No Place at Home: Are Nonresident Students Crowding Out Resident Students at Public Universities?

Abstract

Demand for higher education has grown substantially in recent years. From 2000 to 2015, enrollment in four-year public universities grew from 4.8 million to 6.9 million students, a 44% increase. This paper uses data from the Integrated Postsecondary Education Data System (IPEDS) to examine how this large increase in demand has fundamentally changed the landscape for public higher education in the United States. Using a shift-share instrumental variables strategy, I estimate causal impacts of the rate at which nonresident students crowd-out resident students by constructing an instrument that simulates cross-state student migration. These results find that each additional nonresident student crowds-out 0.55 resident students, driven solely by foreign students. These findings shed new light on the scale and scope of university reactions to increased nonresident demand, and the unintended consequences for resident students.

1 Introduction

Universities serve a vital purpose in the United States' economy. In 2016, 2/3 of the U.S. labor force over the age of 25 had at least some college, while around 40% had at least completed college. In 1992, only 27% of the labor force had completed college (Bureau of Labor Statistics). Put simply, the U.S. labor market demands college-educated workers, and universities are enrolling and educating larger populations to meet demand.

Public universities enroll a considerable number of these future workers. In 1995, degree-granting public universities in the United States enrolled 11.09 million students; in 2016, this number was 14.58 million, a 31% increase (National Center for Education Statistics). Yet public funding for higher education has not grown to accommodate this increased demand for college. From 2000 to 2015, state appropriations for higher education decreased from \$68 billion to \$56 billion in real 2009 dollars, a decrease of roughly 18% (National Center for Education Statistics). This means fewer resources are being spared to serve a larger student body, leaving universities with trade-offs regarding where to put their money to its greatest use.

A key margin that attracts the attention of both policy-makers and researchers is enrollment composition. The determinants of any given student's admission and subsequent enrollment are many: test scores, high school GPA, intended major, and diversity are large components of this process; yet recently financial pressure has led residency status to likely be another component of the admissions criteria. On its own, enrolling nonresident students should be an efficient allocation for the university and the prospective student: for the university, the student contributes to the diverse atmosphere on campus, adds to the prestige of the university, and provides revenue above what resident students pay. For the student, revealed preference clearly shows they value attending the university. This could be due to the academic reputation of the school, location preference, familial legacy reasons, or many other preferences.

The friction that makes this transaction potentially inefficient is the differing objective functions of the state and the university regarding higher education. Individual universities have unique maximization problems, but the average school is probably interested in maximizing academic quality of their student body subject to various constraints like revenue generation, diversity goals, capacity limits, research output, and prestige. All of these constraints would be considered in determining the composition of incoming classes. But in an era where state funding for higher education is declining, universities are increasingly prioritizing revenue generation from their student body (Mathias (2020)). This behavior has also been documented by several news outlets, with articles appearing in the *New York Times* and *The Sacramento Bee* highlighting increased nonresident enrollment in the University of California system.[1]

This behavior by the university is different than the role the state would like the university to fulfill. The state would like the public university to provide an efficient return on their investment of state funding, producing highly-educated workers to contribute to the state economy. This sentiment has percolated to top political talking points, making an appearance during the 2016 Presidential Campaign. The following quote comes from Curs and Jaquette (2017), and neatly summarizes the prevailing opinion of states:

> We have got to get back to using public colleges and universities for what they were

[1] *NYT*: https://www.nytimes.com/2016/07/08/us/public-colleges-chase-out-of-state-students-and-tuition.html. *SacBee*: https://www.sacbee.com/news/local/education/article160029439.html.

intended...If it is in California, for the children of California. If it is in New York, for the children of New York.

–Hilary Clinton, during 2016 Presidential Campaign

The trouble for states is that there is no clear policy solution to realign incentives for universities. Cutting state appropriations tightens the revenue constraint for schools, and could lead to further nonresident enrollment to recoup lost revenue (Deming and Walters (2017); Bound et. al (2017)). Raising out-of-state tuition could decrease nonresident demand, but nonresident applicants become even more attractive due to the higher revenue they provide. The prevailing solution seems to be capping nonresident enrollment at a certain percentage.[2]

Previous work has examined either one of the two nonresident sub-groups. Shih (2017) examines foreign graduate enrollment's effect on domestic graduate enrollment, finding that the former cross-subsidizes the latter. Work by Curs and Jaquette (2017) focused solely on out-of-state students, with particular emphasis on US News Top 50 schools. Li (2018) examines nonresident students' impact resident students at the University of Missouri, finding a negative correlation between foreign enrollment and resident persistence. Of these four papers, only Curs and Jaquette (2017) sought to identify crowd out. Shih's paper concerning graduate students is not really applicable to tell a story of crowd out, as graduate students operate under different enrollment rules than undergraduates. Li's paper identifies a margin other than enrollment where nonresident students could affect resident students, but looking only at one school raises external validity issues.

This paper uses data from the Integrated Postsecondary Education Data System (IPEDS), spanning the years 2000-2015, to quantify the extent to which nonresident students are crowding out resident students at four-year public universities. I employ an instrumental variables identification strategy to find causal estimates of crowd out, creating a distinct IV for both out-of-state and foreign students. I find that two additional nonresident students crowd out one resident. Foreign students are solely responsible for this trend, with no robust evidence of out-of-state crowd out or residents.

This paper extends the literature in meaningful ways. First, it quantifies crowd out on two distinct fronts: out-of-state (but still domestic) students and foreign students. Identifying different marginal effects across these groups is valuable as it can give a primary culprit of crowd out—if one exists at all—informing the discussion and potentially providing targeted solutions to address the

[2]See UW-Madison's recent relaxation of this policy. https://www.insidehighered.com/admissions/article/2019/12/09/uw-madison-shifts-state-enrollment-commitment.

problem. Second, it brings a robust identification strategy to capture causal estimates. The out-of-state student IV creates exogenous cross-state aggregated student flows using past enrollment trends to deal with endogeneity in OLS estimates to give an accurate rate of crowd out. The foreign IV aggregates foreign student stocks at the state-level before dividing them to schools based on their enrollment size. Third, I am the first paper to my knowledge to test the rate at which nonresidents crowd each other in; specifically how foreign students could crowd in out-of-state students.

The rest of the paper is organized as follows: section 2 provides a brief theoretical background, followed by a literature review of the empirical literature. Section 3 describes the data, and section 4 introduces the methodology. Section 5 presents main results. Section 6 checks the robustness of the IV specification. Section 7 concludes.

2 Theoretical Background and Literature Review

2.1 Theoretical Background

The concept that a university would prefer a higher-paying student to a lower-paying student, all else equal, is intuitive. However, universities balance many goals in making their admissions decisions, so this easy-seeming decision becomes more difficult to evaluate. Universities must address issues of revenue generation, school quality and prestige, pressure from state legislators, and various outside factors that make determining a universal objective function for the university near impossible. Danziger (1990) and Kerkvliet and Nowell (2014) have made theoretical contributions to university admissions preferences, presenting models where universities choose to admit and enroll students based on parameters like tuition and state appropriations. Danziger argues that universities have different preferences for students who pay different tuition rates, and Kerkvliet and Nowell assert that universities should manipulate acceptance rates of first-years based on state funding and tuition. Tiffany and Ankrom (1998) explicitly models a university budget constraint to determine price discrimination through tuition rates. Hoxby (2012) takes a different approach, providing a descriptive paper that models a university interested in maximizing social contributions through its endowment.

To more firmly fix the intuition of potential crowd-out, I present a very simple theoretical model. Consider a simple model where a university benefits from the enrollment of a student only through the tuition revenue they provide and the prestige they contribute, which is a positive function of the ability of the student. The costs of enrolling this student are comprised of a fixed cost (think room and board), and a variable cost that is negatively related to student ability.[3]

Define T_r, $r \in \{I, O\}$, to be the tuition paid by the student, where r denotes residency status. I stands for in-state students, and O stands for out-of-state students (which encompass all nonresident students), Assume out-of-state students pay a higher tuition than do in-state students, so $T_O > T_I$. Let γ_j stand for the ability of student j, where γ is i.i.d. across both the in-state and out-of-state student population. Larger values of γ indicate a student is more academically prepared than a student who has a lower value of γ. Define c_j to be the fixed cost of enrolling student j, where $c_i = c_j = c$, $\forall i \neq j$. Lastly, let $\rho(\gamma_j)$ define the variable cost of student j's attendance, where $\rho'(\gamma) < 0$, indicating that higher-ability students have lower variable costs, and vice versa. For simplicity, I will call this function ρ_j in the following equations.

I assume that the university chooses to admit a student if the marginal benefit of that student enrolling is greater than or equal to the marginal cost, which leads to two first-order conditions:

$$\begin{cases} T_O + \gamma_o \geq c + \rho_o, \ \forall o \in O \\ T_I + \gamma_i \geq c + \rho_i, \ \forall i \in I \end{cases} \tag{1}$$

where O and I are the in-state and out-of-state applicant populations, respectively. The two conditions in (1) simply specify that the benefit the university receives—tuition plus ability—must exceed the perceived cost for the university to decide to admit a student.

