

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

Michel Grosz

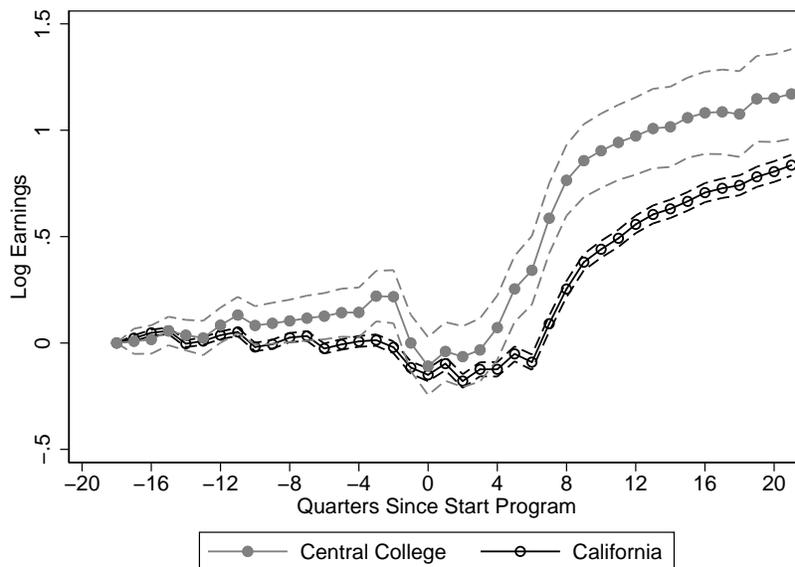
ONLINE APPENDIX

April 22, 2019

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

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

Table A2: Balance, Joint Regressions

	(1)	(2)
	First Lottery	All Lotteries
Female	0.014 (0.010)	0.008 (0.006)
White	0.008 (0.011)	0.004 (0.007)
Hispanic	0.013 (0.011)	0.000 (0.007)
Asian	-0.009 (0.017)	-0.004 (0.009)
Age	0.000 (0.001)	0.000 (0.000)
GPA	-0.004 (0.003)	-0.002 (0.002)
Enrolled in other district	-0.012 (0.010)	-0.006 (0.006)
Had BOG Waiver	0.003 (0.011)	-0.000 (0.007)
Had Pell Grant	-0.009 (0.013)	-0.009 (0.008)
Calgrant	0.000 (0.014)	0.007 (0.008)
Had Loans	-0.009 (0.018)	0.003 (0.011)
Employed > 1 Quarter	0.002 (0.017)	-0.013 (0.010)
Quarters Employed	-0.002 (0.002)	0.001 (0.001)
Employed > 8 Quarters	0.022 (0.021)	-0.008 (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 (0.010)	0.008 (0.006)
Industry is Administrative	0.001 (0.012)	-0.002 (0.007)
Industry is Education	0.001 (0.014)	0.005 (0.008)
Industry is Food Service	-0.005 (0.012)	-0.005 (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. 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)
<u>A. Applications</u>					
Win 1st Lottery	-1.561 (0.0905)	-1.559 (0.0903)	-1.573 (0.0905)	-1.528 (0.0911)	-1.526 (0.0918)
F	297.6	297.7	302.0	281.7	276.3
Students	1730	1730	1730	1730	1730
<u>B. Enroll Immediately</u>					
Win 1st Lottery	0.505 (0.0680)	0.504 (0.0682)	0.504 (0.0683)	0.486 (0.0680)	0.485 (0.0674)
F	55.27	54.59	54.49	50.93	51.70
Students	1730	1730	1730	1730	1730
<u>C. Ever Enroll</u>					
Win 1st Lottery	0.196 (0.0690)	0.194 (0.0690)	0.194 (0.0691)	0.171 (0.0689)	0.175 (0.0688)
F	8.088	7.896	7.903	6.149	6.492
Students	1730	1730	1730	1730	1730
<u>D. Complete ADN</u>					
Win 1st Lottery	0.210 (0.0739)	0.211 (0.0738)	0.218 (0.0733)	0.204 (0.0733)	0.202 (0.0740)
F	8.076	8.156	8.830	7.741	7.439
Students	1730	1730	1730	1730	1730
<u>E. Any Health Award</u>					
Win 1st Lottery	0.193 (0.0739)	0.196 (0.0738)	0.203 (0.0733)	0.190 (0.0733)	0.189 (0.0739)
F	6.856	7.046	7.700	6.732	6.523
Students	1730	1730	1730	1730	1730
<u>F. Any Award</u>					
Win 1st Lottery	0.191 (0.0738)	0.192 (0.0738)	0.198 (0.0735)	0.187 (0.0732)	0.186 (0.0734)
F	6.678	6.804	7.260	6.520	6.443
Students	1730	1730	1730	1730	1730
<u>G. Transfer</u>					
Win 1st Lottery	0.0332 (0.0479)	0.0360 (0.0471)	0.0308 (0.0471)	0.0309 (0.0471)	0.0324 (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)
<u>A. Log Earnings</u>					
Enroll	0.362 (0.143)	0.355 (0.143)	0.385 (0.145)	0.361 (0.149)	0.367 (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
<u>B. Earnings Levels</u>					
Enroll	1623.3 (1873.4)	1587.8 (1889.5)	1796.8 (1910.2)	1531.1 (1946.2)	1596.9 (1933.4)
N	6920	6920	6920	6920	6920
Students	1730	1730	1730	1730	1730
<u>C. Earnings Levels, Winsorized</u>					
Enroll	2639.1 (1742.1)	2633.6 (1753.1)	2861.1 (1773.8)	2615.5 (1810.6)	2657.9 (1797.2)
N	6920	6920	6920	6920	6920
Students	1730	1730	1730	1730	1730
<u>D. Earnings Levels, Censored</u>					
Enroll	3201.9 (1467.1)	3198.9 (1471.3)	3393.2 (1489.4)	3206.4 (1527.9)	3239.3 (1519.1)
N	6695	6695	6695	6695	6695
Students	1706	1706	1706	1706	1706
<u>E. Any Earnings</u>					
Enroll	0.136 (0.123)	0.131 (0.123)	0.150 (0.119)	0.109 (0.113)	0.111 (0.113)
N	6920	6920	6920	6920	6920
Students	1730	1730	1730	1730	1730
<u>F. Employment in Health</u>					
Enroll	0.213 (0.0889)	0.213 (0.0908)	0.221 (0.0902)	0.189 (0.0850)	0.195 (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.

