 2015 state of the system report."

Cellini, Stephanie Riegg, and Nicholas Turner. 2016. "Gainfully Employed? Assessing the Employment and Earnings of For-Profit College Students Using Administrative Data." National Bureau of Economic Research Working Paper 22287.

Cellini, Stephanie Riegg, Fernando Ferreira, and Jesse Rothstein. 2010. "The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design." *The Quarterly Journal of Economics*, 125(1): 215–261.

Clotfelter, Charles T, Helen F Ladd, Clara G Muschkin, and Jacob L Vigdor. 2013. "Success in community college: Do institutions differ?" *Research in Higher Education*, 54(7): 805–824.

Cohen, A.M., F.B. Brawer, and J.R. Lombardi. 2009. *The American Community College.* Wiley.

Coons, Arthur G, A Browne, Howard Campion, G Dumke, T Holy, D McHenry, et al. 1960. "A master plan for higher education in California, 1960-1975." *Sacramento: California State Department of Education*.

Cunha, Jesse M, and Trey Miller. 2014. "Measuring value-added in higher education: Possibilities and limitations in the use of administrative data." *Economics of Education Review*, 42: 64–77.

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

Dall, Timothy M, Yaozhu J Chen, Rita Furst Seifert, Peggy J Maddox, and Paul F Hogan. 2009. "The economic value of professional nursing." *Medical Care*, 47(1): 97–104.

Darolia, Rajeev, Cory Koedel, Paco Martorell, Katie Wilson, and Francisco Perez-Arce. 2015. "Do Employers Prefer Workers Who Attend For-Profit Colleges? Evidence from a Field Experiment." *Journal of Policy Analysis and Management*, 34(4): 881–903.

Deming, David J., Noam Yuchtman, Amira Abulafi, Claudia Goldin, and Lawrence F. Katz. 2016. "The Value of Postsecondary Credentials in the Labor Market: An Experimental Study." *American Economic Review*, 106(3): 778–806.

Doolittle, Fred, and Linda Traeger. 1990. "Implementing the National JTPA Study."

Dustmann, Christian, and Costas Meghir. 2005. "Wages, Experience and Seniority." *The Review of Economic Studies*, 72(1): 77–108.

Executive Office of the President. 2015. "America's College Promise: A Progress Report on Free Community College."

Field, Kelly. 2014. "Obama and Biden Promote $550-Million in New Job-Training Funds." *Chronicle of Higher Education*. [Online; posted 14-April 2014].

Flores-Lagunes, Alfonso, Arturo Gonzalez, and Todd Neumann. 2010. "Learning but not Earning? The Impact of Job Corps Training on Hispanic Youth." *Economic Inquiry*, 48(3): 651–667.

Franchise Tax Board. 2015. "California 540: 2015 Personal Income Tax Booklet."

Gelber, Alexander, Adam Isen, and Judd B. Kessler. 2016. "The Effects of Youth Employment: Evidence from New York City Lotteries." *The Quarterly Journal of Economics*, 131(1): 423–460.

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

Gittleman, Maury, Mark A. Klee, and Morris M. Kleiner. 2015. "Analyzing the Labor Market Outcomes of Occupational Licensing." National Bureau of Economic Research Working Paper 20961.

Goldin, C.D., and L.F. Katz. 2008. *The Race Between Education and Technology.* Harvard University Press.

Grubb, W Norton, and Marvin Lazerson. 2009. *The Education Gospel.* Harvard University Press.

Gurantz, Oded. 2015. "Who Loses Out?: Registration Order, Course Availability, and Student Behaviors in Community College." *The Journal of Higher Education*, 86(4): 524–563.

Hanushek, Eric A., John F. Kain, Steven G. Rivkin, and Gregory F. Branch. 2007. "Charter school quality and parental decision making with school choice." *Journal of Public Economics*, 91(5-6): 823–848.

Heckman, James J, and Jeffrey Smith. 2000. "The Sensitivity of Experimental Impact Estimates (Evidence from the National JTPA Study)." In *Youth Employment and Joblessness in Advanced Countries*. 331–356. University of Chicago Press.

Heckman, James, Neil Hohmann, Jeffrey Smith, and Michael Khoo. 2000. "Substitution and dropout bias in social experiments: A study of an influential social experiment." *Quarterly Journal of Economics*, 651–694.

Heinrich, Carolyn J, and Peter Mueser. 2014. "Training Program Impacts and the Onset of the Great Recession." *Unpublished manuscript, University of Missouri*.

Heinrich, Carolyn J, Peter R Mueser, Kenneth R Troske, Kyung-Seong Jeon, and Daver C Kahvecioglu. 2013. "Do public employment and training programs work?" *IZA Journal of Labor Economics*, 2(1): 1.

Hoekstra, Mark. 2009. "The Effect of Attending the Flagship State University on Earnings: A Discontinuity-Based Approach." *The Review of Economics and Statistics*, 91(4): 717–724.

Hollenbeck, Kevin, and Wei-Jang Huang. 2006. "Net impact and benefit-cost estimates of the workforce development system in Washington state."

Hotz, V Joseph, Guido W Imbens, and Julie H Mortimer. 2005. "Predicting the efficacy of future training programs using past experiences at other locations." *Journal of Econometrics*, 125(1): 241–270.

Institute of Medicine. 2011. *The Future of Nursing: Leading Change, Advancing Health*. National Academies Press.

Jacobson, Louis, Robert LaLonde, and Daniel G. Sullivan. 2005. "Estimating the returns to community college schooling for displaced workers." *Journal of Econometrics*, 125(1-2): 271–304.

Jepsen, Christopher, Kenneth Troske, and Paul Coomes. 2014. "The Labor-Market Returns to Community College Degrees, Diplomas, and Certificates." *Journal of Labor Economics*, 32(1): 95 – 121.

Kane, Thomas J., and Cecilia Elena Rouse. 1995. "Labor-Market Returns to Two- and Four-Year College." *American Economic Review*, 85(3): 600–614.

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

Ketel, Nadine, Edwin Leuven, Hessel Oosterbeek, and Bas van der Klaauw. 2016. "The returns to medical school: evidence from admission lotteries." *American Economic Journal: Applied Economics*, 8(2): 225–254.

Kirkeben, Lars, Edwin Leuven, and Magne Mogstad. 2014. "Field of Study, Earnings, and Self-Selection." National Bureau of Economic Research Working Paper 20816.

Kleiner, Morris M. 2016. "Battling over Jobs: Occupational Licensing in Health Care." *American Economic Review*, 106(5): 165–70.

Kovner, Christine T, Carol S Brewer, Farida Fatehi, and Carina Katigbak. 2014. "Original research: Changing trends in newly licensed RNs." *AJN The American Journal of Nursing*, 114(2): 26–34.

Kuehn, Bridget M. 2007. "No end in sight to nursing shortage." *Journal of the American Medical Association*, 298(14): 1623–1625.

Kurlaender, Michal, Scott Carrell, and Jacob Jackson. 2016. "The Promises and Pitfalls of Measuring Community College Quality." *RSF: The Russell Sage Foundation Journal of the Social Sciences*, 2(1): 174–190.

LaLonde, Robert J. 1986. "Evaluating the Econometric Evaluations of Training Programs with Experimental Data." *The American Economic Review*, 76(4): pp. 604–620.

Leuven, Edwin, Hessel Oosterbeek, and Inge de Wolf. 2013. "The effects of medical school on health outcomes: Evidence from admission lotteries." *Journal of Health Economics*, 32(4): 698–707.

Liu, Vivian YT, Clive R Belfield, and Madeline J Trimble. 2015. "The medium-term labor market returns to community college awards: Evidence from North Carolina." *Economics of Education Review*, 44: 42–55.

Liu, Yuen Ting, and Clive Belfield. 2014. "The Labor Market Returns to For-Profit Higher Education: Evidence for Transfer Students. A CAPSEE Working Paper." *Center for Analysis of Postsecondary Education and Employment*.

Lockard, C Brett, and Michael Wolf. 2012. "Occupational employment projections to 2020." *Monthly Lab. Rev.*, 135: 84.