From (1), it is clear that a university will always admit an out-of-state student if $\gamma_i = \gamma_o$, due the higher tuition received. This is obviously not always the case at public universities, To account for this, while adding minimal complexity to the model, I introduce a scalar $\alpha > 1$ that captures a public university's desire to preferentially enroll resident students.[4] The updated first-order conditions are:

$$\begin{cases} T_O + \gamma_o \geq c + \rho_o, \ \forall o \in O \\ T_I + \alpha\gamma_i \geq c + \rho_i, \ \forall i \in I \end{cases} \tag{2}$$

where the inclusion of α is the only difference from (1). From this, we can combine the equations and solve for the ability parameter of in-state students γ_i. This gives:

[3] Think academic support services. Less-prepared students utilize these services more, costing the university more in the process.

[4] I assign α as a scalar for simplicity; a more realistic extension would be to assume that α is a monotonically increasing function of γ_i.

$$\gamma_i = \frac{1}{\alpha}\bigg(\gamma_o + (\rho_i - \rho_o) + (T_O - T_I)\bigg)$$

From this equation, it is clear that larger tuition differentials between out-of-state and in-state students forces γ_i to be larger; implying that the larger the nonresident-resident tuition gap, the smarter residents need to be to be competitive, making it more difficult for residents to be admitted. On the other hand, larger preference for resident students, represented by α, lower the ability needed for residents to be competitive, making admission easier.

Another way to view this is to solve for tuition differentials between out-of-state and in-state students. Combining equations in (2) and solving yields:

$$T_O - T_I = (\alpha\gamma_i - \rho_i) - (\gamma_o - \rho_o)$$

The respective $\gamma - \rho$ terms on the right-hand side of the equation represent prestige minus cost balances for the university. The intuition of this expression is that the higher a university values resident students (α), or the more net prestige residents contribute over nonresidents, necessitates higher out-of-state tuition rates for this to be viable for the university. Put another way, larger gaps between out-of-state and in-state tuition allow universities to enroll less-qualified out-of-state applicants. As mentioned earlier, this kind of behavior in the University of California system has been noted in multiple news outlets.

With a very simple theoretical model established, I now turn to summarizing empirical work in this subject, which is much larger than the theoretical literature.

2.2 Literature Review

The majority of work in this literature has been empirical. Koshal and Koshal (2000) look at appropriations and tuition and find a negative relationship, but fail to distinguish between in-state and out-of-state tuition. Groen and White (2004) use the sample of colleges from the *College and Beyond* dataset to examine how public and private universities set admissions cutoffs for in-state and out-of-state students, but the data is limited to four public universities and the point estimates are imprecise. Rizzo and Ehrenberg (2004) use IPEDS data to look at how tuition determines enrollment composition at flagship state universities. The authors argue that these universities use resident students to increase revenue, while nonresident students are admitted solely to improve student quality. This result is in contrast to what theory would predict. Adkisson and Peach

(2008) employ IPEDS data to examine the relationship between nonresident tuition and nonresident enrollment at land grant universities. The authors find that there is a positive relationship between the two variables, however their instrument likely violates the exclusion restriction (they instrument with in-state tuition). Curs and Singell (2002) use a unique dataset comprised of admissions and enrollment to the University of Oregon spanning from 1996-2000. The authors find that non-subsidized loans have differential impacts for resident and nonresident students: resident students are more likely to enroll as loans increase, while nonresident students are less likely to enroll.

This paper follows closely follows work by Curs and Jaquette (2017), Li (2018), and Shih (2017). Curs and Jaquette (2017) follow their 2015 paper to look at another implication of cutting state budgets: the possibility of nonresident students crowding out resident students. The authors fail to find crowd out except at prestigious (top 50 ranking) public universities. This paper uses a gravity model IV which is relatively weak, but their lack of precision is due primarily to a relatively small sample size. Li (2018) uses University of Missouri data to examine nonresident crowd-out. He finds no evidence that residents are pushed out by nonresident students, but does find a negative relationship between foreign enrollment and resident persistence. Shih (2017) examines foreign crowd-out for graduate students. He finds that foreign students actually crowd-in residents, where extra revenue from foreign students serves to enroll more residents.

This paper synthesizes aspects from previous papers and extends the literature with a more comprehensive view of the two groups that may crowd-out in-state students—out-of-state but still domestic students and foreign students—while also bringing a more robust methodology. These two contributions build on the prior literature and provide robust causal estimates to further our understanding of crowd-out at public universities.

3 Data

I use data spanning from 2000-2015 collected from the Integrated Postsecondary Education Data System (IPEDS), a series of surveys collected by the National Center for Education Statistics (NCES) for the United States Department of Education. I limit my sample to 4-year public institutions in the United States who are not tribal colleges, religious seminaries, or did not provide a Carnegie classification.

IPEDS structures their data into surveys, with balance sheet variables belonging to the Finance survey, enrollment variables belonging to the Fall Enrollment survey, tuition and admission information in the Institutional Characteristics survey, etc. Each university is assigned a unique identifier that is constant across surveys and years.

The Fall Enrollment surveys provide information about various aspects of the institution's enrollment. I use the "Residence and Migration of First-Time first-year" sub-survey for detailed information about the incoming first-year class. This survey counts the number of first-time first-year students attending university i from each state s, as well as the total first-year first-year enrollment. This particular survey provides yearly data after 2000, but before provides data only in even-numbered years. This feature is why the analysis sample begins in 2000, although I do collect data from earlier to analyze pre-trends of the data for the instrument, described in more detail in the next section.

I choose to classify entering first-year students into three categories: in-state, out-of-state, and foreign. In-state first-years are those who reside in the same state as the university they attend; I create this variable by merging FIPS codes from the directory sub-survey and matching FIPS codes provided in the enrollment survey. I use the term "resident" interchangeably with in-state throughout the paper. Out-of-state students are then those who are neither in-state nor foreign, so they are created by subtracting in-state and foreign students from the total first-year headcount. At times I will group together out-of-state and foreign students, I call this group "nonresident" students.

State-level covariates come from a variety of sources. Annual state unemployment data is collected from the Bureau of Labor Statistics. State GDP by year comes from the Bureau of Economic Analysis (BEA). State population age 17-25 by year comes from the US Census Bureau. Wage income measures come from the IPUMS dataset, specifically the ACS. I deflate dollar values by using the Higher Education Price Index (HEPI).

Table 1 provides summary statistics for enrollment and tuition measures. Average nonresident tuition (\$13,820) is just over 2.5 times as large as average resident tuition (\$5,223). On the other hand, average first-year resident enrollment (1,399) is over four times larger than average first-year nonresident enrollment (308).

Figures 1 and 2 provide additional information regarding the trends of enrollment and tuition over time. Figure 1 plots average tuition and enrollment for resident and nonresident students over the sample. The blue lines indicate resident values, while red lines are nonresident averages. The solid lines are tuition measures, the dashed lines are first-year enrollment averages. Figure 1 shows that tuition for both groups is steadily growing over time, even after deflating values to adjust for the rapidity of price changes in higher education. First-year enrollments are also growing, and there is no immediate evidence of nonresident student enrollment growing at the expense of resident enrollment.

Figure 2 plots tuition revenue growth from resident and nonresident students. I standardize values so that both revenue measures equal one in 2000. Tuition revenue from each group at least doubles by the end of the sample, with nonresident tuition increasing by almost 150% by 2015. Figure 2 shows what figure 1 does not, which is that nonresident students are increasingly replacing residents as the primary revenue source. This trend is contrary to what Rizzo and Ehrenberg (2004) finds.

4 Methodology

4.1 Baseline OLS

I use a panel fixed-effects OLS model to examine the crowd out of nonresident students on resident students. The regressions are:

$$InState_{it} = \beta Nonresident_{it} + \theta X_{it} + \gamma W_{st} + \phi_i + \phi_t + \varepsilon_{it} \tag{3}$$

where i denotes a university and t denotes year. β is the coefficient of interest, ideally giving the causal impact of nonresident student enrollment on resident enrollment.

X is a vector of university-level time-varying controls. This vector contains the number of applications the university receives in a given year, the 25^{th} and 75^{th} percentile of ACT scores for admitted applicants, and the percentage of students at the institution who receive any form of financial aid—whether it be federal, state, or local. Applications provide a measure of demand for the university, and can therefore affect both resident and nonresident enrollment measures. IPEDS does provide admissions information, but admission numbers are likely endogenous in that the university can manipulate this to an extent to influence revenue generation (admissions are also not disaggregated by residency status). The ACT percentiles provide a proxy for student body quality, alluded to in equations (1) and (2) as probable factors in admissions decisions. The ACT measures I use are relatively static, as evidenced by the small standard deviation and inter-quartile range in Table 1, but there is enough variation to identify them even with the inclusion of university fixed effects. The percentage of students who receive any form of aid is a coarse control for student socioeconomic status and ability. IPEDS does not provide a more granular measure of this control, so the “any aid” percentage is the best measure available with the given data. The percentage of students who receive any aid can be correlated with enrollment (through student quality channels or through measures of which demographics they primarily serve), thus warranting its inclusion in the regression.

W is a vector of state-level covariates that includes the unemployment rate, labor force participation rate, state GDP per capita, state minimum wage, median income, average income by educational attainment–less than high school, high school, and college or more–for those aged 18-35, 36-50, and 51-65, poverty rates for ages 18-35, 36-50, and 51-65, the population of neighboring states, and the log of the population aged 17 to 25. These controls are included to account for shocks that may be correlated with nonresident enrollment and the outcome resident enrollment. Income and poverty rates are correlated with the ability of parents to be able to send their children to college and the attractiveness of attending college in a given state in general—thereby being related with determinants of university demand.

ϕ_i are institution fixed effects that capture university-specific, time-invariant unobservables and ϕ_t are year fixed effects, which control for nation-wide trends. ε_{it} is the error term assumed to be orthogonal to the covariates.

