BROWN
Orlando Bravo Center
for Economic Research

Reducing Frictions in College Admissions: Evidence from the Common Application∗

Bravo Working Paper # 2020-001

Brian Knight† Nathan Schiff‡

Abstract: College admissions is decentralized, creating frictions that limit student choice. We study the Common Application (CA) platform, under which students submit a single application to member schools, reducing frictions and increasing student choice. Joining the CA increases the number of applications received by schools, reflecting a reduction in frictions. Joining also reduces the yield on accepted students, reflecting increased choice. The CA also increases out-of-state enrollment, especially from other CA states, consistent with network effects. Finally, joining the CA increases freshmen SAT scores. Given that CA members tend to be selective, the CA has contributed to stratification in higher education.

———————————————————————

Reducing Frictions in College Admissions: Evidence from the Common Application*

Brian Knight[†]   Nathan Schiff[‡]

August 30, 2019

Abstract

College admissions is decentralized, creating frictions that limit student choice. We study the Common Application (CA) platform, under which students submit a single application to member schools, reducing frictions and increasing student choice. Joining the CA increases the number of applications received by schools, reflecting a reduction in frictions. Joining also reduces the yield on accepted students, reflecting increased choice. The CA also increases out-of-state enrollment, especially from other CA states, consistent with network effects. Finally, joining the CA increases freshmen SAT scores. Given that CA members tend to be selective, the CA has contributed to stratification in higher education.

1 Introduction

College admissions in the U.S. is decentralized. After students choose the set of schools to which to apply, schools decide whether or not to make admissions offers to these applicants. Students then select from their choice set, the institutions to which they were offered admissions. Historically, students have completed separate applications for each school, entailing substantial time costs, or frictions.[1]

These frictions in college admissions may limit student applications and ultimately reduce student choice. Since students tend to apply to colleges close to home, these frictions may also lead to a less integrated market from a geographic perspective. By reducing frictions, students might apply more nationally and ultimately attend a school further from home. At the same time, substantial frictions, by reducing student choice sets, might also lead to less sorting across schools according to ability, with schools less stratified than under a more centralized college admissions process. By reducing frictions, high ability students might be disproportionately induced to apply to selective out of state colleges, leading to more stratification.

In this paper, we investigate these issues in the context of the Common Application (CA), a platform allowing students to complete one application for all member institutions. The CA, by allowing students to submit a single application to multiple institutions, may simplify and significantly reduce time costs associated with college admissions. Given this reduction in frictions, the CA might increase student choice, in turn leading to increased geographic integration but also increased stratification according to ability across institutions.

Given this motivation, we ask a series of research questions in this paper. Has the CA reduced frictions, resulting in more college applications at member institutions and increased student choice? Has the CA led to a more geographically integrated market, with more students attending CA institutions far from home? If so, by increasing student choice and integrating the market geographically, has the CA contributed to stratification, a widening of the selectivity gap between more selective and less selective institutions?

We first present a simple model of the college admissions process that includes student application decisions, university admissions offers, and student acceptances of these offers. We then

---

[1]These time-based frictions are in addition to any information frictions facing students as a result of decentralized college admissions.

consider the impact of the CA, modeled as a reduction in the costs of applying to a second CA school. The model predicts that joining the CA leads to an increase in applications due to a reduction in frictions. Likewise, schools joining the CA experience a reduction in their yield and admit more applicants in order to meet their capacity. Thus, the CA increases student choice. The CA also contributes towards geographic integration, with an increase in out-of-state students from other states with CA institutions. Finally, the introduction of the CA increases stratification, with high ability students disproportionately sorting into CA schools.

We then test these predictions using panel data from the College Board (1990 to 2016). We estimate two-way fixed effects models, comparing outcomes for schools before and after joining the CA, and also provide event studies, investigating the timing of any effects associated with entry into the CA. We find that the CA increases applications, consistent with a reduction in frictions, and reduces yield, consistent with enhanced student choice. Schools respond to this reduced yield by admitting more students. We also find that the CA has accelerated geographic integration. In particular, there is strong evidence that entry into the CA is associated with an increase in the fraction of out-of-state students, especially from other states with significant CA penetration, consistent with network effects. Finally, we provide some evidence that entry into the CA is associated with an increase in SAT scores. Given that CA entry has been more common among selective institutions, the CA may have contributed towards increased stratification, a widening of the gap between more selective and less selective institutions.

The paper proceeds as follows. Following the discussion of the relevant academic literature, we offer more details on the CA. This is followed by the presentation of the theoretical model of the college admissions process. We then describe the data, empirical approach, and our key empirical results. The final section concludes.

2   Related Literature

This paper is related to a literature on frictions in college admissions and interventions designed to reduce such frictions.[2] In the context of financial aid policy, Bettinger et al. (2012), for example,

---

[2]Given that the CA reduces the time cost associated with submitting applications to multiple CA schools, we interpret the CA as reducing time-based frictions associated with applications. Yet it is also possible that the CA provides information to applicants through the platform. In this sense, our paper is related to a literature on information frictions

show that assistance filling out the FAFSA, a complex financial aid application, increases aid receipt, college attendance, and persistence. More specifically on the effects of the CA, Smith (2013) studies the effect of the number of applications on enrollment probabilities, using variation induced by the CA, and finds that increasing the number of applications significantly increases enrollment probabilities. Smith et al. (2015) analyze various frictions in the application process, finding some evidence that the CA increases applications.[3] Perhaps the paper most closely related to this one is Liu et al. (2007), who use panel data from the College Board to study how CA membership affects admissions outcomes as well as the composition of enrollees. While their paper has a stronger focus on student characteristics, such as race and SAT scores, we are more interested in the role the CA has played in terms of student choice, geographic integration, and stratification.

Another literature has examined how college admissions, and in particular the recent increase in the number of applications per student, has changed both university behavior and student behavior. Bound et al. (2009) document increases over time in the number of students applying to college, increases over time in applications per student, and reductions in acceptance rates at selective institutions. Blair and Smetters (2019) investigate why colleges, especially elite ones, haven't responded to this increase in applications by expanding capacity. They argue that colleges compete on prestige, providing incentives for each institution to reduce admissions rates. Avery and Levin (2010) document that early applicants are more likely to be admitted, when compared to the regular applicant pool and argue that this finding is consistent with an early application serving as a signal of student enthusiasm for attending the institution. Avery et al. (2013) argue that standard methods of ranking colleges, such as the U.S. News and World Report, provide incentives for institutions to manipulate admissions decisions to reduce acceptance rates and to increase yield, the fraction of admitted students who accept their offers.

This paper is also related to a literature on geographic integration in higher education. His-

---

in college admissions, Hoxby and Avery (2013) document substantial information frictions in the US, with many low-income but high-ability students not applying to selective colleges, despite generous financial aid at these institutions. In a randomized control trial, Hoxby and Turner (2014) and Hoxby and Turner (2015) provide these high-ability but low-income students with information about college admissions and financial aid. This intervention increases applications to, and ultimately attendance at, selective institutions. Gurantz et al. (2019) provided information about selective colleges to 785,000 low-income and middle-income students but find little evidence of changes in college attendance patterns. Bird et al. (2019) conduct a large-scale field experiment using the CA platform and document that providing information about financial aid does not increase college attendance patterns, and similar results are found using a state-sponsored application portal.

[3]Fees and essay requirements, by contrast, decrease applications.

4

torically, the US market for higher education was highly localized, with most students attending universities close to their residence.[4] In 1949, over 90 percent of students attended in-state universities (Hoxby, 2000). During the subsequent decades, the market for higher education became more national, leading to both higher tuition and greater student sorting (Hoxby, 1997, 2000, 2009). Despite this shift, the US market may still be considered local in nature, with roughly 80 percent of first-time students enrolled in a college in their home state.[5] Knight and Schiff (2019) and Cohodes and Goodman (2014) argue that financial incentives, in the form of in-state tuition discounts and financial aid for in-state institutions, have contributed to this high rate of attendance at in-state public institutions. New technologies, such as the CA, may disrupt these patterns. Deming, Lovenheim, and Patterson (2016), for example, argue that less-selective colleges are particularly localized and study the effects of competition from online degree programs in these markets, finding that online competition reduces enrollment at private non-selective institutions.

3   Background

As noted above, college admissions has historically been decentralized in the U.S.[6] As a private effort towards greater centralization, the CA began as a consortium of 15 colleges in 1975. It grew rapidly thereafter, with increases in membership in every year since 1975 and a significant acceleration of membership starting around 2000 (Figure 1). It currently includes over 700 institutions, which, taken together, receive approximately 4 million applications from 1 million students annually.[7]

The CA was founded by a small set of liberal arts colleges but has since expanded to a wide range of public and private institutions, especially more selective institutions. In particular, as shown in Figure 2, membership among the top 50 liberal arts colleges was already very high, over 80 percent, in 1990, the beginning of our analysis, and was universal in this group by the late 1990s.

---

[4]Outside of the US, Gibbons and Vignoles (2012) study student college decisions in England and find that conditional on going to college, distance is the most important factor in explaining choice of college.

[5]As of Fall 2014. See Table 309.20 of the "Digest of Education Statistics," National Center for Education Statistics (NCES), https://nces.ed.gov/programs/digest/d16/tables/dt16_309.20.asp.

[6]This decentralized process is somewhat unique to the U.S. Many countries, such as China, the UK, and Chile, have more centralized admissions procedures, with students typically submitting a single application with ranked preferences and then admitted to at most one institution.

[7]From https://www.commonapp.org/about-us, accessed April 2018.
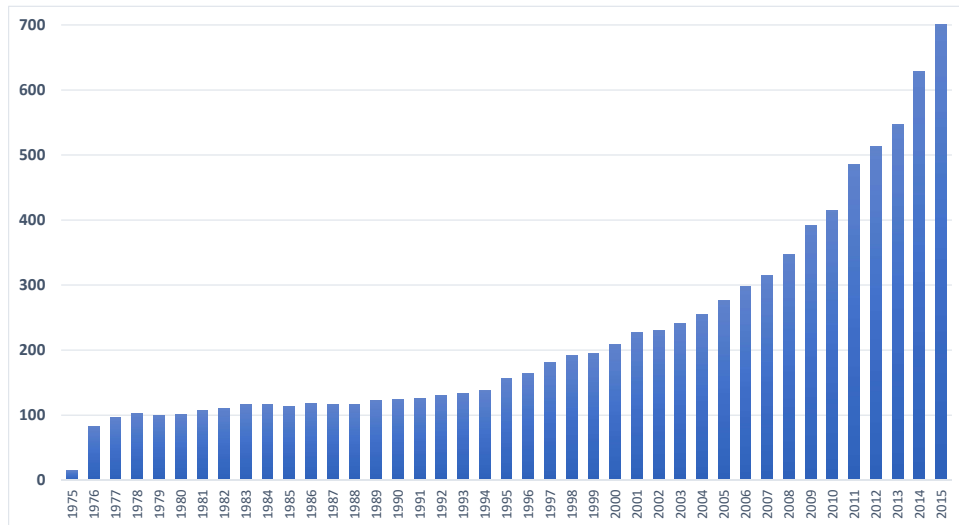
Figure 1: Common Application Membership by Year

During our sample period, membership among top 50 private institutions increased rapidly, from under 40 percent in 1990 to roughly 90 percent by 2016. Membership among less selective liberal arts colleges and other private institutions also increased during our sample period but remained below 50 percent in 2016. The CA was originally closed to public institutions but that ban was lifted in 2002, leading to a rapid increase in membership among the top 50 public institutions. Less selective public institutions, by contrast, joined at a slower rate, with membership still below 20 percent by the end of our sample period.

In addition to this rapid entry overall, the CA has also become more diverse from a geographic perspective. That is, the CA started in the Northeast but is now accepted by colleges in many different states. In particular, Figure 3 plots the locations of CA members in 1986 and 2014, documenting a significantly wider geographic distribution in the latter year, with significant new penetration in states such as California, Oregon, Colorado, Indiana, and Florida. Given this substantial adoption over time and a diverse set of members at current, it is natural that the CA may have led to significant changes in college admissions.
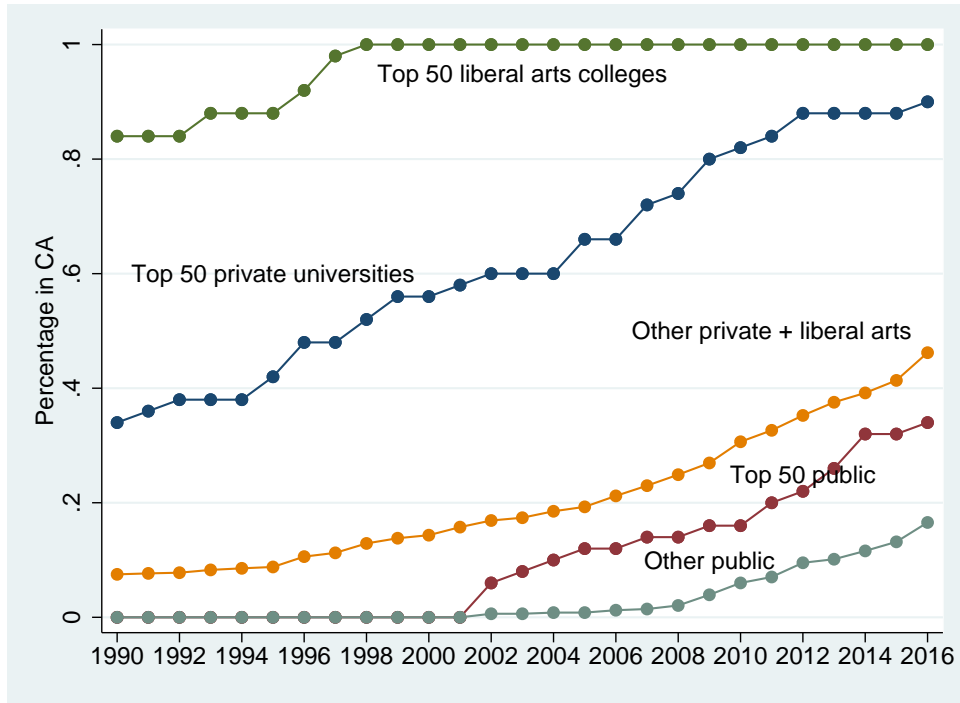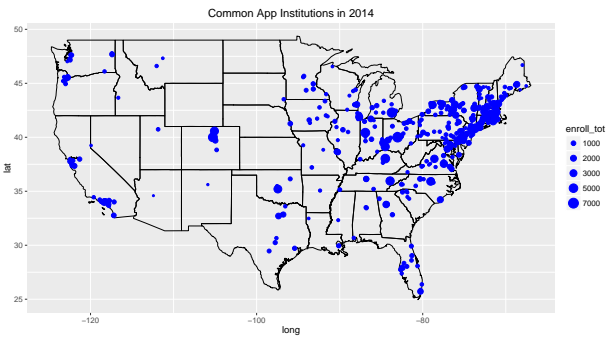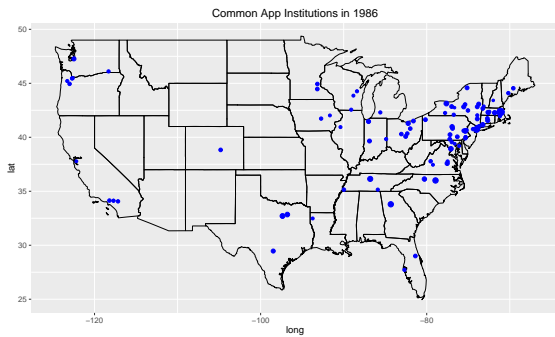
Figure 2: Common Application Membership Rates by Year and Type

Figure 3: Common Application Membership by State

(a) CA in 1986

(b) CA in 2014

4   Theoretical Model

Our model of college admissions has three distinct stages. First, students decide which colleges to apply to. Second, colleges make admissions offers. Finally, given these offers, students decide which college to attend. In the context of this model, we consider the CA, which reduces the costs of applying to a second CA institution.

