The Returns to a Large Community College Program: Evidence from Admissions Lotteries

Michel Grosz*

August 10, 2018

Abstract

In this paper I estimate the labor market returns to a particularly large and important community college 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 leverage 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 44% and the probability of working in the healthcare industry by 19 percentage points. I show that these estimates are similar to ones in models that do not use the lottery variation but do control for individual fixed effects and individual-specific linear time trends. In light of concerns about nursing shortages, I estimate that the economic benefit of expanding an ADN program by one seat far outweighs the costs.

*Abt Associates. 6130 Executive Blvd., Rockville, MD 20852. Email: michel_grosz@abtassoc.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, Joshua Goodman, 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, Abt Associates, Analysis Group, IMPAQ International, the University of Illinois, the University of Nevada-Reno, the US Naval Academy, 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. 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). There are a number of reasons for this increased attention. Community colleges are more accessible and affordable to students than four-year college, offering an alternative in light of postsecondary attainment lagging behind demand for skilled workers (Goldin and Katz, 2008; Cohen, Brawer and Lombardi, 2009). Community colleges 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 as the demand for skills in the labor force changes (Bailey et al., 2003; Acemoglu and Autor, 2011).[1] In recent years, policymakers have focused additional attention on expanding career technical programs.

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 enrolling in an associate's degree in nursing (ADN) program. I leverage the random lottery that assigns admission to a large ADN program in California, and produce causal estimates of the program's effect. Surging demand for

[1] The terms "vocational" and "career technical" are largely interchangeable terms for programs and coursework that train students for specific occupations.

seats has forced many colleges to ration high-demand programs and courses, through methods such as lotteries and waitlists (Gurantz, 2015; Bohn, Reyes and Johnson, 2013; Bound and Turner, 2007). 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.

I rely on data that track community college students through their academic careers and into the labor market. I use detailed individual-level administrative records covering academic and earnings information for all students enrolled in California community colleges between 1992 and 2015. 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 before, during, and after they enroll in a community college.

Winning the admissions lottery substantially increases the likelihood that a student will enroll in and complete the ADN program, relative to other applicants. There is also a strong positive effect on the chances that a student will complete any type of community college degree or certificate at any of the over 110 colleges in the state. This suggests that nursing applicants are on the margin of completing an ADN—one of the lengthiest and most difficult degrees offered in the community college system—and no further postsecondary credential. This is especially striking since applicants must accumulate over half the certificates required for an associate's degree in order to even be eligible to apply for the ADN 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 44 percent increase in earnings five years after enrolling. This is a large effect, especially given standard estimates of the returns to a year of postsecondary education. I also find that students who enrolled in the program were 19 percentage points more likely to work in the healthcare sector five years later. 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. Using the same sample of lottery applicants, the two methods yield almost identical results. While this similarity might be an artifact of the random selection of the lottery, I also show evidence that individual fixed effects estimates using samples of students at other California ADN programs account for bias in similar ways. I also find that heterogeneity across programs is primarily associated with local labor market opportunities in the health industry as well as program characteristics.

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. I find that the private internal rate of return to enrolling ranges between 69 and 101 percent, and a lower bound on the social return is 17 percent. Nevertheless, because colleges in California and many 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 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.

This paper proceeds as follows. Section 2 provides background on the literature and the institutional context. Section 3 describes the dataset and sample selection. Section 4 discusses the methodology for using the admissions lottery to estimate the effects of enrollment, and Section 5 shows the results. Section 6 then compares the results using the lottery to results using individual fixed effects, and also shows heterogeneity in these results across colleges in the state. Section 7 translates the main estimates into internal rates of return and also incorporates information on the costs of program expansion, and Section 8 concludes.

2 Background

2.1 The Returns to Community College

There is much less research on the returns to community college than to other educational sectors such as grade school, high school, and four-year college. This lack of research is perhaps because the multiple missions of community colleges make issues of endogenous selection particularly severe.[2] Research using nationally representative survey data finds large returns to vocational degrees as well as to credit hours (Kane and Rouse, 1995; Gill and Leigh, 2003; Marcotte, 2016).[3] A newer line of research relies on state administrative datasets, pooling the academic and earnings records of many thousands of students and leveraging individual fixed effects models that rely on pre-enrollment labor market experience. A key conclusion from this emerging literature is that the labor market returns to career technical programs, while generally positive, vary by subject and type of degree (Stevens, Kurlaender and Grosz, 2015; Bahr et al., 2015; Liu, Belfield and Trimble, 2015; Jepsen, Troske and Coomes, 2014; Cellini and Turner, 2016). What remains unclear is how well the models typically used in these analyses, which exploit within-individual earnings changes, account for different types of bias. In this paper, I replicate the methods used in these studies and compare them to the results using the admissions lottery.

This paper is the first to use a randomized lottery to evaluate an existing community college program, yet there is a long history of experimental demonstrations in the workforce development literature (Barnow and Smith, 2015). In addition to issues of external validity and scaleability 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

[2]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. 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).

[3]See Belfield and Bailey (2011) for a full overview of this literature.

exceptions (Scrivener and Weiss, 2013; Visher et al., 2012; Peck et al., 2018). 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.[4]

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). 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, the NCLEX-RN, which they may retake multiple times. 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. There is little evidence, though, that BSNs do better in the labor market than ADNs (Auerbach, Buerhaus and Staiger, 2015).

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: 113 campuses and 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, representing one in six vocational associate's degrees across 219 different fields.

Central College[5] is located in California's Central Valley and its ADN program is among the largest in the state. The ADN program is highly structured and takes four semesters to complete.[6] Central College's ADN program has an admissions policy based on a random lottery.

[4] 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).

[5] Anonymized for confidentiality reasons.

[6] Students take a set schedule of courses in a pre-determined order, consistent with standards set by the state's Board of Nursing. Students have access to academic and career support. Beginning in the first semester of the program, students gain hands-on experience, working under the supervision of nurses in nearby hospitals and clinics.

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. In order to become eligible for the Central College lottery, applicants must pass 36 units worth of college-level prerequisite courses, or slightly more than a full year's courseload.[7] One lottery is conducted each fall and spring semester and results are posted online. If they are rejected, students may reapply to the next semester's lottery. Reapplication is easy: rejected applicants must simply click a button on the website within approximately a week of the result. Students who apply for a fifth consecutive time have a higher chance of admission, decided in a non-random process.

3 Data and Summary Statistics

3.1 Data and Sample Construction

I combine two sources of student-level administrative data for my analyses. The first consists of 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. I observe term-level coursework and grades for each student, academic 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.[8] The result is a dataset containing detailed information on each student's experience in the California Community College sytem as well as

[7] These include nine courses in the fields of anatomy, physiology, chemistry, microbiology, and psychology. Some of the courses, such as intermediate algebra, also have prerequisites. The program prerequisites 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. It is difficult to determine when an applicant began the process of preparing to apply, but the median number of years between a student's first community college course and first lottery application is 5.5 years.

[8] 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 never having earnings in California.

quarterly earnings and industry of employment before, during, and after their schooling.

I also use lottery data from Central College's ADN program for each lottery since fall 2005. I can observe the semester and result of each lottery an applicant entered. The data include an applicant's name, gender, date of birth, and an internal identification number. Because the lottery is run at the college level, there is not a perfect match with the statewide academic data system. Instead, I match students in the lottery dataset to students in the statewide administrative dataset based on the few identifying characteristics that exist in both: the first three letters of their first and last name, their birth date, and their gender. I am able to match 83 percent of all 4,726 Central College ADN applicants to student records in the statewide data. A potential concern is that winning the lottery may affect a student's likelihood of being matched and thus of being in the analytic sample. However, because they must take so many prerequisites, few students are likely to have never enrolled in a community college class prior to applying. The difference in match rate between winning and losing applications is 1.1 percentage points with a p-value of 0.46. This suggests that lottery losers who never take any further community college class are matched based on community college enrollment prior to application. Appendix A.2 describes the matching process in more detail, as well as some additional checks.

In addition to the matched student-level dataset, I learned institutional details from visits to Central College, in which I inteviewed administrators, attended an orientation presentation for incoming ADN students, and held a focus group with new students. 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 taken prerequisites and whether they had enrolled in courses associated with the program.

I limit the sample of 3,904 matched applicants to 1,730 who first applied between Spring 2005 (the first lottery in the data) and Spring 2009, in order to have a consistent sample with which to observe long-run post-lottery outcomes. Since I have earnings records up to the last quarter of 2014, this yields 21 quarters of post-lottery earnings data for all applicants in the resulting sample and 20 quarters of pre-lottery earnings records.

3.2 Summary Statistics and Lottery Balance

Column 1 of Table 1 shows summary statistics for all applicants in the analytic sample, using characteristics determined before their first lottery. Applicants were 30 years old on average and predominantly female, which is common for most nursing programs. Most students received some form of financial aid, including a combination of tuition waivers, Pell Grants, state grants, and loans. Applicants had prior labor market attachment; 82 percent had ever been employed in the five years prior to applying. However, applicants were employed in low-paying jobs, with just an average of $4,740 per quarter. A large share, 40 percent, had previously worked in the healthcare industry, consistent with the idea that many applicants are nursing assistants, health aides, or licensed practical nurses looking to upgrade their skills. Appendix Table A1 shows that the student population of both Central College and its ADN program looks qualitatively similar to other colleges and ADN programs across the state.

Column 2 of Table 1 shows the validity of the randomized lottery by reporting the difference in mean characteristics across admitted and rejected applicants within each lottery. These are coefficients of a regression of the baseline characteristic on an admissions dummy and lottery fixed effects. 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. Winning applicants had slightly lower GPA and were less likely to work in the food industry. However, 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 that the lottery is also balanced for first-time applicants. Appendix Table A2 provides further evidence of the randomization of the lottery, and shows that an F-test of the joint significance of all the covariates has values of 0.706 (p=0.844) for all lotteries and 0.594 (p=0.920) for first-time lotteries.

4 Methods

I estimate the effect of enrolling in the ADN program on subsequent labor market outcomes. I assume the following relationship:

$$y_{ict} = \beta_0 + \beta_1 D_{ic} + X_i \beta_2 + \mu_c + \zeta_t + \varepsilon_{ict} \tag{1}$$

where y_{ict} is the labor market outcome at time t for student i in application cohort c, and D_{ic} is a dummy variable taking a value of one if the student enrolled as part of the cohort. Even controlling for observable student characteristics 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.

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 exogenous and also a strong predictor of enrollment, making it a valid instrument. However, the ability for losers to reapply necessitates a departure from this simple estimation strategy.

I estimate a first stage equation of the form:

$$D_{ic} = \gamma_0 + \gamma_1 Admit_i + X_i \gamma_2 + \eta_c + e_{ic} \tag{2}$$

where $Admit_i$ is the result of an applicant's first lottery. The coefficient γ_1 reflects the difference in enrollment among winning and losing compliers in their first lottery. The identifying assumption is that, conditional on X_i and η_c, the result of the first lottery is independent of e_{ic}. As shown earlier, admission seems random within lottery cohorts, supporting the identifying assumption.