Standard errors for all regressions are clustered at the state level to adjust for correlation between observations within a state.

The largest difference between two papers closely related to this one—Jaquette and Curs (2015) and Bound et al. (2016)—is the inclusion (or exclusion) of university-specific controls. Jaquette and Curs (2015) include nearly all cost measures included in the IPEDS data, and also include tuition as controls in looking at enrollment measures. Their rationale is that cost measures are the most accurate measure of university resources, although many of their controls are arguably endogenous. Bound et al. (2016) neglect time-varying institution controls, choosing to instead control only for state-level covariates. They do not discuss their choice of controls, but I would justify it as an extreme caution to not include any potentially endogenous covariates, even if this comes at the expense of model fit and omitted variables bias. I take the middle ground between the two papers in choosing to include university-level covariates, but far more cautiously than Jaquette and Curs (2015). Interestingly, Jaquette and Curs (2015) do not include applications as a control in their model, even though it is a straightforward measure of demand for a university. My selection of only three university-level controls reflects both the need to quantify determinants of university demand, argued for in Jaquette and Curs (2015), while also following the caution of Bound et al.

(2016) in carefully choosing what to include in the model. In this respect, my synthesis of the two prior works contributes to the literature as a framework for choosing covariates in the higher education literature.

4.2 Instrumental Variables

Although I begin with OLS, it is likely that nonresident enrollment is endogenous in the regression above. This stems from the fact that there are very likely unobservable (at least to the researcher) characteristics that co-determine both resident and nonresident enrollment, and thus $\hat{\beta}$ is biased. The direction of the bias is unclear, though.

To begin, $\hat{\beta}$ may be biased upwards–that is, in reality there is no crowd out and the true β is close to zero—due to the researcher imposing a form of capacity constraint for the university. If the university is enrolled below this constraint, then there is no need for an additional nonresident student to force out a resident student, or at least not at a large margin. Worse, if the researcher incorrectly specifies a capacity constraint, then they may be structurally creating one through model mis-specification. Controlling for something like first-year or university enrollment in a regression—aside from both of these being endogenous—would create this collinearity issue, and lead to an artificially high crowd out rate. As seen in equation (3) above, I choose to not include any measure of a capacity constraint for the university. University enrollment is growing over time, so any capacity constraint for a university is a soft cap at best.

The OLS estimates could also be biased downwards due to our inability to control for time-varying university characteristics that impact nonresident inflow and also resident enrollment (aside from what is in equation (3)). Amenities in the nearest city, or a surging academic department could increase demand from both groups, necessitating trade-off from the university, but our inability to quantify this could lead the estimates of crowd out to be artificially low relative to the true β.

Therefore I turn to instrumental variables to address the issue of endogeneity in the regression. The goal is to find a suitable method to isolate exogenous variation in nonresident student enrollment that does not directly effect resident enrollment at the university-level, yielding a causal estimate of crowd out.

I use two variations of a "shift-share" instrument to create exogenous variation for nonresident enrollment (one for foreign enrollment, and one for out-of-state enrollment). To understand the motivation for using a shift-share instrument, it is useful to think of university-level nonresident enrollment as being comprised of two parts: a state-level stock of nonresident students and a university-specific share of that stock. A university's nonresident enrollment is then its share times the state-level stock. In shift-share terms, the nonresident stock that changes yearly is the "shift" element, while the "share" the university is assigned is held fixed. Therefore, the sole source of nonresident variation for a university comes from the changing state-level stock of nonresident students.

The advantage of this approach is that endogenous university behavior, such as a changing propensity to enroll nonresident students, increased university demand allowing greater admission scrutiny, etc., are eliminated by holding the share constant. However, recent literature from Borusyak, Hull and Jaravel (2018) and Goldsmith-Pinkham, Sorkin and Swift (2018) have shown that this share which was once considered orthogonal by design must be validated through testing. Practically, the researcher must show that the chosen share must be uncorrelated with pre-trends of the data (and thus orthogonal), otherwise bias will be exacerbated in the IV estimates (Tables 2 and 3 address this below).

The other threat to causal identification is that the state-level stock of nonresident students may be correlated with unobservable university demand determinants. A practical example of this violation of the exclusion restriction would be a state with only one public university in the sample. In that case, university and state nonresident enrollment would be identical, and therefore endogenous. I perform robustness checks in section 6 to test this assumption of orthogonality of the state-level shift element and the robustness of estimates.

To instrument for foreign enrollment at a given university, I aggregate foreign enrollment at the state level, and then assign each school in the state a share of the total pool. This share is the school's share of the *total state enrollment* in the base year 2000. IPEDS only provides residency status for first-year students, so this estimate of the yearly state pool of foreign students in state s is $StateForeign_{st} = \sum_{k \in s} Foreign_{kt}$, where $Foreign_{kt}$ is the observed foreign first-year enrollment at university k in state s in year t. The share of university i in state s is $s^{enrl}_{i,2000} = \frac{Enroll_{i,2000}}{\sum_{k \in s} Enroll_{k,2000}}$. The instrumented foreign first-year enrollment of university i in year t is then $\widehat{Foreign}_{it} = s^{enrl}_{i,2000} \times StateForeign_{st}$. Unfortunately, IPEDS does not provide country-of-origin information for foreign students, so this aggregation at the state level is the best available measure to create the exogenous "shift" factor before assigning shares to individual universities.[5] As mentioned

[5]Because this aggregate foreign measure comes from observed foreign enrollments, which may or may not be correlated with university-specific shocks, there is a possibility that aggregate foreign enrollment is still endogenous.

earlier, the tests verifying the exogeneity of the chosen share recommended by Borusyak, Hull and Jaravel (2018) and Goldsmith-Pinkham, Sorkin and Swift (2018) must be validated. The results for these tests will be provided after the description of the out-of-state instrument.

The out-of-state instrument uses residency data of out-of-state, but still domestic, first-year students to create a cross-state flow to instrument for observed out-of-state enrollment, instead of a moving stock as in the foreign instrument described above. IPEDS gives information about state residency of first-year cohorts upon enrolling at their chosen school. As an example, for UC Davis, IPEDS provides the number of students from Alabama, Colorado, Utah, etc. enrolled as first-year students during that particular year. Using this state-of-origin information, I pool the number of students from each state who went out of state in a given year by using these counts. Formally, call the number of students from state s who go out of state in year t OUT_{st}. The instrument assigns these students in OUT_{st} to the other states based on past cross-state patterns, but independent of trends in the sample. To assign the number of students in OUT_{st} (from state s) who go to state c, I multiply OUT_{st} by the proportion of students from state s who comprised state c's total out-of-state enrollment in 2000, that is $\widehat{OUT}^c_{st} = OUT_{st} \times \frac{OUT^c_{s,2000}}{\sum_{j \neq c} OUT^c_{j,2000}}$. The assumption is that this initial flow of students from s to c in 2000 is orthogonal to determinants of university-level demand in later years in the sample (which is tested in Table 3). Finally, aggregating over all states who sent students to state c in 2000 gives the imputed state-level out-of-state enrollment for state c in year t as $\widehat{OUT}^c_t = \sum_{j \neq c} \widehat{OUT}^c_{jt}$. The year 2000 was chosen to assign cross-state flows as it is the first year in the sample.

A benefit of using 2000 as the migration determinant is that it is free from subsequent changes in migration behavior over the sample that could bias estimates (such as a state changing its nonresident enrollment policy, etc.). The drawback is that if a particular cross-state assignment did not occur in 2000, i.e., if no students from Tennessee went to Kentucky for some reason, then the instrument will never assign Tennessee students to Kentucky for the duration of the sample. The same logic applies for any outlier flow, whether high or low. This failure to update based on observed flows, like the foreign instrument does with the aggregate foreign student stock, will reduce the power of the instrument and potentially lead to weak instrument issues.

The final step to complete the instrument for university-level out-of-state enrollment is to assign each school its share based on 2000 enrollment, $\widehat{OutState}^c_{it} = s^{enrl}_{i,2000} \times \widehat{OUT}^c_t$.[6]

[6] The $s^{enrl}_{i,2000}$ term is the same as described above. For convenience, I will omit the c in $\widehat{OutState}^c_{it}$ in the regression

The out-of-state instrument requires two tests for its validity. First, the baseline enrollment share chosen must be uncorrelated with pre-trends—this is the Goldsmith-Pinkham, Sorkin and Swift (2018) and Borusyak, Hull and Jaravel (2018) test, which is identical for both the foreign and out-of-state instrument as they use the same share. Second, the flow of students across states in 2000 must also be uncorrelated with pre-trends. Because each state sends students to multiple other states, as well as receives students from multiple states, I use the aggregate measure $\widehat{OUT_{s}}_{,2000}$ to test against the statewide pre-trends.

Lastly, nonresident enrollment, which is the sum of out-of-state and foreign enrollment, is instrumented by using both the out-of-state and foreign measures as excluded instruments in all specifications.

Table 2 presents results for the test of the exogeneity of enrollment share in 2000 to existing pre-trends. The table is divided into two parts: the first two columns test enrollment pre-trends, while columns 3 and 4 test tuition pre-trends. Columns 1 and 3 examine resident components; columns 2 and 4 present nonresident trends. The pre-trends are constructed as percent changes in either tuition or enrollment over the 1990s. For tuition, this is from 1990 to 1999; for enrollment, 1992 to 1999.[7] Each regression includes all controls from equation (3) for the year 2000, as well as 1992 undergraduate enrollment for the university which acts as a pseudo fixed effect to capture time-invariant unobservables that may be correlated with enrollment share in 2000. The results of Table 2 find no correlation between enrollment share and enrollment or tuition pre-trends, confirming the chosen share component is exogenous.