4.1 Setup

There are two colleges: $c = 1$ and $c = 2$. Students receive a payoff $V_c = U_c + \varepsilon_c$ from attending college c. The first term U_c is known prior to applying and may include residency status, reflecting a preference for in-state colleges.[8] The second term, which we assume is distributed type-1 extreme value, is revealed after applying but before choosing a college. Not attending college is also an option, with a payoff normalized to zero. There are two types of students, each representing one-half of the population. The first type prefers college 1 over college 2, the second prefers college 2 over college 1, and we assume symmetry in the expected utility gains from attending the first choice college.[9] When applying, students pay an application fee of F to the first college and a potentially lower fee ($f \leq F$) when applying to a second college. The CA can be interpreted as a reduction in f since the application developed for the first choice college can also be used for the second college.

On the supply side, colleges have a fixed capacity and can serve a fraction $0 < \kappa < 1$ of first-choice students; thus, 2κ represents overall capacity across the two institutions. These capacities, along with student applications, determine admissions rates Q_c.

4.2 Equilibrium

In the third stage, following student application decisions and university admissions offers, there are four possible student choice sets: choosing between both colleges, college 1 or 2 only, or only the outside option. The value of being accepted to both colleges equals $C_{12} = ln[exp(U_1) + exp(U_2) + 1]$.[10] Likewise, we denote $Y_{12} = exp(U_1)/[1 + exp(U_1) + exp(U_2)]$ as the yield for the first-choice college and y_{12} as the yield for the second-choice college for students accepted to both. The value of being accepted to only the first choice equals $C_1 = ln[exp(U_1) + 1]$, and the corresponding yield is $Y_1 = exp(U_1)/[1 + exp(U_1)]$. Similar expressions apply to the value of being accepted to only the second choice (C_2) and the corresponding yield is denoted by Y_2, with $Y_1 > Y_2$.

---

[8]This could be due to either a preference for proximity or lower in-state tuition at public institutions

[9]That is, $U_1 - U_2 = \delta > 0$ for the first type and $U_2 - U_1 = \delta > 0$ for the second type.

[10]This follow the standard formula for consumer surplus in a logit model. Similar derivations apply for type 2 students, given the symmetry of the model.

In the second stage, taking yield as given, schools set their admission rates in order to satisfy capacity. We focus on an equilibrium in which all students apply to their first choice and a fraction b of students also apply to their second-choice college. Then, admissions rates for college 1 equate the number of student acceptances of university admissions offers to university capacity:

$$Q_1 \underbrace{\left[(1-b)Y_1 + bQ_2 Y_{12} + b(1-Q_2)Y_1 \right]}_{\text{yield on first-choice students}} + \underbrace{Q_1 b \left[Q_2 y_{12} + (1-Q_2)Y_2 \right]}_{\text{yield on second-choice students}} = \kappa \tag{1}$$

Among first-choice students, a fraction $1-b$ apply to only their first choice, with yield of Y_1, and a fraction b also apply to their second choice. In the latter case, a fraction Q_2 are also admitted to their second choice, with yield of Y_{12}, and a fraction $1-Q_2$ are denied admission to their second choice, with yield for college 1 thus equal to Y_1. The second term represents yield on second-choice students, with a fraction b applying to both colleges. Among these, a fraction Q_2 are also admitted to their first choice and yield thus equals y_{12}. The remaining fraction $(1-Q_2)$ are not admitted to their first choice and yield on these students equals Y_2.

Then, working backwards to the first stage, applying to both colleges yields a value of $A_{12} = Q_1 Q_2 C_{12} + Q_1(1-Q_2)C_1 + (1-Q_1)Q_2 C_2 - F - f$ for type 1 students. That is, students are accepted to both colleges with probability $Q_1 Q_2$, college 1 only with probability $Q_1(1-Q_2)$, college 2 only with probability $(1-Q_1)Q_2$ and face application costs of $F + f$. For type 1 students applying to only college 1, the value equals $A_1 = Q_1 C_1 - F$. In equilibrium, the fraction of students applying to both colleges increases until the benefits from a second application equal the costs of a second application $(A_{12} - A_1 = f)$. This can be written as:

$$\underbrace{Q_1 Q_2 (C_{12} - C_1)}_{\text{option-value}} + \underbrace{(1-Q_1)Q_2 C_2}_{\text{safety-value}} = f \tag{2}$$

The option value from a second application represents the benefit of being able to attend college 2 when the student has been accepted to both colleges, which occurs with probability $Q_1 Q_2$. This captures the idea that students may learn that college 2 is actually preferred to college 1 throughout the admissions process, following the realization of ε_1 and ε_2. The safety value from a second application represents the benefit of being able to choose college 2 if not admitted to college 1, and this event occurs with probability $(1-Q_1)Q_2$.

9

In a symmetric equilibrium, equations 1 and 2 represent two equations with two unknowns: the fraction of students applying to both colleges (b) and the university admissions rate (Q). For simplicity, we assume here a solution that is unique and interior in both b and Q, and the Appendix details the associated assumptions.[11] We provide a graphical interpretation of this equilibrium in Figure 1 and provide closed form solutions for b and Q in the Appendix. The student indifference constraint is the admissions rate at which students are indifferent between applying to only one college and applying to both colleges, as determined by equation 2. The college feasibility constraint is determined by equation 1 and requires, given capacity, a reduction in admissions rates given an increase in the number of applications received. As shown, the initial equilibrium, in admissions rate Q^* and applications rate b^*, occurs at the intersection between the student indifference condition and the college feasibility constraint.
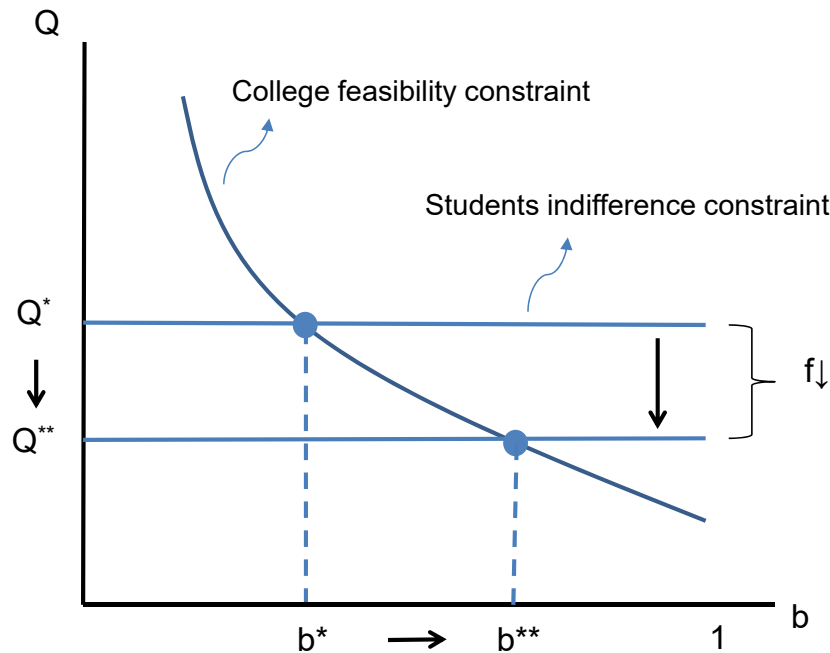

Figure 4: Theoretical Effects of Common Application

---

[11]These require that application costs are small, relative to the benefits of a larger choice set and that college capacity is neither too small nor too large

4.3   Effect of the Common Application

We next consider a comparative static based upon a reduction in the cost of applying to a second college.

Proposition: Consider the introduction of the Common Application, with a marginal reduction in f from f=F. There are four effects: 1) application activity (b) increases, 2) colleges become more selective (a reduction in Q), 3) yield falls and, despite increasing selectivity, universities accept a larger number of applications, 4) students are more likely to attend out-of-state institutions.

We provide graphical interpretation of these results in the context of Figure 4 and refer readers to the Appendix for a Proof. As shown, given a reduction in F, students are no longer indifferent between applying to one institution and applying to both institutions at the initial admissions rate.[12] Given this, admissions rates fall from Q^* to Q^{**} in order to restore student indifference, and the fraction of students applying to both colleges increases accordingly, from b^* to b^{**}. Despite the reduction in admissions rates, universities make more admissions offers in total. This results from the reduction in yield on admitted students, who tend to have larger choice sets. Finally, interpreting first-choice colleges as in-state and second-choice colleges as out-of-state, the CA leads to geographic integration, resulting from more students applying to out-of-state institutions.

4.4   Extensions

We consider two extensions of the model, with details in the Appendix. The first has three colleges, two in the CA and one outside of the CA, allowing for consideration of network effects. Relative to a baseline without the CA, the introduction of the CA (with colleges 1 and 2 as members) leads to an increase in students applying to both colleges 1 and 2 but a reduction in second application activity for students whose ex-ante first choice is college 3, which is outside of the CA. Given this, there are more students from state 1 attending college 2 and, likewise, more students from state 2 attending college 1. Correspondingly, there is a reduction in students from state 3 attending colleges 1 or 2. Thus, the CA leads to network effects, with more migration by students between states connected by the CA.

---

[12]In the absence of any adjustment in admissions rates (Q), the CA in fact drives student application behavior to a corner solution, with all students applying to both colleges.

The second extension also has three colleges, two in the CA and one outside of the CA, but also considers heterogeneous students: low-ability and high-ability. Colleges want to attract as many high-ability students as possible and admit them with probability one. Low-ability students are then admitted at an endogenous admissions rate to fill remaining capacity. In equilibrium, high ability students disproportionately attend CA schools and low-ability students disproportionately attend schools outside of the CA, resulting from increased application activity among high ability students at CA schools.[13] Thus, the CA increases stratification, with high ability students disproportionately attending CA schools.

4.5 Summary

The introduction of the CA increases applications and reduces admissions rates. Likewise, the CA reduces yield and increases the number of admitted students. There is an increase in out-of-state attendance, especially from other CA states. Finally, the CA leads to stratification, with high ability students disproportionately attending CA institutions.

5 Empirical Analysis

To test the model predictions, we next consider how admissions outcomes change when an institution joins the CA. We first describe the data and our empirical approach. We then present the key empirical results.

5.1 Data

Our analysis uses two sources of data on universities, the College Board's Annual Survey of Colleges and the Integrated Postsecondary Education Data System (IPEDS), combined with the year in which each university became a member of the CA.[14] We use the College Board data for most of our analysis because it covers a longer time span (1990-2016) than similar data from the IPEDS.

---

[13]Although there is no safety value motivation, given that these high ability students are admitted to their first choice with certainty, there remains an option value from additional applications.

[14]The entry year for CA members was provided to us by The Common Application organization.

However, the IPEDS data has information on state-to-state student migration conducted biennially from 1986 to 2014, and we use these data to study network effects.

5.2 Specifications

We use two estimating equations in most of our analysis: a two-way fixed effects specification (TWFE) and an event study specification. The TWFE model relates an outcome y_{ct} (e.g., applications) to an indicator for CA membership in year t (CA_t) as follows:

$$ln(y_{ct}) = \beta CA_{ct} + \mu_c + \mu_t + \varepsilon_{ct} \tag{3}$$

where c indexes colleges, μ_c is a college fixed effect, and μ_t is a year fixed effect. Then, given the log specification, the parameter β captures the percent change in outcomes when joining the CA, after controlling for time effects and university effects.

Our event study specification is designed to measure the timing of any effects of entry and is given by:

$$ln(y_{ct}) = \underbrace{\sum_{k=-K}^{k=-1} \beta_{t+k} 1(t - J_c = k)}_{\text{pre-entry}} + \underbrace{\sum_{k=0}^{k=K} \beta_{t+k} 1(t - J_c = k)}_{\text{post-entry}} + \mu_c + \mu_t + \varepsilon_{ct} \tag{4}$$

where J_c is the year college c joined the Common App and $1(t - J_c = k)$ indicates that college c joined the CA k years ago (or will join in the future when k is negative). We normalize β_{t-1} to zero and hence the key parameter β_{t+k} captures the effect of joining the CA at time t on outcomes at time $t + k$, relative to outcomes at time $t - 1$.

In addition, we present structural break analyses in the Appendix. These analyses allow for a more formal determination of whether breaks in outcomes occur during the year of CA entry or in other years surrounding CA entry. That is, we estimate our TWFE model but shift the join year into the past and into the future. We report F-statistics from the test based upon the null hypothesis that $\beta = 0$.

13

5.3 Applications and Yield

We begin our investigation of the effect of joining the CA by estimating equation 3 using the College Board data.[15] As shown in the first column of Table 1, applications are 12 percent higher after a college joins the CA, relative to the period before they joined the CA. This economically and statistically significant result is consistent with the CA reducing frictions in college admissions via a reduction in application costs.

