WORKING PAPERS

The Effect of Tuition-Free Community College

Christopher Lau

WORKING PAPER NO. 347

September 2020

FTC Bureau of Economics working papers are preliminary materials circulated to stimulate discussion and critical comment. The analyses and conclusions set forth are those of the authors and do not necessarily reflect the views of other members of the Bureau of Economics, other Commission staff, or the Commission itself. Upon request, single copies of the paper will be provided. References in publications to FTC Bureau of Economics working papers by FTC economists (other than acknowledgment by a writer that he has access to such unpublished materials) should be cleared with the author to protect the tentative character of these papers.

BUREAU OF ECONOMICS
FEDERAL TRADE COMMISSION
WASHINGTON, DC 20580

The Effect of Tuition-Free Community College

Christopher V. Lau*
Federal Trade Commission

September 1, 2020

Abstract

A popular proposal to address the rising cost of higher education in the United States has been to provide tuition-free access to community colleges. This paper examines the effect of such a policy on college access, consumer welfare, and student outcomes accounting for equilibrium responses from for-profit and four-year competitors. I find that free community college increases enrollment by 26 percent, welfare for all students, and degree completions by 20 percent. I compare these findings to more fiscally practical implementations of free community college. Programs that only cover tuition after accounting for other sources of grants increase enrollment by 10 percent and degree completions by 10 percent, but provide no benefit to low-income students. Need-based programs that make community college free for low-income students increase enrollment by 12 percent and are beneficial to low-income students, but harm middle- and high-income students and only increase degree completions by 4 percent.

1 Introduction

Rising tuition costs and increasing student debt have become major issues for higher education policy in the United States. One proposal to address these issues that has gained significant attention in recent years has been to provide free access to community colleges, often referred to as "promise programs." Many states, including Tennessee, Oregon, California, New York, and others, have passed or put forward legislation that makes enrolling into community colleges free for a large portion of the student population (Trammell, 2019). The overall benefits of these programs, however, have been debated.

Opponents of the free community college have suggested that some students may be negatively impacted by the policy, since they will be drawn away from four-year universities. Other skeptics believe community colleges do not have the infrastructure to handle the short-run increases in enrollment (Quilantan, 2019), or worry that taxpayers would be footing the

*clau@ftc.gov. Bureau of Economics, Federal Trade Commission, 600 Pennsylvania Ave. NW, Washington D.C. 20580. Email: clau@ftc.gov. The opinions expressed are those of the authors and do not represent views of the Institute, the U.S. Department of Education, the Federal Trade Commission, or any Commissioner.

bill for wealthier students (English, 2019). On the other hand, advocates of free community college suggest that its mission is to broaden the base of who starts going to college in the first place (Amour, 2019). For example, early data from the Tennessee Promise Program, the first free community college program in the United States, suggests that more students are enrolling into community college, and more are also persisting in it (Wermund, 2019).

This paper aims to quantify the effect of free community college in terms of access to higher education, consumer welfare, and student outcomes, accounting for equilibrium responses from potential substitutes, for-profit colleges and nonselective four-year universities. In particular, incorporating other sectors of higher education will be important for understanding the policy's effect on welfare and outcomes, since the reduced price of community college could draw students away from higher performing four-year universities, or higher cost for-profit colleges. To do this, I estimate a model of demand and supply for higher education that incorporates a differentiated products market structure. I then use the model to compute counterfactual equilibrium outcomes under a variety of federal free community college proposals.

The advantages of using a structural model are threefold. First, most of the literature on free community college, to the best of my knowledge, has focused on evaluating the tradeoff between increased access to college and diversion from four-year universities in terms of higher education outcomes (e.g., earnings, degree completions, etc.). However, the effect of diversions from higher cost for-profit colleges is often assumed away because of data limitations.[1] A structural model can account for this effect by relying on data sources that have less policy variation, but can accurately measure price sensitivity and student preferences. Second, many active state-wide promise programs, like the Tennessee Promise Program, are not completely free for all students. Instead, more pragmatic implementations are used to reduce the cost of the program. A structural model will allow for the comparison of different applications of free community college, which could provide insight into its practical execution. Finally, the model can estimate effects on consumer welfare, which may be viewed as an alternative evaluation metric for higher education to the traditional metrics of student outcomes and enrollment.

To model demand, I assume students make two choices: First, they choose which college to attend. Specifically, students choose which college to enroll in from a variety of postsecondary options within three sectors of higher education: community colleges, for-profit colleges, and nonselective four-year universities. I particularly use for-profit and nonselec-

[1] Table 1 shows that average tuition at for-profit colleges is over six times larger than that of community colleges in the estimation sample. Furthermore, for-profit colleges, on average, tend to perform worse on certain outcome metrics, like default rates and bachelor-degree completions. (Government Accountability Office, 2011, 2018)

tive four-year universities as potential substitutes for community colleges because they tend to have similar admissions criteria; open or near-open enrollment. At the same time, each institution has its own set of characteristics that may influence which students choose to enroll. As a result, the college choice model resembles a differentiated products model of discrete consumer choice.

Second, students choose an optimal financial aid package for each higher education alternative. In particular, students non-exclusively choose their utility-maximizing choice of federal grants, federal loans, and private loans. This portfolio of grants and loans may differ by the institution the student selects; for example, the student may borrow more when selecting an institution with a higher tuition price. Because grants are basically free money, I assume it will always be accepted when offered. For loans, students balance the tradeoff between receiving money today, and repaying it back tomorrow. There are two differences between federal and private loans. First, federal loans are subject to a maximum limit that can be borrowed, while private loans are not. Second, private loans have interest rates that are set by private lenders, thus will typically differ from federal rates. Together, these differences can help summarize the optimal student loan portfolio: Students will always borrow from the cheaper of the two loans in terms of interest rates. Furthermore, students will only borrow from both sources if federal loans are cheaper than private loans, and the federal loan limit is binding.

To identify and estimate the parameters of the demand model, I make use of a random sample of college students from the restricted-access version of the National Postsecondary Student Aid Study ("NPSAS"), as well as public institution-level data from the Integrated Postsecondary Education Database Survey ("IPEDS"). The NPSAS contains individual-level information, such as which school the student attended, how much financial aid she received, and demographic information. IPEDS contains institution-level information, such as cost of attendance, aggregate enrollments, and school characteristics. I estimate the choice model in two steps. First, I estimate the implied interest rate, or the "price," for federal and private loans using the observed loan borrowing behavior in the NPSAS sample. With these implied interest rates, the loan choice model predicts the amount students would have borrowed had they attended another institution, which is unobserved in the data. In the second step, given the predicted loan borrowing behavior at each institution, I estimate the student's utility-maximizing choice of higher education. In particular, while published tuition prices typically only vary by institution and not by individual, federal financial aid, and thus the *net* price of attendance, varies at both the individual and college level, helping me identify student responsiveness to the net cost of enrolling in an institution. In both steps, I use maximum likelihood.

On the supply side, I naturally assume for-profit colleges set tuition prices to maximize profits. Furthermore, I assume that the objective of a nonprofit institution is to satisfy two factors: (1) financial success as measured by profits, and (2) the value the institution creates for students in its community. To model the latter factor, I borrow from the hospital competition literature and assume the community's value of an institution is given by students' average willingness-to-pay for the college or system of colleges. Nonprofit institutions set tuition prices to maximize a Cobb-Douglas function that considers its two objectives. The weight on the profit maximization objective is estimated using a general method of moments estimator, following Gowrisankaran et al. (2015). I assume colleges compete in a static Nash pricing game such that the equilibrium of this game yields the observed prices found in the data. I then use the equilibrium pricing behavior of the schools to uncover marginal costs at for-profit colleges, and marginal costs and the weight on profit maximization at nonprofit institutions.