Table 3 tests the independence of the constructed out-of-state shift component, $\widehat{OUT_{s}}_{,2000}$, on tuition and enrollment pre-trends. Due to the aggregated nature of $\widehat{OUT_{s}}_{,2000}$, the tuition and enrollment pre-trends are constructed at the state level, creating a cross-section of 50 states.[8] The aggregate enrollment and tuition pre-trends are created by weighting each university's enrollment or tuition by its share of state enrollment in 1992, the first year such data is available. Therefore "state" tuition in 1990 is the weighted average of all university-level resident/nonresident tuition for that year, and "state" enrollment is the weighted average of university-level resident/nonresident first-year enrollment. The pre-trends are percent differences in these measures: for tuition, from

tables.

[7]IPEDS does not provide first-year enrollment data for odd-numbered years before 2000, so the 1999 enrollment measure is the linear interpolation of 1998 and 2000 values. 1992 is the first year that residency status for first-year students is given.

[8]Washington D.C. is excluded.

1990 to 1999; and for enrollment, from 1992 to 1999. Each regression includes all controls from equation 2 for the year 2000. Similar to the structure of Table 2, Table 3 is comprised of resident and nonresident results for both tuition and enrollment. Columns 1 and 2 show enrollment results and columns 3 and 4 are tuition pre-trends. The odd-numbered columns are resident regressions, and the even-numbered columns are nonresident results. The results find there is no correlation between the instrumented cross-state flow and statewide enrollment or tuition trends. This means the constructed shift-component of the out-of-state instrument can more confidently be viewed as an exogenous demand shock, and thus capture causal estimates of crowd-out.

Table 4 shows the first-stage results for the instruments. Columns 1 and 3 use only university and time fixed effects as controls, while columns 2 and 4 include all controls shown in equation 2 above. As can be seen, both are statistically significant with sufficiently strong F statistics. The out-of-state IV is weaker than the foreign IV because it is simulating the flow of students across the states, instead to assigning shares to observed aggregate foreign enrollment. The trade-off is that the out-of-state instrument, although weaker, is less likely to violate the exclusion restriction than the foreign instrument because there is a possibility that aggregate state foreign enrollment is correlated with ε described in equation (3). The interpretation of the coefficients is one-for-one (or student-for-student), so the point estimate in column 2 of Table 4 shows an increase of one student in $\widehat{OutState}_{it}$ leads to a 1.06 student increase in observed first-year out-of-state enrollment. The column 4 estimate finds that an increase of one in instrumented foreign first-year enrollment results in an increase of 1.14 first-year foreign students.

5 Results

Table 5 looks at results from regressions of nonresident first-year enrollment on resident first-year enrollment. Columns 1 and 2 focus on nonresident enrollment, columns 3 and 4 examine the out-of-state component, and columns 5 and 6 foreign enrollment. Coefficient interpretation for this table is student-for-student, but an elasticity is provided below each specification as well. Looking at columns 1 and 2, both the OLS and IV estimates for nonresident crowd-out are statistically insignificant. The corresponding point estimates are also small, so it appears that nonresident first-year students do not push out resident first-years in any capacity.

Columns 3 and 4 focus on the out-of-state component of nonresident enrollment, meaning students who are still US citizens but do not reside in the state in which they attend school. Again the estimates are statistically insignificant, but the point estimate from the OLS specification (column 3) and the IV estimate (column 4) are very different.

Columns 5 and 6 examine foreign first-year results. The OLS estimate in column 5 finds a significant crowd-out of resident first-years by foreign students: one additional foreign student leads to 0.76 fewer resident students. The foreign IV result of column 6 is close to zero and statistically insignificant, which suggests that foreign first-year students do not crowd-out residents.

The results of Table 5 find that nonresidents do not crowd-out residents, at least for the first year of enrollment. These first-year regressions do not account for several factors that could cause crowd-out in aggregate. First, any capacity constraint of the university would be more prevalent at the university-level, as opposed to the first-year level. There could also be transfer activity that subsequently pushes residents out in later years through increased competition for seats in classes past the first year.

To address this, I impute enrollment measures for residents and nonresidents at the university level by adding first-year enrollments for each respective group over the past three years (and including the current year) to create an intermediate variable. Let the imputed intermediate measure of residents attending university i in year t be called $\widehat{RES}_{it}$, and let observed first-year resident enrollment at university i in year t be called res_{it}. Then $\widehat{RES}_{it} = res_{i,t-3} + res_{i,t-2} + res_{i,t-1} + res_{it}$. The imputed nonresident measure $\widehat{NONRES}_{it}$ is constructed in the same fashion, summing up nonresident first-year enrollments for the past three years in addition to current nonresident first-year enrollment. Creating this variable for the years 2000-2003 requires collecting first-year enrollment data prior to 2000, which is only available in 1996, and 1998. I linearly interpolate values for 1997 and 1999 using the surrounding years. I then scale these two intermediate variables so that their sum equals observed university enrollment. This is done by dividing observed university enrollment, $ENRL_{it}$, by the sum of $\widehat{RES}_{it}$ and $\widehat{NONRES}_{it}$, forming a scalar $k_{it} = \frac{ENRL_{it}}{\widehat{RES}_{it} + \widehat{NONRES}_{it}}$. The final imputed measure for university resident enrollment is then $RES_{it} = k_{it} \times \widehat{RES}_{it}$.[9] This is done to approximate the increase in each population received from transfer students, while factoring in decreases to population from attrition or transfer away to another university. The assumption is that this is proportional among the two groups—there is only one k_{it} per university-year pair, not a separate scalar for each sub-group. The respective instruments for this next set of results are identical to above, although the strength of the instruments will change because I am now instrumenting for different populations.

[9]The nonresident imputation is constructed in the same manner.

Table 6 presents results for crowd-out at the university-level. The format of Table 6 identical to that of Table 5: columns 1 and 2 provide estimates for nonresident crowd-out, columns 3 and 4 out-of-state students, and columns 5 and 6 foreign crowd-out. The updated KP F-statistics are below each IV regression, and elasticities are provided for each specification. The coefficient interpretations are student-for-student.

Columns 1 and 2 examine nonresident crowd-out of resident students. The OLS estimate in column 1 is still statistically insignificant, but the IV estimate in column 2 is now much larger and statistically significant. The interpretation of the point estimate in column 2 is that each additional nonresident crowds-out 0.52 residents, or that a 10% increase in nonresident enrollment causes a 1.1% decrease in resident enrollment. To put that point estimate in perspective, Curs and Jaquette (2017) find an elasticity of -0.18 regarding nonresident to resident substitution. Using their summary statistics, this translates to a point estimate -0.699, about 30% larger than what is found here.[10] It should be noted that they find this effect only for prestigious universities (top 50 *U.S. News* ranked) and only at 7% statistical significance.

Columns 3 and 4 look at out-of-state crowd-out. Similar to the OLS estimate from Table 5, the OLS estimate of university-level out-of-state crowd-out is positive and statistically insignificant. The IV estimate of column 4 is much larger but still statistically insignificant.

Columns 5 and 6 present results for foreign student crowd-out. Unlike the results of Table 5, here foreign crowd-out is larger and significant across both the OLS and IV specifications. The point estimates are very similar at around -1, indicating a one-for-one crowd-out of foreign students for resident students. This translates to a very small elasticity of 0.022 because one additional foreign student is a large percentage increase in the foreign population, while a one student decrease in the resident population is a much smaller percent.

Table 7 presents regressions that examine whether foreign students crowd in out-of-state students. This is a check of the relative independence of the enrollment determinants of the two sub-populations, or if the estimates of nonresident enrollment presented in Tables 5 and 6 are skewed by a collinearity of the out-of-state and foreign instruments. As the results of Table 7 show, foreign enrollment does not appear to drive out-of-state enrollment in any way. Looking at the first-year estimates (columns 1 and 2) and the university estimates (columns 3 and 4), the coefficient on foreign student enrollment is close to zero and statistically significant.

The results of Tables 5 and 6 find that residents are being crowded out by nonresidents, but

[10] Their mean of resident students is 2,519 and mean of nonresident students is 648.

this is only occurring in the aggregate and solely by increased foreign student enrollment. First-year enrollment specifications find no evidence of any nonresident sub-group crowding out residents, but examining university-level crowd-out tells a much different story. At the university-level, there is crowd-out coming from foreign students, and at a very high rate. Table 7 adds to these results by finding that foreign enrollment is more or less independent of out-of-state enrollment, and is the sole driver of crowd out. These results show that this increase in demand for higher education has negative consequences for residents who are being pushed out by an ever-increasing population of nonresident students. The next section of the paper will provide robustness checks for the baseline results presented.

6 Robustness Checks

This section will provide robustness checks for the crowd-out instruments. This involves trimming the sample of observations that potentially could be driving the results to see if the coefficients are stable across specifications. The first robustness check drops states that have fewer than three schools, and also universities whose share of state enrollment is greater than 50%. The reasoning behind this is to test the validity of the assumption that the "shift" elements of the two instruments are independent of university-level enrollment decisions. Earlier it was noted that the out-of-state instrument is less likely to violate the exclusion restriction than the foreign instrument, and the results of Table 3 show that the fitted aggregate stock of out-of-state students is orthogonal to tuition and enrollment pre-trends. However, in states where there are a small number of schools, or where one school enrolls a disproportionately large share of the state's undergraduates, this aggregate stock of foreign or out-of-state students is more likely to be correlated with unobservable university-level enrollment determinants. Put another way, the larger a school is—or the fewer schools there are in a state—the higher the likelihood that state-level enrollment measures are directly correlated with university activity. Dropping states with less than three universities still omits 258 observations, while dropping schools with larger than 50% of the state's enrollment drops an additional 166 observations.