To investigate the role of pre-trends and to consider any dynamic effects of joining the CA, we next present results from the corresponding event study specification.[16] As shown in Figure 5, which includes 95 percent confidence intervals, there is a slight downward trend in applications just before a school joins the CA. After joining the CA, by contrast, there is a discontinuous 10 percent increase in the number of applications received. Moreover, the effect grows over time, rising to roughly 25 percent after one decade in the CA. There are at least two possible reasons why the effects might increase over time. First, the effect could be increasing over time due to the design of the platform, with, for example, the internet playing a large role in the success of the CA today, relative to the early days of the CA, when applications were still submitted on paper. Second, there could be network effects, with larger effects associated with joining the CA as the number of other CA members increases over time.[17]

Finally, as shown in Figure 19 in the Appendix, the structural break analysis documents that a break in applications in the join year, or in the years just following, best fits the data. This somewhat delayed response in the break could reflect the increasing effectiveness of the CA, as documented in Figure 5. There is no evidence that the break occurs prior to the join date.

---

[15]Standard errors are clustered at the institution level.

[16]As noted above, we normalize the coefficient one year before joining to zero ($\beta_{t-1} = 0$).

[17]We later investigate network effects more formally via an examination of CA penetration in the source state of students enrolling in CA institutions.

Table 1: Effect of CA Entry on Institutions

	(1) lapp	(2) lyield	(3) ladmit	(4) lenroll	(5) lselect
camember	0.1205***	-0.0872***	0.1149***	0.0386***	-0.0057
	(0.0202)	(0.0131)	(0.0205)	(0.0148)	(0.0111)
N	34584	34425	34626	38686	34532
Clusters	1635	1634	1637	1686	1635

All specifications include institution and year FE.

Standard errors clustered by institution (unitid).

Figure 5: Effect of CA Entry on Applications
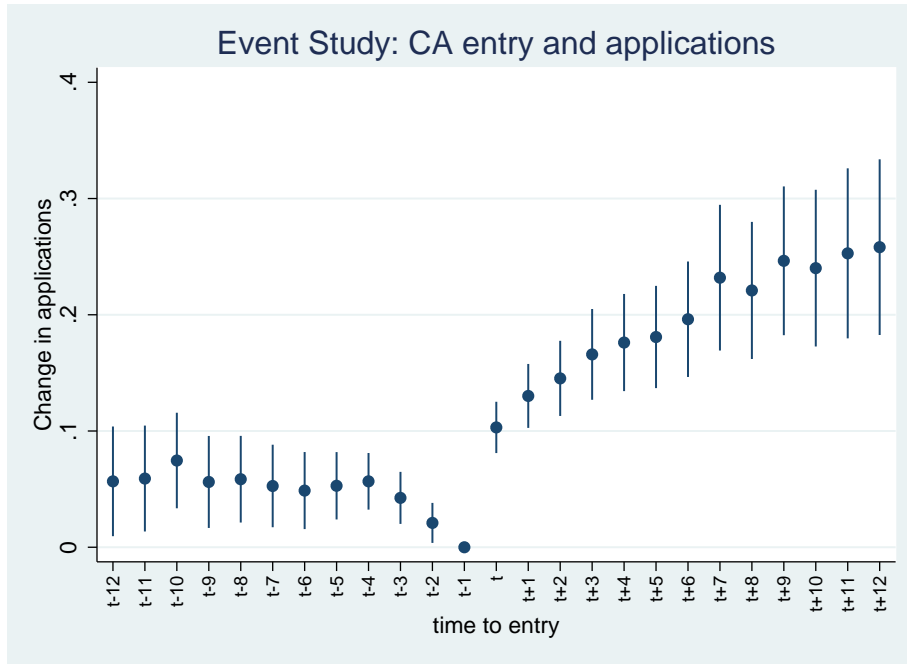
We next investigate whether entry into the CA has led to a decrease in yield, defined as the fraction of admissions offers accepted by students. As shown in column 2 of Table 1, there is a 9 percent reduction in yield after a college joins the CA, relative to the period before they joined the CA, and this effect is statistically significant at conventional levels. In the context of our model, this finding is consistent with the CA increasing student choice via a reduction in frictions associated with submitting college applications.

In order to examine pre-trends and any dynamic effects of joining the CA, we next consider the event study specification for yield. As shown in Figure 6, there is an immediate and discontinuous drop in yield after a college joins the CA, and this effect again becomes more pronounced over

time. This dynamic could again be driven by either the CA becoming more powerful over time or by network effects associated with a increase in the number of CA members. Finally, as shown in Figure 20 in the Appendix, the structural break analysis documents that a break in yield in the join year, or in the following year, best fits the data. There is again no evidence that the break occurs prior to the join date.

Figure 6: CA Entry on Yield
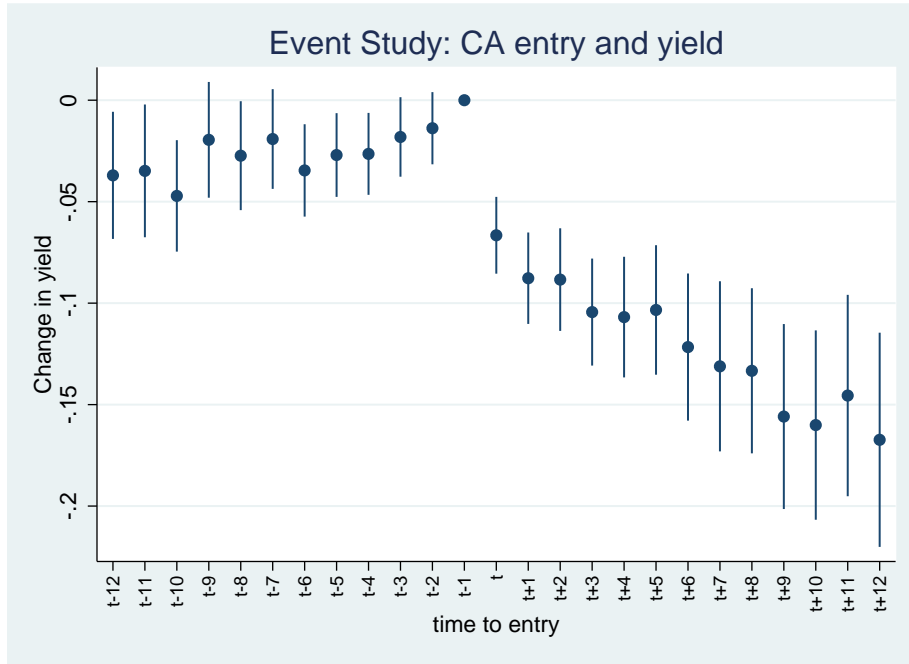
Given this reduction in yield, colleges might need to increase the number of admitted students in order to satisfy their capacity, as discussed in the theoretical model. As shown in column 3 of Table 1, we indeed find a large 12 percent increase in the number of admitted students. As shown in the event study (Figure 7), there is a discontinuous 10 percent increase in admits upon joining, and the effect again increases over time. This break in the series exactly during the join year is also apparent in the structural break analysis in Figure 21 in the Appendix.

Figure 7: Effect of CA Entry on Admissions
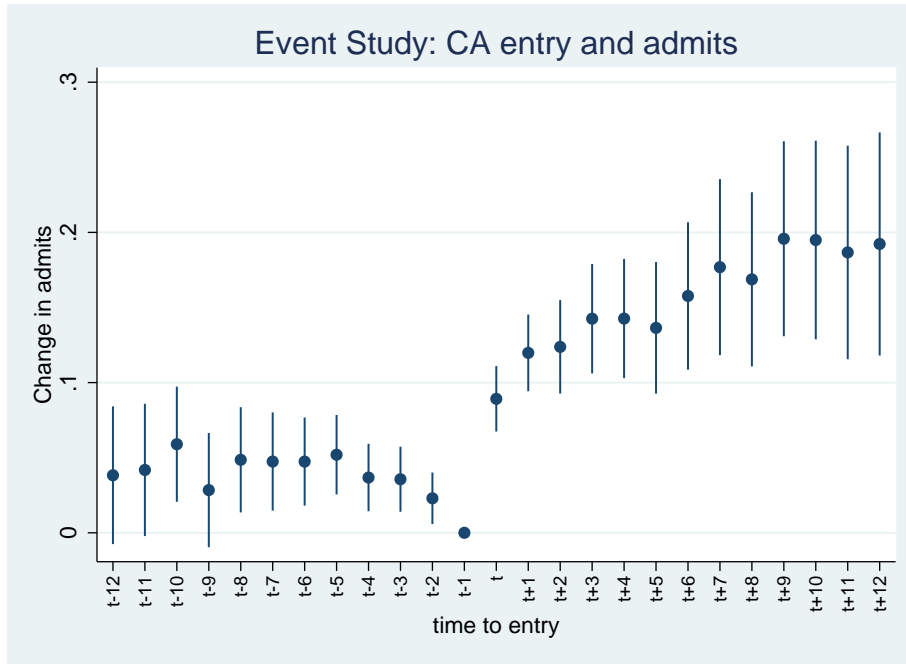
Given that the number of admits is larger in magnitude than the reduction in yield, it must be that enrollment is increasing.[18] As shown in column 4 of Table 1 and in the corresponding event study (Figure 8), we indeed find small increases in enrollment.[19] While schools may have been expanding capacity at the same time that they adopted the CA, it is also possible that some universities were not at full capacity before joining the CA.

---

[18]That is, we can decompose the decrease in yield via the equation $ln(yield) = ln(enollment) - ln(admits)$.

[19]Similar effects are documented in the structural break analysis in Figure 22 in the Appendix.

Figure 8: Effect of CA Entry on Enrollment
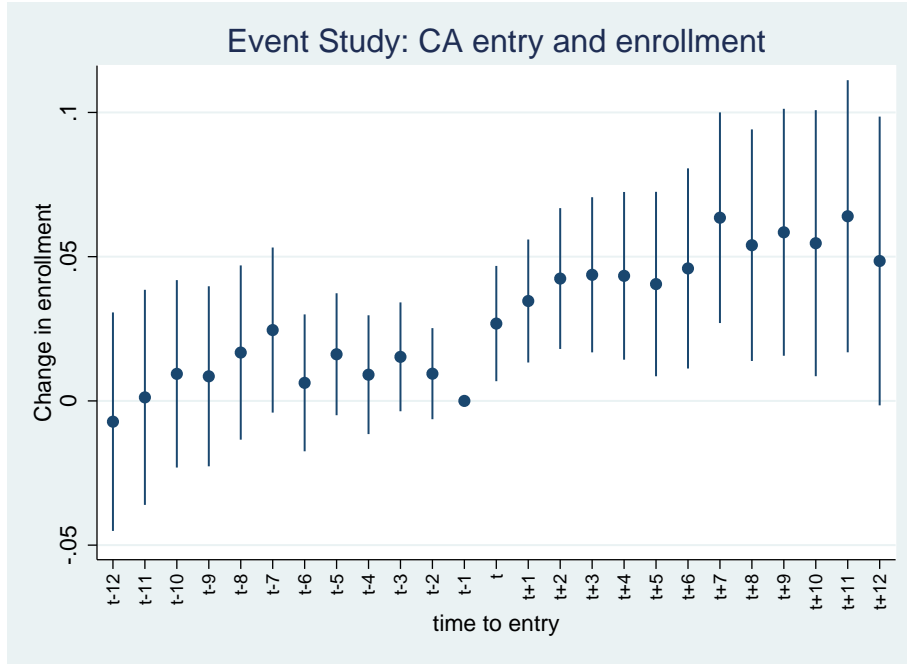
Finally, we investigate whether selectivity has changed, as measured via acceptance rates. Given that both acceptances and admits increase by 12 percent following entry into the CA, there is no change in the log acceptance rate, as documented in the final column of Table 1.[20] The event study in Figure 9, by contrast, provides some evidence of increasing selectivity (i.e., decreasing acceptance rates) following CA adoption.[21] Thus, the evidence with respect to selectivity is mixed, with no effects in the TWFE specifications and increasing selectivity in the event study specification.

---

[20]In particular, we can write the log admissions rates as $ln(selectivity) = ln(admits) - ln(applications)$.

[21]Given the null result in Table 1, there is no structural break during the join year, as documented in Figure 23 in the Appendix.

Figure 9: Effect of CA Entry on Selectivity
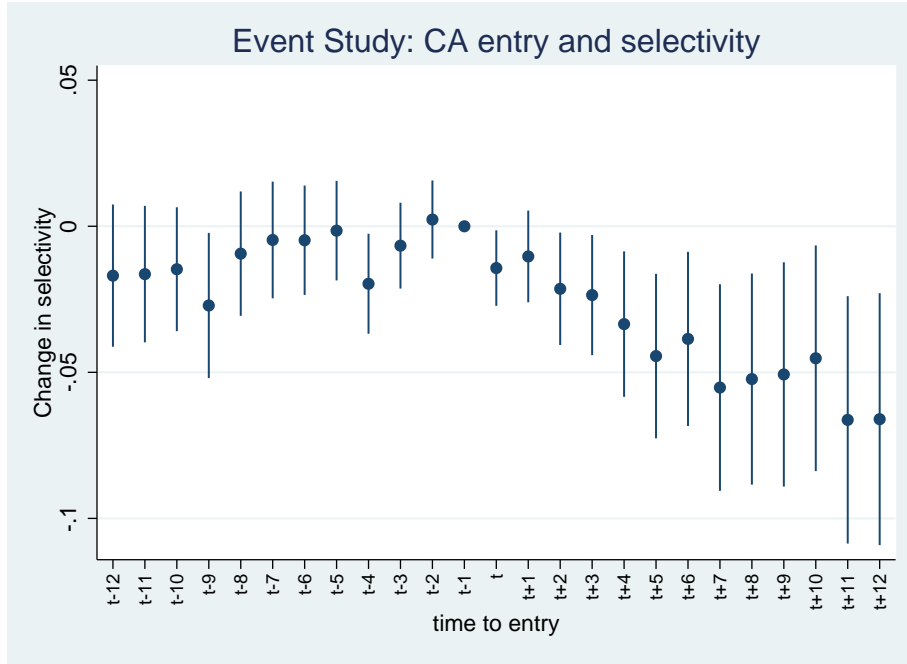
To summarize, we find strong evidence that CA entry increased the number of applications, consistent with reduced frictions, and reduced yield, consistent with large student choice sets. We find only limited evidence that admissions rates fell but strong evidence that the number of admitted students increased.

5.4 Geographic Integration

Given the documented reduction in frictions and increased student choice sets, we next examine geographic integration. In the Appendix, we first provide evidence that the geographic integration documented by Hoxby (2000), covering the period 1949-1994, has continued into our sample period. In particular, we find an increase over time in the average distance traveled by students and an increase in the fraction of out-of-state students.

We next examine, again using College Board data, the extent to which the CA has contributed towards these trends in geographic integration. As shown in column 1 of Table 2, the fraction of out-of-state students rises by 1.4 percentage points in the years after joining, relative to the period prior to joining the CA, a roughly 5 percent increase relative to the sample average of 30 percent out-of-state. Thus, the CA has both contributed towards and accelerated recent trends in

geographic integration.