Using the estimated demand-side and supply-side parameters, I consider the pricing equilibrium under the counterfactual of a "free-for-all" community college plan implemented at the federal level that is supplementary to existing financial aid programs, and examine its impact on access to higher education, consumer welfare, and degree completions. I find that overall enrollment into higher education would increase by 26 percent, with 83 percent of the increase at community college due to students who would have otherwise not enrolled in higher education. In addition, I find that the policy would *increase* average tuition prices at for-profit and nonselective four-year colleges, since low-income, price sensitive students would substitute away, leaving these colleges more price inelastic at the margin. I further use compensating variation to measure the overall welfare impact of free community college finding that students in the NPSAS sample, on average, would pay $498 to have free community college. Compensating variation was also higher for students with incomes between $25,000 and $75,000, suggesting that middle-income students benefit the most from the policy. Specifically, benefits for low-income students are more moderate because free community college just alters their loan borrowing behavior, meaning the welfare improvements are realized tomorrow, rather than today. Overall, I find that the introduction of free community college would increase higher education degree completions by about 22 percent.

I compare these findings to "last dollar" programs, i.e., aid that covers the student's cost of attending a community college *after* accounting for the amount received in federal grants. These programs have been more popular in terms of implementation because they rely on less financial resources. I find that the implementation of a last dollar program would increase enrollment by 10 percent, with 88 percent of the increase at community colleges due to new enrollees. This produces lower levels of access compared to the fully free community college program because the most price sensitive students do not receive any additional aid; these are the students most likely to alter their higher education decisions. Furthermore, the welfare analysis suggests that compensating variation is almost negligible for students with income less than $25,000; this is due to the fact that low income students would likely not benefit from the program, since Pell Grant aid typically covers the cost to tuition at community colleges. Likewise, high income students, as defined by those with incomes greater than $75,000, are the ones who would benefit the most from a last dollar program with the largest compensating variation. I find that the introduction of a last dollar community college program would increase degree completions by 11 percent.

Finally, I examine the counterfactual of placing need-based eligibility restrictions on free community college. In particular, I consider a policy that makes community college free for low-income students. Overall, I find that need-based programs would increase enrollment by 12 percent. However, this would only benefit low-income students, and would actually *harm* middle- and high-income students due to equilibrium price increases at for-profit and nonselective four-year colleges in response to inter-sector substitution from price sensitive students. Furthermore, a need-based program would only increase degree completions by about 4 percent, which is significantly lower than that of the other two free community college programs.

Ultimately, this analysis suggests that all of the free community college programs would increase enrollment and degree completions, with larger effects for a fully free community college scheme. However, last dollar programs disproportionately benefit higher income students, with little to no effect for low-income students. Need-based programs benefit the target population, but may actually harm other students and do little to improve overall completions. I proceed with the rest of this paper by discussing the background of promise programs in the United States, introducing a model of higher education, discussing the data, identification, and estimation used calibrate the model, and finally applying counterfactual analysis to understand the impact of free community college.

2 Background

In 2015, President Barak Obama advocated for America's College Promise, a federal plan to make two-year college "as free and universal as high school." Since then, a number of local and statewide programs were enacted to provide free higher education opportunities for students. Most notably, in 2015, the Tennessee Promise Program was the first statewide effort to implement free community college, and was shortly followed by the Oregon Promise Program beginning in 2016, and other states such as California, New York, and Washington.

Promise programs can be distinguished by three features (Pingel et al., 2016). The first is eligibility criteria. Many active promise programs have eligibility requirements that may include residency, age, and merit factors. For example, the Tennessee Promise Program is only available for recent graduates from a Tennessee high school.[2] The second feature is the definition of free. Specifically, most promise programs cover tuition and fees, but leave other costs, such as textbooks and living expenses, as a burden on the student. Finally, the third feature is the timing of the award. Program funding can be applied either before or after other sources of financial aid, such as Pell Grants, are taken into account. "First-dollar programs" apply aid before other sources and allow students to accumulate additional financial support. On the other hand, "last dollar programs" will only cover the cost of community college after other sources are counted, leaving some federal aid recipients functionally ineligible for the program.

Advocates for free community college have argued that promise programs can make higher education more affordable for many students, and increase access for those who may traditionally not pursue higher education, such as low-income and older students. In addition, promise programs could deter students from attending high-price for-profit colleges that have demonstrated poor outcomes. Critics of free community college have argued that some lower-income, price sensitive students may be attracted away from 4-year colleges, leading to "undermatching." In addition, they could lead to funding and capacity issues at community colleges, and last dollar programs are often criticized for providing little, if any, financial assistance to low-income students, since Pell Grant awards will typically cover tuition and fees.

There have been a handful of studies that have looked into promise programs and their effect on students and communities. Recent studies have examined the effect of statewide promise programs in Tennessee (Carruthers, 2019; Bell, 2018) and Oregon (Gurantz, 2020), finding overall increases in college enrollment, with moderate decreases at four-year colleges. Earlier studies have explored the effect of free college endorsed by local governments and communities, such as in Kalamazoo, MI (Andrews et al., 2010; Bartik et al., 2019), Pittsburgh, PA, (Bozick et al., 2015; Page et al., 2019), and Knox, TN (Carruthers and Fox, 2016), finding increases in enrollment and completions.

Avery et al. (2019) used a simulation study to examine the impact of four different higher education policies, finding that free community college was the least cost effective at improving college completions. This was in part because too many students would divert away from higher performing four-year colleges. Some of the empirical literature corroborates

[2]Tennessee has also introduced the Tennessee Reconnect Program, which is similar to Tennessee Promise, but for adult learners.

this notion that financial incentives induce substitution to poorer performing institutions; for example, Cohodes and Goodman (2014) found that students were willing to divert from private universities to lower performing public institutions for relatively little money. Furthermore, Goodman et al. (2017) found that increased access to four-year colleges improved degree completion rates for students who, on the margin, would have otherwise attended a two-year college. This suggests that increased access to community college may leave some students worse off.

Mountjoy (2019), on the other hand, found that increased access to community college would result in net increases in attainment and earnings, with two-year entry from students who would have otherwise not pursued postsecondary education significantly dominating the effect of students diverting from four-year colleges. Denning (2017) further found some evidence that community college attendance increases bachelor degree attainment, even for students who are on the margin of attending community college and a four-year university. For students attaining an associate degree from a community college, the literature suggests a return of between 22 percent (Zimmerman, 2014) and 44 percent (Grosz, 2020).

Finally, an important line of research has examined the impact of financial aid beyond community colleges, with particular focuses on four-year nonprofit universities, and for-profit colleges. Specifically, many studies have investigated the effect of state and federal financial aid on student enrollment and outcomes (Darolia, 2013; Castleman and Long, 2016; Deming and Walters, 2017; Bettinger et al., 2019), net tuition prices (Singell Jr and Stone, 2007; Cellini and Goldin, 2014; Gibbs and Marksteiner, 2016; Turner, 2017), and inter-sector substitution (Cohodes and Goodman, 2014; Cellini et al., 2019).

3 Empirical model

In this section, I develop a static model of college and financial aid choice (demand), as well as for-profit and nonprofit tuition pricing (supply). In the student choice model, I suppose students make two sequential decisions: Students make a discrete choice in which college they want to enroll in, and a continuous choice of the financial aid package used to afford their chosen institution. The pricing model assumes for-profit colleges set tuition prices to maximize profits, while nonprofit colleges optimize over two factors: (1) its profits, and (2) the value it provides to the market.

3.1 Demand for higher education

Suppose there exists M markets that contain n_m potential students looking to enroll in higher education. Within each market m, students can choose from C_m community colleges, F_m

for-profit colleges, and N_m nonselective four-year nonprofit institutions. I define community colleges as degree-granting institutions that take two- or fewer years to complete, for-profit colleges as degree-granting institutions designated as for-profit, and nonselective four-year nonprofit colleges as degree-granting institutions that take four years to complete and are open enrollment, e.g. do not require standardized testing, letters of recommendations, etc. for admission. Let $\mathbf{S}_m$ be the set of higher education institutions available to students in market m, with an outside option of not enrolling in higher education indexed by 0.