The second trimming of the sample drops all observations from Colorado from the sample. The reasoning behind this is that in 2006 Colorado eliminated state appropriations as a budget line item for universities, potentially confounding how universities in Colorado behaved for the majority of the sample. The effect this policy change would have on university crowd-out is ambiguous, but could meaningfully impact results. Dropping Colorado schools eliminates 191 university-year observations.

Table 8 uses the trimmed sample to test the robustness of the university-level results. Column 1 finds that the nonresident estimate of crowd-out after eliminating larger schools and states with few universities is slightly larger than the baseline results of Table 6. Column 1 states that every additional nonresident student crowds-out 0.56 residents, or that a 10% increase in nonresident enrollment causes a 1.1% decrease in resident enrollment. The biggest change is that the point-estimate in column 2 for out-of-state student crowd-out is higher than in Table 6, and is now statistically significant. This seems to suggest that smaller universities are more reliant on out-of-state students to generate revenue, or that out-of-state students are the primary lever to pull to increase nonresident enrollment. The foreign student specification in column 3 is very similar to the baseline results, with an almost identical elasticity as well. These results suggest that larger universities that enroll the majority of students in the state and states with few universities—a real threat to the exogeneity assumption for aggregate foreign enrollment—are not driving the crowd-out results.

Table 9 presents crowd out results after dropping Colorado from the sample. The format of Table 9 is similar to Table 8: column 1 provides the nonresident specification, column 2 the out-of-state regression, and column 3 the foreign regression. The nonresident specification in column 1 is also similar to the baseline results, finding every additional nonresident students causes 0.53 less residents to be enrolled. The elasticity here is identical to Table 6. The out-of-state point estimate is once again statistically insignificant. The foreign estimate in column 3 is almost unchanged after dropping Colorado, with an elasticity still stable at around -0.03. Taken together, these results also confirm that Colorado's enrollment profile after its policy change to eliminate appropriations in 2006 did not radically alter how universities there substituted residents for nonresident students.

6.1 Robustness Checks in the Appendix

As an additional test of the sensitivity of the instrumental variables, Table 10 in the appendix bootstraps progressively larger samples of states to test the sensitivity of the baseline model and how robust estimates are to the exclusion of certain states. In both tables, I present my preferred specifications in bold in the top row. The rows below present average coefficients from 100 repetitions of progressively larger samples of states. The "10 State Coefficient" row in each table randomly chooses ten states as its sample, and runs the associated regression (equation (3)) and saves the regression coefficient. This is repeated one hundred times to obtain a distribution of coefficients for the ten state samples, for which I present the mean. This process is repeated taking random samples of twenty, thirty, and forty states. There are associated figures that show kernel densities of the distributions of the coefficients, with vertical red lines to indicate the baseline coefficient.

Table 10 and its associated figures 3-5 show bootstrap results for the three crowd out specifications. Table 10 shows that the baseline results in all specifications tend to overstate the crowd out margin relative to the bootstrapped samples. The bootstrap coefficient distributions in Figures 3-5, while centered around the baseline coefficient, are far more variable than the state appropriations instrument. This highlights two issues with the crowd out instrument: first, the validity of the instruments clearly depend on the inclusion of all states in the sample, and are relatively fragile to the exclusion of particular states. This is particularly true for the out-of-state instrument, which has the lowest Kleibergen-Paap F-statistic and correspondingly the largest variance of coefficients of the three instruments. Second, the variability of the bootstrapped distributions suggests that certain states are key contributors of crowd out in the sample. States like California are known for preferentially enrolling nonresident students at the expense of residents.[11] Other states may not have the amenity appeal of California or other warm-weather states, and thus have no trace of crowd out and may even have nonresidents crowd in additional residents.

The second test of the robustness of the IVs is to add cost and revenue measures for the university and observe how the estimates change. Although the focus of this paper is to simulate nonresident demand shocks for the university, the propensity of an individual university to enroll more nonresident students is closely linked to their financial situation and a potential need to increase revenue. I regress all revenue and cost measures on the instruments, and begin by including as controls only ones that are uncorrelated or weakly correlated with nonresident enrollment. I then include all cost and revenue measures to see how much the point estimates change.

Tables 11-16 regress the nonresident crowd out instrument on revenue and cost measures. I choose to only use the nonresident specification because it includes both the out-of-state and foreign instruments. The estimates are not described here for brevity, but the uncorrelated revenue and cost measures that are added as additional controls are local and state operating and non-operating

[11] This behavior has also been documented by several news outlets, with articles appearing in the *New York Times* and *The Sacramento Bee* highlighting increased nonresident enrollment in the University of California system. *NYT*: https://www.nytimes.com/2016/07/08/us/public-colleges-chase-out-of-state-students-and-tuition.html. *SacBee*: https://www.sacbee.com/news/local/education/article160029439.html.

grants, federal operating grants, gifts, local appropriations, independent costs, and "other" costs. Table 17 presents results for all three instruments with the additional covariates. Results are robust to the inclusion of these additional controls, with no real change in coefficient magnitude or statistical significance. Table 18 adds all revenue and cost measures as controls, and these included measures change results significantly. Crowd out margins increase significantly across all three specifications, and each regression is significant at the 1% level. The large change in results is likely due to endogeneity in the some of the added covariates that inflate the estimates. It is possible that these results are in fact closer to the real measure of crowd out—and therefore that my baseline estimates are biased downwards—but caution is advised in interpreting these larger estimates as being truly unbiased.

Table 19 presents results with the inclusion of a capacity constraint. I detailed earlier that the logic in not including any measure of a capacity constraint for the university was to not falsely overstate the rate of crowd out. Inclusion of controls that approximate a capacity constraint can create collinearity problems that would bias results upwards. As an extreme example, including university enrollment as a control in aggregate crowd out regressions would lead crowd out estimates to be one by definition. Table 19 tests this theory by including a more static measure of a capacity constraint by controlling for the maximum number of dorm rooms at the university.[12] Baseline estimates are in columns 1, 3, and 5. Columns 2, 4, and 6 include the room cap as a control. The nonresident and out-of-state estimates are both larger with the inclusion of the room cap control, and the out-of-state coefficient becomes significant at the 10% level. The foreign estimate is unchanged. The results from Table 19 show that inclusion of an approximation of a capacity constraint for a university do inflate estimates of crowd out, even a relatively static measure like a dorm room cap. Similar to my discussion of results from Table 18, perhaps the room cap estimates are closer to the true rate of crowd out than my baseline estimates. Whether dorm room construction is endogenous is a difficult question to address. I prefer the more cautious estimates, with the tentative interpretation of my results as a lower bound of crowd out for universities.

7 Conclusion

This paper examined the rate of crowd-out of resident students by nonresident students in four-year public universities in the United States. I use IPEDS data spanning the years 2000-2015

[12]This variable, called room cap, is included in the "Institutional Characteristics" survey in IPEDS.

and employed two distinct instruments to find causal estimates of both out-of-state and foreign crowd out. The out-of-state instrument created exogenous cross-state aggregate student flows by using historical enrollment patterns, while the foreign instrument aggregated foreign enrollment at the state-level before assigning shares to schools based on their enrollment size. Using these instruments, I find that every 2 additional nonresidents crowd out 1 resident student. Foreign students are the sole contributor to this measure, with no evidence of out-of-state students crowding out residents. These results are robust to the exclusion of large schools and states with few schools, which could potentially drive results due to the aggregate nature of the instruments.

There is still much research to be done in this area. IPEDS is a wonderful data source, but efforts should be made to move to microdata. Quantifying crowd-out, particularly within state university systems like the University of California, can paint a much more detailed picture of the scale and scope of crowd-out. The ability to analyze effects of nonresident enrollment on student-body quality and graduation rates are interesting future avenues of research.

The persistent discussion of the role of public higher education in America makes this a prescient work. National political figures and state legislators alike will argue that public universities should supply an educated workforce to the state, and find crowd-out anathema to the mission of public colleges. The schools have different incentives, instead seeking to maximize student quality while navigating budget issues due to declining state support. While the goals of universities and the state governments that fund them are not mutually exclusive, the misaligned incentives create friction due to the dependent nature of the relationship. Quantifying the nature and extent of crowd-out is an important measure then as it can better inform the discussion between state and university. The pervasiveness and intensity of crowd out found in this paper suggests that states and their universities should quickly find solutions to this problem before there truly is no place at home for resident students.

References

Adkisson, Richard V and James T Peach. 2008. "Non-resident enrollment and non-resident tuition at land grant colleges and universities." *Education Economics* 16(1):75–88.

Bartik, Timothy. 1991. *Who Benefits from State and Local Economic Development Policies?* W.E. Upjohn Institute for Employment Research.

Basso, Gaetano and Giovanni Peri. 2015. The Association between Immigration and Labor Market Outcomes in the United States. Technical Report No. 9436 IZA Discussion Paper.

Borusyak, Kirill, Peter Hull and Xavier Jaravel. 2018. "Quasi-Experimental Shift-Share Research Designs." *NBER Working Paper 24997* .