The treatment D_{ic} is defined as enrollment in the cohort for which the applicant first applied. Thus, the coefficient γ_1 represents the fraction of compliers, for whom winning the first lottery leads to enrolling in the ADN program that semester. There are two main types of non-compliers. The first are students who are admitted but do not take up the offer.[9] The second consists of

[9] In theory this would also include students who are admitted in their first lottery but take up the offer in a later cohort.

students who gain admission outside the lottery process, which is rare. I define the treatment narrowly, as immediate attendance following a lottery win, as opposed to ever enrolling in the program. This is because first-time losers who ultimately enroll do so after a delay, either in a subsequent lottery or in their fifth application. This means that the lottery affects earnings other than just through students enrolling in the program, which weakens the exclusion restriction. For completeness, I do show estimates of the effects of ever enrolling in Appendix A.3; they essentially rescale the main effect by a smaller first stage, and thus yield larger estimates. This is similar to the approach taken by Ketel et al. (2016), who use the result of an applicant's first lottery as an instrument for medical degree completion.

Figure 2 shows quarterly mean earnings for first-time lottery winners and first-time lottery losers, net of age, year, and quarter effects, and controlling for concurrent enrollment in at least a half-time load of eight community college credits. The panel is balanced: the same 1,730 students are represented at all points. The figure shows relatively flat earnings trajectories prior to application, with declines in the quarter immediately preceding application. The difference between the two curves represents the reduced form effect, which is large especially at later quarters. In fact, earnings for first-time lottery winners only begin to rapidly grow approximately 11 quarters after application, which corresponds to a few quarters after a student who enrolled in the program would be expected to have completed it.

5 Results

5.1 First Stage Results and Academic Outcomes

Table 2 displays the effect of an applicant winning the first lottery on academic outcomes. The regressions control for cohort, demographics (age, gender, race), academic background (prior GPA, prior number of units), prior financial aid receipt (Pell grants, tuition waivers), and labor market experience (mean prior earnings, any prior employment in health). Appendix Table A5 shows that these results are not sensitive to the inclusion of additional control variables, which is not

Because students cannot defer admission, this is an empirically non-existent group.

surprising since the lottery is random.

On average, applicants submitted almost 3 applications; winning the first lottery reduced the number of applications by more than half that amount, or almost a year of waiting. Winning the first lottery increases the probability of enrolling in the Central College ADN program that semester by 0.49 percentage points. Few losing applicants enroll that semester, so the coefficient being lower than one is driven by admitted students choosing not to enroll. The result is highly statistically significant, with a large F statistic. This is the first stage for the instrumental variables estimates of the effects on earnings. Because losing applicants can reapply, and many are admitted on their fifth application, the effect of winning the first lottery on ever enrolling in the program is lower, only 20 percentage points. Approximately half of all applicants (46 percent) ever enroll in the program, though only 13 percent enroll after their first application.

There is also a strong first stage effect on completing the ADN program. The Central College ADN lottery pushes students to complete an ADN at Central College, but it is not necessarily the case that it should have a strong effect on overall community college completion. Losing the lottery might merely steer applicants away from nursing and into other health fields or even other non-health fields. In fact, by completing 36 units of prerequisites, ADN applicants are already more than halfway towards the typical requirement for an associate's degree. However, Table 2 shows that ADN admission has a strong effect on the receipt of any health-related degree or certificate at any California community college (column 5) as well as the receipt of any type of community college certificate or degree whatsoever (column 6).[10] Of course, lottery losers might instead decide to enroll in four-year colleges. If so, then the effect of the lottery on four-year college enrollment would be negative. Although I cannot observe four-year college completion, the final column of the table shows that the lottery has a slightly positive but not statistically significant effect on subsequent enrollment in four-year colleges. I cannot observe whether students enroll in for-profit colleges, which also offer programs in health. However, overall, these results support the idea that ADN program applicants are on the margin between pursuing an ADN and no further

[10]Approximately a third of students (29 percent) who are never admitted the program ultimately earn an ADN, presumably because they enroll at other colleges throughout the state. Only an additional three percent of these never-admitted students earn any type of degree or certificate.

postsecondary credentials.

5.2 Lottery Results for Labor Market Outcomes

Panel a) of Figure 3 shows quarter-by-quarter estimates of the effect of immediate enrollment on log earnings.[11] Appendix Table A3 shows the coefficients and standard errors of these results. Earnings effects in the first years following application—while enrolled students are still in school—are generally slightly positive but not statistically significant. However, starting in the tenth quarter after application the results become positive and statistically significant. The earnings effects stabilizes at approximately 0.40 log points four years after application.

Because the estimates vary slightly quarter by quarter, Table 3 displays the main instrumental variables estimates for the 18th to 21st quarters after application, which are the last four quarters of earnings data I have available for all students. All the specifications control for calendar year and quarter effects, first application cohort, demographics, academic background, and prior financial aid receipt. For all these labor market outcomes, Appendix Table A6 shows that the results are not sensitive to the incremental inclusion of individual controls. Column 1 of Table 3 shows that immediate enrollment following application leads to an earnings effect of 0.367 log points, or 44 percent.

As mentioned earlier, I do not instrument for degree completion, since doing so would assume that students who enroll in the program but do not complete do not see any earnings impact. However, an estimate of the effect of completing an ADN on earnings is to scale the enrollment effect by the completion rate. The instrumental variables estimate of the effect of enrollment on completion is 0.498 (0.09), which implies that the completion effect is double the enrollment effect subject to this strong assumption.

The next three columns of Table 3 show the estimates for earnings levels in quarters 18 to 21 after application. The point estimate is an effect of $1,597, but is not statistically significant. One of the reasons for using log earnings is to avoid decreased precision due to outliers in quarterly earn-

[11] For log earnings regressions I impute $1 of earnings for students who have zero earnings. I show in a later section that results are robust to this imputation.

ings. Column 3 of Table 3 top-codes quarterly earnings above the 95th percentile—approximately $30,000 in quarterly earnings—and column 4 drops these observations entirely. This leads to more precise estimates that are consistent with the log results.

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.[12] On the other hand, the program 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 5 of Table 3 shows estimates of the effect of enrolling in the program on having non-zero earnings. The estimate suggests that enrolling in the program leads to an 11 percentage point increase in the likelihood of having non-zero quarterly earnings, but the coefficient is not statistically significant. Similarly, panel b) of Figure 3 shows a quarterly effect on employment that is consistently positive but not statistically significant. Despite the fact that this result is not statistically significant, it is relatively large and economically meaningful.

The last column of Table 3, however, shows that there is a 19.5 percentage point effect of enrolling in the program on being employed in the health industry.[13] 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 are mostly driven by earnings conditional on employment. That is, the program seems to lead to significant occupational sorting. This is evidence that the program drives participants to more lucrative occupations rather than increasing their likelihood of employment or improving their hours.

[12] 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.

[13] 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.

5.3 Robustness Checks

Table 4 shows a series of robustness checks for the main log earnings coefficient. Column 1 excludes all students who applied for a fifth time. The result is almost identical to the main estimate. Column 2 excludes students who enrolled in the program despite never winning a lottery. In conversations with program administrators I learned that these students are often military veterans or students with a special arrangement from a local hospital. In the main analyses I code these students as not winning the lottery. These students will be non-compliers because they are almost always admitted after their first application but do not win a lottery. Because they are a small enough group, however, excluding them from the analysis altogether does not significantly affect the estimates.

In the third column I limit the sample to students who had non-zero earnings prior to first application. This is similar to the set of students who will serve to identify the effects in the individual fixed effects specification in a later section. The estimated coefficient is similar to the preferred estimate.

One potential concern is that the cause of the large returns may be from students transferring to four-year colleges, making the ADN itself just an intermediary step. Column 4 excludes students who transferred to a four-year institution, and still reveals a large and almost unchanged coefficient.

Column 5 excludes data points with missing earnings. In the specifications so far I have taken the common practice of imputing \$1 where earnings are in fact zero. Taking these observations out of the analysis—that is, changing their earnings to \$0— has a negligible effect on the coefficient.

The last column of Table 4 is an attempt to investigate the mechanism by which the earnings effects accrue to students. I add controls for concurrent work in health, retail, administrative services, and education, which are the most popular industries for these students. Including these controls reduces the coefficient by 0.08 log points. While doing so is akin to conditioning on an outcome, this regression is at least suggestive evidence that the industry shift is not the sole driver of the large earnings effects.

Appendix A.2 discusses four additional specifications. In two specifications I leverage all the lottery information beyond just the first lottery. I also estimate the effects of ever enrolling in the

ADN program. Finally, I combine the instrumental variables approach with a specification that includes individual fixed effects.

6 Individual Fixed Effects

6.1 Method and Main Results

The results so far suggest a large effect of the program on earnings. In this section I compare the estimates to individual fixed effects models, which Jacobson, LaLonde and Sullivan (2005) note can produce valid estimates when students have considerable 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.

Because individual fixed effects models do not account for time-varying shocks that may affect individuals, 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.

I estimate a model of the form:

$$y_{it} = \alpha_i + \gamma Enroll_{it} + \Phi Z_{it} + \mu_t + \xi_i * t + u_{it} \tag{3}$$

For student i in quarter t, $Enroll_{it}$ takes a value of one after enrolling in an ADN program. The matrix Z_{it} consists of time-varying individual characteristics, including dummies for age and whether the student was taking at least eight community college courses that quarter. The individual fixed effect, α_i, accounts for time-invariant characteristics, so that the coefficient of interest, γ, is identified off within-individual changes in earnings. I also include calendar year and quarter fixed effects, μ_t. Individual-specific linear time trends $\xi_i * t$ account for unobserved factors that may be correlated with enrollment and that change at a constant rate over time.

Table 5 shows results of this exercise. The first column repeats the preferred lottery instrumental variables result. Columns 2 and 3 show estimates of equation 3 using the same sample of students.[14] I use data for each student from 20 quarters prior to their first application to 21 quarters afterwards. Following the approach commonly used in the literature, I drop any quarters of earnings for students when they were under 18 years old. Column 3 adds individual-specific linear time trends. All three estimates—the instrumental variables and both with individual fixed effects—are quite similar in magnitude and statistically indistinguishable. Figure 4 shows estimates of equation 3 where γ is allowed to vary by quarter, and these coefficients are displayed in Table A4.

The similarity between the instrumental variables and fixed effects estimates is perhaps not surprising since enrollment at Central College is driven in large part by the lottery process. In other words, if the lottery had perfect compliance the lottery information would not be necessary to produce causal estimates: a simple comparison of earnings between enrollees and non-enrollees would suffice.

In the next analyses I benchmark the validity of the individual fixed effects approach in cases where admission is not based on a lottery. The challenge is that estimates that leverage random lotteries cannot be produced at colleges without lotteries. Instead I show that there is a pattern in the difference between fixed effects and OLS estimates that exists at Central College as well as other ADN programs in the state, including those that do not use lotteries for admissions.