3.1.1 Utility

Suppose students live for two periods: "today" and "tomorrow." Student i's utility from attending school j today and tomorrow are given by:

$$v_{ij} = \alpha \log(c_{ij}) + \mathbf{x}_j^T \boldsymbol{\beta} + \sum_n x_{jn} \mathbf{d}_i^T \boldsymbol{\gamma}_n + \psi_j + \varepsilon_{ij} \tag{1}$$

$$v'_{ij} = \alpha \log(c'_{ij}), \tag{2}$$

where c_{ij} is student i's consumption today if she attends institution j, $\mathbf{x}_j$ is a vector of observable school characteristics, and $\mathbf{d}_i$ is a vector of observable student characteristics. Furthermore, c'_{ij} is student i's consumption tomorrow from attending institution j. The coefficient α represents the marginal utility of consumption, $\boldsymbol{\beta}$ represents the mean preferences for observed school characteristics, $\boldsymbol{\gamma}_n$ represents heterogenous preferences for school characteristics by student characteristics, ψ_j represents an average unobserved preference for school j, and ε_{ij} is unobserved idiosyncratic preferences for school j.

The student's objective is to maximize her present value of utility for attending school j, which is given by

$$V_{ij} = v_{ij} + \delta E(v'_{ij}),$$

where δ is the discount rate for future utility and $E(v'_{ij})$ is the expected value of utility tomorrow. In other words, students do not perfectly observe their future utility, but rather have some expectation of what their utility can be when they choose to attend institution j. When maximizing her present value of utility, the student is subject to binding budget constraints, such that

$$c_{ij} + p_j = y_i + g_{ij} + l^f_{ij} + l^p_{ij} \tag{3}$$

$$c'_{ij} + R^f_i \cdot l^f_{ij} + R^p_i \cdot l^p_{ij} = y'_{ij}, \tag{4}$$

where equation (3) represents the budget constraint in the today period, and (4) represents the budget constraint in the tomorrow period. In the today period, y_i is the student's income today, p_j is the price paid to attend school j, g_{ij} is the amount of federal grants student i receives for attending school j, and $l_{ij}^f$and l_{ij}^p are the amount of federal and private loans student i borrows when attending school j, respectively.[3] In the tomorrow period, y'_{ij} is the income student i earns from attending school j, and R_i^f and R_i^p are the accrued interest owed when taking out an additional dollar of federal and private loans, respectively.

To derive the distribution of v'_{ij}, I solve the binding budget constraint for consumption tomorrow, given by equation (4), and assume that students believe their future income is log-normally distributed, such that $y'_{ij} \sim \log\mathcal{N}(\mu_j, \sigma_j)$.[4] From equation (2), this implies that the distribution of v'_{ij} is normal, such that

$$v'_{ij} \sim \mathcal{N}\bigg(\alpha(\mu_j - R_i^f l_{ij}^f - R_i^p l_{ij}^p), \alpha^2\sigma_j^2\bigg).$$

Plugging in the expectation of v'_{ij} into the present value of utility and applying the binding budget constraints, student i's indirect utility for attending institution j can be written as

$$V_{ij} = \alpha\left(\log(y_i - n_{ij}) - \lambda_i^f l_{ij}^f - \lambda_i^p l_{ij}^p\right) + \mathbf{x}_j^T\boldsymbol{\beta} + \sum_n x_{jn}\mathbf{d}_i^T\boldsymbol{\gamma}_n + \xi_j + \varepsilon_{ij}$$

where $n_{ij} = p_j - g_{ij} - l_{ij}^f - l_{ij}^p$ is the net price of attendance, $\lambda_i = \delta R_i$ represents the discounted accrued interest from borrowing an extra dollar of each type of loan, and $\xi_j = \psi_j + \delta\alpha\mu_j$ represents a school-specific utility term that includes unobserved features of school j, as well as the utility from future discounted expected income.[5] I assume that λ_i^f and λ_i^p are distributed according to the distributions F_f and F_p, respectively.[6]

To maximize her present value of utility, the student has four choice variables:[7] the

[3]For each student, income in the NPSAS sample is defined by the parent's income for dependent students, and the student's income for independent students. The extent to which dependency status is endogenous with college choice, e.g. parents would pay for particular colleges but not others, is not explored here.

[4]Specifically, because y'_{ij} is log normal, then c'_{ij} is a shifted log-normal, and v'_{ij} is normally distributed.

[5]This assumes that expected returns to education only vary at the institution-level, but not the individual-level. Individual-level variation can be incorporated if it is assumed to be separable from institution-level effects; in this case, the individual-level returns will be linear in the indirect utility function. This will shift utilities in parallel (since it is the same for all alternatives for a given individual), thus will not affect the student's college and financial aid choice.

[6]This formulation assumes students view the interest rate as a constant, while in reality interest rates can vary over time for some types of loans. In addition, federal loans are subject to forgiveness programs and alternative repayment plans that could affect the stock interest rate or the student's ability to pay back their loan. I view the accrued interest parameters as the *expected* discounted marginal accrued interest that implicitly averages over this temporal variation.

[7]In principle, students also have the ability to make additional choices, such as college major. Because of

institution j to enroll in, the amount of grants g_{ij} to accept, and the amount of federal and private loans, l_{ij}^f and l_{ij}^p, to borrow.[8] For federal loans, students are subject to borrowing limits, such that they are only allowed to borrow up to $\bar{L}$. Private loans, on the other hand, are assumed to not be subject to a maximum. The student chooses these variables to maximize her indirect utility V_{ij}. Because grants and loans can vary by institution, the student's optimization problem can be solved in two parts. First, the student can derive her optimal financial aid package for each institution. Then, she can select the utility maximizing college, given that she knows her optimal financial aid package for each alternative.

3.1.2 Federal grant aid

For most students, the main source of federal grants is Pell Grants. The amount the student receives depends on financial need. In particular, every student planning on enrolling into an institution of postsecondary education must fill out a Free Application for Federal Student Aid ("FAFSA"). Here, the student indicates their demographic information, as well as dependence status and income level. Using this information, the government calculates the student's Expected Family Contribution ("EFC"), a measure of how much the student can contribute towards higher education, and uses this, as well as a school's cost of attendance, to determine the amount of the award. The award is disbursed according to a function determined by the U.S. Department of Education:

$$g_{ijm} = G(p_{jm}, EFC_i)$$

where p_{jm} is the cost of attendance and EFC_i is the student's expected family contribution. The financial aid function $G(\cdot, \cdot)$ is described by an award chart that represents a step function in both arguments.[9] Because grants are essentially free money, I assume students will always accept them when offered.

While other grant programs exist, Pell grants are by far the most prominent; in the NPSAS sample, approximately 97.3 percent of all federal grants were from the Pell program. In addition, some institutions, especially four-year nonprofits, will offer their own financial assistance through need-based grants and merit scholarships. However, the NPSAS data also

the wide array of major types offered by community and for-profit colleges, I decided to abstract away from this choice.

[8] The notion that financial aid is a choice variable for the student can be supported by the fact that the US Department of Education issues guidance on which type of aid to accept, and how much of it to receive. For example, they suggest to "borrow only what you need! If your living expenses are not going to be as high as the amount estimated by your school, you have the right to turn down the loan or to request a lower loan amount." See https://studentaid.gov/complete-aid-process/accept-aid.

[9] An example of the award chart: https://ifap.ed.gov/dpcletters/attachments/GEN1502Attach.pdf

shows that this is rare for for-profit colleges, community colleges, and nonselective nonprofits; the third quartile of institutional aid is zero within all three sectors.