Bound, John, Breno Braga, Gaurav Khanna and Sarah Turner. 2016. A passage to america: University funding and international students. Technical report National Bureau of Economic Research.

Curs, Bradley and Larry D Singell. 2002. "An analysis of the application and enrollment processes for in-state and out-of-state students at a large public university." *Economics of Education Review* 21(2):111–124.

Curs, Bradley R and Ozan Jaquette. 2017. "Crowded Out? The Effect of Nonresident Enrollment on Resident Access to Public Research Universities." *Educational Evaluation and Policy Analysis* p. 0162373717704719.

Danziger, Leif. 1990. "A model of university admission and tuition policy." *The Scandinavian Journal of Economics* pp. 415–436.

Deming, David and Christopher Walters. 2017. "The Impact of Price Caps and Spending Cuts on U.S. Postsecondary Attainment." *Working Paper* .

Goldsmith-Pinkham, Paul, Isaac; Sorkin and Henry Swift. 2018. "Bartik Instruments: What, When, Why, and How." *NBER Working Paper 24408* .

Groen, Jeffrey A and Michelle J White. 2004. "In-state versus out-of-state students: The divergence of interest between public universities and state governments." *Journal of Public Economics* 88(9):1793–1814.

Hoxby, Caroline M. 2012. "Endowment Management Based on a Positive Model of the University.".

Jaquette, Ozan and Bradley R Curs. 2015. "Creating the out-of-state university: Do public universities increase nonresident freshman enrollment in response to declining state appropriations?" *Research in Higher Education* 56(6):535–565.

Kerkvliet, Joe and Clifford Nowell. 2014. "Public subsidies, tuition, and public universities choices of undergraduate acceptance and retention rates in the USA." *Education Economics* 22(6):652–666.

Koshal, Rajindar K and Manjulika Koshal. 2000. "State Appropriation and Higher Education Tuition: what is the relationship?" *Education Economics* 8(1):81–89.

Li, Diyi. 2018. "The Effect of Non-Resident Enrollment Growth on In-State College Students." *Working Paper* .

Mathias, Max. 2020. "Cents and Students: How State Appropriations Impact Public Universities.".

Rizzo, Michael and Ronald G Ehrenberg. 2004. Resident and nonresident tuition and enrollment at flagship state universities. In *College choices: The economics of where to go, when to go, and how to pay for it.* University of Chicago Press pp. 303–354.

Shih, Kevin. 2017. "Do international students crowd-out or cross-subsidize Americans in higher education?" *Journal of Public Economics* 156:170–184.

Tiffany, Frederick G and Jeff A Ankrom. 1998. "The Competitive Use of Price Discrimination by Colleges." *Eastern Economic Journal* 24(1):99–110.

Figure 1: Tuition and Enrollment Trends

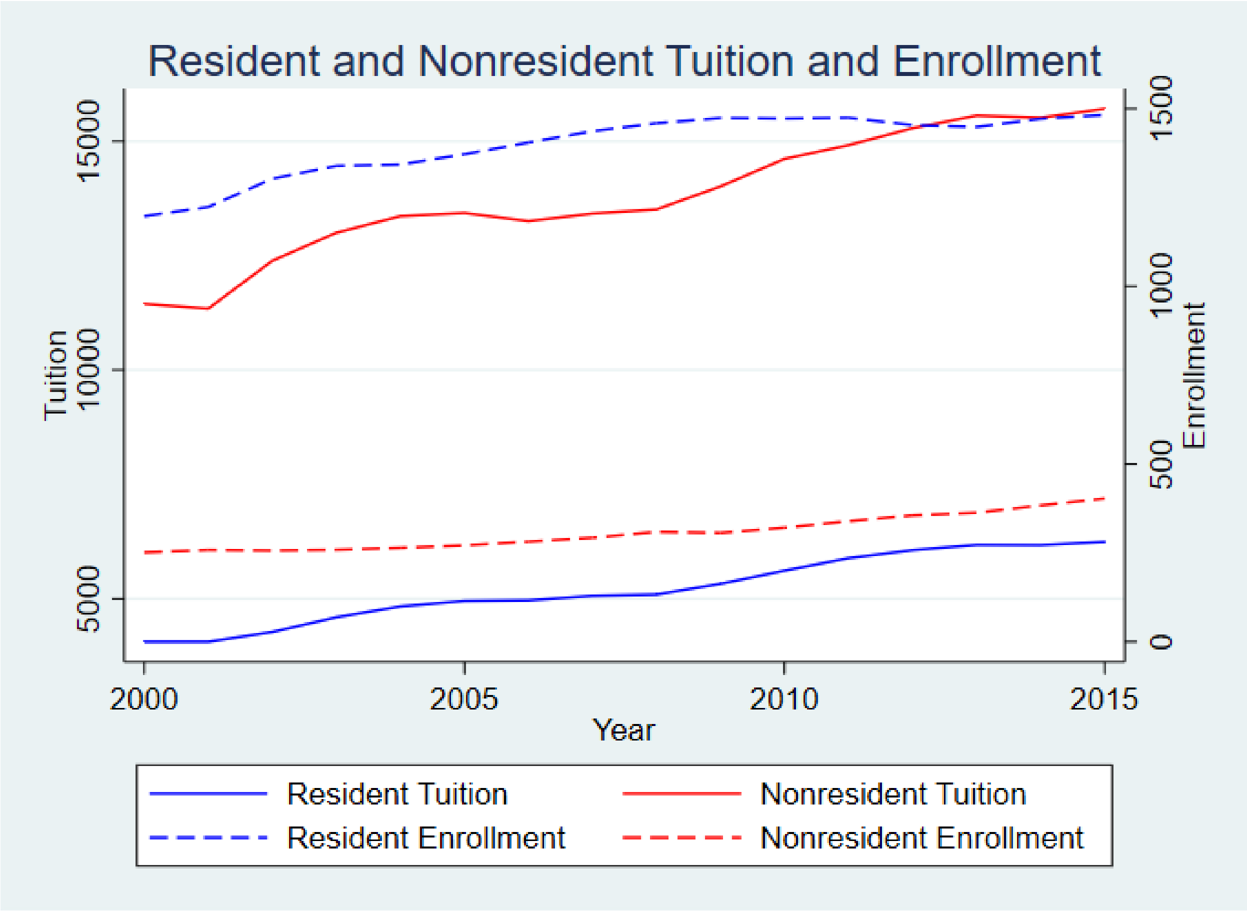
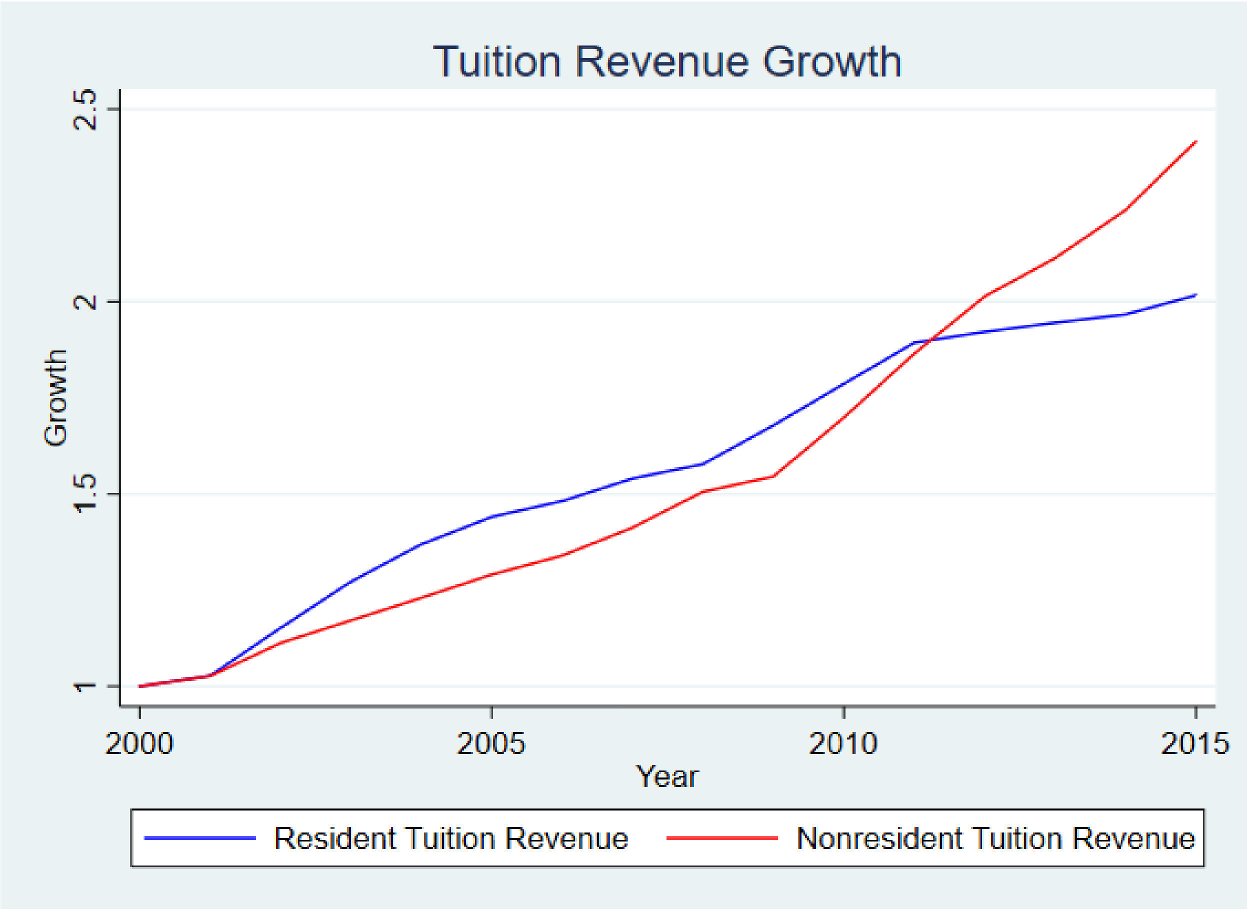


Figure 2: Tuition Revenue by Source



Tuition revenues are defined so that 2000 values equal 1. Average resident tuition revenue in 2000 is $4,882,289; average nonresident tuition revenue in 2000 is $3,404,237.