First I limit the sample of Central College applicants to only those who ultimately enrolled in the Central College ADN program: the ability to identify applicants who never enroll is a unique feature of the Central College lottery data, and I do not have such information for other programs in the state. A sample limited to only students who enroll in the Central College program is sufficient, however, since the timing of the different enrolling cohorts helps identify the effects. Column 4 of Table 5 shows an estimate of equation 3 without controlling for individual fixed effects and trends. It is a naive estimate that merely compares earnings before and after enrollment, with added controls for time-invariant characteristics such as demographics, academic performance, and financial aid receipt. The estimate in column 5 adds individual fixed effects and trends. The

[14] 135 students are dropped from the fixed effects estimates because they have zero earnings throughout the entire period and thus do not help identify the effect. As shown in the previous section, though, the IV result is similar when dropping students with no pre-application earnings.

difference in the coefficient between columns 4 and 5 shows the extent of the bias in the pre-post estimate. The specification with individual fixed effects and trends is also almost identical to the one in Column 3. The results in Columns 4 and 5 demonstrate two key points. First, the fixed effect estimate is similar to the lottery estimate. Second the fixed effects estimate also accounts for a considerable amount of bias compared to pre-post regressions.

The point of benchmarking the pre-post and fixed effects estimates in Table 5 is to then estimate the effects at the other colleges that offer ADN programs. I have academic and earnings records but no application information for colleges other than Central College, so I can only observe who enrolls in nursing programs, not who applies. I construct the statewide sample in a parallel way as the Central College lottery sample: I limit the sample to students who started an ADN program at any California community college between the Spring 2005 and Spring 2009 cohorts. Again I use earnings data from 20 quarters prior to enrollment to 21 quarters after enrollment. Panel A of Table 6 shows estimates of equation 3, with and without individual fixed effects and trends, for all ADN students at all California programs. A similar pattern exists as for Central College in the previous table: the estimates that do not control for individual fixed effects and trends are much larger than those that do. Panel B repeats this process, this time in a subsample of Central College enrollees.[15] Even controlling for individual fixed effects and trends, the estimates for Central College are much larger than for the state as a whole. A comparison in Appendix Figure A1 shows that, while Central College and other California ADN students have similar earnings trajectories prior to enrollment, Central College students see a larger increase following enrollment. As I show below, however, there is a wide range of estimates across the 66 programs in the state, with Central College towards the top but not an outlier.

A lingering concern is that the lottery and fixed estimates for Central College might be similar only because of the random variation from lottery. I cannot compare lottery and fixed effects estimates for colleges other than Central College; however, panels C and D of Table 6 suggest

[15]The sample and estimates are different here than in previous analyses because I constructed the sample for this exercise in exactly the same manner as for other colleges. The sample in this table includes all ADN enrollees, not just those matched to the lottery data. It also includes students who may have first applied prior to 2005, while the lottery analyses limit the data to applicant cohorts since 2005. Nevertheless, the estimates for Central College in Tables 5 and 6 are quite similar.

that the individual fixed effects models control for bias even in cases where there is no lottery. Panel C limits the sample to students at the 27 colleges where admission was decided by a random lottery, while panel D limits the sample to students at all other colleges, where admission might be decided by waitlist or by a non-random selection process.[16] The large drop in the coefficient when controlling for individual fixed effects and trends is apparent in both panels, and the coefficients are similar. There is still the possibility that in non-lottery colleges the individual fixed effects models may still be biased; however, the similarity of the two approaches suggests that the variation from the lottery is not the only reason that the lottery and fixed effects coefficients are similar in the Central College case.

Figure 5 shows that the pattern between the pre-post and fixed effects estimates is apparent at almost all the colleges. Panel a) of Figure 5 plots estimates of the return to enrollment in an ADN program from equation 3, calculated for each individual college, arranged in ascending order of the estimated coefficient. The dashed line gives the overall mean. The majority of estimates are large and positive, yet there is a considerable range. This is surprising, given that all ADN programs offer a similar curriculum, have similar prerequisites, and are overseen by a state board. As mentioned earlier, the coefficient of 0.437 for Central College is larger than the average, but not an outlier. Panel b) of Figure 5 compares each college's coefficient controlling and not controlling for individual fixed effects and trends, with the Central College estimate highlighted as a larger black square. For all but one college the individual fixed effects coefficient is smaller, further providing support for the use of this approach.

6.2 Heterogeneity in Individual Fixed Effects Results

In this section I explore the heterogeneity across programs shown in Figure 5. In particular, I focus on characteristics of the program and characteristics of the surrounding labor market, defined as the college's county. Table 7 shows estimates of the difference in the coefficient from Equation 3

[16] 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. 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)

when splitting the sample into groups based on demographic or college characteristics. In results not shown I instead interact the main coefficient of interest with an indicator for membership in the particular group, and find similar patterns of heterogeneity. It is important to note that the heterogeneity results presented here are correlational and meant to be exploratory, since selection into particular programs is likely not exogenous.

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. The first row of Panel A of Table 7 shows that students in programs that were above the median size, in terms of numbers of enrolled students, saw earnings increases 0.11 log points greater than students in small programs. The χ^2 statistic from a test of the equality of the two coefficients is highly statistically significant. Similarly, students in programs with high completion rates also saw larger returns. Perhaps puzzlingly, students in programs that had high first-time pass rates on the NCLEX-RN, the national licensing exam, saw significantly lower earnings returns. However, this test is perhaps not a good indicator of program quality: first-time pass rates were high across almost all colleges, with more than 80 percent of colleges having pass rates above 80 percent. Following Gill and Leigh (2004), I also examine differences in earnings estimates across different types of colleges. Students at colleges that had an above-median share of their programs in career technical fields saw substantially larger earnings effects than other students. This is perhaps because these types of colleges are able to specialize in their connections with local labor market opportunities or because they might have better non-instructional supports for CTE students. Students at colleges with high transfer rates, as defined by the CCCCO, had slightly larger earnings increases.

Heterogeneity in earnings returns might also be due to differences in local conditions in the labor market for registered nurses. As a measure of the density of job opportunities for nurses, I compiled data from the California Office of Statewide Health Planning and Development on the capacity of hospitals in the state. Students in counties with a higher density of hospital beds per capita had earnings returns that were approximately nine percent higher than other students. As another rough estimate of demand for healthcare, I categorized counties by median share of the population age 60 and over in the 2000 Census. This measure was not correlated with the earnings return. 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 employment and earnings. Students in counties where medical assistants made lower wages relative to nurses saw higher returns.

Another key 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. There were 12 programs whose policies featured a lottery among all eligible applicants, like at Central College. There were 40 "competitive" programs, at which admission decisions were based in part on the applicant's qualifications. In an additional nine programs admission was based on a waitlist or first-come-first-served.[17] Students in programs with lotteries saw earnings returns that were slightly lower than students in competitive and wait-list programs, while students in competitive programs saw slightly larger earnings effects.

Panel D shows differences in the estimated returns by certain individual demographic characteristics. African American and Hispanic/Latino students have an earnings estimate 0.05 log points lower than other students. Estimates for women and older students are also slightly larger and statistically significant.

7 Private and Social Returns to ADN Program Expansion

With ADN programs oversubscribed and a growing demand for healthcare workers, a crucial question is whether it is cost-effective to increase capacity to ADN programs. In this section I calculate the internal rate of return (IRR) implied by the estimates. I assume no general equilibrium effects, which would likely decrease the estimated return.

[17] 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.

I calculate the IRR using both the quarterly lottery estimates from Table A3 and comparable quarterly fixed effects estimates for all Central College applicants from Table A4. I assume a 30 year career after enrollment. Because I only estimate effects up to 21 quarters (5.25 years) after application, I assume that the earnings effect stabilizes at the average of the final four quarters of estimated effects.[18] I convert the log earnings effect into a dollar amount by using the mean pre-application quarterly earnings of \$4,740. An important component of the IRR is the earnings growth rate of the comparison group. Empirically, I find that earnings of students who do not enroll in the program rise rapidly in the first few years and then level off, with overall growth rates of between two and five percent. In Table 8 I give a range of estimates of the IRR for assumed comparison group earnings growth rates of zero, three, and 10 percent. Another important factor for the private return to enrollment is whether the student pays tuition. The program costs \$2,100 over six calendar year quarters, but approximately half of students have their tuition waived. I present estimates of the IRR for students who pay full tuition and students who do not. I assume that all students incur an upfront cost of \$5,700 in supplies, immunizations, and books, which is what the Central College catalog estimates.

The IRR estimates for the lottery estimates range from 69 to 101 percent, while those for the fixed effects specification are smaller. The difference between the estimates for lottery and fixed effects specifications is due in large part to the fixed effects estimates being negative in the early quarters after application. Nevertheless, all the IRR estimates are large, especially compared to standard estimates of the returns to a year of post-secondary education. I cannot measure non-earnings benefits, so these estimates are still likely a lower bound.

An often cited reason for the lack of expansion of nursing programs is the prohibitive cost of adding a new seat. A reasonable empirical question, then, is to estimate the social return to expansion. There is no prior study to my knowledge that explicitly estimates expansion costs though. I compiled data from sources to separately estimate variable and fixed costs of program

[18] Because the IRR is so large, changes in the assumptions about the earnings effects in later years and about career length have a negligible effect on the IRR. For example, assuming that the earnings effect decays by 20 percent each quarter after the 21st quarter only reduces the IRR by between one and two percentage points depending on the specification. Similarly, shortening or lengthening the assumed 30 year career length by ten years has negligible effects.

expansion. I obtained per-student operating costs from two nursing programs in the state, which ranged from \$7,600 to \$9,200 per year per student. A similar estimate comes from a California legislature initiative that granted expansion funds from between \$6,500 and \$9,100 per new full-time equivalent student (California Community College Chancellor's Office, 2015, 2016*a*). A conservative estimate, rounding up, is that variable costs are approximately \$10,000 per year per student. Much less information is available about capital, infrastructure, and equipment costs that a college would need to pay upfront. Healthcare programs require specialized machinery and teaching equipment, and instruction often occurs in dedicated facilities. In the past five years two community colleges completed construction for new nursing buildings in California. Both projects cost approximately \$8,000 per new student. In addition, I used one program's inventory list to estimate that teaching equipment, such as practice maniquins, cost \$1,000 per additional student. Thus, the fixed infrastructure and capital spending associated with adding an additional student is approximately \$9,000.

To calculate a back-of-the-envelope estimate of the social return to program expansion, I include \$9,000 in upfront costs and \$10,000 in operating costs split over two years into the IRR calculations, and omit tuition costs. The resulting social return ranges from 21 to 33 percent, depending on the specification and control group earnings growth rate. Not included in these calculations are the obvious spillover benefits to training new nurses and filling vacant positions. 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. My estimates, while large, should thus be taken as a lower bound of the social return to expanding a nursing program.

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. 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.

8 Conclusion

In this paper I leverage the random assignment of admissions to a large community college ADN program to estimate the effect of enrollment on later earnings, 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 show that these estimates are consistent with methods that are more common in the literature.

I find large earnings effects and estimate that the value of expanding an ADN program far outweighs the costs. Despite the large economic benefit, there are limited incentives to community college administrators to expand enrollment. 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. Potential solutions include adopting differential pricing or funding of programs, or performance-based funding schemes that allocate additional funds for increased graduation rates in certain costly fields (Stange, 2012; Smith, 2016; Long, 2016). Policies in North Carolina and Texas, for example, have received recent attention.