3.1.3 Student loan choice

There are two types of loans the student can non-exclusively borrow from: federal and private. The student's objective is to select the amount of each loan to borrow such that her present value of utility is maximized. Because she is allowed to borrow a different amount at each school, the student can solve for her optimal loan profile condition on attending each college. Define latent variables

$$\ell_{ij}^f = (1/\lambda_i^f) + (p_j - y_i - g_{ij}); \tag{5}$$

$$\ell_{ij}^p = (1/\lambda_i^p) + (p_j - y_i - g_{ij}). \tag{6}$$

Observe that these represent the first-order conditions from maximizing indirect utility V_{ij} over federal and private loans, respectively, without any loan limits. These latent variables give an intuitive representation of the optimal loan choice decision; the student will borrow enough to cover the cost of attending college j, plus an additional amount that will depend on the price of each loan (one over its accrued interest).

The relationship between interest rates for federal and private loans will determine the optimal loan profile. Specifically, the composition of federal and private loans will depend on which is less expensive. Consider two cases. First, suppose private loans are more expensive than federal loans, i.e. $\lambda_i^f \leq \lambda_i^p$. The student will borrow her optimal amount of federal loans ℓ_{ij}^f, unless it is below zero or above the federal limit $\bar{L}$. If the latent variable is below zero, she will not borrow anything. If the latent variable is above the federal limit, she will borrow the federal limit, then consider borrowing from private loans accounting for the fact she already acquired $\bar{L}$ in federal loans. The choice in private loans is then given by $\ell_{ij}^p - \bar{L}$, as long as this value is above zero.[10] As a result, the optimal loan profile in this case can be summarized as a censored function of the latent variables: $l_{ij}^f = \min\{\max\{\ell_{ij}^f, 0\}, \bar{L}\}$ and $l_{ij}^p = \max\{\ell_{ij}^p - \bar{L}, 0\}$.

Second, suppose private loans are less costly than federal student loans, i.e. $\lambda_i^p < \lambda^f$. In this case, the student borrows everything in private loans, as long as the private loan latent variable ℓ_{ij}^p is above zero. As a result, the optimal loan profile when private loans are less expensive will be given by $l_{ij}^f = 0$ and $l_{ij}^p = \max\{\ell_{ij}^p, 0\}$. Overall, the optimal loan portfolio

[10]Observe that when federal loans are cheaper, the student will only borrow from a private lender when she reaches the loan maximum in federal loans; if the student can borrow an additional dollar in federal loans, she will be better off doing so relative to taking that dollar out in private loans.

choice can succinctly be written as:

$$l_{ij}^{f} = I[\lambda_i^f \leq \lambda_i^p] \times \min\{\max\{\ell_{ij}^f, 0\}, \bar{L}\} \tag{7}$$
$$l_{ij}^{p} = \max\{\ell_{ij}^p - l_{ij}^f, 0\}, \tag{8}$$

such that $I[\cdot]$ is the indicator function.

3.1.4 College choice

Given the optimal loan profile for each college, the student then selects the school that maximizes her expected utility. I normalize utility tomorrow in the outside option of not attending college as zero. Assuming ε_{ij} is identically and independently distributed Type I Extreme Value, the probability college j is chosen by student i is given by:

$$s_{ij} = \frac{\exp(\alpha\Delta_{ij} + \mathbf{x}_j^T\boldsymbol{\beta} + \sum_n x_{jn}\mathbf{d}_i^T\boldsymbol{\gamma}_n + \xi_j)}{\sum_{k\in\mathbf{S}}\exp(\alpha\Delta_{ik} + \mathbf{x}_k^T\boldsymbol{\beta} + \sum_n x_{kn}\mathbf{d}_i^T\boldsymbol{\gamma}_n + \xi_k)}, \tag{9}$$

where $\Delta_{ij} = \log(y_i - n_{ij}) - \lambda_i^f l_{ij}^f - \lambda_i^p l_{ij}^p$. The number of enrolled students at college j, i.e. the demand for college j, can be defined by integrating over all individual-level characteristics and multiplying by the number of students in the market:

$$D_j(p_j, \mathbf{p}_{-j}) = n_m \int s_{ij}\, dF_f dF_p dF_{\mathbf{d}} \tag{10}$$

such that $F_{\mathbf{d}}$ represents the cumulative distribution function of student characteristics.

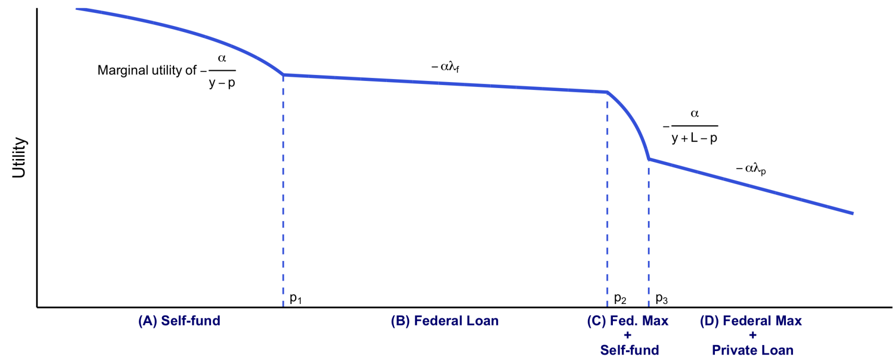
3.1.5 Discussion

To analyze the impact of free community college, it is necessary to understand students' price sensitivity. In other words, the extent to which the policy leads to significant changes depends on the price elasticity of demand. Under the demand model discussed above, the elasticity of demand is given by:

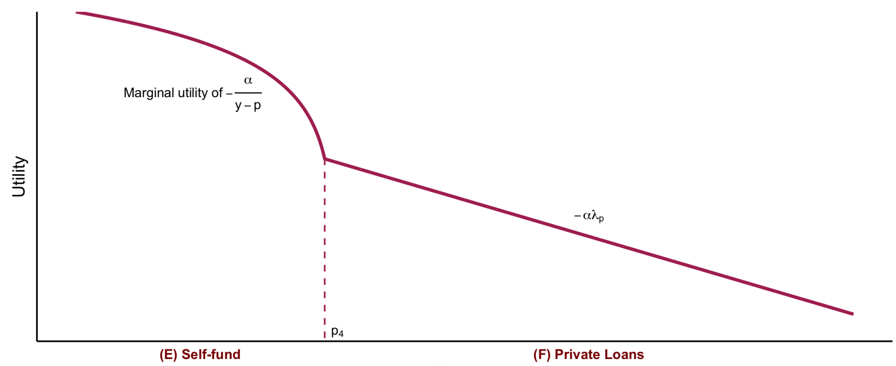
$$\eta_{jk} = \begin{cases} \dfrac{p_j}{D_j} \times n_m \displaystyle\int \alpha \frac{\partial\Delta_{ij}}{\partial p_j}\ s_{ij}(1 - s_{ij})\ dF_f dF_p dF_{\mathbf{d}} & \text{if } j = k, \\ -\dfrac{p_j}{D_j} \times n_m \displaystyle\int \alpha \frac{\partial\Delta_{ik}}{\partial p_k}\ s_{ij}s_{ik}\ dF_f dF_p dF_{\mathbf{d}} & \text{if } j \neq k, \end{cases}$$

such that $\alpha\frac{\partial\Delta_{ij}}{\partial p_j}$ represents the marginal utility of price. The marginal utility of price directly measures how the utility for each higher education option changes when there is a small

Figure 1: *Plot of indirect utility as a function of tuition price for customers with lower interest for federal loans (top), and lower interest for private loans (bottom).*



(a) $\lambda_f < \lambda_p$



(b) $\lambda_f > \lambda_p$

increase in tuition, and will dictate substitution patterns. In particular, it is written as follows:

$$\frac{\partial \Delta_{ij}}{\partial p_j} = \left(-\frac{1}{y_i - n_{ij}} (1 - z_{ij}^f - z_{ij}^p) - \lambda_i^f z_{ij}^f - \lambda_i^p z_{ij}^p \right),$$

where z_{ij}^f and z_{ij}^p are indicators for whether or not student i is borrowing from federal and private loans, respectively, and is below the maximum.