Table 2: Effect of CA Entry on Student Profiles

|	(1) out-of-state%	(2) foreign%	(3) SAT 25	(4) SAT 75
camember	1.3654***	0.3105*	4.4301	9.8704***
	(0.4462)	(0.1774)	(2.7824)	(2.5368)
N	38019	29209	28543	28549
Clusters	1648	1563	1433	1433

All specifications include institution and year FE.

In cols 1+2 percentages multiplied by 100.

Standard errors clustered by institution (unitid).

The corresponding event study (Figure 10) documents an immediate increase in the fraction of out-of-state enrollment of roughly 1 percentage point following a school joining the CA, and this effect roughly doubles 10 years after joining the CA. The structural break analysis, as reported in the Figure 24, also documents a break in the series exactly during the join year, with no evidence of a break in the series prior to the school joining the CA.

Figure 10: Effect of CA Entry on Out-of-State Enrollment
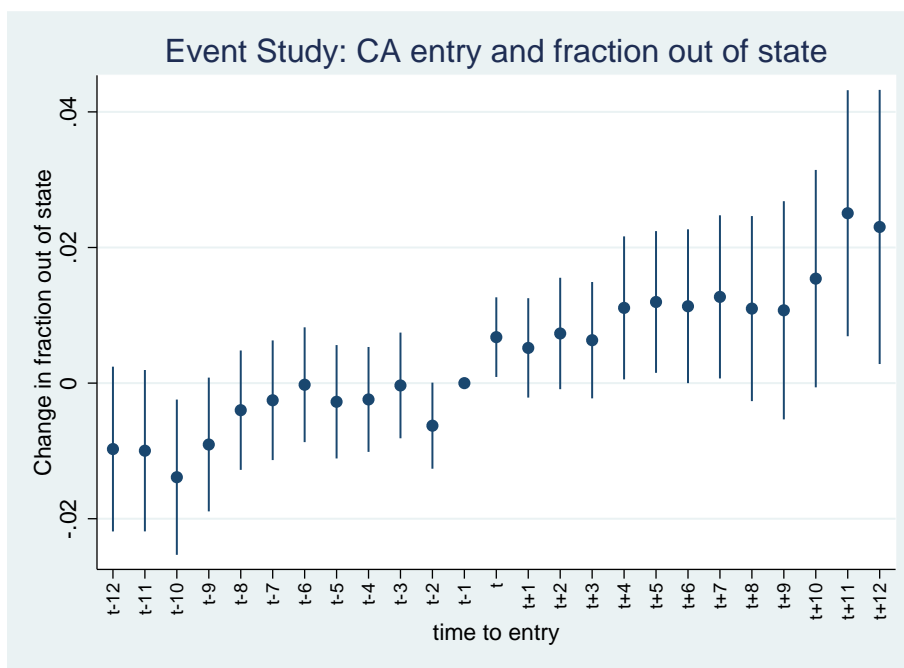
We next consider the effects of joining the CA on foreign enrollment. As shown in column 2 of Table 2, the fraction of foreign students rises by 0.3 percentage points in the years after

joining, relative to the period prior to joining the CA. The corresponding event study (Figure 11) also documents an increase but the effects are noisy in this case. Likewise, the structural break analysis, as reported in the Figure 25 in the Appendix, also provides some evidence of a break in the join year; but the results are again noisy, with corresponding evidence of breaks in the years following the join year.

Figure 11: CA Entry on Foreign Enrollment
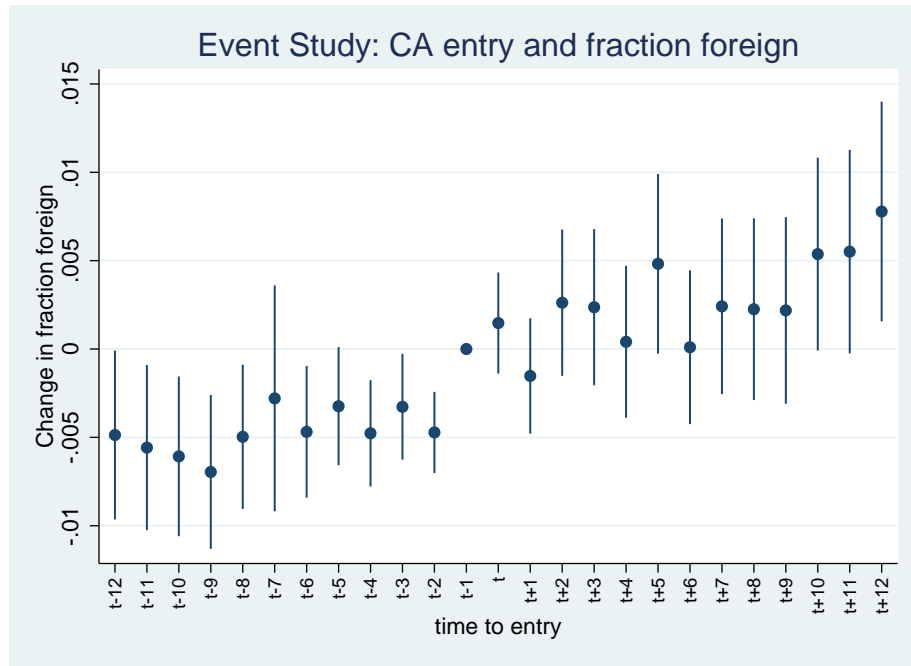
In the Appendix, we document similar results using data from IPEDS. In particular, CA entry leads to an increase in out-of-state students and an increase in foreign students. In addition, IPEDS includes information on state-to-state migration of college students, and we use this information to measure the average distance that students travel to attend college. We find that entry into CA does increase distance traveled, and this effect largely comes from an increase in attendance from nearby states.

To summarize, we find, using both IPEDS and College Board data, that the CA has contributed towards geographic integration, with an increase in the fraction of out-of-state students, with noisier results for the fraction of foreign students. We next consider the source of these out-of-state students via an investigation of network effects associated with the CA.

5.5 Network Effects and Sorting

If the CA is driving these geographic integration results, then the documented increase in out-of-state enrollment should be driven by students applying to other CA institutions. This suggests that institutions that join the CA are likely to see a greater increase in applications from students in states that already have a significant number of CA colleges. For example, if New York has high CA penetration (i.e. many New York schools in the CA), then we might expect that UW-Madison will attract more New York students after joining the CA since these New York students are already using the platform to apply to CA colleges in New York.

To examine these issues, we use IPEDS biennial migration data and measure CA penetration (P_{st}) as the fraction of colleges in source state s at time t that are members of the CA. We then add this penetration measure and an interaction with the CA entry indicator to our two-way fixed effects specification, where the dependent variable is the number of freshmen (N_{sct}) from source state s attending college c at time t. This interaction term provides a test of whether enrollment from high CA penetration states increases when college c joins the CA. In our specification, the unit of observation is now a college by source state by year, and we thus include college by source state fixed effects (μ_{sc}):

$$ln(N_{sct}) = \beta_1 CA_{ct} + \beta_2 P_{st} + \beta_3 CA_{ct} \times P_{st} + \mu_{sc} + \mu_t + \varepsilon_{sct} \tag{5}$$
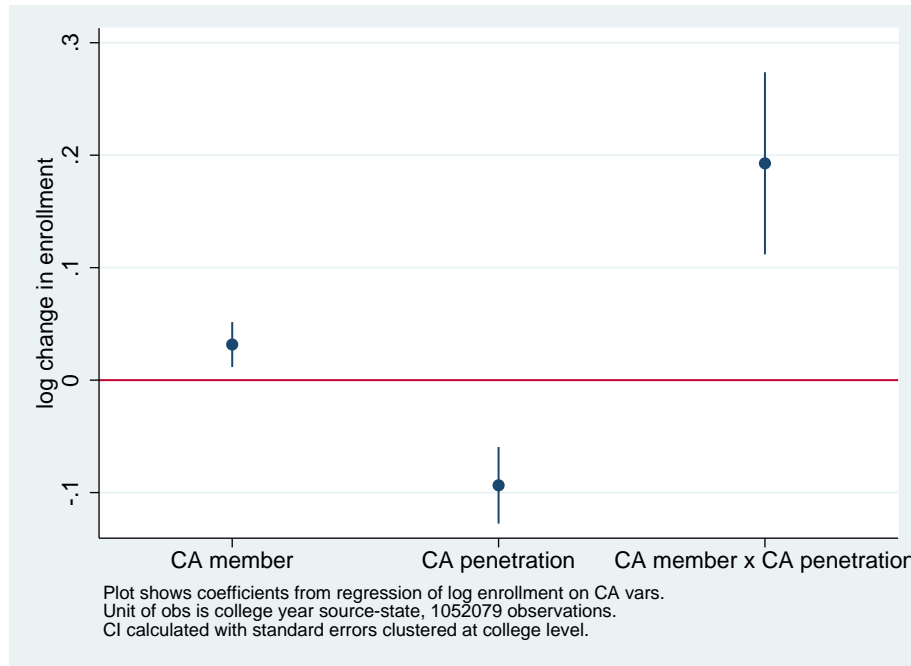
where our key parameter of interest, β_3, captures the increase in enrollment from states with high CA penetration after a college joins the CA, after controlling for differences across states according to CA penetration and overall differences across colleges in CA membership.

In Figure 12 we plot the main effects and the interaction from the above fixed effects specification.[22] Interestingly, the coefficient on CA penetration (P_{st}), which can be interpreted as the effect of CA penetration before the college joins the CA, is significantly negative. This is consistent with network effects in the sense that non-CA schools can be disadvantaged by students applying to CA schools. The coefficient on the interaction term is large, positive, and statistically significant, again suggesting that the increase in applications upon joining the CA is derived from students applying to other CA schools. This evidence is consistent with network effects, as predicted by

---

[22]Confidence intervals are calculated with standard errors clustered by college
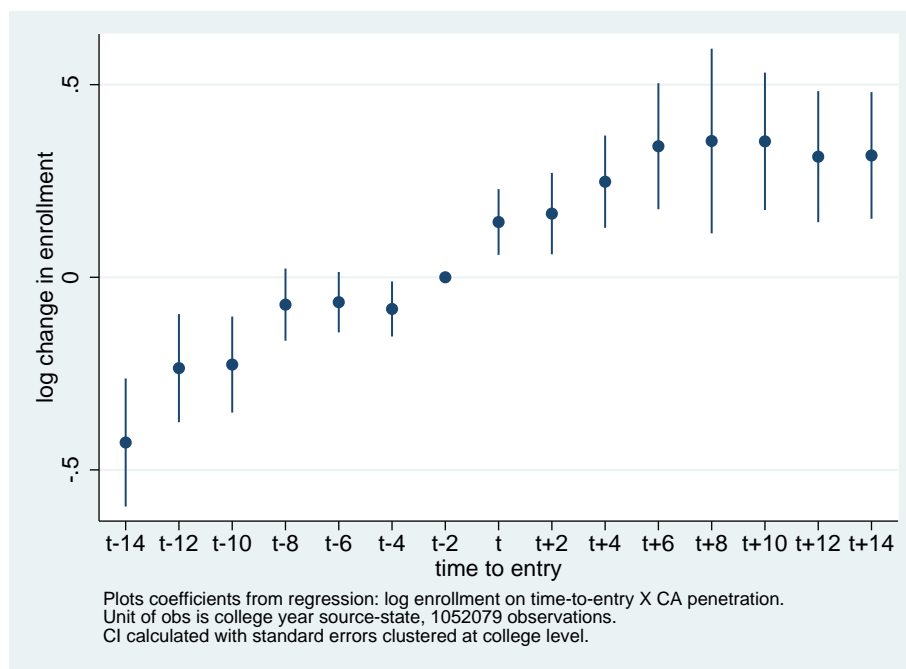
our theoretical model.

Figure 12: Sources of Additional Students

Plot shows coefficients from regression of log enrollment on CA vars.
Unit of obs is college year source-state, 1052079 observations.
CI calculated with standard errors clustered at college level.

Lastly, in order to investigate pre-trends and any dynamic effects, we investigate network effects using our event-study specification. As shown in in Figure 13, there are negative coefficients on source state penetration prior to joining, followed by a sharp jump at the time of entry, and then positive and increasing coefficients thereafter. This type of pattern is consistent with a negative network effect before joining that becomes positive after the school becomes a CA member. Taken together, these results provide evidence that the increase in out-of-state enrollment after joining the CA is driven by migration from states with significant CA penetration.

Figure 13: Effects by Source State CA Penetration

Plots coefficients from regression: log enrollment on time-to-entry X CA penetration.
Unit of obs is college year source-state, 1052079 observations.
CI calculated with standard errors clustered at college level.

5.6 Stratification

Our final research question involves whether the CA has contributed towards stratification in higher education. That is, given the reduction in frictions, the increased student choice sets, and geographic integration, has the CA contributed towards a widening of the gap between more selective and less selective institutions? This could be the case if, for example, high ability students primarily apply to local colleges in the absence of the CA but apply more nationally, and to more selective institutions, following the introduction of the CA. Then, these high ability students might travel further to attend more selective institutions as a result of the CA.

Before investigating the role of the CA in stratification, we first document general trends in SAT scores at different types of institutions. To do so, and in parallel with Figure 2, we classify schools into five categories: top 50 liberal arts, top 50 private, top 50 public, other private and liberal arts, and other public. As shown in Figure 14, there is a large and increasing gap in SAT/ACT scores at the 75th percentile between selective schools (top 50 liberal arts, top 50 private, and top 50 public), and less selective institutions over our sample period. Thus, there is evidence of increasing stratification in general during our sample period.