My results also show that, for ADN applicants, the alternative to a degree in nursing seems to be no degree whatsoever. This is in line with growing work on differences in payoffs across college majors, where the differences may be driven in large part by selection into majors (Kirkeboen, Leuven and Mogstad, 2016). My findings suggest that selection into field of study can be highly specific and not necessarily driven by anticipated earnings returns. Rejected ADN applicants do not enter associate degree programs in similar fields, such as dental hygiene or radiologic technology, even though these programs have similar starting wages, are less selective, and are often offered at the same college.

My findings also have implications for researchers using individual fixed effects models to estimate earnings returns to educational and training interventions. Starting with Jacobson, LaLonde and Sullivan (2005), these methods have increasingly been used. This paper shows that these models produce estimates that are similar to those generated through a random lottery.

An open question is what the mechanism for the large earnings effects is. A large portion of this effect likely comes 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. On the other hand, part of the return is likely due to the highly structured nature of the program, which also has many non-academic supports for students. A growing field of research has increasingly shown the benefits of a structured curriculum and non-academic supports for community college students (Scott-Clayton, 2015; Gardiner et al., 2018).

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.

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

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, 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.

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.

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.

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.

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.

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

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.

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

Gardiner, Karen, Howard Rolston, David Fein, and Sung-Woo Cho. 2018. "Pima Community College Pathways to Healthcare Program: Implementation and Early Impact Report, OPRE Report No. 2017-10." *Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services*.

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.

Gill, Andrew Matthew, and Duane E Leigh. 2004. *Evaluating academic programs in California's community colleges.* Public Policy Institute of California San Francisco, CA.

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.

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.

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.

Kirkeboen, Lars J., Edwin Leuven, and Magne Mogstad. 2016. "Field of Study, Earnings, and Self-Selection." *The Quarterly Journal of Economics*, 131(3): 1057–1111.

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

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.

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."

Marcotte, Dave E. 2016. "The Returns to Education at Community Colleges: New Evidence from the Education Longitudinal Survey." Institute for the Study of Labor (IZA) IZA Discussion Papers 10202.

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

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*.

Peck, Laura R., Alan Werner, Eleanor Harvill, Daniel Litwok, Shawn Moulton, Alyssa Rulf Fountain, and Gretchen Locke. 2018. "Health Profession Opportunity Grants (HPOG 1.0) Impact Study Interim Report: Program Implementation and Short-Term Impacts, OPRE Report 2018-16." *Research and Evaluation, Administration for Children and Families, US Department of Health and Human Services*.

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*.

Scott-Clayton, Judith. 2015. "The Shapeless River: Does a Lack of Structure Inhibit Students Progress at Community Colleges." In *Decision Making for Student Success: Behavioral Insights to Improve College Access and Persistence..* , ed. Sandy Baum, Ben Castleman and Saul Schwartz. London:Routledge.

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.

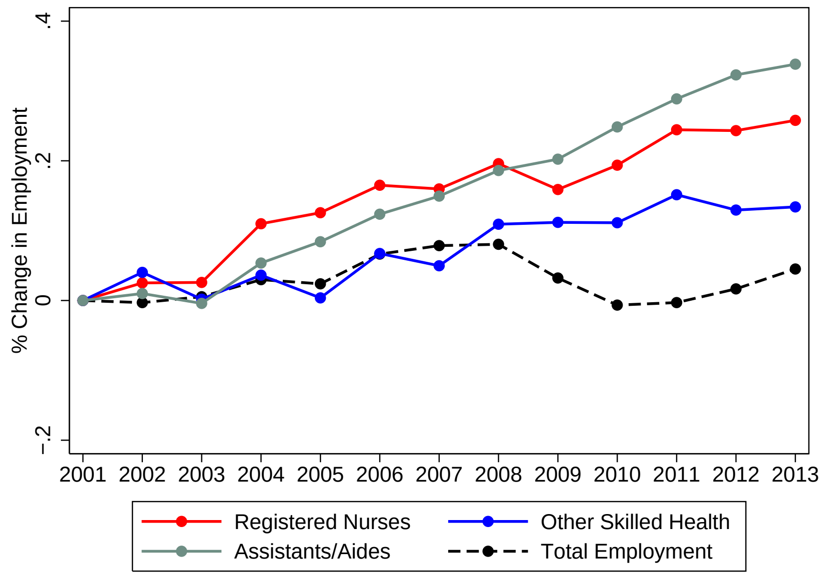
Stevens, Ann Huff, Michal Kurlaender, and Michel Grosz. 2015. "Career Technical Education and Labor Market Outcomes: Evidence from California Community Colleges." National Bureau of Economic Research, Inc NBER Working Papers 21137.

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.

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.

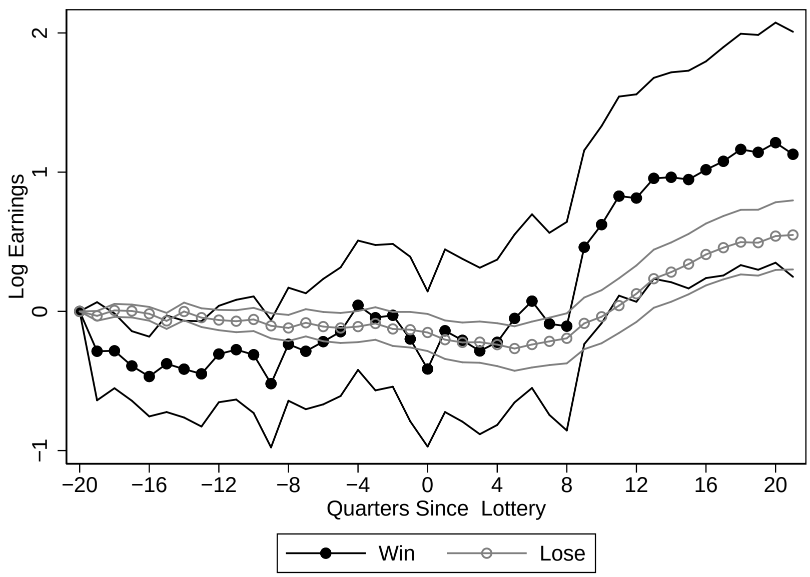
9 Tables and Figures

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



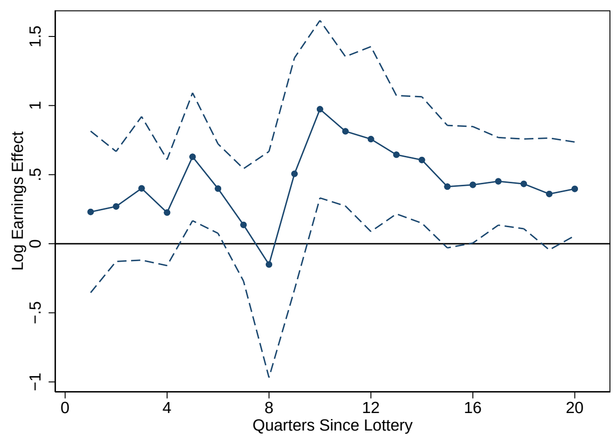
Notes. This graph shows employment growth, relative to 2001 levels, for healthcare occupations that require a sub-baccalaureate degree or a certificate. Data come from the Occupational Employment Statistics. The category of Other Skilled Health professions include LPN, radiologic technicians, dental hygienists, respiratory care therapists, and surgical technicians.

Figure 2: Earnings Trajectories, by 1st Lottery Result

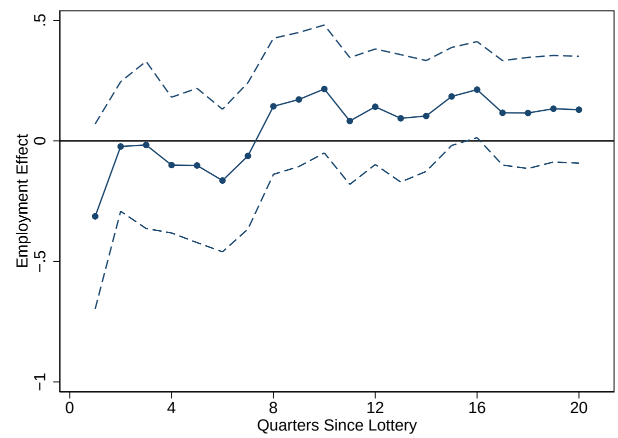


Notes. Sample consists of 1,730 students who applied to Central College's ADN program between Spring 2005 and Spring 2009, separated by whether the student was admitted upon first lottery application. Point estimates come from a regression of log earnings on dummies of quarter since first application (omitting quarter 20 prior to application), age dummies, concurrent community college enrollment, and calendar time effects. Point estimates and 95% confidence interval shown, with standard errors clustered at individual level.

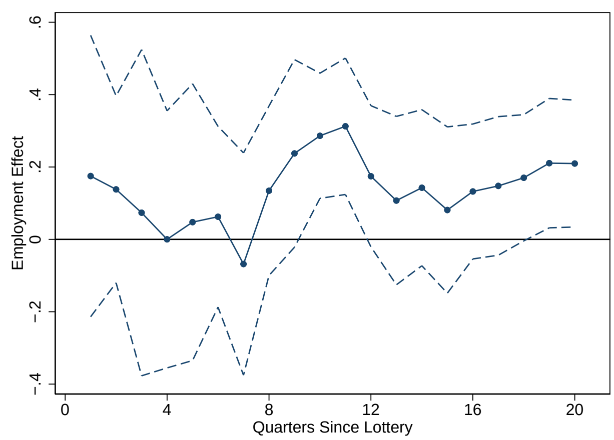
Figure 3: Instrumental Variables Estimates, Quarter-by-Quarter