Thus, the extent to which students are responsive to changes in the tuition price depends on whether they are taking out a student loan, and if they can borrow more. Consider the case where the student is not borrowing from either federal or private loans, so that $z_{ij}^f = z_{ij}^p = 0$. Suppose, tuition increases by a small amount, say \$1. In this case, the small increase will likely not alter the students decision to borrow or not borrow. Thus to attend institution j, the student will pay for the additional dollar by reducing her consumption today by \$1. As a result, the student's utility will change by $-\alpha/(y_i - n_{ij})$.

Next, consider the case where the student is taking out federal loans, but not at the maximum, so that $z_{ij}^f = 1$ and $z_{ij}^p = 0$. In this case, the student's optimal response to a \$1 increase in price is to borrow an additional dollar of federal loans. Thus, the student's consumption today will be unaffected, since she can consume the same amount, but her consumption tomorrow will decrease by the accrued interest she must pay for that additional dollar. The marginal utility of price is then $-\alpha\lambda_i^f$. Finally, if the student is taking out a private loan, such that $z_{ij}^p = 1$ and $z_{ij}^f = 0$, her consumption tomorrow will decrease by the accrued interest she must pay for that additional dollar, and the marginal utility of price is $-\alpha\lambda_i^p$.[11]

Figure 1 plots a hypothetical example of how indirect utility changes as tuition increases for students with $\lambda^f < \lambda^p$, and $\lambda^f \geq \lambda^p$. Panel (a) represents how utility changes with tuition for students in which federal loans are cheaper ($\lambda^f < \lambda^p$). For prices less than p_1, the student will not take out any loans and will self-fund her education yielding a marginal utility of $-\alpha/(y-p)$. In other words, she will forgo consumption today to pursue higher education. For prices above p_1, but below p_2, the student will only borrow from federal loans, where at price p_2 the federal loan limit will be reached. Thus, between p_1 and p_2, the student will forgo consumption tomorrow for her education, and will have a marginal utility of $-\alpha\lambda^f$. Between p_2 and p_3, the student will borrow the federal limit, and self-fund the difference between tuition and the federal maximum. In this range, the student will self-fund because private loans are too expensive, but she cannot borrow any more from federal loans. For prices above p_3, the student will turn to borrowing private loans, and will have a marginal utility of $-\alpha\lambda^p$. Panel (b) represents the analogous case for students in which private loans are cheaper ($\lambda^p < \lambda^f$); in this case, the student will self-fund until the price reaches p_4, in which case she will turn to private loans.

3.2 Higher education institutions

For-profit institutions are assumed to be profit maximizers. Suppose each firm f operates a set of institutions $\mathbf{S}_f^{FP}$ in market m. The firm sets the tuition price of institution $j \in \mathbf{S}_f^{FP}$ by maximizing its joint profits:

$$\underset{p_j}{\text{argmax}} \quad \sum_{k \in \mathbf{S}_f^{FP}} (p_k - c_j) D_k(p_j, \mathbf{p}_{-j}),$$

where p_j is the tuition price, c_j is the marginal cost of enrolling an additional student, and $D_k(\cdot)$ is the number of enrollees in institution k as a function of institution j's tuition, as well as the tuition of all other institutions in the market $\mathbf{p}_{-j}$. The first order conditions of

[11] Recall that the optimal loan profile will either take out only private loans, or will take out private loans if the student has exhausted all federal loans. Thus, z_{ij}^p and z_{ij}^f can never be equal to 1 at the same time.

the for-profit firm's optimization problem is given by:

$$D_j(p_j, \mathbf{p}_{-j}) + \sum_{k \in \mathbf{S}_f^{FP}} (p_k - c_k) \frac{\partial D_k(p_j, \mathbf{p}_{-j})}{\partial p_j} = 0. \tag{11}$$

Nonprofit colleges, unlike for-profit colleges, do not necessarily maximize profits.[12] Instead, nonprofits, especially community colleges and public nonselective 4-year institutions, were established to provide affordable access to higher education. To model the objective function of nonprofit colleges, I assume that institutions set prices to achieve two goals: (1) maximizing profits, to optimize the amount the institution can reinvest on student services and amenities, and (2) the value it provides to students in the market, as measured by students' average willingness-to-pay for college j.[13]. Let $\mathbf{S}_n^{NP}$ represent the set of community colleges owned by nonprofit entity n. The pricing problem is to set a tuition for college j that maximizes the system's joint objectives:

$$\underset{p_j}{\operatorname{argmax}} \quad \left(\sum_{k \in \mathbf{S}_n^{NP}} (p_k - \kappa_k) D_k(p_j, \mathbf{p}_{-j}) \right)^{\omega_j} \left(V(p_j, \mathbf{p}_{-j}, \mathbf{S}) - V(\mathbf{p}_{-j}, \mathbf{S}/\mathbf{S}_n^{NP}) \right)^{1-\omega_j}, \tag{12}$$

such that κ_j is the marginal cost of nonprofit college j net of any incremental subsidies received from non-revenue sources (e.g. state governments, alumni donations, etc.), ω_j is the weight institution j places on profit maximization, and $V(p_j, \mathbf{p}_{-j}, \mathbf{S})$ is the average value potential students have from being able to choose from schools in set $\mathbf{S}$ at prices $\mathbf{p}$. Given the demand model, this value is given by the expected utility of the student's utility maximizing choice:

$$V(p_j, \mathbf{p}_{-j}, \mathbf{S}) = \int \log \left(\sum_{j \in \mathbf{S}} \exp(\alpha \Delta_{ij} + \mathbf{x}_j' \boldsymbol{\beta} + \sum_n x_{jn} \mathbf{d}_i' \boldsymbol{\gamma}_n + \xi_j) \right) dF_f dF_p dF_{\mathbf{d}}.$$

The second component of equation (12) yields the difference in the value of college choice set $\mathbf{S}$ and the value of the same set without college j's system $\mathbf{S}_n^{NP}$, which I interpret as an estimate of the average value the nonprofit system provides to the market.

The optimal tuition price for nonprofit colleges can be solved by taking the (log) first-

[12]The notion that nonprofit colleges and for-profit colleges have a different objective function is common in the structural higher education literature. For example, Fu (2014) and other earlier work assume that the objective function of prestigious nonprofit universities is comprised of the quality of enrolled students (as measured by ability) and revenue.

[13]A nonprofit college may also be interested in the value it provides to other parts of the community beyond students. However, because education is a college's primary function, I assume nonprofit colleges set prices based on its impact to current and prospective students.

order conditions and rearranging; the first-order condition is given by:

$$D_j(p_j, \mathbf{p}_{-j}) + \sum_{k \in \mathbf{S}_n^{NP}} \left[\frac{\partial D_k}{\partial p_j} + \frac{1-\omega_j}{\omega_j} \left(\frac{D_k(p_j, \mathbf{p}_{-j}) A_j}{B_j} \right) \right] (p_k - \kappa_k) = 0 \tag{13}$$

where

$$\begin{aligned} A_j &= \frac{\partial V(p_j, \mathbf{p}_{-j}, \mathbf{S})}{\partial p_j} = \int \alpha \frac{\partial \Delta_{ij}}{\partial p_j} s_{ij} dF_f dF_p dF_{\mathbf{d}} \\ B_j &= V(p_j, \mathbf{p}_{-j}, \mathbf{S}) - V(\mathbf{p}_{-j}, \mathbf{S}/\mathbf{S}_n^{NP}). \end{aligned}$$

As a result, the first-order conditions for nonprofit colleges resembles that of for-profit colleges, with an additional term that represents the school's preference for maximizing its value to the market. The objective function of nonprofit colleges, given by equation (12), resembles a Nash bargaining problem similar to the models used in the health economics literature, where competition models assume hospital systems and health insurers bargain over reimbursement rates (Town and Vistnes, 2001; Capps et al., 2003; Gowrisankaran et al., 2015).[14] Here, a nonprofit colleges can be seen as "negotiating" with itself between two opposing goals: financial success and the value it provides to the community.