Table 1: Summary Statistics

	Mean	SD	IQR
Resident Tuition ($)	5223	2520	2683
Nonresident Tuition ($)	13820	5199	5958
Resident 1^{st} Year Enrollment	1399	1221	1457
Nonresident 1^{st} Year Enrollment	308	486	281
Undergraduate Enrollment	9608	8186	10227

N=7586. All dollar values are deflated using the Higher Education Price Index (HEPI). Interquartile range (IQR) is the distance between the 75^{th} and 25^{th} percentile observations.

Table 2: Enrollment Share Pre-trend

	Enrollment		**Tuition**	
	(1)	(2)	(3)	(4)
	Resident	Nonresident	Resident	Nonresident
$s_{i,2000}^{enrl}$	27.35	1.441	0.0958	-0.0907
	(27.03)	(4.951)	(0.159)	(0.152)
N	546	530	499	499

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Standard errors are in parentheses and clustered at the state level. Each column is a separate regression. All controls from equation 2 are included, as well as 1992 university enrollment. The enrollment pre-trend is the percent change in first-year enrollment from 1992 to 1999. The tuition pre-trend is the percent change in tuition from 1990 to 1999.

Table 3: Out-of-State Instrument Pre-trend

	Enrollment		**Tuition**	
	(1)	(2)	(3)	(4)
	Resident	Nonresident	Resident	Nonresident
$\widehat{OUT}_{s,2000}$	-0.0000264	0.000133	0.000217	-0.0000188
	(0.000199)	(0.000212)	(0.000129)	(0.0000958)
N	50	50	50	50

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Standard errors are in parentheses and adjusted for heteroskedasticity. Each column is a separate regression. All controls from equation 3 are included. The enrollment pre-trend is the percent change in aggregate state enrollment from 1992 to 1999. The tuition pre-trend is the percent change in a constructed average state tuition from 1990 to 1999. The average is constructed by weighting university tuition in 1990 and 1999 by baseline enrollment in 1992.

Table 4: Crowd Out First Stage

	(1)	(2)	(3)	(4)
	$OutState_{it}$	$OutState_{it}$	$Foreign_{it}$	$Foreign_{it}$
$\widehat{OutState}_{it}$	1.238***	1.062***	-	-
	(0.268)	(0.229)		
$\widehat{Foreign}_{it}$	-	-	1.494***	1.143***
			(0.161)	(0.172)
N	7586	7576	7586	7576
KP F-stat	21.32	21.50	86.22	44.28
University FE	X	X	X	X
Year FE	X	X	X	X
Controls		X		X

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Standard errors are in parentheses and clustered at the state level. Controls in columns 2 and 4 are described in equation 3.

Table 5: First-year Crowd Out

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	IV	OLS	IV	OLS	IV
$Nonresident_{it}$	-0.106	-0.258	-	-	-	-
	(0.0688)	(0.200)				
$OutState_{it}$	-	-	0.0645	-0.294	-	-
			(0.0576)	(0.266)		
$Foreign_{it}$	-	-	-	-	-0.761***	-0.318
					(0.132)	(0.331)
N	7576	7576	7576	7576	7576	7576
KP F-stat		22.21		21.50		44.28
Elasticity	-0.0031	-0.020	0.030	-0.087	-0.013	-0.0011

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Standard errors are in parentheses and clustered at the state level. Each column is a separate regression. Coefficient interpretation is 1 student $\rightarrow$ β students. All controls from equation 3 are included.

Table 6: Crowd Out University Enrollment

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	IV	OLS	IV	OLS	IV
$Nonresident_{it}$	-0.164	-0.521**	-	-	-	-
	(0.126)	(0.209)				
$OutState_{it}$	-	-	0.128	-0.434	-	-
			(0.114)	(0.282)		
$Foreign_{it}$	-	-	-	-	-0.979***	-1.139***
					(0.151)	(0.320)
N	7576	7576	7576	7576	7576	7576
KP F-stat		16.93		15.60		33.94
Elasticity	-0.020	-0.11	0.043	-0.14	-0.022	-0.022

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Standard errors are in parentheses and clustered at the state level. Each column is a separate regression. Coefficient interpretation is 1 student $\rightarrow \beta$ students. All controls from equation 3 are included.

Table 7: Foreign Crowd-In of Out-of-State Students

	First-Year Out-of-State Enrollment		**University Out-of-State Enrollment**	
	(1)	(2)	(3)	(4)
	OLS	IV	OLS	IV
$Foreign_{it}$	-0.042	-0.0191	0.022	-0.0460
	(0.097)	(0.295)	(0.065)	(0.299)
N	7576	7576	7576	7576
KP F-stat		44.28		33.94

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Standard errors are in parentheses and clustered at the state level. Each column is a separate regression. Coefficient interpretation is 1 student $\rightarrow \beta$ students. All controls from equation 3 are included.

Table 8: Crowd Out: Dropping Big Schools

	(1)	(2)	(3)
	$InState_{it}$	$InState_{it}$	$InState_{it}$
$Nonresident_{it}$	-0.563***		
	(0.198)		
$OutState_{it}$		-0.651**	
		(0.301)	
$Foreign_{it}$			-1.185***
			(0.337)
N	7152	7152	7152
KP F-stat	14.90	13.35	29.00
Elasticity	-0.11	-0.11	-0.033

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Standard errors are in parentheses and clustered at the state level. Each column is a separate regression. All controls from equation 2 are included. Coefficient interpretation is 1 student $\rightarrow \beta$ students. States that have fewer than 3 schools are dropped as well as any university that has larger than 50% of the aggregate enrollment share.

Table 9: Crowd Out: Dropping Colorado

	(1)	(2)	(3)
	$InState_{it}$	$InState_{it}$	$InState_{it}$
$Nonresident_{it}$	-0.533***	-	-
	(0.205)		
$OutState_{it}$	-	-0.452	-
		(0.277)	
$Foreign_{it}$	-	-	-1.159***
			(0.321)
N	7385	7385	7385
KP F-stat	16.68	15.41	33.92
Elasticity	-0.11	-0.080	-0.033

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Standard errors are in parentheses and clustered at the state level. Each column is a separate regression. All controls from equation 2 are included. Coefficient interpretation is 1 student $\rightarrow$ β students. Colorado schools are dropped from each regression.

A Appendix

Table 10: Crowd Out: Bootstrapped Samples for Bias

	(1) Nonresident	(2) Out of State	(3) Foreign
Baseline Coefficient	**-0.52**	**-0.43**	**-1.14**
10 State Coefficient	-0.24	-0.76	-0.79
20 State Coefficient	-0.45	-0.35	-0.90
30 State Coefficient	-0.46	-0.38	-0.99
40 State Coefficient	-0.46	-0.43	-1.00

Row 1 presents results from the baseline specifications. The rows below show the average regression coefficient for each bootstrapped sample. Each bootstrap group was run 100 times.