Figure 14: Stratification in Higher Education
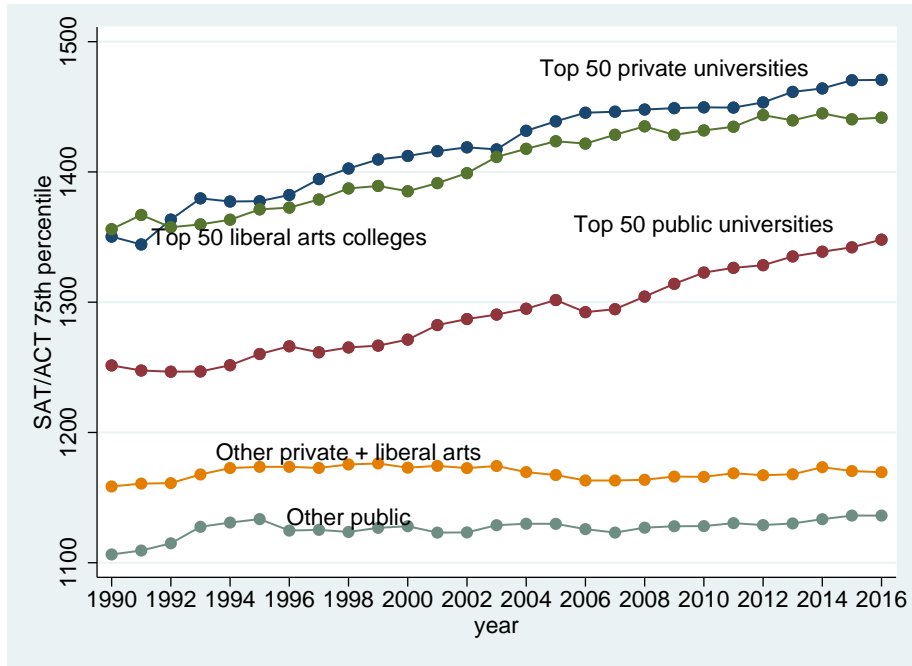
Given that the CA is disproportionately selective institutions, as documented above, we next investigate the degree to which the CA has contributed to this widening of the gap between more selective and less selective institutions. As shown in columns 3 and 4 of Table 2, there is a general increase in SAT scores of enrolled freshman following entry into the CA. In particular, SAT scores at the 25th percentile increase by 4 points and SAT scores at the 75th percentile increase by 9 points, although only the latter effect is statistically significant.[23]

In order to examine pre-trends and any dynamic effects associated with joining the CA, we examine this effect of CA entry on SAT scores via an event study approach. As shown in Figure 15, the effects for the 25th percentile are noisy and, consistent with the results from the fixed effects regression, the effects are small in magnitude. For the 75th percentile (Figure 16), there is a small increase in SAT scores following CA entry, and the effect rises to over 20 points after 10 years in the CA. These effects should be interpreted with caution, however, as there do appear to be some pre-trends in these figures. Nonetheless, the trends are small in the three years just before entry, and, moreover, the 20 point increase in the decade following CA entry is larger than the roughly

---

[23]One interpretation of the difference in effects between the 25th and 75th percentile, in the context of our theoretical extension to heterogeneous student ability, is that the fraction of high ability students is small. In this case, only the top of the distribution of SAT scores would change following entry into the CA, and the bottom of the distribution would be unaffected since it is composed of low-ability students regardless of CA membership.

10 point increase in the decade prior to CA entry. Consistent with this, we find evidence of a break in SAT score at the 75th percentile in the join year or in the following two years (Figure 27 of the Appendix) but no evidence of a break in the series prior to the join year. To conclude, we find mixed evidence regarding the effects of CA entry on SAT scores, with no evidence of increases for the 25th percentile but stronger evidence for increases for the 75th percentile.
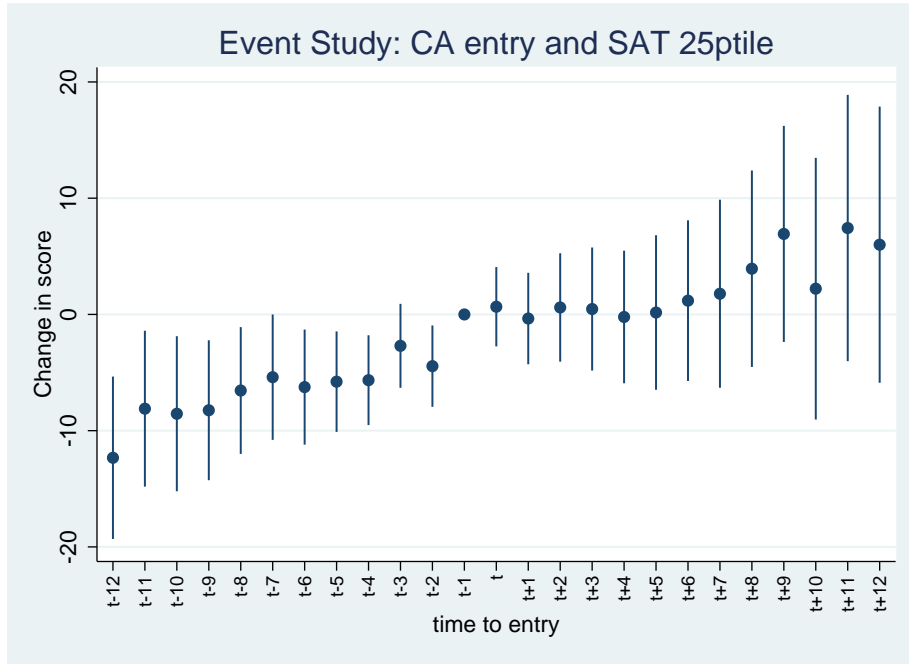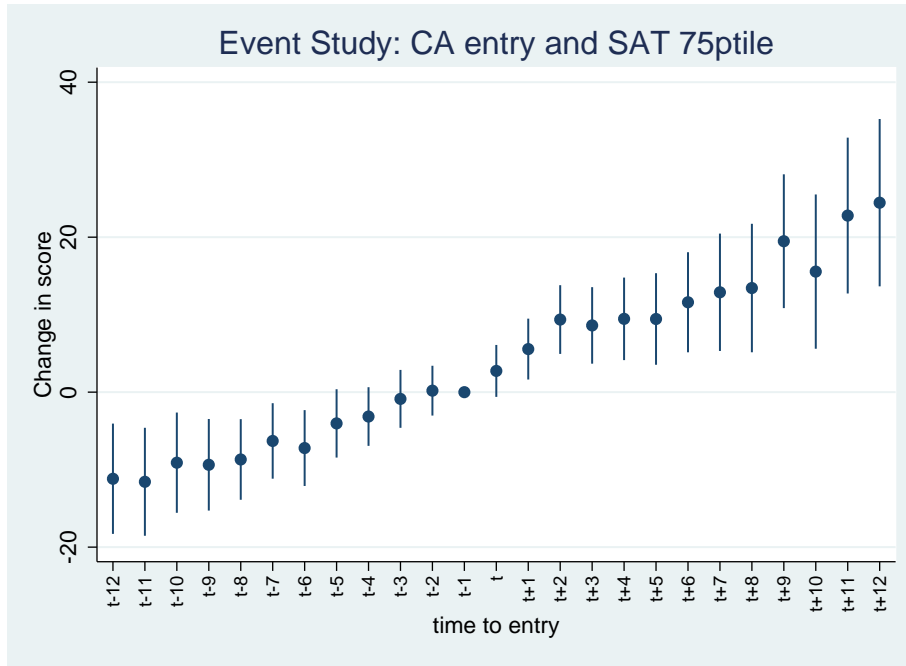
Figure 15: SAT 25 and CA Entry

Figure 16: SAT 75 and CA Entry

6 Conclusion

Consistent with model predictions, we find that the CA has significantly altered college admissions. In particular, after joining the CA, institutions experience an increase in the number of applications, consistent with a reduction in frictions. There is also a significant reduction in yield, consistent with increased student choice due to the CA. We also provide evidence of an effect of the CA on geographic integration, with more out-of-state and foreign students. Moreover, these out-of-state students tend to come from other states with significant CA penetration, patterns consistent with network effects in the CA. Taken together, these results suggest that the CA, by reducing application costs, has reduced frictions and increased student choice sets in college admissions, resulting in a more integrated market. Finally, we provide some evidence that CA entry is associated with an increase in freshmen SAT scores. If so, and given that the CA is disproportionately composed of selective institutions, the CA has contributed towards the rising gap between more selective and less selective institutions.

References

Avery, Christopher and Jonathan Levin, "Early admissions at selective colleges," *American Economic Review*, 2010, *100* (5), 2125–56.

Avery, Christopher N, Mark E Glickman, Caroline M Hoxby, and Andrew Metrick, "A revealed preference ranking of US colleges and universities," *The Quarterly Journal of Economics*, 2013, *128* (1), 425–467.

Bettinger, Eric P, Bridget Terry Long, Philip Oreopoulos, and Lisa Sanbonmatsu, "The role of application assistance and information in college decisions: Results from the H&R Block FAFSA experiment," *The Quarterly Journal of Economics*, 2012, *127* (3), 1205–1242.

Bird, Kelli A, Benjamin L Castleman, Jeffrey T Denning, Joshua Goodman, Cait Lamberton, and Kelly Ochs Rosinger, "Nudging at a national scale: Experimental evidence from a FAFSA completion campaign," 2019. NBER working paper 26158.

Blair, Peter and Kent Smetters, "Why Don't Elite Colleges Expand Supply?," 2019. University of Pennsylvania Working Paper.

Bound, John, Brad Hershbein, and Bridget Terry Long, "Playing the admissions game: Student reactions to increasing college competition," *Journal of Economic Perspectives*, 2009, *23* (4), 119–46.

Cohodes, Sarah R and Joshua S Goodman, "Merit aid, college quality, and college completion: Massachusetts' Adams scholarship as an in-kind subsidy," *American Economic Journal: Applied Economics*, 2014, *6* (4), 251–285.

Deming, David J, Michael Lovenheim, and Richard W Patterson, "The Competitive Effects of Online Education," Technical Report, National Bureau of Economic Research 2016.

Gibbons, Stephen and Anna Vignoles, "Geography, choice and participation in higher education in England," *Regional Science and Urban Economics*, 2012, *42* (1-2), 98–113.

Gurantz, Oded, Jessica Howell, Michael Hurwitz, Cassandra Larson, Matea Pender, and Brooke White, "Realizing Your College Potential: Impacts of College Board's RYCP Campaign on Postsecondary Enrollment," 2019. working paper.

Hoxby, Caroline and Christopher Avery, "The Missing 'One-Offs': The Hidden Supply of High-Achieving, Low-Income Students.," *Brookings Papers on Economic Activity*, 2013.

__ **and Sarah Turner**, "Expanding college opportunities for high-achieving, low income students (SIEPR Discussion Paper No. 12-014)," *Stanford Institute for Economic Policy Research, Stanford University*, 2014.

Hoxby, Caroline M, "How the changing market structure of US higher education explains college tuition," *NBER WP 6323*, 1997.

__ , "The effects of geographic integration and increasing competition in the market for college education," 2000.

__ , "The changing selectivity of American colleges," *Journal of Economic Perspectives*, 2009, *23* (4), 95–118.

__ **and Sarah Turner**, "What high-achieving low-income students know about college," *American Economic Review*, 2015, *105* (5), 514–17.

Knight, Brian and Nathan Schiff, "The Out-of-State Tuition Distortion," *American Economic Journal: Economic Policy*, 2019, *1* (11), 317–50.

Liu, Albert Yung-Hsu, Ronald G Ehrenberg, and Jesenka Mrdjenovic, "Diffusion of Common Application membership and admissions outcomes at American colleges and universities," Technical Report, National Bureau of Economic Research 2007.

Smith, Jonathan, "The effect of college applications on enrollment," *The BE Journal of Economic Analysis & Policy*, 2013, *14* (1), 151–188.

__ **, Michael Hurwitz, and Jessica Howell**, "Screening mechanisms and student responses in the college market," *Economics of Education Review*, 2015, *44*, 17–28.

A Appendix (For Online Publication)

A.1 Conditions for an Interior Solution

Regarding equation 2, the key condition for a unique solution is that the upper solution to the quadratic equation implies an admissions rate in excess of one. To ensure that only the lower solution is feasible requires that application costs be small, relative to the benefits of a larger choice set:

$$F < C_{12} - C_1 \tag{6}$$

That is, the cost of a second application must be less than the option value of also being admitted to one's second choice. The requirement that F is small also guarantees that a solution exists, in the sense that the discriminant is positive.

Regarding equation 1, we require the following condition for an interior solution:

$$QY_1 < \kappa < Q^2(Y_{12} + y_{12}) + Q(1-Q)(Y_1 + Y_2) \tag{7}$$

where Q is set at its equilibrium value and is thus a function of model parameters. The left hand side of the inequality requires that college capacity is more than sufficient to accommodate accepted students when all students apply to only their first choice, given equilibrium admissions rates. The right hand side requires that the college capacity is not sufficient to accommodate the situation when all students apply to both colleges, given equilibrium admissions rates. Thus, capacity can be neither too small nor too large.

A.2 Equilibrium Solution

We first solve equation 2 for the equilibrium admissions rate. While this equation is quadratic in Q and thus has two solutions in principle, the upper solution implies an admissions rate in excess of 1, under the assumptions outlined in the Appendix above, and we thus focus on the lower solution:

$$Q^* = \frac{C_2 - \sqrt{C_2^2 - 4f(C_1 + C_2 - C_{12})}}{2(C_1 + C_2 - C_{12})} \tag{8}$$

Given this equilibrium admissions rate, one can then calculate the equilibrium fraction of students applying to both colleges via equation 1, yielding:

$$b^* = \frac{\kappa - QY_1}{Q^2[Y_{12} + y_{12} - Y_1] + Q(1-Q)Y_2} \tag{9}$$

where Q is set at equilibrium levels.

A.3 Proof of Proposition 1

Parts 1) and 2): In Equation 8, it is clear that equilibrium admissions rates are increasing in F. Thus, a marginal reduction in F leads to a reduction in equilibrium admissions rates. This effect is illustrated in Figure 17 below.
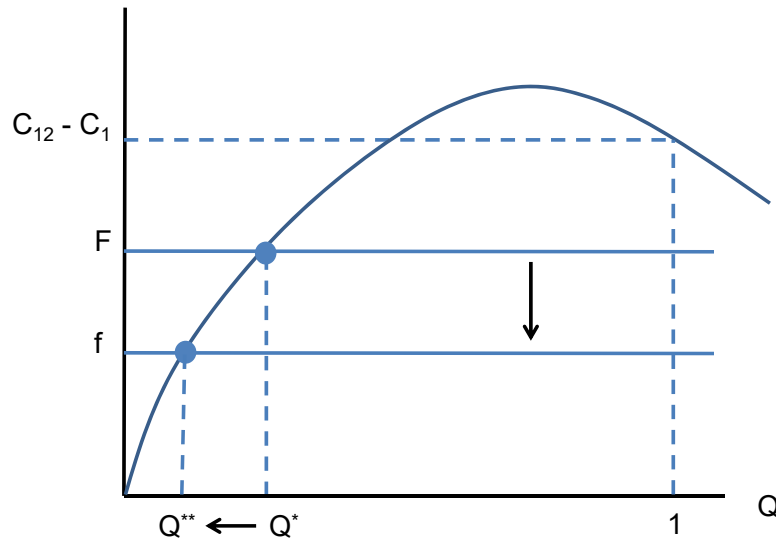

Figure 17: Effects on admissions rates

Given that Q declines under the CA, we must next show that b is decreasing in Q. Taking the derivative of equation 9 with respect to Q, we have:

$$\frac{db}{dQ} = \frac{-Y_1}{D} - \frac{(\kappa - QY_1)[2Q(Y_{12} + y_{12} - Y_1 - Y_2) + Y_2]}{D^2} \tag{10}$$

where the denominator equals $D = Q^2[Y_{12} + y_{12} - Y_1] + Q(1-Q)Y_2$. This denominator is positive

31

since $Y_{12} + y_{12} > Y_1$.