The solution concept for tuition pricing at for-profit and nonprofit colleges is a Nash equilibrium. Each for-profit college solves (11) and each nonprofit institution solves (13) given the prices of all other schools $\mathbf{p}_{-j}$. Within each market, equilibrium tuition prices solve the system of equations given by (11) and (13), for all schools j.

4 Identification and Estimation

I estimate the parameters of the model in three steps. First, I use the student loan choice model to estimate the distribution of federal and private accrued interest, λ_i^f and λ_i^p. Second, I estimate the parameters of the demand model, given the accrued interest distributions. Finally, I use the demand model to estimate the marginal cost of for-profit colleges and the net marginal cost and profit weight of nonprofit colleges. This section proceed as follows: I begin by discussing market definition and the data sources used for estimation. I then describe the identification and estimation of each step in turn: the loan parameters, the demand parameters, and the supply-side parameters.

[14]In the hospital setting, willingness-to-pay is used to approximate the hospital system's value to the insurer's profits, e.g. marketability to potential beneficiaries. In this setting, I assume willingness-to-pay is a direct objective of a nonprofit college.

4.1 Market definition

To define each market, I consider two components: (1) a product component and (2) a geographic component. In terms of the product component, I include community colleges, degree-granting for-profit colleges, and nonselective four-year nonprofit institution. Degree-granting for-profit colleges refer to those that have a wide range of programs offered that are comparable to the programs offered by community colleges. This excludes very narrow vocational certificates and degrees, such as cosmetology or culinary arts, since students pursuing these fields may only consider a narrower set of colleges related to their desired vocation. Nonselective four-year institutions include public and private nonprofit colleges that are designated as open enrollment, i.e. do not require test scores, high school grade point average, or letters of recommendation for admission. In terms of the geographic component, I assume markets are defined by core-based statistical areas ("CBSA").[15] While previous studies, particularly for for-profit colleges, have defined a market as a county (Cellini, 2010; Cellini et al., 2016), CBSAs (which are groups of counties) more appropriately capture students willingness to travel for nonselective four-year universities.

4.2 Data

The data used for analysis comes from two primary sources: The National Postsecondary Student Aid Study ("NPSAS") and the Integrated Postsecondary Education Database System ("IPEDS"). The NPSAS is a restricted-use, nationally representative random sample of first-time college students obtained through the National Center for Education Statistics. It is used to study characteristics of students in postsecondary education, with a focus on finances and financial aid decisions. This includes information on income, expected family contribution, grants received, and student loans borrowed. IPEDS, on the other hand, is a public-access survey of aggregate-level information directly from postsecondary institutions, including tuition prices, enrollment totals, and other institutional characteristics.

The IPEDS sample sample contains 1,167 institutions that includes 500 community colleges, 545 for-profit colleges, and 122 nonselective four-year institutions. Table 1 presents descriptive statistics for both samples. Panel A presents information about the IPEDs data, broken down by higher education sector. The table suggests that community colleges have the largest student population, and the lowest tuition prices across the three sectors. Mean-

[15] To determine students who chose the outside option in the market share calculations, I used estimates from the American Community Survey ("ACS") of individuals between the ages of 18 and 35 with only a high school diploma or equivalent, and not already enrolled in college. Other conditions were placed to approximate for students who preferred colleges outside of the product component of the market definition. I simulate demographics by drawing from the NPSAS sample for enrolled students and the ACS for students not in higher education, separately by market, in proportions implied by the market shares.

Table 1: Descriptive statistics from the IPEDS and NPSAS sample

	All Sectors	Community College	For-profit College	Nonselective 4-year
	A. IPEDS			
Observations	1,167	500	545	122
Avg. enrollment	2,522	4,039	889	3,600
Avg. tuition ($)	5,515	2,685	16,622	6,280
Avg. student-faculty ratio	23.5	23.8	23.6	21.8
Offer life credits (%)	62.3	67.4	36.9	67.4
Offer distance learning (%)	91.1	99.8	49.9	96
Offer evening classes (%)	72.3	74.5	68	67
Offer placement services (%)	87.3	87.3	82	93.2
	B. NPSAS			
Observations	18,650	8,760	8,210	1,680
Avg. Pell Grant ($)	2,615	2,371	2,909	2,444
Avg. Federal Loan ($)	3,854	1,579	6,363	3,451
Avg. Private Loan ($)	552	67	1,143	193
Avg. Income ($)	38,965	41,788	34,374	46,682
Avg. EFC ($)	4,810	5,285	3,992	6,335
% Dependent	53.4	60.7	43.7	62.7
% Female	52	53	50.9	52.8
% Minority	30	31.7	29	26.7
% Older than 25	29.3	23.9	36.2	23.7

NOTE: Observations in IPEDS sample represents number of institutions, while observations in NPSAS sample represents number of students. Tuition is the published price in IPEDS.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), Fall 2011; and 2011-12 National Postsecondary Student Aid Study (NPSAS:12).

while, for-profit colleges had an average tuition price that was several times larger than that of community colleges, and nonselective four-year colleges had enrollments and tuition between the other two sectors. The last five rows of Panel A displays institutional characteristics; nonselective colleges tend to have smaller class sizes, while for-profit colleges are less likely to offer credits for life experience, distance learning, and placement services than the other sectors. In addition, community colleges are more likely to offer evening classes.

The NPSAS sample contains 18,650 observations of students who enrolled in one of the IPEDS institutions. Panel B of Table 1 displays descriptive statistics for the NPSAS sample, again broken down by higher education sector. Overall, the table suggests that the NPSAS sample yields a relatively strong representation of the IPEDS sample. In particular, the volume of students in each sector broadly follows the respective proportion of enrollment from the IPEDS sample. Furthermore, the table suggests that for-profit students receive more Pell Grants, while simultaneously borrowing more from both federal and private loans. Finally, the student characteristics displayed in the last six rows of the panel corroborate the notion that students attending community and for-profit colleges are less traditional; these are students that tend to come from backgrounds that are lower income, are less dependent on their parents, are more likely to be minorities, and are more likely to be older. These features are significantly more pronounced at for-profit colleges.

4.3 Financial aid parameters

As a first step, I identify the distribution of the accrued interest parameters, λ_i^f and λ_i^p, using the student-level NPSAS data. In particular, I assume that the distribution of accrued interest on federal loans F_f is given by a log-normal distribution with parameters $\mathbf{d}_i^T\boldsymbol{\tau}^f$ and σ_f. In other words, I assume

$$\lambda_i^f = \exp(\mathbf{d}_i^T\boldsymbol{\tau}_f + \sigma_f\eta_i),$$

where η_i is a standard normal random variable; each student's accrued interest on federal loans may depend on their demographics $\mathbf{d}_i$, and an unobserved shock η_i. Furthermore, I assume the accrued interest on private loans λ_i^p takes the form:

$$\lambda_i^p = \lambda_i^f \exp(\mathbf{d}_i^T\boldsymbol{\tau}_p + \sigma_p\eta_i^p),$$

where η_i^p is a standard normal random variable independent of η_i. In other words, the term $\exp(\mathbf{d}_i^T\boldsymbol{\tau}_p+\sigma_p\eta_i^p)$ represents the percent difference in the marginal accrued interest of private loans relative to federal loans and is log-normally distributed with mean $\mathbf{d}_i^T\boldsymbol{\tau}_p$ and standard deviation σ_p.

Under the parametric assumptions outlined above, the parameters $\boldsymbol{\theta}_1 = (\boldsymbol{\tau}_f, \sigma_f, \boldsymbol{\tau}_p, \sigma_p)$ can be identified by observing students' behavior when borrowing from federal and private lenders. To see this, consider Table 2, which displays the five cases in which loan portfolios can be observed in the NPSAS data. In case 1, the student does not take out any loans. In case 2, the student borrows only federal loans, but not at the maximum. In case 3, the student borrows only the maximum amount of federal loans. In case 4, the student borrows the maximum amount of federal loans, and some private loans. In the last case, the student only has private loans. The columns represent the conditions on the latent variables ℓ^f and ℓ^p in which each case can occur, depending on the relative accrued interest values.