Figure 3: Nonresident Crowd Out Distributions

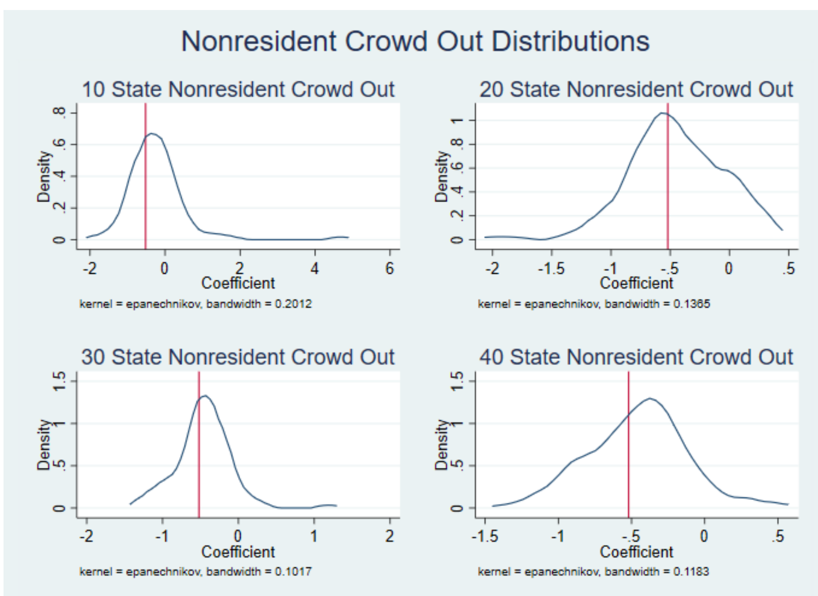


Figure 4: Out-of-State Crowd Out Distributions

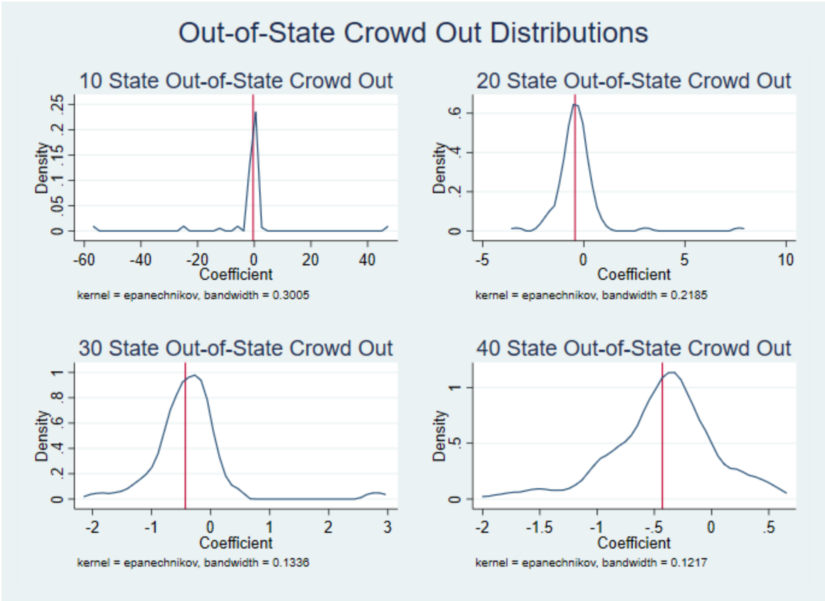
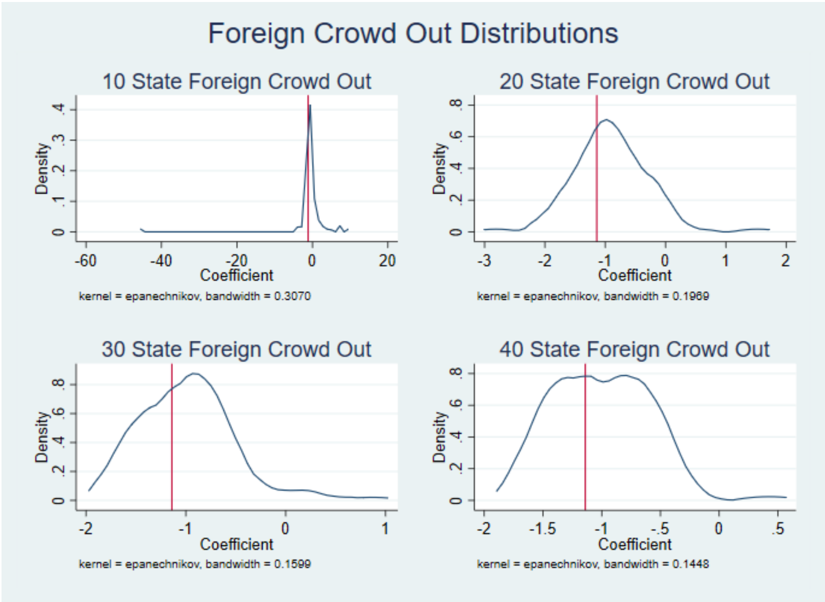


Figure 5: Foreign Crowd Out Distributions



A.1 Crowd Out: Sensitivity to Budget Items and Capacity Constraint

Table 11: Crowd Out: Revenue 1

	Local Grants		**State Grants**		**Federal Grants**	
	(1)	(2)	(3)	(4)	(5)	(6)
	Operating	Non-Operating	Operating	Non-Operating	Operating	Non-Operating
$Nonresident_{it}$	0.00384*	-0.0000150	0.000422	0.000703	0.00168	0.00856**
	(0.00207)	(0.000455)	(0.00126)	(0.00137)	(0.00331)	(0.00411)
N	7123	7123	7123	7123	7158	7158

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Standard errors are in parentheses and clustered at the state level. Each column is a separate regression. All regressions use the nonresident enrollment instrument. All controls from equation 2 are included. Coefficient interpretation is 1 student $\rightarrow$ \$$\beta$ million.

Table 12: Crowd Out: Revenue 2

	(1)	(2)	(3)	(4)	(5)	(6)
	Federal Approp.	Local Approp.	Pell Grants	Gifts	Investment Inc.	Other Rev.
$Nonresident_{it}$	-0.000370*	0.0000348	0.00469**	0.000805	0.00521*	-0.00966**
	(0.000210)	(0.0000664)	(0.00185)	(0.00120)	(0.00278)	(0.00473)
N	7123	7123	7576	7123	7123	7123

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Standard errors are in parentheses and clustered at the state level. Each column is a separate regression. All regressions use the nonresident enrollment instrument. All controls from equation 2 are included. Coefficient interpretation is 1 student $\rightarrow$ \$$\beta$ million.

Table 13: Crowd Out: Revenue 3

	(1) Sales & Services	(2) Hospital Services	(3) Independent Operations	(4) Total Revenue
$Nonresident_{it}$	0.00832***	0.0232	0.00390*	0.0774***
	(0.00321)	(0.0143)	(0.00232)	(0.0227)
N	7123	7123	7123	7576

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Standard errors are in parentheses and clustered at the state level. Each column is a separate regression. All regressions use the nonresident enrollment instrument. All controls from equation 2 are included. Coefficient interpretation is 1 student $\rightarrow$ \$$\beta$ million.

Table 14: Crowd Out: Costs 1

	(1) Instruction	(2) Research	(3) Public Service	(4) Academic Support
$Nonresident_{it}$	0.0348***	0.0133**	0.00572***	0.0156***
	(0.00936)	(0.00519)	(0.00191)	(0.00464)
N	7171	7123	7123	7123

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Standard errors are in parentheses and clustered at the state level. Each column is a separate regression. All regressions use the nonresident enrollment instrument. All controls from equation 2 are included. Coefficient interpretation is 1 student $\rightarrow$ \$$\beta$ million.

Table 15: Crowd Out: Costs 2

	(1) Student Services	(2) Institutional Support	(3) Scholarships	(4) Auxiliary Enterprises
$Nonresident_{it}$	0.00390**	0.00600***	0.00521***	0.0106***
	(0.00195)	(0.00233)	(0.00194)	(0.00358)
N	7171	7171	7123	7123

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Standard errors are in parentheses and clustered at the state level. Each column is a separate regression. All regressions use the nonresident enrollment instrument. All controls from equation 2 are included. Coefficient interpretation is 1 student $\rightarrow$ \$$\beta$ million.

Table 16: Crowd Out: Costs 3

	(1) Hospital Costs	(2) Independent Costs	(3) Other Costs	(4) Total Costs
$Nonresident_{it}$	0.0252**	0.00266*	-0.00224	0.0745***
	(0.0125)	(0.00153)	(0.00240)	(0.0182)
N	7123	7123	7123	7123

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Standard errors are in parentheses and clustered at the state level. Each column is a separate regression. All regressions use the nonresident enrollment instrument. All controls from equation 2 are included. Coefficient interpretation is 1 student $\rightarrow$ \$$\beta$ million.

Table 17: Crowd Out: Robustness with Unrelated Controls

	(1)	(2)	(3)	(4)	(5)	(6)
	Baseline	Added Controls	Baseline	Added Controls	Baseline	Added Controls
$Nonresident_{it}$	-0.462**	-0.464**	-	-	-	-
	(0.198)	(0.195)				
$OutState_{it}$	-	-	-0.429	-0.412	-	-
			(0.270)	(0.282)		
$Foreign_{it}$	-	-	-	-	-1.108***	-1.110***
					(0.326)	(0.304)
N	7123	7123	7123	7123	7123	7123

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Standard errors are in parentheses and clustered at the state level. Each column is a separate regression. All regressions use the associated instrumental variable. All controls from equation 2 are included. The additional uncorrelated controls are local and state operating and non-operating grants, federal operating grants, gifts, local appropriations, independent costs, and other costs.

Table 18: Crowd Out: Robustness with All Controls

	(1)	(2)	(3)	(4)	(5)	(6)
	Baseline	Added Controls	Baseline	Added Controls	Baseline	Added Controls
$Nonresident_{it}$	-0.462**	-1.481***	-	-	-	-
	(0.198)	(0.405)				
$OutState_{it}$	-	-	-0.429	-1.251***	-	-
			(0.270)	(0.393)		
$Foreign_{it}$	-	-	-	-	-1.108***	-3.347***
					(0.326)	(0.825)
N	7123	7123	7123	7123	7123	7123

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Standard errors are in parentheses and clustered at the state level. Each column is a separate regression. All regressions use the associated instrumental variable. All controls from equation 2 are included. All revenue and cost measures are included except for total revenue and total cost due to multicollinearity.

Table 19: Crowd Out: Adding a Capacity Constraint

	(1)	(2)	(3)	(4)	(5)	(6)
	Baseline	Room Cap	Baseline	Room Cap	Baseline	Room Cap
$Nonresident_{it}$	-0.521**	-0.626***	-	-	-	-
	(0.209)	(0.223)				
$OutState_{it}$	-	-	-0.434	-0.592*	-	-
			(0.282)	(0.310)		
$Foreign_{it}$	-	-	-	-	-1.139***	-1.133***
					(0.320)	(0.327)
N	7576	7260	7576	7260	7576	7260

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Standard errors are in parentheses and clustered at the state level. Each column is a separate regression. All regressions use the associated instrumental variable. All controls from equation 2 are included. The maximum number of dorm rooms is also included as a control to approximate a capacity constraint.