a) Log Earnings



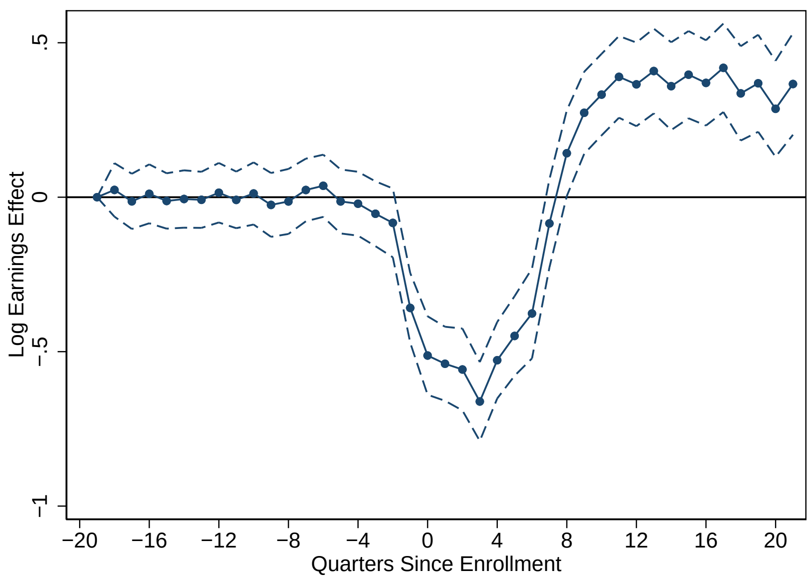
a) Had Any Earnings



c) Employed in Health Industry

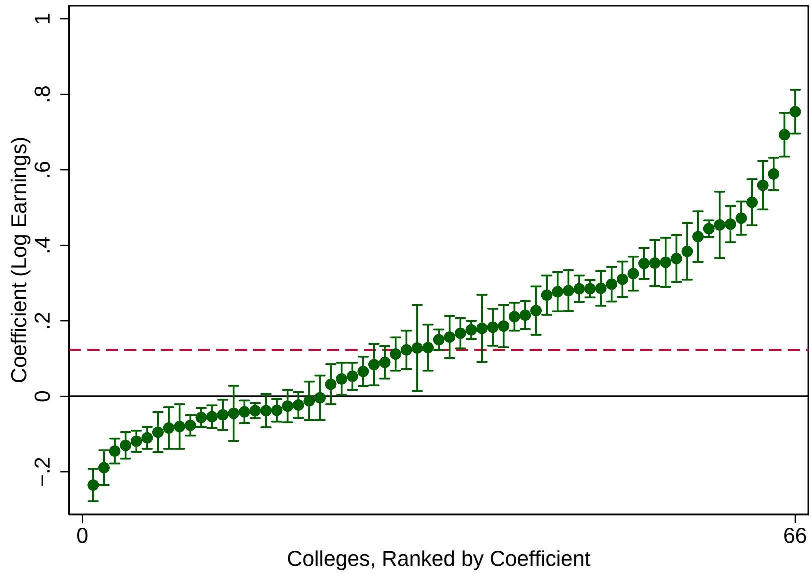
Notes. Figure shows point estimate and 95% confidence interval for instrumental variables estimates of immediate enrollment in Central College ADN program, instrumented with result of first lottery. Effects at each quarter come from a separate regression, with 1,730 students at each point. Outcomes are quarterly log earnings, having non-zero earnings, and employment in health professions. Employment in Health defined as employment in the two-digit NAICS industry code 62: Health Care and Social Assistance. Regressions control for calendar year, application cohort, demographics (age, gender, race), academic background (prior GPA, prior number of units), prior financial aid receipt (Pell grants, tuition waivers), and prior labor market experience (mean prior earnings, any prior employment in health). Standard errors clustered at the individual level.

Figure 4: Quarterly Individual Fixed Effects Estimates, Central College

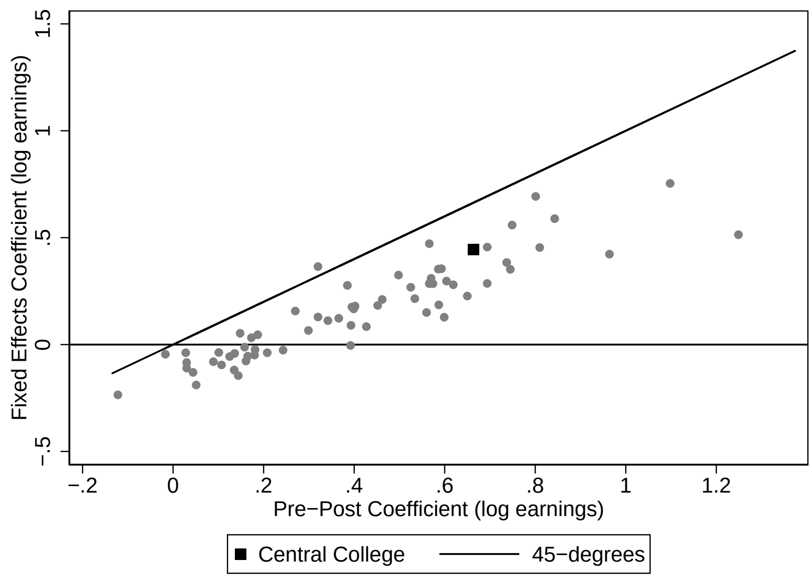


Notes. Figure shows point estimates and 95% confidence interval for a single regression of log earnings effects at each quarter relative to enrollment at the Central College ADN program. Omitted category is quarter 20 prior to enrollment. Sample consists of 1,595 applicants to the Central College ADN program between Spring 2005 and Spring 2009. Regressions control for calendar time effects, age dummies, full-time community college enrollment, and individual fixed effects. Standard errors clustered at the individual level.

Figure 5: Individual Fixed Effects Returns, Heterogeneity by College



a) Coefficients and Confidence Intervals, Ranked



b) Comparison of Fixed Effects and Pre-Post Design

Notes. Panel a) shows coefficients and 95% confidence intervals for college-specific regressions of log earnings on post-enrollment, controlling for individual fixed effects, individual-specific linear time trends, calendar year, age dummies, and concurrent full-time community college enrollment, and clustering at the individual level. Sample consists of all students who enrolled in ADN programs at any community college in California between Spring 2005 and Spring 2009. Coefficients are shown in ascending order of point estimate. The dashed line shows the overall system-wide coefficient when aggregating all colleges together. Panel b) shows the coefficients of the regressions from panel a) on the vertical axis, and the coefficient from an equivalent regression that does not control for individual fixed effects or trends on the horizontal axis. The black diagonal line is the 45-degree line. The large black square is the coefficient for Central College

Table 1: Applicant Characteristics and Lottery Balance

		Admit-Reject Difference	
	Mean	All Lotteries	First Lottery
Female	0.786	0.0571	0.0788
		(0.0429)	(0.0762)
White	0.342	0.0412	-0.0186
		(0.0837)	(0.0936)
Hispanic	0.385	-0.00976	0.0481
		(0.0505)	(0.0558)
Asian	0.144	-0.0328	-0.0867
		(0.0459)	(0.0387)
Age	29.35	1.003	0.191
		(0.770)	(0.925)
GPA	2.793	-0.207	-0.214
		(0.0645)	(0.119)
Enrolled in other district	0.275	-0.0248	-0.0508
		(0.0616)	(0.0656)
Had BOG Waiver	0.540	-0.0294	-0.0192
		(0.0301)	(0.0361)
Had Pell Grant	0.327	-0.0494	-0.0568
		(0.0507)	(0.0620)
Calgrant	0.139	0.00235	-0.0265
		(0.0276)	(0.0395)
Had Loans	0.0540	-0.00168	-0.0231
		(0.0189)	(0.0190)
Employed > 1 Quarter	0.818	-0.0416	0.0131
		(0.0354)	(0.0624)
Quarters Employed	9.161	-0.352	0.353
		(0.508)	(0.693)
Employed > 8 Quarters	0.626	-0.0322	0.0622
		(0.0436)	(0.0557)
Mean Quarterly Earnings	4740.5	-65.66	282.3
		(629.5)	(1019.8)
Industry is Health	0.398	0.0577	0.115
		(0.0556)	(0.0615)
Industry is Retail	0.198	0.0133	-0.00539
		(0.0332)	(0.0579)
Industry is Administrative	0.104	-0.0152	0.00252
		(0.0258)	(0.0518)
Industry is Education	0.0891	0.0108	0.00804
		(0.0257)	(0.0157)
Industry is Food Service	0.141	-0.0439	-0.0304
		(0.0244)	(0.0260)
N	1730	4082	1730

Notes. First column shows mean characteristics for applicants in the Spring 2005 to Spring 2009 Central College ADN lotteries, measured at term of first application. GPA measures grades in prerequisites prior to application. Enrollment at other district defined as ever having taken a course at a community college outside Central College's district. BOG waiver is a full tuition waiver. Calgrant is state-specific financial aid. 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. Mean quarterly earnings measured in four years prior to application. Employment by industry defined by two-digit NAICS industry codes: Health is NAICS code 62; Retail is NAICS codes 44 and 45; Administrative is NAICS code 56; Education is NAICS code 61; and Food Service is NAICS code 72. Second and third columns show results of regressing each characteristic on lottery admission and cohort fixed effects. Second column includes all applications that were decided by random lottery, including up to four lotteries per applicant. The final column only includes the first lottery a student entered. Standard errors clustered at individual level.

Table 2: Effect of Lottery Result on Academic Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Applica-tions	Enroll Im-mediately	Ever Enroll	Complete Program	Any Health Award	Any Award	Transfer
Win 1st Lottery	-1.524	0.485	0.176	0.203	0.189	0.187	0.0326
	(0.0903)	(0.0676)	(0.0678)	(0.0741)	(0.0739)	(0.0735)	(0.0470)
F	285.2	51.42	6.716	7.473	6.569	6.469	0.480
Students	1730	1730	1730	1730	1730	1730	1730
Y-Mean	2.898	0.134	0.461	0.355	0.370	0.390	0.0936

Notes. Table shows estimates of the effect of a student being admitted to the Central College ADN on the first application. Sample consists of students who first applied between Spring 2005 and Spring 2009. Applications is the number of applications ever submitted. Enrolled immediately is enrollment in the Central College ADN program the following semester. Ever enrolled is ever having enrolled in the Central College ADN program. Complete programs is earn an ADN from Central College. Any Health Award is earn any associate's degree or certificate in a health field from any California community college. Any award is earn any associate's degree or certificate in any field from any California community college. Transfer is whether the student ever later enrolled in a four-year college. Regressions control for calendar year, application cohort, demographics (age, gender, race), academic background (prior GPA, prior number of units), prior financial aid receipt (Pell grants, tuition waivers), and prior labor market experience (mean prior earnings, any prior employment in health). Standard errors clustered at the individual level.

Table 3: IV Estimate of Effect of Enrollment on Labor Market Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)
		Earnings Levels			Employment	
	Log Earnings	All	Top-code	Censor	Any Earnings	Health Industry
Enroll	0.367	1596.9	2657.9	3239.3	0.111	0.195
	(0.148)	(1933.4)	(1797.2)	(1519.1)	(0.113)	(0.0848)
N	6920	6920	6920	6695	6920	4926
Students	1730	1730	1730	1706	1730	1316
Mean Earnings	11741.7	11741.7	11424.1	10805.0	0.736	0.818
First stage F	44.79	44.79	44.79	51.88	44.47	40.60

Notes. Table shows estimates of the effect of immediate enrollment in the Central College ADN program, instrumented with result of first application. Sample consists of students who first applied between Spring 2005 and Spring 2009. There are four quarters of data for each student, corresponding to quarters 18 through 21 following first application to the program. Top-coded earnings levels recoded quarters of earnings above $30,000, approximately the 95th percentile, as $30,000. Censored earnings levels drop these quarters of high earnings altogether. Health industry employment measures whether the individual had earnings in the two-digit NAICS code 62. Regressions control for calendar time, application cohort, demographics (age, gender, race), academic background (prior GPA, prior number of units), prior financial aid receipt (Pell grants, tuition waivers), and prior labor market experience (mean prior earnings, any prior employment in health). Standard errors clustered at the individual level.