Long, Katherine. 2016. "New funding formula will mean cuts at some community colleges."

Maguire, Sheila, Joshua Freely, Carol Clymer, Maureen Conway, and Deena Schwartz. 2010. "Tuning in to Local Labor Markets: Findings from the Sectoral Employment Impact Study." *Public/Private Ventures*.

Moltz, David. 2010. "Vying for limited slots." *Inside Higher Ed*.

Munasinghe, Lalith, Tania Reif, and Alice Henriques. 2008. "Gender gap in wage returns to job tenure and experience." *Labour economics*, 15(6): 1296–1316.

Needleman, Jack, Peter I Buerhaus, Maureen Stewart, Katya Zelevinsky, and Soeren Mattke. 2006. "Nurse staffing in hospitals: is there a business case for quality?" *Health Affairs*, 25(1): 204–211.

Noy, M Van, J Jacobs, S Korey, T Bailey, and K.L. Hughes. 2008. "Noncredit Enrollment in Workforced Education: State Policies and Community College pRactices." *American Association of Community Colleges and Community College Research Center*.

Puma, Michael J, and Nancy R Burstein. 1994. "The national evaluation of the food stamp employment and training program." *Journal of Policy Analysis and Management*, 311–330.

Ross, Martha, Nicole Svajlenka, and Jane R. Williams. 2014. "Part of the Solution: Pre-Baccalaureate Healthcare Workers in a Time of Health System Change." *Metropolitan Policy Program, Brookings Institution*.

Schochet, Peter Z, John Burghardt, and Sheena McConnell. 2008. "Does job corps work? Impact findings from the National Job Corps Study." *The American Economic Review*, 98(5): 1864–1886.

Scrivener, Susan, and Michael J Weiss. 2013. "More Graduates: Two-Year Results from an Evaluation of Accelerated Study in Associate Programs (ASAP) for Developmental Education Students." *Available at SSRN 2393088*.

Smith, Ashley. 2016. "Funding Edge for Workforce Programs." *Inside Higher Ed*.

Stange, Kevin. 2012. "Ability sorting and the importance of college quality to student achievement: Evidence from community colleges." *Education Finance and Policy*, 7: 74–105.

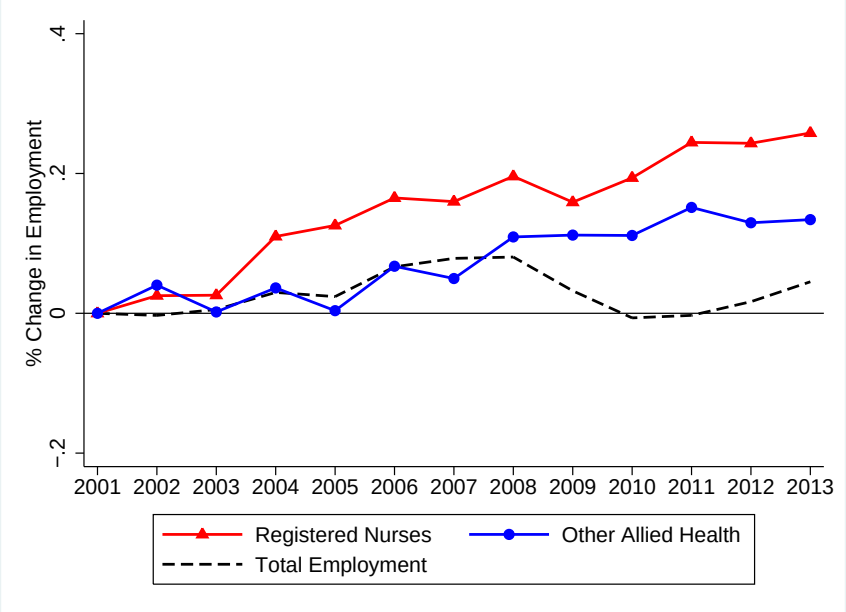
Stevens, Ann Huff, Michal Kurlaender, and Michel Grosz. 2015. "Career Technical Education and Labor Market Outcomes: Evidence from California Community Colleges." National Bureau of Economic Research, Inc NBER Working Papers 21137.

Visher, Mary G., Michael J. Weiss, Evan Weissman, Timothy Rudd, and Heather D. Wathington. 2012. "The Effects of Learning Communities for Students in Developmental Education: A Synthesis of Findings from Six Community Colleges." National Center for Postsecondary Research.

Xu, Di, and Madeline Trimble. 2015. "What about certificates? Evidence on the labor market returns to nondegree community college awards in two states." *Educational Evaluation and Policy Analysis*.

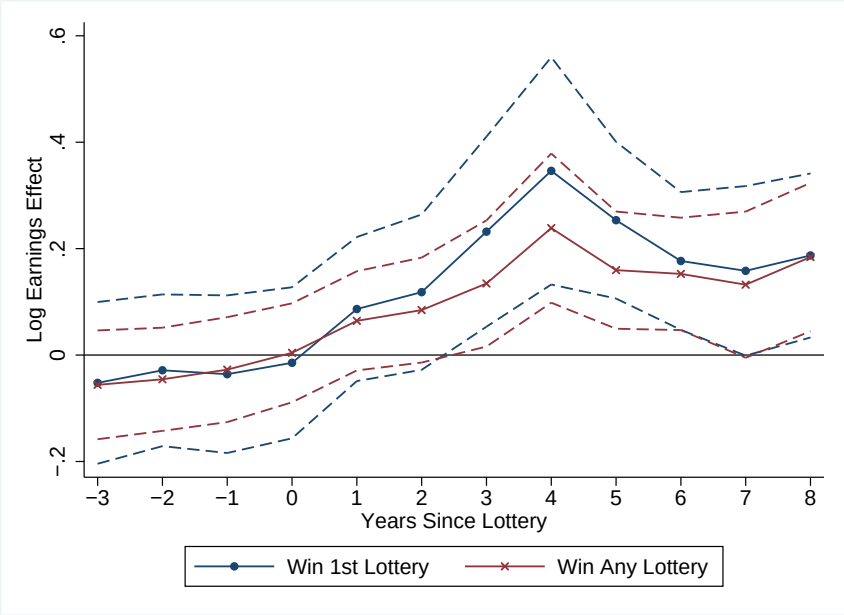
Zeidenberg, Matthew, Marc Scott, and Clive Belfield. 2015. "What about the non-completers? The labor market returns to progress in community college." *Economics of Education Review*, 49: 142 – 156.

Figure 1: Employment Growth for Healthcare Occupations, 2001-2013



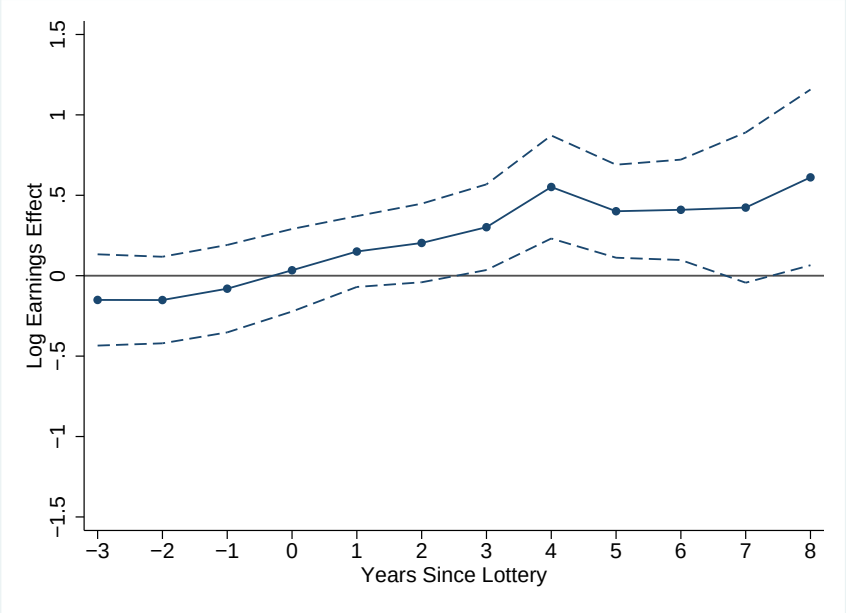
Notes. This graph shows employment growth, relative to 2001 levels, for healthcare occupations that require an Associate's Degree in California. Data come from the Occupational Employment Statistics. The category of Other Health professions include LPN, radiologic technicians, dental hygienists, respiratory care therapists, and surgical technicians.