Table A7: Sensitivity of Earnings Levels Calculations

	(1) Winsorize	(2) Censor
<u>A. Quarterly Earnings Above \$30,000 Earnings</u>		
Enroll	3239.3 (1519.1)	2657.9 (1797.2)
N	6695	6920
Students	1706	1730
<u>B. Quarterly Earnings Above \$40,000</u>		
Enroll	3353.2 (1923.8)	2183.3 (1916.9)
N	6838	6920
Students	1719	1730
<u>C. Quarterly Earnings Below \$100</u>		
Enroll	1596.9 (1933.4)	1596.3 (1933.4)
N	6920	6920
Students	1730	1730
<u>D. Quarterly Earnings Above \$30,000 or Below \$100</u>		
Enroll	3380.6 (1692.2)	2657.4 (1797.2)
N	4858	6920
Students	1333	1730
<u>E. Quarterly Earnings Above \$40,000 or Below \$100</u>		
Enroll	3370.8 (2150.2)	2182.7 (1916.9)
N	5001	6920
Students	1347	1730
Demographics	X	X
Academic	X	X
Labor Market	X	X
Financial Aid	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. 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. Winsorized data in column 1 code earnings above and/or below the stated amount as the stated amount. Censored data in column 2 drop observations above and/or below the stated amount. Standard errors clustered at the individual level.

Table A8: Individual Fixed Effects Estimates at All California ADN Programs, Not Controlling for Concurrent Enrollment

	(1) Pre-Post	(2) Individual Fixed Effects	(3)
<u>A. Central College</u>			
Start Program	0.659 (0.0305)	0.509 (0.0241)	0.444 (0.0222)
N	49710	49710	49710
Students	1535	1535	1535
<u>B. Colleges with Lottery Admissions</u>			
Start Program	0.344 (0.0168)	0.181 (0.0124)	0.112 (0.0113)
N	234096	234096	234096
Students	8424	8424	8424
<u>C. Colleges with Non-Lottery Admissions</u>			
Start Program	0.367 (0.00836)	0.215 (0.00648)	0.134 (0.00572)
N	981842	981842	981842
Students	36292	36292	36292
<u>D. All Colleges</u>			
Start Program	0.363 (0.00748)	0.208 (0.00575)	0.130 (0.00511)
N	1215938	1215938	1215938
Students	44716	44716	44716
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, and calendar time effects. Column 2 adds individual fixed effects, and Column 3 adds individual-specific linear time trends. Panel A is for students who ever enrolled in the Central College ADN program, Panel B is for students who ever enrolled in an ADN program with lottery-based admissions, and Panel C is for students who ever enrolled in a program that did not have lottery-based admissions. Panel D includes all programs. Standard errors clustered at the individual level.