Table 4: IV Estimates: Robustness

	(1) Exc. 5x Applicants	(2) Exc. Other Admits	(3) Exc. No Pre-Earn	(4) Exc. Transfer	(5) Exc. Zeros	(6) Industry Control
Enroll	0.348	0.372	0.367	0.338	0.398	0.294
	(0.147)	(0.130)	(0.161)	(0.147)	(0.162)	(0.145)
N	4692	6060	6164	6532	5109	6920
Students	1173	1515	1541	1633	1364	1730

Notes. Table shows estimates of the effect of immediate enrollment in the Central College ADN program, instrumented with result of first application. Sample consists of students who first applied between Spring 2005 and Spring 2009. There are four quarters of data for each student, corresponding to quarters 18 through 21 following first application to the program. Column 1 omits all students who applied for a fifth consecutive time. Column 2 omits students who were admitted to the program but not through the lottery (e.g. special programs for veterans). Column 3 omits students who had no quarters of non-zero earnings prior to first application. Column 4 omits students who ever transferred to a four-year institution at any time after the first lottery. Column 5 excludes all observations with zero earnings, which in the preferred estimates are coded as zero. Column 6 includes additional controls for concurrent industry of employment, measured at the two-digit NAICS code level. All regressions control for calendar time, application cohort, demographics (age, gender, race), academic background (prior GPA, prior number of units), prior financial aid receipt (Pell grants, tuition waivers), and prior labor market experience (mean prior earnings, any prior employment in health). Standard errors clustered at the individual level.

Table 5: Comparison of Lottery and Observational Estimates

	(1)	(2)	(3)	(4)	(5)
	All Central College Applicants			ADN Program Enrollers Only	
	IV	FE	FE+Trends	Pre-Post	FE+Trends
Enroll	0.367	0.339	0.327	0.666	0.361
	(0.148)	(0.0351)	(0.0372)	(0.0397)	(0.0372)
N	6920	46268	46268	22806	22806
Students	1730	1595	1595	754	754

Notes. Sample for the first three columns consists of students who first applied between Spring 2005 and Spring 2009. Column 1 shows estimates of the effect of immediate enrollment in the Central College ADN program, instrumented with result of first application. There are four quarters of data for each student, corresponding to quarters 18 through 21 following first application to the program. Regression controls for calendar time, application cohort, demographics (age, gender, race), academic background (prior GPA, prior number of units), prior financial aid receipt (Pell grants, tuition waivers), and prior labor market experience (mean prior earnings, any prior employment in health). Column 2 shows estimates of equation 3; include quarterly data 20 quarters prior and 21 quarters after enrollment; and control for age dummies, concurrent community college enrollment, calendar time effects, and individual fixed effects. Column 3 repeats Column 2 specification and includes individual-specific linear time trends. Columns 4 and 5 limit the sample to only students who enrolled in the Central College ADN program. Column 4 includes quarterly data 20 quarters prior and 21 quarters after enrollment, and controls for demographics, academic background, financial aid receipt, age dummies, calendar time effects, and concurrent enrollment. Column 5 specification adds individual fixed effects and trends. Standard errors clustered at the individual level.

Table 6: Individual Fixed Effects Estimates at All California ADN Programs

	(1) Pre-Post	(2) Individual Fixed Effects	(3)
A. All Colleges			
Start Program	0.358	0.199	0.123
	(0.00748)	(0.00572)	(0.00509)
N	1215938	1215938	1215938
Students	44716	44716	44716
B. Central College			
Start Program	0.649	0.499	0.437
	(0.0305)	(0.0240)	(0.0221)
N	49710	49710	49710
Students	1535	1535	1535
C. Colleges with Lottery Admissions			
Start Program	0.342	0.173	0.107
	(0.0168)	(0.0123)	(0.0113)
N	234096	234096	234096
Students	8424	8424	8424
D. Colleges with Non-Lottery Admissions			
Start Program	0.362	0.205	0.127
	(0.00835)	(0.00644)	(0.00570)
N	981842	981842	981842
Students	36292	36292	36292
Individual fixed effects		X	X
Individual-specific linear time trends			X

Notes. Sample consists of students who ever enrolled in an ADN program at any California community college between Spring 2005 and Spring 2009. Data include quarters between 20 quarters prior and 21 quarters after enrollment. Outcome is log earnings. Main coefficient is on a dummy variable with value of one after enrollment and zero otherwise. Column 1 controls for demographics, academic background, financial aid receipt, age dummies, calendar time effects, and concurrent enrollment. Column 2 adds individual fixed effects, and Column 3 adds individual-specific linear time trends. Panel A is all colleges, Panel B is for students who ever enrolled in the Central College ADN program, Panel C is for students who ever enrolled in an ADN program with lottery-based admissions, and Panel D is for students who ever enrolled in a program that did not have lottery-based admissions. Standard errors clustered at the individual level.

Table 7: Individual Fixed Effects, Heterogeneity

	(1)	(2)	(3)	(4)	(5)
				Students	
	Diff	χ^2	p-value	Group 1	Group 2
A. Program Characteristics (Above Median)					
Program Size	0.11	110.83	0.00	27034	17682
Program Completion Rate	0.03	7.60	0.01	22112	22604
NCLEX First-Time Pass Rate	-0.08	63.01	0.00	21171	23545
College's CTE Share of Awards	0.09	68.41	0.00	25555	19123
College's Transfer Rate	0.03	11.08	0.00	19802	24786
B. County-Level Characteristics (Above Median)					
Number of Hospital Beds	0.09	72.51	0.00	24911	19677
Population Share Older than 60	-0.01	0.62	0.43	19650	24938
RN Employment	0.03	8.81	0.00	21941	22775
Wage Ratio of RN to Medical Assistant	0.07	33.64	0.00	13352	31364
C. Admissions Type					
Lottery	-0.02	2.57	0.11	8424	36292
Competitive	0.02	4.48	0.03	25618	19098
Waitlist or First-Come-First-Served	-0.01	0.33	0.56	4920	39796
D. Individual Characteristics					
African American or Hispanic	-0.05	20.94	0.00	13591	31125
Older than 30	0.02	5.03	0.02	19446	25270
Female	0.04	8.18	0.00	34870	8409

Notes. Sample consists of students who ever enrolled in an ADN program at any California community college between Spring 2005 and Spring 2009. Data include quarters between 20 quarters prior and 21 quarters after enrollment. Outcome is log earnings. Each row shows the difference in the coefficient on enrollment from two separate regressions. Main coefficient is on a dummy variable with value of one after enrollment and zero otherwise. Column 1 shows the difference, columns 2 and 3 show the χ^2 test and associated p-value of the test of equality of the coefficients. Columns 4 and 5 show the number of students in groups 1 and 2 for the particular regression. In Panel A group 1 consists of students at colleges with above the median level of the listed attribute, and group 2 is below the median. Program size refers to the number of new students; program completion rate is the share of students who complete the program; NCLEX First-Time Pass Rate is the share of students who pass the licensing exam on their first attempt; and college transfer rate comes from the CCCCO estimate of transfer velocity. In Panel B the groups are also high (group 1) and low (group 2) relative to the median. Information on hospital beds comes from the California Office of Statewide Health Planning and Development, 2014, and is expressed a share of county population from the 2010 Census. RN employment is the number of RN's as a share of total employment, and the wage ratio also comes from the 2010 Census. For Panel C, lottery programs (27 programs) had randomization in their admission process. 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. For Panel D, age is measured at first date of enrollment. Standard errors clustered at the individual level.

Table 8: Internal Rate of Return Calculations for Central College

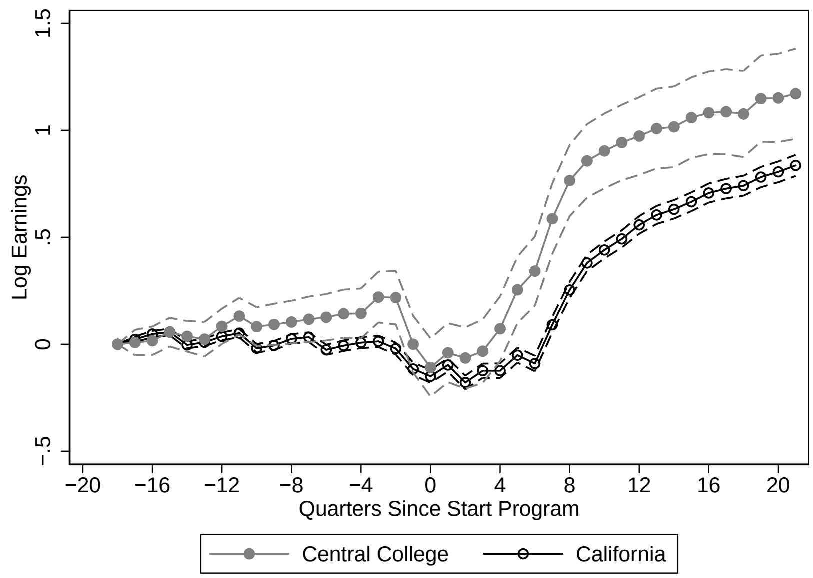
	(1)	(2)	(3)	(4)	(5)	(6)
	No Tuition Waiver			Full Tuition Waiver		
Control Group Earnings Growth Rate	None	3%	10%	None	3%	10%
Lottery Instrument	69	72	84	84	91	101
Fixed Effects	46	52	63	52	57	66

Notes. Table shows calculations of the internal rate of return, expressed as percentages. Lottery instruments use quarterly estimates from Appendix Table A3 and fixed effects use estimates from Appendix Table A4. Earnings benefits are the log estimate converted to a percent, multiplied by the counterfactual earnings mean. Counterfactual earnings in the first quarter are $4,740, and the columns of the table show whether there is zero, 3%, or 10% subsequent annual earnings growth. Earnings effects are calculated up to 30 years; log earnings effects after quarter 21 are assumed to remain consistent at the mean of quarters 18 to 21. The first three columns of the table show estimates where students are assumed to pay $350 in tuition each quarter for the first six quarters, while the second set of three columns assume the students have their tuition waived. All students are assumed to pay $5,700 upfront in costs and supplies.

A Appendices

A.1 Additional Tables and Figures

Figure A1: Earnings Trajectories, Central College and Statewide ADN Enrollees



Notes. Sample consists of all students who enrolled at ADN programs in California between Spring 2005 and Spring 2009. Log quarterly earnings displayed since quarter of first enrollment, and net of calendar time effects, age dummies, and concurrent community college enrollment. Point estimates shown relative to earnings at 20 quarters prior to enrollment. Standard errors clustered at the individual level.