Figure 2: Reduced Form Effect of Winning Lottery on Log Earnings

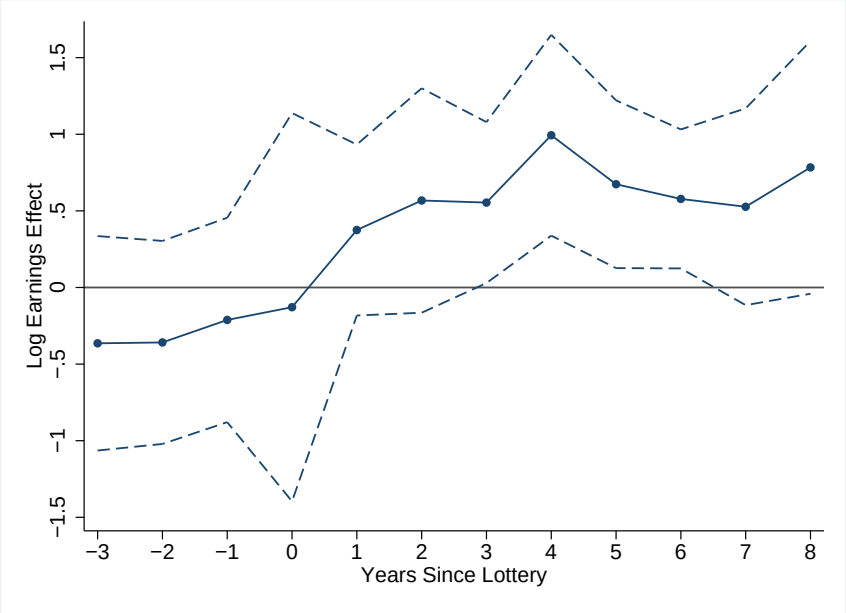


Notes. Figure shows point estimate and 95% confidence interval for reduced form effect of a student winning the first entered lottery, as well as the student winning any of the first four entered lotteries. Outcome is quarterly log earnings, with outcomes grouped by years since the lottery. Both sets of regressions control for year and cohort. The "win any lottery" specification also controls for lottery-instance and is weighted as described in text. Standard errors clustered at the individual level.

Figure 3: Instrumental Variables Estimates on Log Earnings



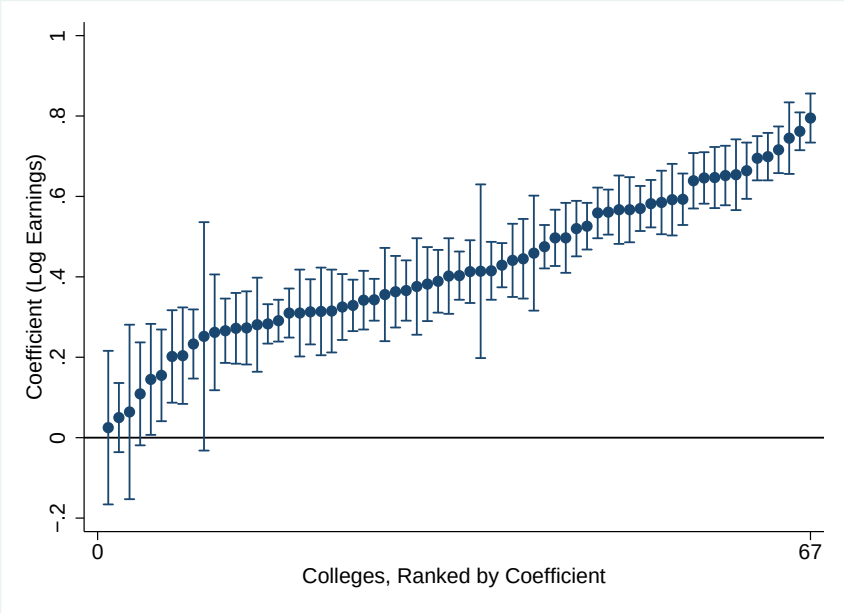
a) Enroll



b) Complete

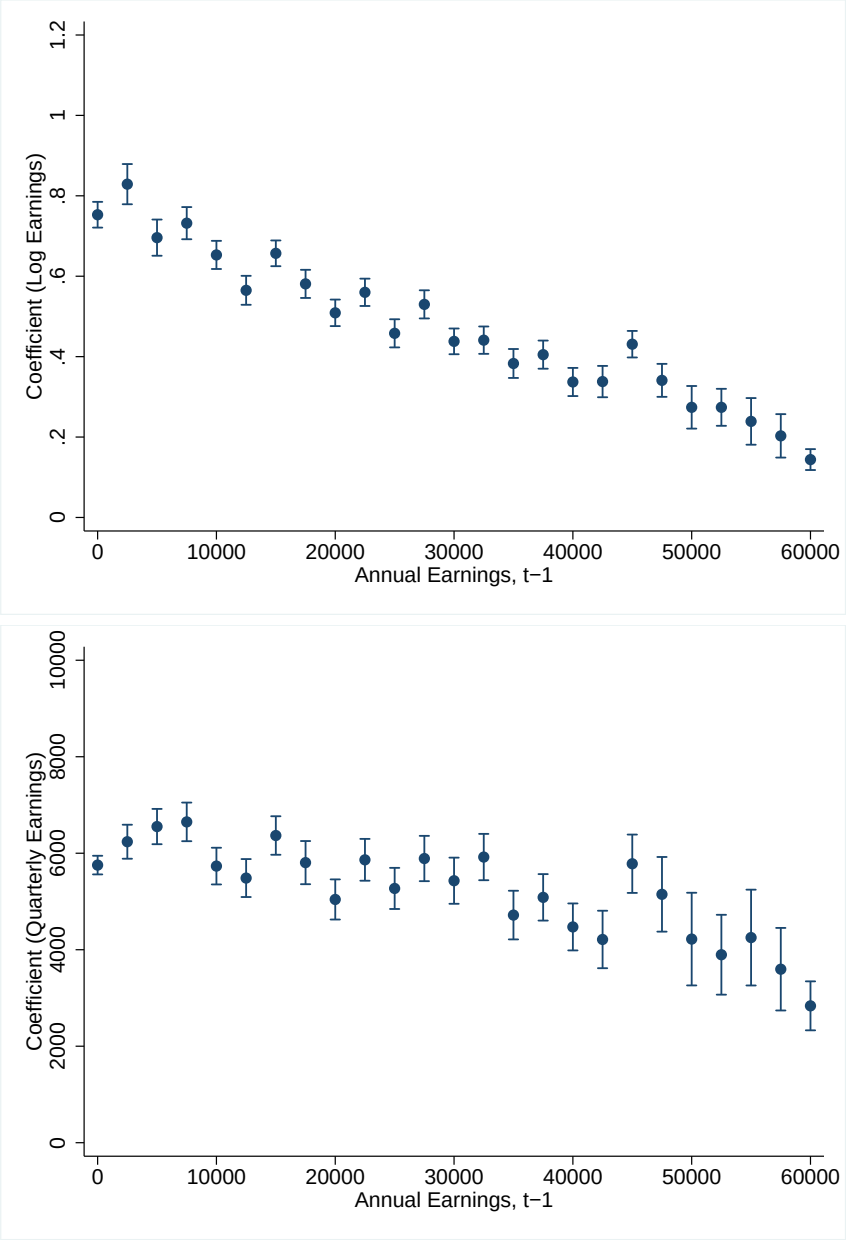
Notes. Figure shows point estimate and 95% confidence interval for instrumental variables estimates of effect of enrolling and completing an ADN program. Outcome is quarterly log earnings, with outcomes grouped by years since the lottery. Both sets of regressions control for year, cohort, and lottery-instance, and are weighted as described in text. Standard errors clustered at the individual level.