Substituting back in the definition of b, we have that:

$$\frac{db}{dQ} = \frac{-Y_1 - b[2Q(Y_{12} + y_{12} - Y_1 - Y_2) + Y_2]}{D} \tag{11}$$

Re-arranging the numerator, this relationship can be written as follows:

$$\frac{db}{dQ} = \frac{-2bQ(Y_{12} + y_{12} - Y_1) + (bQY_2 - Y_1) - bY_2(1 - Q)}{D} \tag{12}$$

Each of these three terms in the numerator are negative. In particular, the first term is negative since $Y_{12} + y_{12} > Y_1$. The second term is negative since $Y_2 < Y_1$, $b < 1$, and $Q<1$. Finally, the third term is negative since $Q < 1$ in equilibrium. Since the denominator must be positive for b to be positive, the slope is negative. This change in application rates is illustrated in Figure 18 below.
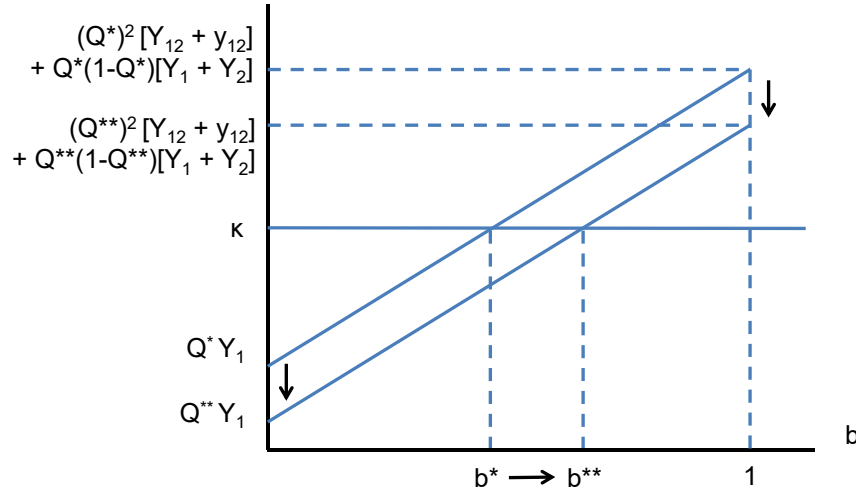

Figure 18: Effects on applications

Part 3): Note that the number of admitted students is equal to $Q(1 + b)$, the product of the admissions rate and the number of applications received. Using the closed form solution for b, This can be written as:

$$Q(1 + b) = Q + Qb = Q + \frac{\kappa - QY_1}{Q[Y_{12} + y_{12} - Y_1] + (1 - Q)Y_2} \tag{13}$$

32

Taking the derivative, we have that:

$$\frac{dQ(1+b)}{dQ} = 1 - \frac{Y_1}{D} - \frac{\kappa - QY_1}{D^2}[Y_{12} + y_{12} - Y_1 - Y_2] \tag{14}$$

where the denominator equals $D = Q[Y_{12} + y_{12} - Y_1] + (1 - Q)Y_2$.

Using the fact that $Qb = [\kappa - QY_1]/D$, the slope can be re-written as:

$$\frac{dQ(1+b)}{dQ} = 1 - \frac{Y_1}{D} - \frac{Qb}{D}[Y_{12} + y_{12} - Y_1 - Y_2] \tag{15}$$

This can be re-written as:

$$\frac{dQ(1+b)}{dQ} = \frac{D - Y_1 - Qb[Y_{12} + y_{12} - Y_1 - Y_2]}{D} \tag{16}$$

Since D is positive, we simply need to show that the numerator is negative. Since the term $Y_{12} + y_{12} - Y_1 - Y_2$ is negative, the numerator is increasing in b. Thus, to show that it is negative for all b between 0 and 1, we simply need to show that it is negative when $b = 1$. In this case, and canceling terms, the numerator can be written as $Y_2 - Y_1$, which is negative.

Part 4): The increase in out-of-state students follows directly from the increase in b resulting from the reduction in F.

A.4   Extension to 3 Colleges

We next consider the case in which two colleges ($c = 1$ and $c = 2$) join the CA but a third college ($c = 3$) does not join. In this case, there are three types of students, corresponding to the ex-ante ranking of the third college. Type 1 students have ex-ante preferences that rank college 3 last (there are two sub-types: either $U_1 > U_2 > U_3$ or $U_2 > U_1 > U_3$). Type 2 students have ex-ante preferences that rank college 3 in the middle (either $U_1 > U_3 > U_2$ or $U_2 > U_3 > U_1$). Type 3 students have ex-ante preferences that rank college 3 first (either $U_3 > U_1 > U_2$ or $U_3 > U_2 > U_1$). Given all of this, we can write the ex-ante preferences of the three different types (six different sub-types) of students as follows:

$$U_1 > U_2 > U_3, 1.1$$

$$U_2 > U_1 > U_3, 1.2$$

$$U_1 > U_3 > U_2, 2.1$$

$$U_2 > U_3 > U_1, 2.2$$

$$U_3 > U_1 > U_2, 3.1$$

$$U_3 > U_2 > U_1, 3.2$$

Let Q_{CA} and Q_N denote admissions rates at the CA colleges and the non-CA college, respectively. Capacities are symmetric and equal κ.

We focus here on the case in which students do not apply to all three colleges.[24] Let b_1 be the fraction of type 1 students applying to their first and second choice and likewise for b_2 and b_3. The feasibility constraint for college 1 (college 2 is analogous) is now given by:

$$Q_{CA}[\underbrace{(1-b_1)Y_1}_{type\,1.1} + \underbrace{(1-b_2)Y_1}_{type\,2.1} +$$

$$\underbrace{b_1 Q_{CA}Y_{12} + b_1(1-Q_{CA})Y_1}_{type\,1.1} + \underbrace{b_2 Q_N Y_{12} + b_2(1-Q_N)Y_1}_{type\,2.1}] + \tag{17}$$

$$Q_{CA}[\underbrace{b_1 Q_{CA}y_{12} + b_1(1-Q_{CA})Y_2}_{type\,1.2} + \underbrace{b_3 Q_N y_{12} + b_3(1-Q_N)Y_2}_{type\,3.1}] = \kappa \tag{18}$$

And for college 3 it is:

---

[24]This could formalized by having a large difference in preferences between the second and third choice. For type 1.1 students, for example, $U_2 - U_3$ would be large.

$$2Q_N\underbrace{[(1-b_3)Y_1+b_3Q_{CA}Y_{12}+b_3(1-Q_{CA})Y_1]}_{types\,3}+$$

$$2Q_N\underbrace{[b_2Q_{CA}y_{12}+b_2(1-Q_{CA})Y_2]}_{types\,2}=\kappa \tag{19}$$

Prior to the CA, the three relevant indifference conditions are given, similarly to before, by:

$$Q^2(C_{12}-C_1)+(1-Q)QC_2=F \tag{20}$$

With the introduction of the CA, the conditions become as follows for type 1 students:

$$Q_{CA}^2(C_{12}-C_1)+(1-Q_{CA})Q_{CA}C_2=f \tag{21}$$

However, for types 2 and 3, now they incorporate the two different admission rates. For type 2 students, taking the case of 2.1, we have that:

$$Q_{CA}Q_N(C_{12}-C_1)+(1-Q_{CA})Q_NC_2=F \tag{22}$$

For type 3 students, taking the case of 3.1, we have that:

$$Q_NQ_{CA}(C_{12}-C_1)+(1-Q_N)Q_{CA}C_1=F \tag{23}$$

Claim: these three indifference conditions cannot be simultaneously satisfied

Proof: The introduction of the CA causes the right hand side in equation 20 to fall from F to f. This implies that, given the admissions rate at CA schools, more type 1 individuals find it profitable to apply to a second school, increasing b_{CA}. However, under a higher b_{CA}, CA colleges will have excess demand, violating their feasibility constraint (equation 18). Thus, Q_{CA} must decrease until type 1 students are indifferent between applying to a second school or not.[25]

Now, note that the fall in Q_{CA} causes the left hand side in equation 22 to increase, implying that more type 2 students want to apply to a second college. However, the opposite happens with type

---

[25]Note that an increase in b_{CA} must accompany the fall in Q_{CA}. Else, if only Q_{CA} were to fall, colleges would not meet the feasibility constraint, as they would have open vacancies given the smaller admission rate.

3 applicants. The fall in Q_{CA} pushes down the left hand side in equation 23, implying that fewer type 3 students want to apply to a second college.

Due to these opposing effects of a decrease in Q_{CA} for type 2 and type 3 students, both conditions cannot be simultaneously satisfied, meaning that either b_2 or b_3 must be at a corner solution. More formally, comparing the conditions for type 2 and type 3, we have that:

$$Q_{CA}Q_N(C_{1,2} - C_1) + (1 - Q_{CA})Q_N C_2 = Q_N Q_{CA}(C_{1,2} - C_1) + (1 - Q_N)Q_{CA}C_2$$

This is only satisfied when $Q_{CA} = Q_N$. However, under this condition, the left hand side of the three conditions are equal. But this is a contradiction with the fact that the first equation equals f, the second and third equations equal F, with $f < F$.

Claim: There is no equilibrium with $b_2 = 1$ and b_3 interior.

Proof: Assuming that b_1 is interior, and imposing symmetry, this would require the following:

$$Q_{CA}^2(C_{1,2} - C_1) + (1 - Q_{CA})Q_{CA}C_2 = f$$

$$Q_{CA}Q_N(C_{1,2} - C_1) + (1 - Q_{CA})Q_N C_2 > F$$

$$Q_N Q_{CA}(C_{1,2} - C_1) + (1 - Q_N)Q_{CA}C_2 = F$$

Comparing the conditions for type 1 and type 3 and using the fact that $f < F$, we have that:

$$Q_{CA}^2(C_{1,2} - C_1) + (1 - Q_{CA})Q_{CA}C_2 < Q_N Q_{CA}(C_{1,2} - C_1) + (1 - Q_N)Q_{CA}C_2$$

Re-arranging, this can be written as:

$$Q_{CA}(Q_{CA} - Q_N)(C_{1,2} - C_1 - C_2) < 0$$

Since $C_{1,2} - C_1 - C_2 < 0$, this requires $Q_{CA} > Q_N$.

Comparing types 2 and 3, we have that:

$$Q_{CA}Q_N(C_{1,2} - C_1) + (1 - Q_{CA})Q_N C_2 > Q_N Q_{CA}(C_{1,2} - C_1) + (1 - Q_N)Q_{CA}C_2$$

Re-arranging, this can be written as:

$$(1 - Q_{CA})Q_N > (1 - Q_N)Q_{CA} \tag{24}$$

This requires that $Q_N > Q_{CA}$. This contradicts that earlier requirement that $Q_{CA} > Q_N$.

Summary: The introduction of the CA leads to an increase in b_1, the fraction applying to both CA schools and a reduction in Q_{CA}. Given this, it must be case that b_2 increases to 1 or that b_3 decreases to zero since both cannot be interior. However, we have shown that b_2 cannot equal 1, meaning that $b_3 = 0$. Thus, there is a reduction, all the way to zero, in the fraction applying to a school outside of the CA and a school inside the CA. Given this, there are network effects with more type 1.1 students attending college 2 and more type 1.2 students attending college 1. Likewise, there are fewer type 3.1 students attending college 1 and fewer type 3.2 students attending college 2.

Quantitative analysis: To provide further evidence on the three college case, we choose parameters that guarantee interior solutions before the policy change. In particular, we set $U_1 = 2$, $U_2 = 1$, $\kappa = 0.55$, and $F = 0.3$. Prior to the CA, colleges and students are symmetric, with $Q = 0.2985$ and $b = 0.0776$. Introduction of the CA lowers F to $f = 0.29$. Under the CA, imposing that $b_3 = 0$ in the new equilibrium, the admission rate of the CA schools falls from $Q = 0.2985$ to $Q'_{CA} = 0.2844$, while the admission rate of the non-CA school also falls, but by a smaller degree, to $Q'_N = 0.2942$. These changes reflect the direct and indirect effects of the decrease in F on the different types of students. Type 1 students are the only ones that benefit directly by the policy and strongly increase their applications to a second school from $b = 0.0776$ to $b'_1 = 0.3346$. On the other hand, type 2 students see their application rate grow only marginally, to $b'_2 = 0.0908$. The reason is that type 2 students have as first choice a CA college followed by the non-CA college, so they do not enjoy the lower application fee but do face the lower admission rate from the CA school, providing them with incentives to apply to their second choice.

A.5 Extension to Ability Types

We consider three colleges (c) and two ability types: low-ability and high-ability. We assume that colleges want to attract as many high ability students as possible and thus admit them with probability one. Low ability students are then admitted at a lower rate in order to fill any remaining capacity. Given our interest in stratification, we can then simply study the behavior of high ability students. Given that high ability students are admitted with certainty, the model plays out differently in this case. In particular, students will not be indifferent when choosing their application sets, and corner solutions are now relevant for these high ability students. Given these corner solutions, we focus on non-marginal changes in application costs.

We focus here on two cases. In the first case, application costs are sufficiently high that high ability students only apply to their first choice in the absence of the CA. For students with the preference order $U_1 > U_2 > U_3$, this requires:

$$C_{12} - C_1 < F \tag{25}$$

Suppose now that colleges 1 and 2, but not college 3, join the CA. Further, suppose that application costs fall sufficiently such that $C_{12} - C_1 > f$, and likewise for students with preference ordering $U_2 > U_1 > U_3$. Then, these two sets of students will apply to both colleges, and students with other preference orderings are unaffected. Since these two sets of students are now more likely to attend college (recall that $Y_{12} + y_{12} > Y_1$), the fraction of high ability students at CA colleges increases. The fraction of high ability students at colleges outside of the CA is unchanged.