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

	4-Year Public	2-Year Public	All Students California	All Students Central	All Health Awards California	All Health Awards Central	ADN Graduates California	ADN Graduates 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: Balance, Joint Regressions

	(1) First Lottery	(2) All Lotteries
Female	0.014	0.008
	(0.010)	(0.006)
White	0.008	0.004
	(0.011)	(0.007)
Hispanic	0.013	0.000
	(0.011)	(0.007)
Asian	-0.009	-0.004
	(0.017)	(0.009)
Age	0.000	0.000
	(0.001)	(0.000)
GPA	-0.004	-0.002
	(0.003)	(0.002)
Enrolled in other district	-0.012	-0.006
	(0.010)	(0.006)
Had BOG Waiver	0.003	-0.000
	(0.011)	(0.007)
Had Pell Grant	-0.009	-0.009
	(0.013)	(0.008)
Calgrant	0.000	0.007
	(0.014)	(0.008)
Had Loans	-0.009	0.003
	(0.018)	(0.011)
Employed > 1 Quarter	0.002	-0.013
	(0.017)	(0.010)
Quarters Employed	-0.002	0.001
	(0.002)	(0.001)
Employed > 8 Quarters	0.022	-0.008
	(0.021)	(0.013)
Mean Quarterly Earnings	0.000	0.000
	(0.000)	(0.000)
Industry is Health	0.014	0.009
	(0.009)	(0.005)
Industry is Retail	0.004	0.008
	(0.010)	(0.006)
Industry is Administrative	0.001	-0.002
	(0.012)	(0.007)
Industry is Education	0.001	0.005
	(0.014)	(0.008)
Industry is Food Service	-0.005	-0.005
	(0.012)	(0.007)
N	1730	4082
F	0.594	0.706
p	0.920	0.844

Notes. Outcome in both columns is admission to Central College ADN program. Sample consists of applications in the Spring 2005 to Spring 2009 Central College ADN lotteries. Column 1 shows just the first applications, and Column 2 shows all applications. Regressions control for lottery cohort. GPA measures grades in prerequisites prior to application. Enrollment at other district defined as ever having taken a course at a community college outside Central College's district. BOG waiver is a full tuition waiver. Calgrant is state-specific financial aid. 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. Mean quarterly earnings measured in four years prior to application. Employment by industry defined by two-digit NAICS industry codes: Health is NAICS code 62; Retail is NAICS codes 44 and 45; Administrative is NAICS code 56; Education is NAICS code 61; and Food Service is NAICS code 72. Standard errors clustered at individual level.

Table A3: Quarter-by-Quarter IV Estimates for Central College

Quarter Since Lottery	Log Earnings		Any Earnings		Employment in Health	
1	-0.217	(0.23)	-0.058	(0.10)	-0.098	(0.13)
2	0.230	(0.30)	0.175	(0.20)	-0.313	(0.20)
3	0.270	(0.20)	0.138	(0.13)	-0.023	(0.14)
4	0.401	(0.26)	0.074	(0.23)	-0.017	(0.18)
5	0.226	(0.20)	0.000	(0.18)	-0.100	(0.14)
6	0.629	(0.24)	0.048	(0.20)	-0.102	(0.16)
7	0.399	(0.17)	0.063	(0.13)	-0.164	(0.15)
8	0.137	(0.21)	-0.068	(0.16)	-0.062	(0.15)
9	-0.150	(0.42)	0.134	(0.12)	0.144	(0.14)
10	0.506	(0.43)	0.237	(0.13)	0.172	(0.14)
11	0.974	(0.33)	0.286	(0.09)	0.216	(0.14)
12	0.814	(0.28)	0.313	(0.10)	0.083	(0.13)
13	0.757	(0.34)	0.174	(0.10)	0.142	(0.12)
14	0.644	(0.22)	0.107	(0.12)	0.094	(0.13)
15	0.606	(0.23)	0.143	(0.11)	0.104	(0.12)
16	0.413	(0.23)	0.081	(0.12)	0.185	(0.10)
17	0.426	(0.22)	0.132	(0.10)	0.213	(0.10)
18	0.452	(0.16)	0.148	(0.10)	0.117	(0.11)
19	0.433	(0.17)	0.170	(0.09)	0.116	(0.12)
20	0.360	(0.21)	0.211	(0.09)	0.134	(0.11)
21	0.397	(0.17)	0.210	(0.09)	0.130	(0.11)

Notes. Table shows estimates of the effect of immediate enrollment in the Central College ADN program, instrumented with result of first application. Sample consists of 1,730 students who first applied between Spring 2005 and Spring 2009. There are four quarters of data for each student, at each quarter relative to first application to the program. Each cell corresponds to an individual regression. Regressions control for calendar time, application cohort, demographics (age, gender, race), academic background (prior GPA, prior number of units), prior financial aid receipt (Pell grants, tuition waivers), and prior labor market experience (mean prior earnings, any prior employment in health). Standard errors clustered at the individual level.

Table A4: Quarter-by-Quarter Fixed Effects Estimates for Central College

Quarter Since Enrollment		
1	-0.539	(0.060)
2	-0.558	(0.066)
3	-0.661	(0.064)
4	-0.528	(0.062)
5	-0.449	(0.064)
6	-0.377	(0.073)
7	-0.085	(0.071)
8	0.142	(0.069)
9	0.273	(0.066)
10	0.332	(0.066)
11	0.390	(0.066)
12	0.365	(0.068)
13	0.408	(0.069)
14	0.360	(0.071)
15	0.397	(0.071)
16	0.370	(0.069)
17	0.419	(0.072)
18	0.336	(0.076)
19	0.369	(0.079)
20	0.286	(0.078)
21	0.366	(0.082)

Notes. Table shows estimates of the effect of immediate enrollment in the Central College ADN program. Outcome is log earnings. Sample consists of 1,730 students who first applied between Spring 2005 and Spring 2009, with data up to 20 quarters prior and 21 quarters after enrollment. Omitted quarter is 20 quarters prior to enrollment Regressions control for calendar time, age dummies, concurrent community college enrollment, individual fixed effects, and individual-specific linear time trends. Standard errors clustered at the individual level.

Table A5: Sensitivity of Academic Outcome Estimates to Inclusion of Controls

	(1)	(2)	(3)	(4)	(5))
A. Applications					
Win 1st Lottery	-1.561	-1.559	-1.573	-1.528	-1.526
	(0.0905)	(0.0903)	(0.0905)	(0.0911)	(0.0918)
F	297.6	297.7	302.0	281.7	276.3
Students	1730	1730	1730	1730	1730
B. Enroll Immediately					
Win 1st Lottery	0.505	0.504	0.504	0.486	0.485
	(0.0680)	(0.0682)	(0.0683)	(0.0680)	(0.0674)
F	55.27	54.59	54.49	50.93	51.70
Students	1730	1730	1730	1730	1730
C. Ever Enroll					
Win 1st Lottery	0.196	0.194	0.194	0.171	0.175
	(0.0690)	(0.0690)	(0.0691)	(0.0689)	(0.0688)
F	8.088	7.896	7.903	6.149	6.492
Students	1730	1730	1730	1730	1730
D. Complete ADN					
Win 1st Lottery	0.210	0.211	0.218	0.204	0.202
	(0.0739)	(0.0738)	(0.0733)	(0.0733)	(0.0740)
F	8.076	8.156	8.830	7.741	7.439
Students	1730	1730	1730	1730	1730
E. Any Health Award					
Win 1st Lottery	0.193	0.196	0.203	0.190	0.189
	(0.0739)	(0.0738)	(0.0733)	(0.0733)	(0.0739)
F	6.856	7.046	7.700	6.732	6.523
Students	1730	1730	1730	1730	1730
F. Any Award					
Win 1st Lottery	0.191	0.192	0.198	0.187	0.186
	(0.0738)	(0.0738)	(0.0735)	(0.0732)	(0.0734)
F	6.678	6.804	7.260	6.520	6.443
Students	1730	1730	1730	1730	1730
G. Transfer					
Win 1st Lottery	0.0332	0.0360	0.0308	0.0309	0.0324
	(0.0479)	(0.0471)	(0.0471)	(0.0471)	(0.0472)
F	0.480	0.585	0.428	0.430	0.473
Students	1730	1730	1730	1730	1730
Demographics		X	X	X	X
Academic			X	X	X
Labor Market				X	X
Financial Aid					X

Notes. Table shows estimates of the effect of a student being admitted to the Central College ADN after the first application. Sample consists of students who first applied between Spring 2005 and Spring 2009. Enrolled immediately is enrollment in the Central College ADN program the following semester. Ever enrolled is ever having enrolled in the Central College ADN program. Complete program is earn an ADN from Central College. Any Health Award is earn any associate's degree or certificate in a health field from any California community college. Any award is earn any associate's degree or certificate in any field from any California community college. Transfer is whether the student ever later enrolled in a four-year college. All regressions control for calendar year and application cohort. Demographics include age, gender, race; academic background includes prior GPA prior number of units; prior financial aid receipt includes receipt of Pell grants and tuition waivers; and prior labor market experience includes mean prior earnings, any prior employment in health. Standard errors clustered at the individual level.

Table A6: Sensitivity of Labor Market Estimates to Inclusion of Controls

	(1)	(2)	(3)	(4)	(5))
A. Log Earnings					
Enroll	0.362	0.355	0.385	0.361	0.367
	(0.143)	(0.143)	(0.145)	(0.149)	(0.148)
N	6920	6920	6920	6920	6920
Students	1730	1730	1730	1730	1730
First stage F	47.04	46.57	46.30	44.08	44.79
B. Earnings Levels					
Enroll	1623.3	1587.8	1796.8	1531.1	1596.9
	(1873.4)	(1889.5)	(1910.2)	(1946.2)	(1933.4)
N	6920	6920	6920	6920	6920
Students	1730	1730	1730	1730	1730
C. Earnings Levels, Top-Coded					
Enroll	2639.1	2633.6	2861.1	2615.5	2657.9
	(1742.1)	(1753.1)	(1773.8)	(1810.6)	(1797.2)
N	6920	6920	6920	6920	6920
Students	1730	1730	1730	1730	1730
D. Earnings Levels, Censored					
Enroll	3201.9	3198.9	3393.2	3206.4	3239.3
	(1467.1)	(1471.3)	(1489.4)	(1527.9)	(1519.1)
N	6695	6695	6695	6695	6695
Students	1706	1706	1706	1706	1706
E. Any Earnings					
Enroll	0.136	0.131	0.150	0.109	0.111
	(0.123)	(0.123)	(0.119)	(0.113)	(0.113)
N	6920	6920	6920	6920	6920
Students	1730	1730	1730	1730	1730
F. Employment in Health					
Enroll	0.213	0.213	0.221	0.189	0.195
	(0.0889)	(0.0908)	(0.0902)	(0.0850)	(0.0848)
N	4926	4926	4926	4926	4926
Students	1316	1316	1316	1316	1316
Demographics		X	X	X	X
Academic			X	X	X
Labor Market				X	X
Financial Aid					X

Notes. Table shows estimates of the effect of immediate enrollment in the Central College ADN program, instrumented with result of first application. Sample consists of students who first applied between Spring 2005 and Spring 2009. There are four quarters of data for each student, corresponding to quarters 18 through 21 following first application to the program. All regressions control for calendar year and application cohort. Demographics include age, gender, race; academic background includes prior GPA prior number of units; prior financial aid receipt includes receipt of Pell grants and tuition waivers; and prior labor market experience includes mean prior earnings, any prior employment in health. Standard errors clustered at the individual level.

A.2 Matching Between Lottery and Academic Data

A.2.1 Description of Match

Information on the result of each application to the Central College ADN lottery comes from a spreadsheet that includes student names, date of birth, gender, an identification number, the semester of the application, and the application result. There are 4,726 applications in the full Central College lottery file. All other information, such as course-taking, demographics, financial aid, and earnings, comes from the California Community College data system. There is no one-for-one crosswalk between the two datasets: the student identification number in the lottery data is used for internal Central College purposes and does not match the student identification numbers in the academic data. This appendix describes the process I implement to match between the Central College lottery data and the system-wide academic data.