Figure 4: Individual Fixed Effects Returns, Heterogeneity by College



Notes. Coefficients from individual fixed effects regressions of log earnings as explained in the text. Vertical bars correpond to 95 percent confidence intervals of the coefficient. Sample consists of all students who enrolled in ADN programs at any community college. Students grouped by college of first enrollment in an ADN program. Coefficients ranked by the point estimate. Regressions control for age, year, and concurrent community college enrollment. Regressions cluster at the individual level.

Figure 5: Individual Fixed Effects Returns, Heterogeneity by Pre-Enrollment Earnings



Notes. Coefficients from individual fixed effects regressions of log earnings as explained in the text. Vertical bars correpond to 95 percent confidence intervals of the coefficient. Sample consists of all students who enrolled in ADN programs at any community college. Students grouped by annual earnings in the year prior to first enrollment, rounded to the nearest $2,500. Category for $60,000 includes all students with pre-enrollment earnings above that amount. Regressions cluster at the individual level.

Table 1: Lottery Balance

		Admit-Reject Difference	
	Mean	All	1st Lottery
Female	0.794	0.0393	0.0548
		(0.0225)	(0.0337)
White	0.264	0.00189	0.0375
		(0.0349)	(0.0459)
Black	0.0463	-0.0251**	-0.00181
		(0.00759)	(0.0188)
Hispanic	0.301	0.0795	0.0826
		(0.0476)	(0.0583)
Asian	0.118	-0.0184	-0.0186
		(0.0225)	(0.0293)
Age	29.82	-0.216	-0.104
		(0.444)	(0.679)
GPA	2.391	0.0434	0.0200
		(0.0651)	(0.109)
Enrolled at other college	0.734	0.0246	-0.0316
		(0.0193)	(0.0385)
Enrolled in other district	0.407	-0.0110	0.0121
		(0.0279)	(0.0343)
Had BOG Waiver	0.720	0.0442	-0.0412
		(0.0386)	(0.0648)
Had Pell Grant	0.437	-0.00919	-0.103*
		(0.0376)	(0.0442)
Had Calgrant	0.185	0.0145	-0.0500
		(0.0312)	(0.0538)
Had Loans	0.0726	0.00125	-0.0250
		(0.0221)	(0.0198)
Employed > 1 Quarter	0.819	-0.0160	-0.00440
		(0.0187)	(0.0345)
Quarters Employed	9.168	-0.0788	0.327
		(0.379)	(0.555)
Employed > 8 Quarters	0.628	0.00672	0.0503
		(0.0365)	(0.0522)
Mean Quarterly Earnings	4717.3	-210.3	246.4
		(252.4)	(452.1)
Industry is Health	0.407	0.0281	0.00638
		(0.0338)	(0.0475)
Industry is Health: Hospitals	0.203	0.0233	0.0671
		(0.0195)	(0.0328)
Industry is Health: Ambulatory	0.140	0.0236	-0.0225
		(0.0238)	(0.0254)
Industry is Health: Nursing/Residential Care	0.0886	0.0193	0.0205
		(0.0163)	(0.0270)
Industry is Retail	0.223	0.00259	-0.0174
		(0.0299)	(0.0432)
Industry is Administrative	0.116	0.00214	-0.0179
		(0.0144)	(0.0238)
Industry is Education	0.101	0.00418	-0.00171
		(0.0217)	(0.0264)
Industry is Food Service	0.158	0.00805	0.00476
		(0.0238)	(0.0358)
N	3506	8870	3506

Notes. First column shows mean characteristics measured at term of first application. Enrollment at other college defined as ever having taken a course at another community college, with similar definition by district. Employment defined as nonzero quarterly earnings. Quarters employed defined as the number of quarters with nonzero earnings in the four years prior to application, with maximum 16. Consistent employment defined as employment in at least eight of the 16 quarters before first application. Employment in Health defined as employment in the two-digit NAICS industry code 62: Health Care and Social Assistance. Second and third columns show results of regressing mean characteristics on lottery admission and cohort fixed effects. Second column includes all applications that were decided by random lottery: the first through fourth. The final column only includes first lottery. $*p < 0.05, **p < 0.01, ***p < 0.001$

Table 2: Lottery Dynamics and Reapplication

Lottery Number	1	2	3	4	5	Ever
Applicants	3463	2527	1756	1306	1030	3463
Share Win Lottery	0.0924	0.0641	0.0450	0.0436	0.878	0.412
Share of Losers Reapply	0.824	0.732	0.765	0.818	0	0.824

Notes. Table displays means. Lottery number refers to how many times an applicant has applied.

Table 3: Academic and Labor Market Outcomes

	Ever Win		Never Win
	Enroll	Do Not Enroll	
Panel A. Academic			
Units Attempted, next 3 years	37.74	7.350	13.65
	(14.84)	(15.16)	(19.88)
Start Program	1	0	0.138
	(0)	(0)	(0.345)
Finish Program	0.464	0.0115	0.0784
	(0.499)	(0.107)	(0.269)
Any Degree in Nursing	0.464	0.0443	0.146
	(0.499)	(0.206)	(0.353)
Any Degree in Health	0.464	0.0607	0.156
	(0.499)	(0.239)	(0.363)
Any Degree	0.478	0.0803	0.164
	(0.500)	(0.272)	(0.370)
Years to Nursing Degree	3.159	2.769	2.662
	(1.020)	(0.714)	(1.064)
Years to Any Degree	3.080	2.837	2.570
	(1.063)	(1.010)	(1.074)
Transfer within 6 Years	0.0740	0.0459	0.0521
	(0.262)	(0.209)	(0.222)
Panel B. Labor Market			
Mean Earnings	16389.2	13965.3	15873.4
	(7070.5)	(8482.9)	(9513.7)
Employed	0.874	0.648	0.729
	(0.333)	(0.480)	(0.445)
Health	0.747	0.381	0.561
	(0.436)	(0.488)	(0.497)
Health: Hospitals	0.596	0.314	0.419
	(0.492)	(0.466)	(0.494)
Health: Ambulatory	0.101	0.0381	0.102
	(0.302)	(0.192)	(0.303)
Health: Nursing/Residential	0.0606	0.0286	0.0530
	(0.239)	(0.167)	(0.224)
Observations	743	610	2167

Notes. Starting program measured as enrolling in the courses in the first semester of the sequence, regardless of completion or performance. Finishing program includes earning an associate's degree or certificate in Registered Nursing from Central College. "Any degree" completion refers to earning any certificate or associate's degree. Health defined as the two-digit Taxonomy of Programs (TOP) code "Health." Any degree in nursing defined as a degree or certificate in the six-digit TOP code for "Registered Nursing." Enrollment defined as enrolling in at least 3 credits in any term in the time period noted. Years to degree based on years since first lottery attempt. Labor market outcomes measured five to seven years after first lottery attempt. Employment defined as non-zero earnings in at least half of the calendar year quarters encompassing the noted years since the first lottery. Mean earnings do not include zeros for individuals with missing earnings. Industry employment defined by NAICS industry codes. Standard deviations in parentheses.