In the second case, suppose that application costs are sufficiently low that high ability students apply to their top two choices, but not their third choice, in the absence of the CA. For students with the preference order $U_1 > U_3 > U_2$, not applying to the third college requires:

$$C_{123} - C_{13} < F \tag{26}$$

where C_{123} represents the value from having a full choice set of all three colleges. Suppose now that colleges 1 and 2, but not college 3, join the CA, and application costs fall sufficiently such that $C_{123} - C_{13} > f$, and likewise for all students that have college 1 or 2 as their third choice. Then, all students except those with preference orderings $U_1 > U_2 > U_3$ and $U_2 > U_1 > U_3$ will apply to

all three colleges. Thus, there is an increase in applications for colleges 1 and 2 and no increase in applications for college 3. Given that the yield on students accepted to college 3 now falls (resulting from more college 3 applicants also applying to colleges 1 and 2), this implies that colleges 1 and 2 will now draw some high ability students who would have attended college 3 in the absence of the CA. Thus, as in the first case, the fraction of high ability students at CA colleges increases. The new effect here is that the fraction of high ability students falls at schools outside of the CA.

A.6   Structural Break Exhibits

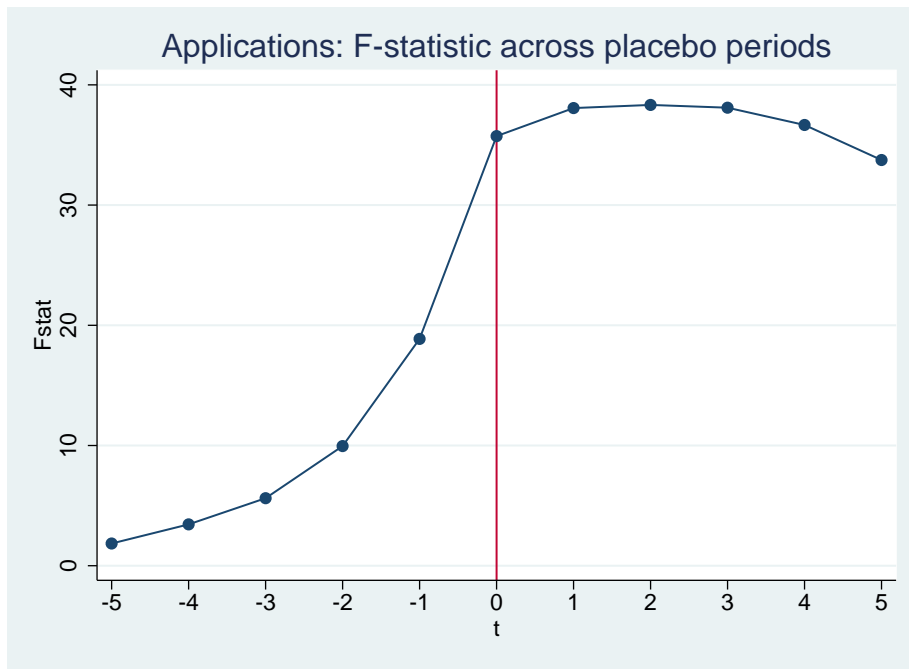Figure 19: Structural break analysis: applications

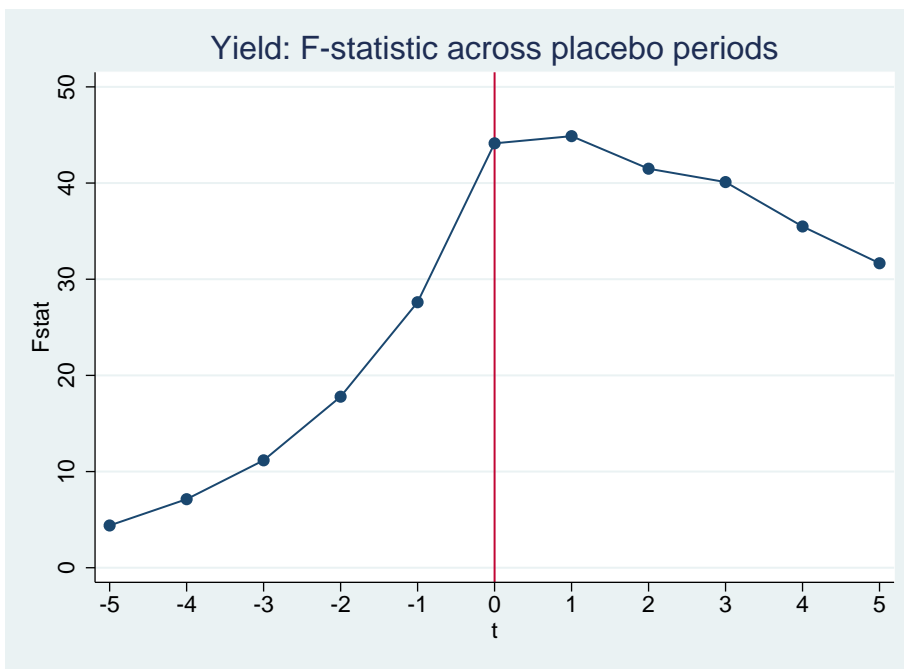Figure 20: Structural break analysis: yield

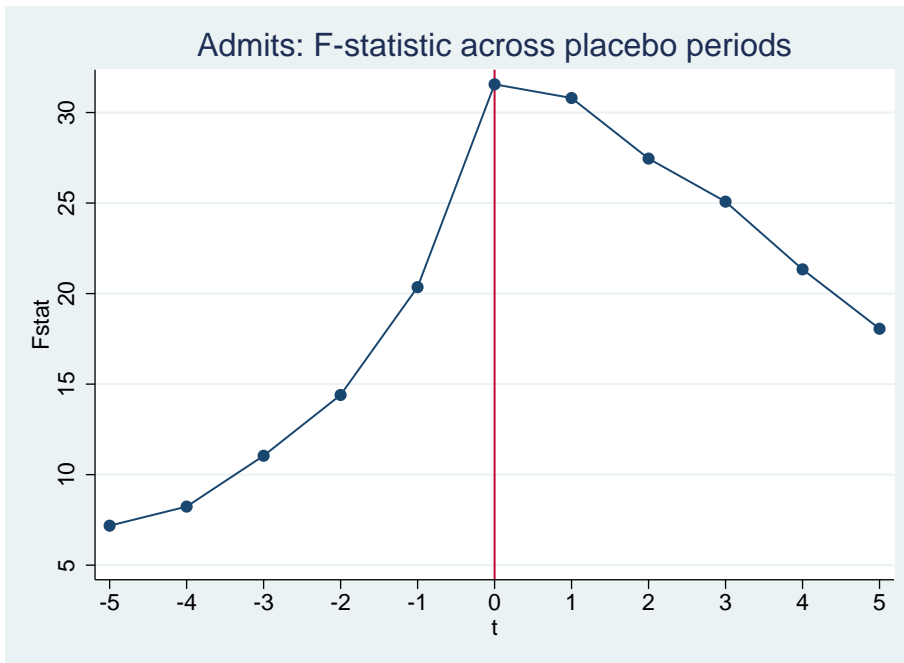Figure 21: Structural break analysis: admits

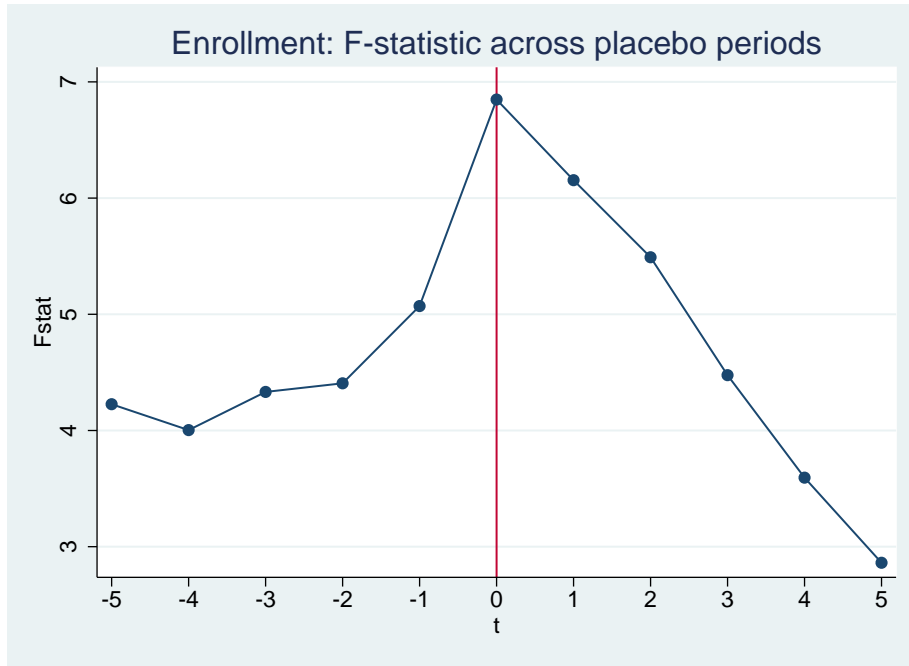Figure 22: Structural break analysis: enrollment

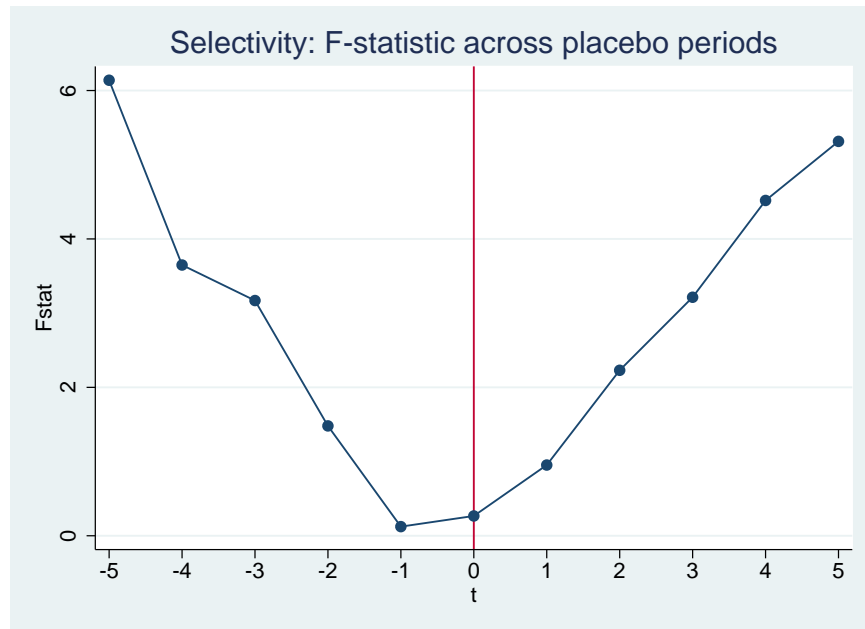Figure 23: Structural break analysis: selectivity

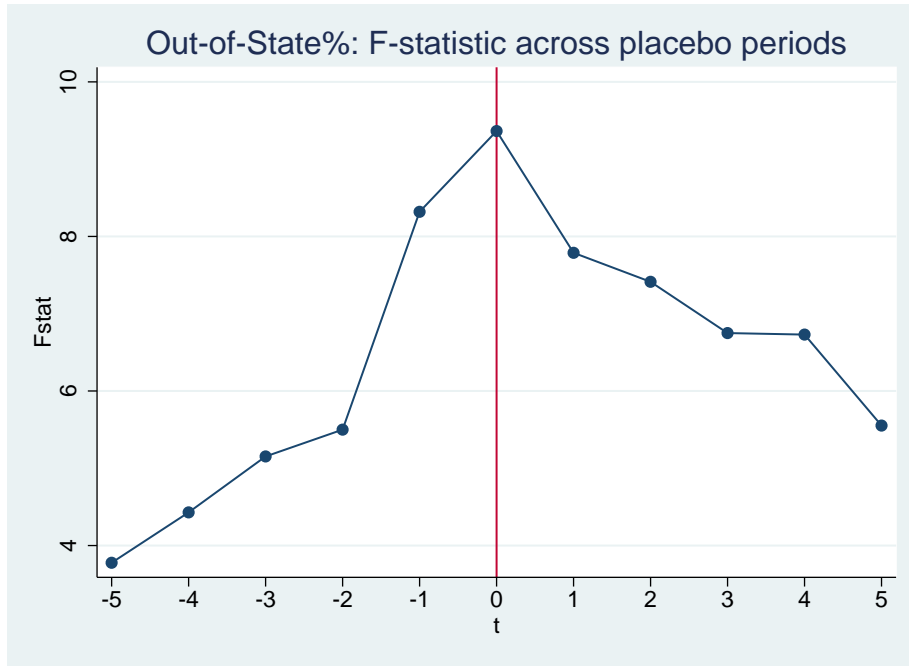Figure 24: Structural break analysis: out-of-state

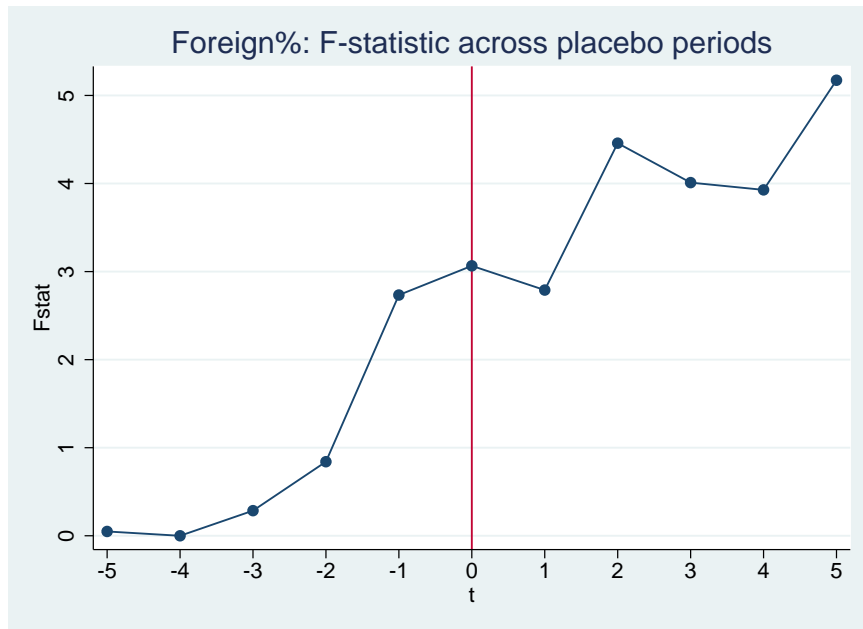Figure 25: Structural break analysis: foreign