In the first step of the process, I matched based on the sets of identifying information that were common to the two datasets. The lottery data has first and last name, date of birth, and gender. The academic data has date of birth and gender, but only the first three letters of each student's first and last names. Therefore, I used date of birth, gender, and the first three letters of first and last names to match. Two records in the application file were exact duplicates on these four identifying characteristics, so I drop both of them from the match. Likewise, four percent of all 26,559,940 students in the full academic file were not unique on these four variables, so I also drop these students as potential matches to the applications. I was able to match 3,473 (73 percent) of the 4,724 non-duplicate Central College lottery applicants to a unique student record in the statewide academic file.

To improve the match rate, I then did a second round of matching for the 1,251 still unmatched Central College applicants. This time, I limited the sample to 386,513 students in the larger academic file who had ever enrolled in a course at Central College and were not already matched to an applicant record. Of these students, 372,728 (96 percent) had unique values on the identifying information. This match yielded an additional 431 applicants matched to academic records. This means that, overall, I was able to match 3,904 of 4,724 applicants to academic records, for an overall match rate of 83 percent.

A.2.2 Match Diagnostics

The main concern with the match process is that it might be non-random. In other words, it may be the case that applicants I am able to match to the system-wide academic data are systematically different than students I am not able to match. This would be particularly problematic if matched students were more likely to be admitted to the program or to enroll in it. I regress a dummy for being admitted on the match outcome and find a coefficient of 0.004 (s.e.=0.018, p=0.82). A similar regression where I regress admission status on a stricter version of the match outcome (i.e. matched in the initial process, without accounting for college) yields a coefficient of -0.011 (s.e.=0.016, p=0.461). This suggests that the match does not seem to be causing differential selection on the lottery outcome.

Another potential concern is if a substantial number of applicants did not take their prerequisites at a community college in California. If this were the case, then admitted students would be more likely to be matched: some non-admitted students would never appear in the system-wide academic data, having never taken a California community college course. Students could potentially take their courses in for-profit in-state institutions, four-year colleges, out-of-state colleges, or in high school. In conversations with administrators, including the dean of Central College's health sciences department, I learned that out-of-state applications are rare and prerequisites from for-profits are also rarely accepted. Moreover, high school classes with college credit such as AP's are not accepted as fulling prerequisites. Empirically, I cannot observe whether unmatched applicants took their prerequisites out of state. However, I do find that 90 percent of students had taken community college coursework prior to applying, with no substantial differences between students who enrolled and those who did not.

As a final check, Table A7 shows the main results, limiting the sample to only applicants who were matched in the first type of matching. That is, it does not include students who were matched based on college. These results are quite similar to the main results from Table 3.

Table A7: IV Estimates Using Conservative Match Method

	(1) Log Earnings	(2) Any Earnings	(3) Health Industry
Enroll	0.309	0.0879	0.162
	(0.134)	(0.107)	(0.0709)
N	6060	6060	4308
Students	1515	1515	1150
First stage F	61.46	60.50	65.60
Demographics	X	X	X
Academic	X	X	X
Labor Market	X	X	X
Financial Aid	X	X	X

Notes. Table shows estimates of the effect of immediate enrollment in the Central College ADN program, instrumented with result of first application. Sample consists of students who first applied between Spring 2005 and Spring 2009, in the restricted matching approach described in this section. There are four quarters of data for each student, corresponding to quarters 18 through 21 following first application to the program. Health industry employment measures whether the individual had earnings in the two-digit NAICS code 62. Regressions control for calendar time, application cohort, demographics (age, gender, race), academic background (prior GPA, prior number of units), prior financial aid receipt (Pell grants, tuition waivers), and prior labor market experience (mean prior earnings, any prior employment in health). Standard errors clustered at the individual level.

A.3 Additional Robustness Exercises

A.3.1 Multiple Lotteries

The first two columns of Table A8 shows estimates where I utilize variation from the up to four random lotteries a student can enter. Each lottery a student applies to is a valid instrument for immediate enrollment. For example, among all students in their second lottery attempt, admission is random and also a valid instrument for enrollment. I estimate the following first-stage equation:

$$D_{icg} = \delta_0 + \delta_1 Admit_{ig} + X_i \gamma_3 + \pi_c + \theta_g + e_{icg} \tag{4}$$

where $Admit_{ig}$ is a dummy variable taking a value of one for a student winning their gth lottery, with $g \in 1,2,3,4$. Each student is represented up to four times in this setup. When $g = 1$ equation 4 is equivalent to equation 2. In other words, the coefficient δ_1 yields the average effect of winning a lottery on subsequent enrollment. I include lottery instance fixed effects θ_g and lottery term fixed effects π_c in order to separately identify the effect of each individual lottery pool. I cluster standard errors at the individual level.

A potential concern in leveraging all four potential lotteries a student enters is that there may be selection in who reapplies among the set of lottery losers. The local average treatment effect of each lottery would be different if, for example, first-time applicants were systematically different than third-time applicants. However, the cost of reapplying, which only involves clicking a button on a computer screen, is relatively low, and most students do reapply. This makes it less likely that using all four lotteries to estimate the effects will introduce bias. Appendix Table A9 shows that observable characteristics do not strongly predict reapplication among lottery losers, meaning that the pool of applicants is quite similar across lottery instances. Moreover, a test that the coefficients across the first four columns of the table are equal yields a χ^2 statistic with a p-value of 0.53.

Column 1 shows the resulting coefficient. There are more than four observations per student because each student can be represented with up to four applications, with four years of earnings data per application. The coefficient is slightly smaller than that using just the first lottery, but marginally so.

Since all applicants apply for a first time but not necessarily in subsequent lotteries, students with multiple applications are overly represented, so in column 2 I weight the regressions by $w_i = \frac{1}{max_i(k)}$ where k takes values one through four. This weighting approach does not make a substantial difference on the estimate.

A.3.2 One-Step Dynamic Regression

The third column of Table A8 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:

$$y_{ict} = \alpha + \sum_{\tau=0}^{\bar{\tau}} (\theta_\tau Admit_{i,t-\tau} + \phi_\tau Apply_{i,t-\tau}) + X_{itc}\Psi + \eta_c + \nu_t + u_{itc} \tag{5}$$

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 coefficient is similar in magnitude to the preferred estimate, but less precisely estimated.

A.3.3 Any Enrollment

The main lottery estimates instrument for immediate enrollment following application. Instrumenting for ever enrolling in the program is less clean than the preferred specification because some students who ultimately enroll are admitted through the non-random fifth application. This approach will lead to an additional group of non-compliers, those who were not admitted in their first lottery attempt, but were admitted in a future lottery. I estimate the effect of ever enrolling in the Central College ADN program in two ways. In the first I note that the instrumental variables estimate of immediate enrollment on any enrollment is 0.58 (0.070). If I scale my preferred estimate of the effect of immediate enrollment on log earnings at quarters 18-21, this gives me a coefficient of 0.63. I can also explicitly estimate the effect of any enrollment on earnings through equation 2. The fourth column of Table A8 shows this estimate, which is larger than my preferred estimate.

A.3.4 Individual Fixed Effects and Instrumental Variables

The final specification combines the individual fixed effects approach described in section 6 with the instrumental variables from the lottery. I estimate equation 3, but treat the timing of enrollment as the endogenous regressor to be instrumented with the lottery result. In this case, I run the following first stage for equation 3:

$$Enroll_{it} = \pi_i + \delta Post\_Admit_{it} + \Gamma Z_{it} + \lambda_t + \phi_i * t + e_{it} \tag{6}$$

where $Post\_Admit_{it}$ is a dummy variable that takes on a value of one in quarters after a student has been admitted to the Central College ADN program through a random lottery. Thus, the enrollment effects are identied by the interaction between the lottery and time relative to enrollment. These are comparable to the estimates of the effect of ever enrolling in the program, as there is no way to separate out immediate enrollment, which is the preferred estimate. The coefficient is in the final column of Table A8 and is similar to the coefficient in the previous column.

Table A8: Additional Robustness Exercises

	(1)	(2)	(3)	(4)	(5)
	Multiple Stacked Lotteries				
	Unweighted	Weighted	"One-Step"	Any Enrollment	IV & Fixed Effects
Ever Enroll	0.308	0.314	0.476	0.641	0.716
	(0.108)	(0.129)	(0.319)	(0.285)	(0.252)
N	11903	11903	38060	6920	47303
Students	1730	1730	1730	1730	1603

Notes. Sample consists of students who first applied between Spring 2005 and Spring 2009. Outcome is log earnings. Column 1 shows estimates of the effect of immediate enrollment in the Central College ADN program, instrumented with result of each of up to four applications a student submitted. There are four quarters of data for each application, corresponding to quarters 18 through 21 after the application date. Column 2 weights each observation by the inverse of the number of applications the student submitted. Column 3 shows estimates from equation 5. Column 4 limits the sample to the first application, but endogenous regressor is ever enrolling in the Central College ADN program, as opposed to immediate enrollment. Regressions in Columns 1 through 4 control for calendar time, application cohort, demographics (age, gender, race), academic background (prior GPA, prior number of units), prior financial aid receipt (Pell grants, tuition waivers), and prior labor market experience (mean prior earnings, any prior employment in health). Standard errors clustered at the individual level.

Table A9: Determinants of Lottery Reapplication Among Lottery Losers

	1st	2nd	3rd	4th	Any
Female	0.0335	0.0233	-0.0281	0.0645	0.0314
	(0.0266)	(0.0336)	(0.0366)	(0.0371)	(0.0176)
White	0.0262	0.00582	-0.00866	0.0300	0.0130
	(0.0279)	(0.0352)	(0.0388)	(0.0411)	(0.0186)
Hispanic	0.0755	-0.0103	0.00762	0.0239	0.0282
	(0.0292)	(0.0361)	(0.0401)	(0.0410)	(0.0191)
Asian	0.0899	-0.00494	0.113	-0.0452	0.0418
	(0.0411)	(0.0491)	(0.0540)	(0.0535)	(0.0262)
Age	0.00161	0.000743	0.00336	-0.000730	0.00140
	(0.00127)	(0.00157)	(0.00177)	(0.00181)	(0.000835)
GPA	0.0186	0.0108	0.0124	0.0210	0.0152
	(0.00856)	(0.0105)	(0.0116)	(0.0120)	(0.00558)
Enrolled in other district	-0.122	-0.147	-0.0465	-0.103	-0.124
	(0.0254)	(0.0332)	(0.0385)	(0.0408)	(0.0175)
Had BOG Waiver	0.0301	-0.0199	0.0179	0.0199	0.00672
	(0.0285)	(0.0358)	(0.0398)	(0.0401)	(0.0189)
Had Pell Grant	-0.0450	-0.00200	-0.0334	-0.0139	-0.0274
	(0.0300)	(0.0376)	(0.0417)	(0.0422)	(0.0198)
Employed >1 Quarter	0.0228	-0.00660	0.0633	0.0132	0.0250
	(0.0290)	(0.0362)	(0.0399)	(0.0422)	(0.0193)
Share Persist	0.837	0.777	0.811	0.844	0.793
N	1266	1052	779	623	3305
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, for lotteries between Spring 2005 and Spring 2009. Regressions control for lottery cohort. GPA measures grades in prerequisites prior to application. Enrollment at other district defined as ever having taken a course at a community college outside Central College's district. BOG waiver is a full tuition waiver. Calgrant is state-specific financial aid. Employment defined as nonzero quarterly earnings.