Table 4: First Stages

	(1) Enroll	(2) Earn ADN	(3) Applications	(4) Years Enrolled	(5) Ever Admitted
A. 1st Lottery					
Win 1st Lottery	0.241***	0.0846*	-1.578***	-0.207	0.739***
	(0.0427)	(0.0365)	(0.0868)	(0.107)	(0.0202)
F	31.75	5.380	331.0	3.764	1333.5
N	3459	3459	3459	3459	3459
B. 1st-4th Lotteries					
Win 1st Lottery	0.241***	0.0846*	-1.578***	-0.207	0.739***
	(0.0428)	(0.0365)	(0.0869)	(0.107)	(0.0203)
Win 2nd Lottery	0.329***	0.233***	-1.571***	0.0834	0.636***
	(0.0504)	(0.0513)	(0.0688)	(0.124)	(0.0240)
Win 3rd Lottery	0.327***	0.0893	-1.301***	0.244	0.442***
	(0.0619)	(0.0585)	(0.0713)	(0.161)	(0.0356)
Win 4th Lottery	0.108	0.0905	-0.730***	0.212	0.355***
	(0.0728)	(0.0762)	(0.0488)	(0.240)	(0.0476)
F	26.53	7.552	332.3	1.809	489.2
N	8693	8693	8693	8693	8693
C. 1st-4th Lotteries, Win Any Lottery					
Win Lottery	0.265***	0.132***	-1.417***	0.0157	0.605***
	(0.0270)	(0.0256)	(0.0462)	(0.0713)	(0.0176)
F	96.69	26.56	938.8	0.0482	1180.1
N	8693	8693	8693	8693	8693

Notes. Sample for Panel A consists of one observation per student, with the key independent variable being the result of the first lottery a student entered. Sample in panels B and C consist of one observation per lottery, with potentially multiple observations per student. Key Panel B independent variables are an indicator the result of each individual lottery, and in Panel C the effects of each lottery are constrained to all be the same. All regressions control for academic term of lottery. Panels B and C control for lottery instance (ie. 1st, 2nd, 3rd, 4th). Standard errors clustered at the individual level. $*p < 0.05, **p < 0.01, ***p < 0.001$

Table 5: IV Estimate of Effect of Enrollment and Completion on Earnings

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Start Program	0.420**	0.424**	0.444**	0.445**	0.446**	0.403**	
	(0.157)	(0.160)	(0.161)	(0.166)	(0.164)	(0.150)	
Earn ADN							0.667*
							(0.293)
Number of Applications						-0.0746	
						(0.135)	
N	20885	20885	20885	20885	20885	20885	20885
Mean Earnings	12686.3	12686.3	12686.3	12686.3	12686.3	12686.3	12686.3
First stage F	20.04	18.35	18.19	18.70	17.36	24.34	3.775
Demographics		X	X	X	X	X	X
Academic			X	X	X	X	X
Labor Market				X	X	X	X
Financial Aid					X	X	X

Notes. Dependent variable is quarterly log earnings five to seven years after focal lottery. Starting program defined as enrolling in the first course of the Central College ADN sequence. The four instruments include the result of an applicant's first four lotteries. All regressions control for year, cohort, and lottery-instance. Demographic variables include race, gender, and age when first applied. Academic covariates include pre-lottery GPA and units attempted. Labor market covariates include earnings and employment, measured in the two years prior to first application. All covariates interacted with lottery instance. Regressions weighted as described in text. Standard errors clustered at the individual level. $*p < 0.05, **p < 0.01, ***p < 0.001$

Table 6: IV Estimate of Effect of Enrollment and Completion on Employment

	(1)	(2)	(3)	(4)	(5)	(6)
	Any Emp	Quarters Employed	Health Industry	Any Emp	Quarters Employed	Health Industry
Start Program	0.0606	1.080	0.234*			
	(0.118)	(1.394)	(0.0994)			
Earn ADN				0.0829	1.392	0.328*
				(0.173)	(2.103)	(0.150)
N	20885	20885	15134	20885	20885	15134
Mean Y	0.818	7.006	0.827	0.818	7.006	0.827
Demographics	X	X	X	X	X	X
Academic	X	X	X	X	X	X
Labor Market	X	X	X	X	X	X

Notes. Dependent variable is quarterly log earnings five to seven years after focal lottery. Starting program defined as enrolling in the first course of the Central College ADN sequence. The four instruments include the result of an applicant's first four lotteries. All regressions control for year, cohort, and lottery-instance. Demographic variables include race, gender, and age when first applied. Academic covariates include pre-lottery GPA and units attempted. Labor market covariates include earnings and employment, measured in the two years prior to first application. All covariates interacted with lottery instance. Regressions weighted as described in text. Standard errors clustered at the individual level. $*p < 0.05, **p < 0.01, ***p < 0.001$

Table 7: IV Estimates: Robustness

	(1) Un-Weighted	(2) Exc. 5x Applicants	(3) Exc. Other Admits	(4) Exc. No Pre-Earn	(5) Exc. Transfer	(6) Last Lottery	(7) "One-Step"	(8) Levels	(9) Exc. Men
Start Program	0.398**	0.463**	0.355**	0.462**	0.399*	0.361**	0.461*	3694.9	0.391*
	(0.149)	(0.179)	(0.129)	(0.176)	(0.164)	(0.122)	(0.210)	(3313.8)	(0.156)

Notes. Dependent variable is quarterly log earnings five to seven years after focal lottery unless otherwise specified. Starting program defined as enrolling in the first course of the Central College ADN sequence. Except for in columns 6 and 7 the four instruments include the result of an applicant's first four lotteries. All regressions control for year, cohort, and lottery-instance. Regressions weighted as described in text, unless otherwise specified. Standard errors clustered at the individual level. See text for more details on individual specifications. See appendix A.3 for in-depth explanation of "one-step" procedure. $*p < 0.05, **p < 0.01, ***p < 0.001$

Table 8: Instrumental Variables and Individual Fixed Effects Estimates

	(1)	(2)	(3)	(4)	(5)	(6)
Sample	Central College ADN Applicants			Central College Start ADN	California Start ADN	
Method	IV	OLS	FE	FE	OLS	FE
Started Program	0.446**	0.146***	0.433***			
	(0.164)	(0.0434)	(0.0388)			
	20885	8057	64759			
Post-Degree	0.667*	0.155***	0.566***	0.455***	0.288***	0.504***
	(0.293)	(0.0446)	(0.0397)	(0.0497)	(0.0099)	(0.0094)
	20885	8057	64759	22660	539772	2857336

Notes. Outcome variable in all regressions is quarterly log earnings. Sample for columns (1)-(3) includes all applicants to the Central College program since 2005. IV and OLS specifications are at five to seven years since lottery, and control for demographics, prior labor market experience, financial aid, and GPA, as well as year and term dummies. FE specification has earnings measured between two years prior and eight years after first application. Sample for columns (4) includes all applicants to the Central College program who also started the program, regardless of lottery result. Sample for columns (5)-(6) include all students who enroll in an associate's degree in Registered Nursing program at any California community college, defined as taking a course in the six-digit Taxonomy of Programs code for Registered Nursing "1230.10." Standard errors clustered at the individual level. See text for more details on specifications. $*p < 0.05, **p < 0.01, ***p < 0.001$

Table 9: Individual Fixed Effects, Heterogeneity

	(1)	(2)
Program Quality		
NCLEX Pass Rate	-0.025	-0.062**
	(0.020)	(0.019)
ADN Graduates in 2004	0.016	-0.011
	(0.025)	(0.024)
Program Completion Rate	0.016	0.015
	(0.027)	(0.027)
Program Admissions		
Lottery	-0.059*	-0.059*
	(0.030)	(0.029)
Other Randomized	-0.002	-0.030
	(0.028)	(0.027)
Competitive	-0.002	0.026
	(0.027)	(0.027)
Labor Market Characteristics		
Hospital Beds	0.083**	0.090***
	(0.031)	(0.026)
Long-Term Care Beds	0.044	-0.031
	(0.029)	(0.026)
Share Older than 60	0.035	-0.004
	(0.033)	(0.033)
RN Employment	-0.130***	-0.000
	(0.035)	(0.031)
Med. Asst. Relative Wage	-0.126**	
	(0.039)	
RN Relative Wage	0.257***	
	(0.032)	
Nurse-Med Assistant Relative Wage		0.048*
		(0.023)
Individual		
Minority	-0.127***	-0.131***
	(0.020)	(0.020)
Older than 30	-0.103***	-0.101***
	(0.019)	(0.020)
Female	-0.014	-0.013
	(0.018)	(0.018)
	2850447	2850447
Main Effect	0.562***	0.584***
	(0.044)	(0.045)
Year-Qtr FE	X	X
Age Dummies	X	X
Enrolled	X	X