The structure of the observed cases in Table 2 helps identify the relationship between federal and private loan interest. In particular, if a student *ever* takes out any federal loans, it is known that $\lambda_i^f < \lambda_i^p$, while if the student only takes out private loans, it is known that $\lambda_i^p \leq \lambda_i^f$. As a result, the parameters $\boldsymbol{\tau}_p$ can be identified by comparing student demographics

Table 2: Federal and Private Loan Observation Cases

Case	Observation	$\lambda_i^f < \lambda_i^p$	$\lambda_i^p \leq \lambda_i^f$	# of Obs.
1	$(0, 0)$	$\ell^f \leq 0$	$\ell^p \leq 0$	7,790
2	$(f, 0)$	$0 \leq \ell^f = f < \bar{L}$	–	5,100
3	$(\bar{L}, 0)$	$\ell^p < \bar{L} \leq \ell^f$	–	4,110
4	$(\bar{L}, z - \bar{L})$	$\bar{L} < \ell^p = z$	–	1,510
5	$(0, z)$	–	$0 \leq \ell^p = z$	150

NOTE: The variables f and z are positive real numbers that represent the observed amount of federal and private loans in each case, respectively. Number of observations are rounded to the nearest ten.
SOURCE: U.S. Department of Eduction, National Center for Education Statistics, 2011-12 National Postsecondary Student Aid Study (NPSAS:12)

in situations where the student ever takes out a federal loan (cases 2-4), with those that only borrowed from private loans (case 5). Furthermore σ_p is identified by capturing the remaining variance not explained by the demographics.

While the observation of cases can help identify the private loan interest parameters $\boldsymbol{\tau}_p$ and σ_p, the actual level of loans borrowed can be used to identify the federal interest parameters, $\boldsymbol{\tau}_f$ and σ_f. In particular, recall that the optimal loan profile without any constraints takes the form of the latent variables given in equations (5) and (6). That is, the optimal amount of borrowing can be described by the sum of the inverse of their accrued interest and the amount tuition exceeds income and grants. In the case with no constraints, the difference of observed loans and the amount needed to cover tuition would yield a function of the unknown parameters:

$$loan_i - (p_j - y_i - g_{ij}) = \frac{1}{\lambda_i(\boldsymbol{\tau}_f, \sigma_f)},$$

where λ_i represents either λ_i^f or λ_i^p, which is a function of the parameters $\boldsymbol{\tau}_f$ and σ_f, and $loan_i$ is the observed amount of loans borrowed. In particular, the entire left-hand side is completely observed in the data. This relationship illustrates how implied interest rates are identified; the larger the amount the student observably borrows above what is needed to attend college, the lower the implied interest rate for the given type of loan.

Accounting for the fact that federal loans are censored above and below, and that private loans are censored from below, variation in the difference of observed loans and the amount tuition exceeds income and grants in the NPSAS sample can be used to identify $\boldsymbol{\tau}_f$ and σ_f, given that $\boldsymbol{\tau}_p$ and σ_p are already identified. Generally, this will take a form that resembles a censored regression. Furthermore, because the observed loans will depend on which type of loan is less expensive, the likelihood of each cases resembles a finite mixture of censored regressions, such that the mixing probability is the probability that private interest exceeds federal interest, which is known given that the relationship between λ_i^p and λ_i^f can be identified.

Thus, variation in student loan borrowing behavior between federal and private sources can be used to identify the set of parameters $\boldsymbol{\theta}_1 = (\boldsymbol{\tau}_f, \sigma_f, \boldsymbol{\tau}_p, \sigma_p)$. To estimate the loan choice model, let $P_{ik}(\boldsymbol{\theta}_1)$ represent the probability that student i is in observation case k, according to Table 2. These probabilities will be simulated because the probability is integrated over the unobservable portion of private loan interest; a detailed description is outlined in Appendix A. Let z_{ik}^{ℓ} be an indicator for whether the loan profile of student i belongs in case k. I estimate the parameters $\boldsymbol{\theta}_1$ by maximizing the simulated log-likelihood function:

$$\widehat{\boldsymbol{\theta}}_1 = \arg\max_{\boldsymbol{\theta}_1} \sum_i \sum_k z_{ik}^{\ell} \log\left(P_{ik}(\boldsymbol{\theta}_1)\right).$$

4.4 College choice parameters

The second step estimates the marginal utility of income and the coefficients on student-school interaction terms, $\boldsymbol{\theta}_2 = (\alpha, \boldsymbol{\gamma})$, which are identified by combining the NPSAS and IPEDS data. In particular, because the model includes institution-specific fixed effects, identification is coming from variation in student-school covariates within the group of students that chose a given college. As an example, consider the coefficient on net income α. Here, there exists individual-level variation in net income through (1) idiosyncratic student income before paying for college, and (2) differences in the amount of federal and private loans each student borrows. The coefficient on net income is then identified by the extent to which students with a higher net income when attending a given college are more likely to attend that college relative to students with less net income.

Unfortunately, the NPSAS sample alone cannot identify the parameters. Specifically, the NPSAS is a random sample of students who decided to attend a higher education institution; those who chose the outside option of not attending college are not represented in the sample. As a result, the individual-level data cannot be used to understand preferences for not attending college, which is particularly important since utility is measured relative to the outside option. To identify the college choice parameters, I follow the discrete choice literature (Petrin, 2002; Berry et al., 2004) and form moments that set quantities predicted by the model equal to their empirical counterparts. In particular, I set predicted demand defined by equation (10) equal to observed enrollments from the aggregate-level IPEDS data (or equivalently, predicted market shares equal to observed market shares). The moment conditions identify the institution-specific fixed effects since there exists unique values of the unobserved institution-specific terms ξ_j that satisfies the conditions, as described by the contraction mapping in Berry (1994). Given that the institution-specific fixed effects are identified through the moment conditions, variation in the NPSAS data can then identify the the remaining college choice parameters $\boldsymbol{\theta}_2$.

Estimation is done using constrained simulated maximum likelihood. The estimates of $\boldsymbol{\theta}_2$ and $\boldsymbol{\xi} = (\xi_1, ...\xi_n)$ are the solution to the following constrained optimization problem:

$$(\widehat{\boldsymbol{\theta}}_2, \widehat{\boldsymbol{\xi}}) = \arg\max_{\boldsymbol{\theta}_2, \boldsymbol{\xi}} \sum_{i=1}^{N} \sum_{j} z_{ij} log(\widehat{s}_{ij}(\boldsymbol{\theta}_2, \boldsymbol{\xi}))$$

$$s.t. \quad D_j(\boldsymbol{\theta}_2, \boldsymbol{\xi}) = enroll_j, \ \forall j$$

such that z_{ij} is an indicator for customer i choosing to attend college j, $\widehat{s}_{ij}(\boldsymbol{\theta}_2, \boldsymbol{\xi})$ is the choice probability as a function of the parameters given by equation (9) integrated over the distribution of federal and private accrued interest, $D_j(\boldsymbol{\theta}_2, \boldsymbol{\xi})$ is the demand of college j as a function of the parameters given by equation (10), and $enroll_j$ represents the observed enrollment of college j from the IPEDS data.[16]

Simulations are used for both the log-likelihood and to compute demand. In the likelihood function, the choice probabilities are integrated over the distribution of the accrued interest variables F_f and F_d. The demand function is also integrated over the accrued interest variables, as well as the distribution of student characteristics $F_{\mathbf{d}}$. While simulated maximum likelihood can be inconsistent for a small number of draws, I use 200 draws for each distribution, which I believe is sufficient. For the demand computation, because each draw is less computationally burdensome, I use 1000 draws from each distribution.