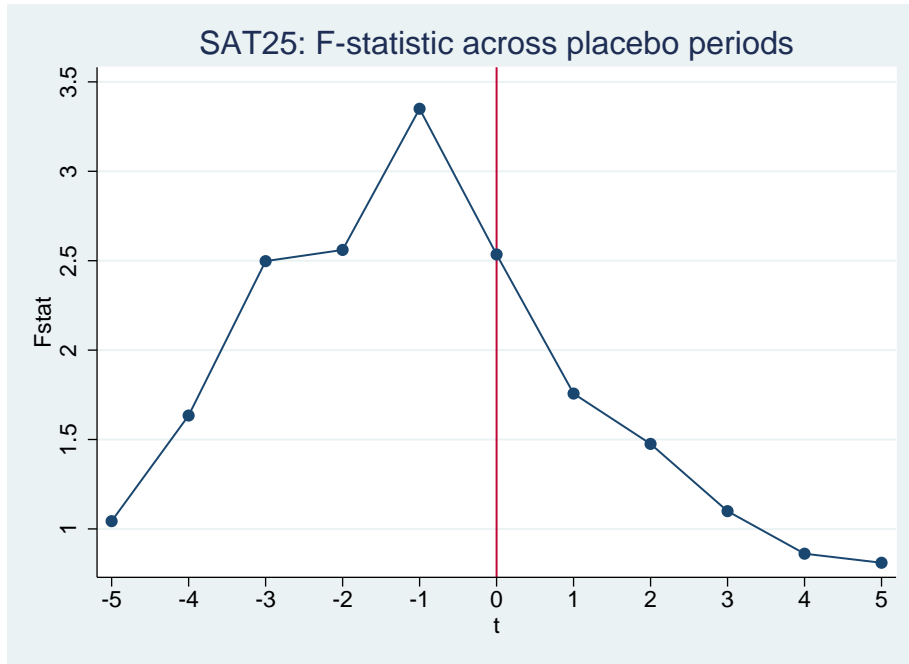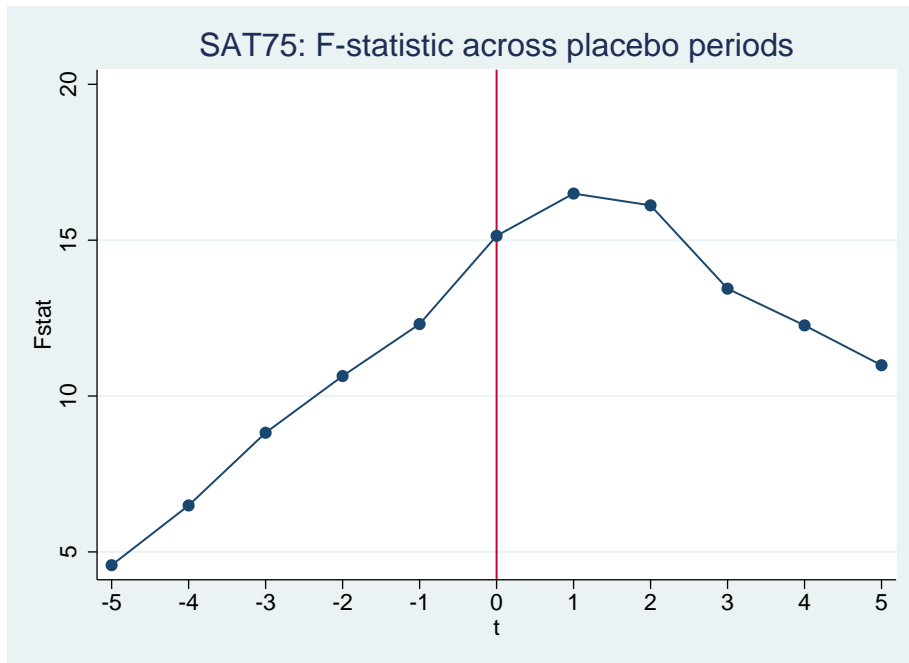Figure 26: Structural break analysis: SAT25

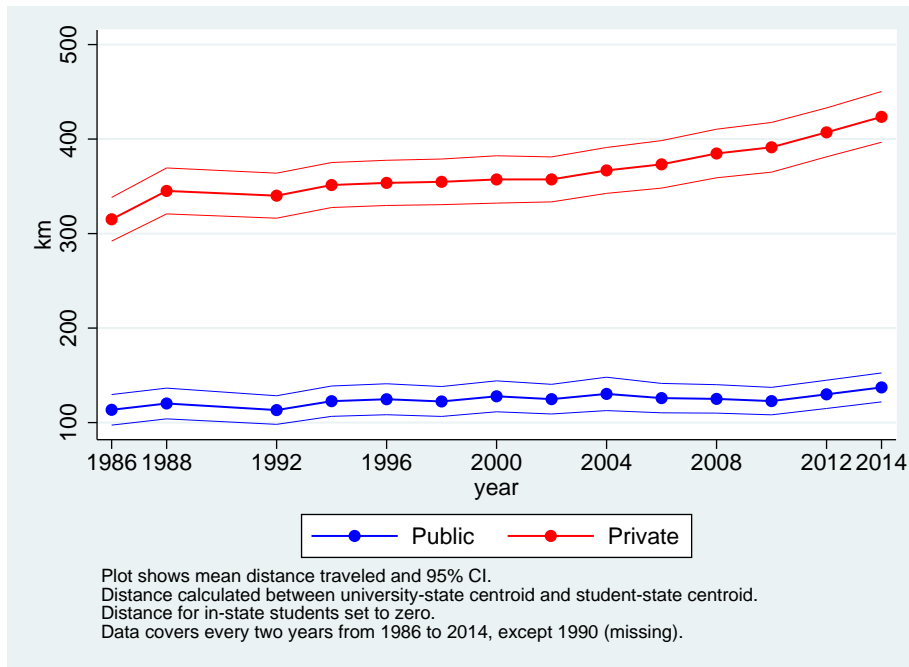

Figure 27: Structural break analysis: SAT75

A.7 Trends in Geographic Integration

Hoxby (2000) documents that the percentage of students attending in-state institutions fell consis-

tently from 1949 to 1994 and that the role of distance in explaining college choice decreased as well. We extend this study of geographic integration into our sample period by measuring trends in distance traveled from a student's home state to the state of their university using IPEDS data over the period from 1986 to 2014.[26] In Figure 28 we plot the mean distance traveled in each year, along with 95 percent confidence intervals.[27] As shown, there is a clear increase in distance traveled over this time period, with the average distance traveled increasing by over 100 kilometers for private universities and roughly 40 kilometers for public universities. Thus, the trends towards greater geographic integration documented by Hoxby between 1949 and 1994 also appear over our sample period 1986-2014.

Figure 28: Distance Traveled per Student

Plot shows mean distance traveled and 95% CI.
Distance calculated between university-state centroid and student-state centroid.
Distance for in-state students set to zero.
Data covers every two years from 1986 to 2014, except 1990 (missing).

In Table 3 we calculate the average increase in geographic integration over time for public and private institutions, using the specification $y_{ct} = \beta_1 years_t + \beta_2 years_t \times public_c + \mu_c + \varepsilon_{ct}$. In

---

[26]In particular, we measure the great circle distance in kilometers between state centroids, defining the distance for all in-state students as zero

[27]Letting the subscript s denote a student's state, we define the mean distance traveled by all students from US states at college c in year t as $avdist_{c,t} = (1/nat\_enroll_{c,t})\sum_{s \in S} enroll_{c,s,t} * dist_{c,s}$. The variable nat_enroll is total enrollment from the 50 US states and D.C. The home location of foreign students and students from US territories is usually not available, and therefore we excluded these groups from the total. However, students from these groups are counted in total enrollment when calculating percentage of students attending in-state.

column 1 we find that distance traveled increases by about 3 kilometers per year for private institutions and 1 kilometer per for public institutions, while column 2 specifies average distance in logs and shows that both types of institutions have roughly the same percentage increase over time of 1.4 percent. In columns 3 and 4 we examine the average distance traveled by out-of-state students only to separate the effects of a change in the percentage of out-of-state students from a change in the geographic composition of the out-of-state students. Interestingly, the results show that while out-of-state students at private institutions are traveling further each year, there is essentially no increase in distance for out-of-state students at public institutions (the interaction effect is the same magnitude as the main effect). This implies that the increasing distance traveled by public university students comes entirely from an increase in the out-of-state percentage, which increases at about 0.15 percentage points each year for both types of institutions (the in-state percentage is multiplied by 100 in column 5). Lastly, in column 6 we show that the percentage of total enrollment from foreign students (multiplied by 100) increases by about 0.04 percentage points each year for private institutions and 0.01 percentage points for public institutions.

Table 3: Geographic Integration by Institution Type

	(1) dist	(2) ln(dist)	(3) dist	(4) ln(dist)	(5) instate%	(6) foreign%
years	3.187***	0.014***	5.136***	0.007***	-0.145***	0.043***
	(0.280)	(0.001)	(0.328)	(0.000)	(0.021)	(0.006)
years X public	-1.958***	0.000	-5.511***	-0.006***	0.007	-0.031**
	(0.357)	(0.002)	(0.605)	(0.001)	(0.028)	(0.011)
Constant	236.937***	4.679***	876.442***	6.585***	70.587***	2.013***
	(2.909)	(0.013)	(4.271)	(0.005)	(0.223)	(0.077)
Observations	20685	20236	20236	20236	20685	20685
Clusters	1708	1688	1688	1688	1708	1708

Dep var in cols 1 and 2 is distance per student.
Dep var in 3,4 is distance per out-of-state student.
In cols 5+6 shares are multiplied by 100.
All spec include institution FE; std errors clustered by institution.

A.8 CA and Geographic Integration (IPEDS data)

As shown in Table 4, universities, after joining the CA, experience a significant increase in the average distance students travel to attend, with column 1 documenting an increase of 30 kilometers, a roughly 10 percent increase (column 2). Restricting to only out-of-state students, the distance increases by 55 kilometers (column 3), an increase of 7 percent for this population (column 4). In addition to out-of-state students traveling further, joining the CA also decreases the fraction of in-state students by about 2.3 percentage points (column 5) and increases the fraction of foreign students by 0.6 percentage points (column 6). Generally, the magnitudes of the effects in Table 4 are large. Comparing each coefficient in Table 4 to its counterpart in Appendix Table 3, the effect of joining is about 10 times larger than the yearly trend for distance measures and about 15 times larger for in-state and foreign percentage.

Table 4: Effect of Common Application on Geographic Integration

	(1)	(2)	(3)	(4)	(5)	(6)
	dist	ln(dist)	dist	ln(dist)	instate%	foreign%
camember	30.165***	0.104***	55.123***	0.070***	-2.332***	0.621***
	(6.195)	(0.025)	(8.553)	(0.010)	(0.480)	(0.169)
Observations	20629	20176	20176	20176	20629	20629
Clusters	1652	1628	1628	1628	1652	1652

Dep var in cols 1 and 2 is distance per student.
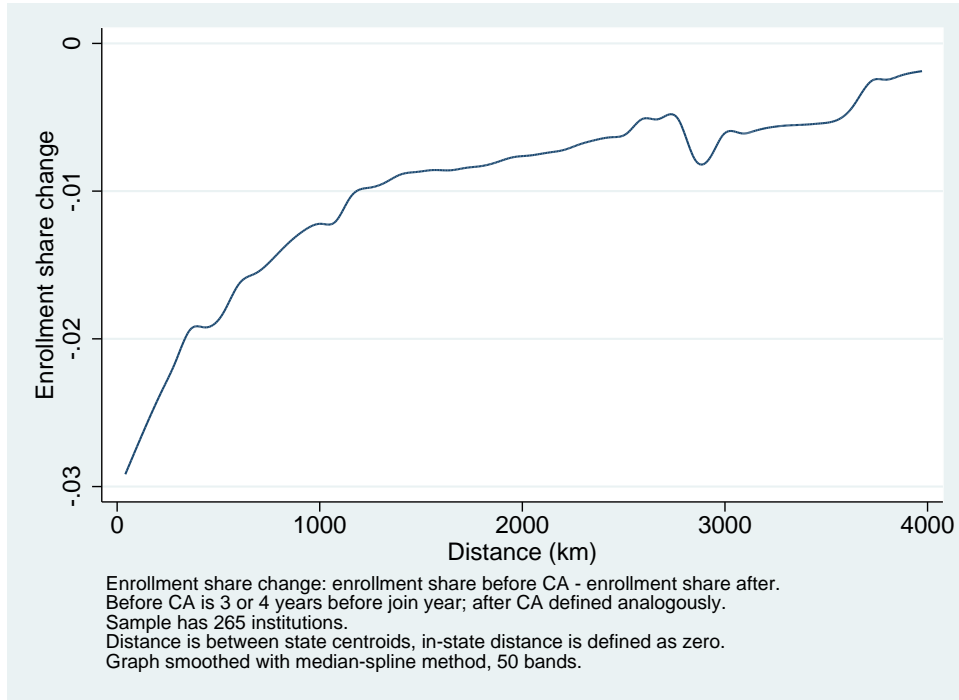Dep var in 3,4 is distance per out-of-state student.
In cols 5+6 shares are multiplied by 100.
All spec include institution FE; std errors clustered by institution.

As a further analysis of the effect of joining CA on distance traveled, we now consider the change in the entire distribution of a college's enrollees over distance. To do so, we restrict our sample to only those institutions which joined the CA and which have migration data both three or fours years before and three or four years after joining, depending on whether the institution joined in an odd or even year. We then sum the enrollees across all universities in each period (pre,post) and calculate the percentage coming from each state-to-state distance. This allows us to calculate two cumulative distribution functions (CDF), where each bin of the CDF represents a given state-to-state distance and we are calculating the percentage of students traveling that distance to all universities in a period. We then plot the difference between the before and after CDF in Figure

46

29.[28] The largest difference occurs at zero, indicating that most of the effect comes from a 3 percentage point decrease in the in-state percentage. The slope of this differenced CDF increases sharply and approaches zero, so that a distance of 1200 kilometers the change is less than 1 percent, and then flattens. This shape suggests that CA entry increases the distance traveled by enrollees by mostly increasing the number of enrollees from nearby states.

Figure 29: Effect of CA Entry on Enrollment Share Change by Distance

Enrollment share change: enrollment share before CA - enrollment share after.
Before CA is 3 or 4 years before join year; after CA defined analogously.
Sample has 265 institutions.
Distance is between state centroids, in-state distance is defined as zero.
Graph smoothed with median-spline method, 50 bands.

---

[28]We truncate the graph at 4000km since the share of students coming from a greater distance is very small in both periods.