Table A9: Internal Rate of Return Calculations for Central College up to 21 Quarters

	(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	60	66	75	81	84	94
Fixed Effects	31	34	44	36	41	49

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 21 quarters and are assumed to be zero afterwards. 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.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, Appendix Table A10 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 A10: IV Estimates Using Conservative Match Method

	(1)	(2)	(3)
	Log Earnings	Any Earnings	Health Industry
Enroll	0.309 (0.134)	0.0879 (0.107)	0.162 (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 Appendix Table A11 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:

$$D_{icg} = \delta_0 + \delta_1 Admit_{ig} + X_i \gamma_3 + \pi_c + \theta_g + e_{icg} \quad (1)$$

where $Admit_{ig}$ is a dummy variable taking a value of one for a student winning their g th lottery, with $g \in 1, 2, 3, 4$. Each student is represented up to four times in this setup. When $g = 1$ equation 1 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 A12 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 Appendix Table A11 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:

$$y_{ict} = \alpha + \sum_{\tau=0}^{\bar{\tau}} (\theta_{\tau} Admit_{i,t-\tau} + \phi_{\tau} Apply_{i,t-\tau}) + X_{itc} \Psi + \eta_c + \nu_t + u_{itc} \quad (2)$$

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 Appendix Table A11 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:

$$Enroll_{it} = \pi_i + \delta Post_Admit_{it} + \Gamma Z_{it} + \lambda_t + \phi_i * t + e_{it} \quad (3)$$

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 identified 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 Appendix Table A11 and is similar to the coefficient in the previous column.

Table A11: Additional Robustness Exercises

	(1)	(2)	(3)	(4)	(5)
	Multiple Stacked Lotteries				
	Unweighted	Weighted	“One-Step”	Any Enrollment	IV & Fixed Effects
Ever Enroll	0.308 (0.108)	0.314 (0.129)	0.476 (0.319)	0.641 (0.285)	0.716 (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 A12: Determinants of Lottery Reapplication Among Lottery Losers

	1st	2nd	3rd	4th	Any
Female	0.0335 (0.0266)	0.0233 (0.0336)	-0.0281 (0.0366)	0.0645 (0.0371)	0.0314 (0.0176)
White	0.0262 (0.0279)	0.00582 (0.0352)	-0.00866 (0.0388)	0.0300 (0.0411)	0.0130 (0.0186)
Hispanic	0.0755 (0.0292)	-0.0103 (0.0361)	0.00762 (0.0401)	0.0239 (0.0410)	0.0282 (0.0191)
Asian	0.0899 (0.0411)	-0.00494 (0.0491)	0.113 (0.0540)	-0.0452 (0.0535)	0.0418 (0.0262)
Age	0.00161 (0.00127)	0.000743 (0.00157)	0.00336 (0.00177)	-0.000730 (0.00181)	0.00140 (0.000835)
GPA	0.0186 (0.00856)	0.0108 (0.0105)	0.0124 (0.0116)	0.0210 (0.0120)	0.0152 (0.00558)
Enrolled in other district	-0.122 (0.0254)	-0.147 (0.0332)	-0.0465 (0.0385)	-0.103 (0.0408)	-0.124 (0.0175)
Had BOG Waiver	0.0301 (0.0285)	-0.0199 (0.0358)	0.0179 (0.0398)	0.0199 (0.0401)	0.00672 (0.0189)
Had Pell Grant	-0.0450 (0.0300)	-0.00200 (0.0376)	-0.0334 (0.0417)	-0.0139 (0.0422)	-0.0274 (0.0198)
Employed >1 Quarter	0.0228 (0.0290)	-0.00660 (0.0362)	0.0633 (0.0399)	0.0132 (0.0422)	0.0250 (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.

References

- 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.
- 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.