Notes. See notes for table 8 for notes on sample construction. Reported coefficients are the estimates of the interaction term γ from equation 6. Minority students are Hispanic and African American. Over 30 years old defined as age at first application for Central College samples, and age at enrollment in the program for the statewide sample. County-level information comes from 2010 Census. Relative wages and employment are relative to overall mean wages and share of employment, respectively. Medical assistants include nurse assistants and other medical assistants, as well as licensed vocational nurses. Information on hospital beds comes from the California Office of Statewide Health Planning and Development, 2014. All community variables expressed as a dummy variable interaction with being above or below the median county. Lottery programs (15 programs) had no selection based on student characteristics. Any Randomization programs (27 programs) had any element of randomization. Competitive programs (43 programs) had admission based on student characteristics including but not limited to coursework, work experience, references, and essays. The rest of the colleges (9 programs) had waitlists or first-come-first-served lottery systems. Standard errors clustered at the individual level.. $*p < 0.05, **p < 0.01, ***p < 0.001$

Table 10: Present Discounted Value of Earnings Stream

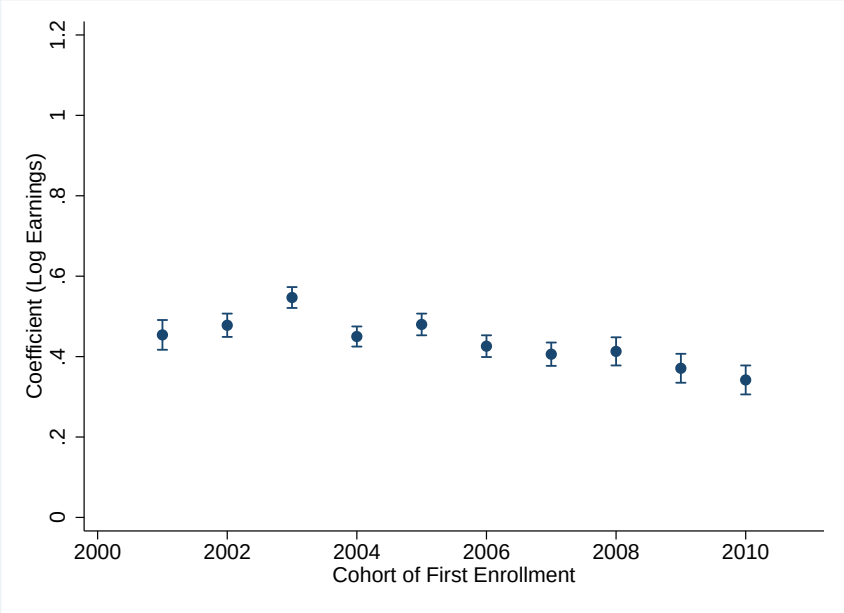
	(1)	(2)	(3)	(4)
	Estimate	Earnings PDV over 20 Year Career		
		Full Price	2%Earn Growth	Half-Time Work
IV Estimate of Enrollment	0.402	$141,986	$167,491	$180,928
FE Estimate of completion, Central College	0.566	$210,913	$266,251	$269,840

Notes. Preferred estimates in column (1) come from tabes 5 and 8. PDV calculated over 2 years of coursework and 20 years of post-program career. Pre-enrollment mean earnings are assumed to be $5,757 per quarter, which is the mean for program applicants. In column (2) students are assumed to spend two years to complete the program, foregoing all earnings, and have no earnings growth over time. Estimated price of books and supplies is $2,855 per year, estimates of additional costs from program websites and catalogs. In column (3) students again incur tuition costs but also experience 2% annual wage growth. In column (4) students are assumed to forego only half their earnings while studying. In all cases, the real interest rate is 3%.

A Appendices

A.1 Additional Tables and Figures

Figure A1: Individual Fixed Effects Returns, by Cohort of First Nursing Course



Notes. Coefficients from individual fixed effects regressions of quarterly log earnings. Vertical bars correspond to 95 percent confidence intervals of the coefficient. Sample consists of all students who enrolled in ADN programs at any community college. Cohort of nursing course refers to the academic year the student first enrolled in a course in Registered Nursing, defined by the six-digit TOP code "123010." Regressions control for age, year, and concurrent community college enrollment. Regressions cluster at the individual level.

Table A1: Summary Statistics, California and Central College Students and Health Degree Recipients

			All Students		All Health Awards		ADN Graduates	
	4-Year Public	2-Year Public	California	Central	California	Central	California	Central
N	6,721,861	6,625,141	2310170	30360	17008	505	4990	367
Female	0.56	0.56	0.53	0.51	0.73	0.79	0.81	0.82
Race								
White	0.56	0.53	0.29	0.24	0.39	0.25	0.40	0.23
Black	0.15	0.15	0.07	0.07	0.05	0.08	0.04	0.10
Hispanic	0.20	0.21	0.40	0.49	0.25	0.34	0.22	0.32
Asian	0.05	0.07	0.11	0.13	0.12	0.15	0.12	0.15
Other Race	0.05	0.04	0.12	0.07	0.18	0.18	0.21	0.20
Age								
19 or less	0.32	0.30	0.24	0.26	0.28	0.41	0.24	0.39
20-24	0.48	0.32	0.32	0.35	0.25	0.23	0.27	0.22
25-29	0.09	0.13	0.14	0.15	0.17	0.15	0.21	0.17
30-34	0.04	0.08	0.08	0.08	0.11	0.07	0.13	0.08
35-39	0.02	0.05	0.05	0.05	0.07	0.06	0.08	0.06
40-49	0.03	0.07	0.08	0.07	0.09	0.07	0.07	0.07
50 plus	0.02	0.05	0.09	0.04	0.03	0.01	0.01	0.01

Notes. National-level data from 2013 NCES Digest of Education Statistics. Data on students compiled from California Community College Chancellor's Office Datamart and cover 2013 academic year. Data on awards compiled from administrative sources. Data count each award separately, not taking into account multiple awards per student.

Table A2: Determinants of Lottery Reapplication Among Lottery Losers

	1st	2nd	3rd	4th	Any
Female	0.0150	0.0675*	0.00669	0.0315	0.0292*
	(0.0184)	(0.0288)	(0.0281)	(0.0305)	(0.0128)
Hispanic	0.0238	0.0243	0.0386	0.0210	0.0266*
	(0.0191)	(0.0288)	(0.0312)	(0.0316)	(0.0134)
Asian	-0.0300	0.0124	0.00892	0.0352	0.00133
	(0.0275)	(0.0391)	(0.0402)	(0.0450)	(0.0185)
Other Race	-0.0286	-0.0197	0.0127	-0.0325	-0.0166
	(0.0210)	(0.0322)	(0.0341)	(0.0356)	(0.0148)
Age	0.00109	0.00353*	0.00277	-0.00180	0.00146*
	(0.00106)	(0.00155)	(0.00170)	(0.00175)	(0.000731)
GPA	0.00267	0.00488	-0.00567	-0.0225	-0.00193
	(0.00789)	(0.0118)	(0.0130)	(0.0135)	(0.00543)
Earnings (1000s), t-1	0.00172	0.00455	0.00324	0.00856***	0.00408***
	(0.00145)	(0.00233)	(0.00218)	(0.00205)	(0.00100)
Emp. Health, t-1	-0.00480	0.0371	-0.00674	0.0170	0.00974
	(0.0162)	(0.0240)	(0.0251)	(0.0242)	(0.0109)
ymean	0.830	0.719	0.776	0.825	0.788
N	2043	1510	991	833	5377
Cohort FE's	X	X	X	X	X
Lottery FE's					X

Notes. Dependent variable is reapplication conditional on losing the lottery in question. Sample consists of all non-admitted students in each lottery. Earnings, GPA, and employment measured one year prior to first lottery. Employment in health measured as non-zero earnings in any quarter in the year prior to first application. $*p < 0.05, **p < 0.01, ***p < 0.001$

Table A3: First Stages with covariates

	Ever Admitted	Enroll	Earn ADN	Number of Applications	Years Enrolled
A. 1st-4th Lotteries					
Win 1st Lottery	0.244***	0.0550	-1.525***	-0.195	0.750***
	(0.0446)	(0.0369)	(0.0954)	(0.119)	(0.0216)
Win 2nd Lottery	0.286***	0.163**	-1.524***	0.0284	0.664***
	(0.0585)	(0.0558)	(0.0826)	(0.163)	(0.0276)
Win 3rd Lottery	0.307***	0.0584	-1.294***	0.195	0.462***
	(0.0666)	(0.0607)	(0.0775)	(0.182)	(0.0401)
Win 4th Lottery	0.0917	0.0303	-0.718***	0.308	0.374***
	(0.0762)	(0.0818)	(0.0618)	(0.271)	(0.0547)
F	19.30	2.967	241.4	1.282	433.4
N	7374	7374	7374	7374	7374
B. 1st-4th Lotteries, Win Any Lottery					
Win Lottery	0.247***	0.0822**	-1.378***	0.00250	0.626***
	(0.0293)	(0.0265)	(0.0523)	(0.0826)	(0.0193)
F	71.26	9.603	694.4	0.000916	1046.2
N	7374	7374	7374	7374	7374

Notes. All regressions control for term of lottery. Panels B and C control for lottery instance (ie. 1st, 2nd, 3rd, 4th). Standard errors clustered at the individual level. $*p < 0.05, **p < 0.01, ***p < 0.001$

Table A4: Reduced Form Effect on Earnings

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Win 1st Lottery	0.164**	0.162**	0.170**	0.158**	0.160**		
	(0.0616)	(0.0627)	(0.0606)	(0.0602)	(0.0597)		
Win 2nd Lottery	0.0712	0.0859	0.0847	0.109	0.108		
	(0.108)	(0.110)	(0.110)	(0.106)	(0.107)		
Win 3rd Lottery	0.283*	0.265*	0.260*	0.207	0.235		
	(0.114)	(0.115)	(0.112)	(0.123)	(0.135)		
Win 4th Lottery	0.290	0.271	0.271	0.378	0.368		
	(0.224)	(0.220)	(0.220)	(0.212)	(0.217)		
Win Lottery						0.152**	0.154**
						(0.0530)	(0.0518)
N	20885	20885	20885	20885	20885	20885	20885
Demographics		X	X	X	X		X
Academic			X	X	X		X
Labor Market				X	X		X
Financial Aid					X		X

Notes. Dependent variable is quarterly log earnings five to seven years after focal lottery. All regressions control for year, cohort, and lottery-instance. Demographic variables include race, gender, and age when first applied. Academic covariates include pre-lottery GPA and units attempted. Labor market covariates include earnings and employment, measured in the two years prior to first application. All covariates interacted with lottery instance. Standard errors clustered at the individual level. $*p < 0.05, **p < 0.01, ***p < 0.001$

Table A5: Main Estimates using Alternate Instruments

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Instrument	1st Lottery				Win Any Lottery			
	Earnings	Any Emp	Quarters Employed	Health Industry	Earnings	Any Emp	Quarters Employed	Health Industry
Start Program	0.500	0.0574	-0.184	0.592**	0.514**	0.0670	0.820	0.284*
	(0.269)	(0.169)	(2.216)	(0.205)	(0.199)	(0.133)	(1.693)	(0.126)
Earn ADN	0.622	0.0710	-0.228	0.553**	0.706*	0.0867	1.068	0.332*
	(0.362)	(0.207)	(2.760)	(0.211)	(0.315)	(0.178)	(2.250)	(0.150)
Demographics	X	X	X	X	X	X	X	X
Academic	X	X	X	X	X	X	X	X
Labor Market	X	X	X	X	X	X	X	X

Notes. Dependent variable is quarterly log earnings five to seven years after focal lottery unless otherwise specified. Instrumented variable is enrollment in the ADN program in the top panel and completion of an ADN in the bottom panel. Each cell represents results from an individual regression. Instruments used are either the first lottery sample with the result of the first lottery or the stacked sample with winning any lottery as the instrument. All regressions control for year, cohort, and lottery-instance. Regressions weighted as described in text, unless otherwise specified. Standard errors clustered at the individual level. See text for more details on individual specifications. $*p < 0.05, **p < 0.01, ***p < 0.001$

Table A6: Selection on Observables and Academic Outcomes, by Admission Type

	(1) Pre-Earnings	(2) Pre-Earnings (Log)	(3) Pre-Employment	(4) Earn ADN	(5) Earn Any Degree
Lottery	5248.9	0.106	0.0134	0.00899	0.0116
	(3401.2)	(0.0679)	(0.00801)	(0.0154)	(0.0181)
Some Randomization	-279.2	0.0147	0.0112	0.0303	0.0300
	(2695.9)	(0.0541)	(0.00857)	(0.0248)	(0.0263)
N	66623	59541	66623	66623	66623
R-squared	0.112	0.0806	0.0143	0.0537	0.0537
Mean Y	61488.3	10.51	0.894	0.0844	0.100
Cohort FE's	X	X	X	X	X

Notes. Sample consists of all students who enrolled in an ADN program. Pre-earnings and employment are determined up to four years before enrollment in a program. All regressions include controls for demographics: age, gender, race. Standard errors clustered at the college level. $*p < 0.05, **p < 0.01, ***p < 0.001$

Table A7: Heterogeneity in Individual Fixed Effects

Minority	-0.133***
	(0.020)
Main Effect	0.543***
	(0.011)
	2857336
Older than 30	-0.102***
	(0.020)
Main Effect	0.544***
	(0.012)
	2857336
Female	-0.012
	(0.018)
Main Effect	0.515***
	(0.019)
	2857336
NCLEX Pass Rate	-0.038*
	(0.018)
Main Effect	0.523***
	(0.013)
	2857336
ADN Graduates in 2004	0.023
	(0.019)
Main Effect	0.489***
	(0.016)
	2857336
Program Completion Rate	0.007
	(0.026)
Main Effect	0.498***
	(0.024)
	2857336
Hospital Beds	0.057**
	(0.018)
Main Effect	0.471***
	(0.014)
	2850447
RN Employment	-0.000
	(0.018)
Main Effect	0.504***
	(0.013)
	2857336
Med. Asst. Relative Wage	0.073***
	(0.019)
Main Effect	0.458***
	(0.015)
	2857336
RN Relative Wage	0.166***
	(0.019)
Main Effect	0.449***
	(0.011)
	2857336
Nurse-Med Assistant Relative Wage	0.037
	(0.020)
Main Effect	0.494***
	(0.011)
	2857336

Notes. See notes for table 8 for notes on sample construction. Reported coefficients are the estimates of the interaction term γ from equation 6, run as individual regressions. Minority students are Hispanic and African American. Over 30 years old defined as age at first application for Central College samples, and age at enrollment in the program for the statewide sample. County-level information comes from 2010 Census. Relative wages and employment are relative to overall mean wages and share of employment, respectively. Medical assistants include nurse assistants and other medical assistants, as well as licensed vocational nurses. Information on hospital beds comes from the California Office of Statewide Health Planning and Development, 2014. All community variables expressed as a dummy variable interaction with being above or below the median county. Lottery programs (15 programs) had no selection based on student characteristics. Any Randomization programs (27 programs) had any element of randomization. Competitive programs (43 programs) had admission based on student characteristics including but not limited to coursework, work experience, references, and essays. The rest of the colleges (9 programs) had waitlists or first-come-first-served lottery systems. Standard errors clustered at the individual level. $*p < 0.05, **p < 0.01, ***p < 0.001$

A.2 Further Evidence of Lottery Randomization

I discuss the validity of the Central College lottery in section 3. In this appendix I provide further evidence through additional tests. First, I regress admission on a set of individual characteristics, defined in the year prior to the application, as shown in Table A8. No variable is statistically significant. In another test of lottery balance, I regress the outcome variables defined in the pre-lottery period on the lottery results, while controlling for demographic and academic variables. The results of this excercise are shown in Table A9. Here, again, is evidence of the balance of the lottery. One way to interpret these are as reduced form estimates of the outcomes in the pre-lottery period.

Table A8: Balance, Joint Regressions, by Lottery Instance

	All	Lot 1	Lot 2	Lot 3	Lot 4
Female	0.0137	0.0276	-0.00575	0.0102	0.0120
	(0.00520)	(0.00830)	(0.00890)	(0.0145)	(0.0156)
Hispanic	0.00695	0.000677	0.0244	-0.0142	0.0282
	(0.00719)	(0.00948)	(0.0141)	(0.0140)	(0.0145)
Asian	-0.0160	-0.0185	-0.0161	-0.0523	0.0578
	(0.00803)	(0.00897)	(0.0115)	(0.0196)	(0.0197)
Other Race	-0.0329*	-0.0624*	0.0139	-0.0334	-0.0287
	(0.00451)	(0.00804)	(0.00827)	(0.0174)	(0.0113)
Age	0.00698	0.0147	0.00135	0.0304	-0.0265
	(0.000212)	(0.000375)	(0.000400)	(0.000708)	(0.000710)
GPA	-0.0159	-0.000125	-0.0404*	-0.0353	0.0272
	(0.00247)	(0.00404)	(0.00271)	(0.00373)	(0.00715)
Enrolled at other college	0.00207	-0.0164	0.00982	0.0785**	-0.0433
	(0.00362)	(0.00908)	(0.00749)	(0.00989)	(0.0101)
Had BOG Waiver	0.0180	0.00372	0.0585**	-0.00853	0.0203
	(0.00525)	(0.00952)	(0.00725)	(0.00761)	(0.0201)
Had Pell Grant	-0.0209	-0.0322	-0.0133	0.0619	-0.117*
	(0.00529)	(0.00818)	(0.00834)	(0.0139)	(0.0164)
Had Calgrant	0.0128	-0.00397	0.00936	0.0639	0.000810
	(0.00666)	(0.0116)	(0.0128)	(0.0176)	(0.0113)
Had Loans	0.00119	-0.00725	-0.0193	-0.0156	0.0927
	(0.00870)	(0.0108)	(0.0150)	(0.0254)	(0.0360)
Consistent Employment	0.0156	0.0776	-0.0104	0.00694	-0.120
	(0.0128)	(0.0175)	(0.0195)	(0.0187)	(0.0335)
Employed in Health	-0.0113	-0.0519	0.0373	-0.00551	0.00416
	(0.00783)	(0.0124)	(0.0136)	(0.0113)	(0.0178)
Quarterly (log) Earnings)	-0.0269	-0.0532	-0.0422	-0.0388	0.115
	(0.00127)	(0.00137)	(0.00229)	(0.00214)	(0.00441)
N	7316	2885	2130	1280	1021
Sh. Admitted	0.0284	0.0308	0.0272	0.0266	0.0264
F	1.333	1.757	1.094	1.918	1.226
p < F	0.179	0.0394	0.357	0.0211	0.250

Notes. First column shows mean characteristics measured at term of first application. Enrollment at other college defined as ever having taken a course at another community college, with similar definition by district. Employment defined as nonzero quarterly earnings. Quarters employed defined as the number of quarters with nonzero earnings in the four years prior to application, with maximum 16. Consistent employment defined as employment in at least eight of the 16 quarters before first application. Employment in Health defined as employment in the two-digit NAICS industry code 62: Health Care and Social Assistance. Second and third columns show results of regressing mean characteristics on lottery admission and cohort fixed effects. Second column includes all applications that were decided by random lottery: the first through fourth. The final column only includes first lottery. $*p < 0.05, **p < 0.01, ***p < 0.001$

Table A9: Lottery Balance, Outcome Regressions

	All Lotteries	1st Lottery
Earnings	-100.0	571.0
	(333.8)	(603.5)
N	78109	31238
Log Earnings	-0.0250	0.0317
	(0.0491)	(0.0854)
N	78109	31238
Employed	-0.00721	0.0373
	(0.0258)	(0.0393)
N	126154	51206
Quarters Employed	0.266	0.164
	(0.633)	(0.954)
N	2929	1155
Employed in Health	0.0108	0.0243
	(0.0243)	(0.0368)
N	126154	51206
Employed in Health if Employed	0.0244	0.0148
	(0.0346)	(0.0520)
N	74180	29664

Notes. Coefficients are the results of the second of the following two steps. First, a regression of the variable of interest in the two years prior to first application on a set of demographic and academic characteristics. The second is a regression of lottery admission on the residual from the first step and cohort dummies. Variables included in the first step are age, race, gender, GPA, enrollment in another district, and units earned, as well as year and quarter dummies. Employment defined as nonzero quarterly earnings. Employment in Health defined as employment in the two-digit NAICS industry code 62: Health Care and Social Assistance. Results for the number of quarters employed includes one observation for each student-lottery combination,with the outcome variable being the number of quarters with nonzero earnings. $*p < 0.05, **p < 0.01, ***p < 0.001$

A.3 One-Step Dynamic Regression

Column 5 of Table 7 shows estimates of the "one-step" regression as used by Gelber, Isen and Kessler (2016) and Cellini, Ferreira and Rothstein (2010) in scenarios where applicants may reapply. One concern is that reapplication itself may have an effect on later earnings. In the case of the Central College lottery, losing a lottery increases the likelihood of participating in a future lottery. This is similar to the case of Cellini, Ferreira and Rothstein (2010), where a district failing to pass a bond is more likely to consider a similar bond in a later year than a district that succeeded in passing a bond. The "one-step" estimator Cellini, Ferreira and Rothstein (2010) propose takes this added effect into account. I adapt this estimator using the following equation of the reduced form:

$$y_{ict} = \alpha + \sum_{\tau=0}^{\bar{\tau}} (\theta_\tau Admit_{i,t-\tau} + \phi_\tau Apply_{i,t-\tau}) + X_{itc}\Psi + \eta_c + \nu_t + u_{itc} \tag{7}$$

The coefficient of interest, θ_τ, represents the effect of winning the lottery on earnings at year τ regardless of the effect of losing the lottery on future lottery participation and admission to the program. The reduced form estimate of θ_τ is 0.141 (0.07), which is quite similar to the results shown using the more conventional approach.

A.4 Individual Fixed Effects Estimates, Sensitivity

Table A10 shows that the individual fixed effects estimates from Table 8 are robust to a variety of specifications. The first column only includes time fixed effects in addition to the individual fixed effect. Addition of age dummies is important, since it accounts for different earnings profiles over time. Column 3 adds a dummy variable for whether a student was enrolled in at least eight units—a halftime load—that quarter. Column 3 is my preferred specification. As two additional tests of robustness, I first exclude a full year of data for each student: their first year prior to application or enrollment in a program. This accounts for the "Ashenfelter's Dip", which is that students entering an educational program tend to experience job loss or a decline in earnings immediately prior. In the final column I exclude from the comparison group all students who earned any other degree, making the comparison between ADN earners and students who did not earn a single other degree or certificate. The results are broadly similar across specifications.

Table A10: Individual Fixed Effects Estimates, Sensitivity

	(1)	(2)	(3)	(4)	(5)
Central College Applicants					
Started Program	0.417***	0.459***	0.433***	0.439***	0.477***
	(0.0417)	(0.0391)	(0.0388)	(0.0397)	(0.0396)
	64759	64759	64759	56060	59389
Post-Degree	0.584***	0.590***	0.566***	0.575***	0.495***
	(0.0434)	(0.0400)	(0.0397)	(0.0404)	(0.0427)
	64759	64759	64759	56060	59389
Central College ADN Enrollees					
Post-Degree	0.450***	0.474***	0.455***	0.469***	0.452***
	(0.0551)	(0.0501)	(0.0497)	(0.0505)	(0.0498)
	22660	22660	22660	19728	22386
Statewide Enrollees					
Post-Degree	0.607***	0.595***	0.504***	0.532***	0.502***
	(0.0101)	(0.00955)	(0.00940)	(0.00947)	(0.00940)
N	2857336	2857336	2857336	2595104	2812914
Year-Qtr FE	X	X	X	X	X
Age Dummies		X	X	X	X
Enrolled			X	X	X
Exclude t-1				X	
Exclude other degree-earners					X

Notes. See notes for Table 8 for notes on sample construction. Column 4 excludes one year of data. For Central College applicant samples this year is the second year prior to the last lottery a student entered. For the Central College enrollees and the Statewise enrollees samples, this excluded year is the year prior to enrollment in the program. Column 5 excludes from the comparison group any student who earned another degree or certificate from a community college in California. Standard errors clustered at the individual level. $*p < 0.05, **p < 0.01, ***p < 0.001$