Finally, the mean utility parameters on school characteristics $\boldsymbol{\beta}$ are not estimated through maximum likelihood because they are absorbed into the institution-specific fixed effects. However, they are recovered by applying a minimum distance procedure (Nevo, 2000).

4.5 Supply-side parameters

In the final step, I estimate the supply-side parameters for for-profit colleges, marginal costs c_j, by inverting the first-order conditions given by equation (11). In matrix notation,

[16] In practice, I use the nested fixed point algorithm described in Berry et al. (1995). Additional details can be found in Appendix A.

marginal costs are given by:

$$\mathbf{c} = \mathbf{p} + \boldsymbol{\Omega}^{-1}\mathbf{D},$$

where $\mathbf{c}$ is a vector of marginal costs, $\mathbf{p}$ is a vector of tuition prices, $\mathbf{D}$ is a vector of enrollment quantities, and $\boldsymbol{\Omega}$ is a matrix that takes a value $\partial D_k/\partial p_j$ when row j and column k belong to the same firm f, and zero otherwise.

For nonprofit colleges, I borrow from the hospital bargaining literature and estimate the supply-side parameters, profit weight ω_j and net marginal cost κ_j, following Gowrisankaran et al. (2015). In particular, I assume net marginal costs at nonprofit colleges can be decomposed as:

$$\kappa_j = \boldsymbol{\gamma}\mathbf{v}_j + \iota_j,$$

where $\mathbf{v}_j$ is a vector of state indicator variables, and ι_j is an econometric error. In other words, there exists common subsidies by state that factor into the institution's net marginal cost. Furthermore, I suppose the profit weight varies by nonprofit sub-sector: public university, private university, and community college. The identifying moment condition is that the expectation of the econometric error ι_j, conditional on a set of exogenous covariates $\mathbf{z}_j$ is equal to zero. In matrix notation, the moment condition can be obtained by inverting the nonprofit's first order condition to solve for net marginal cost:

$$E(\boldsymbol{\iota}|\mathbf{z}) = E\left(-\boldsymbol{\gamma}\mathbf{v} + \mathbf{p} + (\boldsymbol{\Omega} + \boldsymbol{\Lambda}(\boldsymbol{\omega}))^{-1}\mathbf{D}|\mathbf{z}\right) = 0,$$

such that $\boldsymbol{\Lambda}(\boldsymbol{\omega})$ is a matrix that takes on the value $(1-\omega)/\omega \cdot A_j/B_j \cdot D_k$ when row j and column k belong to the same nonprofit system, and is equal to zero otherwise.

As instruments $\mathbf{z}_j$, I include state and nonprofit sub-sector indicators, as well as total enrollment and the predicted value of college j to the market. Identification of the parameters $\boldsymbol{\gamma}$ are through a linear instrumental variables regression conditional on marginal costs that can be recovered from the nonprofit's first-order conditions. As discussed in Gowrisankaran et al. (2015), identification of the parameters $\boldsymbol{\omega}$ likely have similar equilibrium implications to fixed effects, and thus cannot easily be identified at the same level of the marginal cost parameters $\boldsymbol{\gamma}$. As a result, $\boldsymbol{\omega}$ varies only by institutional sub-sector, while $\boldsymbol{\gamma}$ varies only at the state level. Finally, I estimate the parameters $\boldsymbol{\gamma}$ and $\boldsymbol{\omega}$ using the general method of moments.

Table 3: *Loan estimation results*

	Federal loans	Private loans
σ	1.101	0.668
	(0.010)	(0.016)
Constant	-6.258	1.531
	(0.052)	(0.082)
Non-white	0.238	-0.034
	(0.021)	(0.032)
Over 25	-0.492	0.020
	(0.028)	(0.042)
Female	-0.005	0.079
	(0.019)	(0.030)
Dependent	-0.454	0.178
	(0.027)	(0.042)
Income (log)	-0.345	-0.015
	(0.005)	(0.007)

NOTE: Standard errors are in parentheses. Estimation done using simulated maximum likelihood with 200 draws. The number of total observations used is 18,650 and maximized likelihood is -21359.6 with a pseudo R-squared of 0.18.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), Fall 2011; and 2011-12 National Postsecondary Student Aid Study (NPSAS:12).

5 Results

Table 3 presents estimates of the financial aid parameters $\widehat{\boldsymbol{\theta}}_1$. The student characteristics used as covariates in the accrued interest model includes race, age, gender, dependence status, and (log) income. The first column presents the parameters for federal interest, while the second column presents the parameters for the difference in federal and private interest. The federal interest parameters suggest that younger, minority students that were not dependent on their parents tend to face higher levels of accrued interest. Furthermore, the coefficient on income was large and negative, suggesting that low income students also faced a higher cost of borrowing student loans. Overall, this implies that conditional on the amount needed to attend a given institution, students that are more traditionally well-off face lower costs to borrowing a loan, and will borrow more all else equal. This is consistent with the literature that suggests students from more nontraditional backgrounds tend to have more trouble paying back loans (Dynarski, 1994; Flint, 1997). The second column of Table 3 shows a large and positive constant for the private loan parameters, suggesting private loans were on

Table 4: *College enrollment estimation results*

		Student Characteristics			
	Mean	Minority	Over 25	Female	Income $\leq$ \$25k
Price (α)	15.465				
	(0.082)				
For-profit	-0.628	0.376	0.099	0.588	4.513
	(0.382)	(0.098)	(0.132)	(0.089)	(0.137)
Student-faculty ratio (log)	-0.459	0.003	-0.796	-0.085	-1.264
	(0.331)	(0.063)	(0.079)	(0.058)	(0.076)
Life credits	0.076	-0.078	0.058	0.333	-0.510
	(0.291)	(0.082)	(0.116)	(0.074)	(0.091)
Distance learning	-1.646	-0.498	-0.866	0.382	-1.037
	(0.615)	(0.152)	(0.176)	(0.143)	(0.167)
Evening courses	0.007	-0.483	0.371	0.368	-0.380
	(0.308)	(0.091)	(0.125)	(0.080)	(0.100)
Placement services	-0.626	-0.025	-1.269	-0.294	0.640
	(0.468)	(0.123)	(0.143)	(0.109)	(0.135)
Urban location	0.517	0.579	0.130	0.075	0.235
	(0.262)	(0.078)	(0.109)	(0.070)	(0.087)

NOTE: Standard errors are in parentheses. Estimation done using simulated maximum likelihood with 200 draws. The number of total observations used is 18,650, and maximized likelihood is -48096.6 with a pseudo R-squared of 0.25. Estimates for mean institutional characteristics (except for price) are estimated using a minimum distance estimator.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), Fall 2011; and 2011-12 National Postsecondary Student Aid Study (NPSAS:12).

average more expensive. Furthermore, female and dependent students faced higher private interest, suggesting these students borrowed less from private institutions relative to other students.

Table 4 presents estimates of the parameters to the college choice model. Standard errors are corrected for the two-step procedure, following Murphy and Topel (2002). The first column presents the mean utility estimates for eight institutional characteristic covariates, including net cost of attendance, for-profit status, the (log) student-faculty ratio, whether the institution accepts life experiences as credits, whether the institution offers distance/online learning opportunities, whether the institution offers evening courses, the existence of placement services, and whether the college is located in an urban locale. The coefficient on net cost of attendance is large and positive, suggesting that students are very price sensitive. The mean utility estimates also suggests that students on average receive a disutility from attending a for-profit college, possibly influenced by a contemporaneous wave of negative press. Columns 2-5 provide interactions between the covariates and student characteristics, which include race, age, dependence status, and income (whether the student's income is below $25,000). While students dislike for-profit colleges on average, low-income students have a strong preference for them. In addition, students generally have a disutility for large class sizes, while older students have a preference for evening classes, likely because some may hold full-time jobs.

Furthermore, the median implied own-price elasticity for community colleges is 1.26, with 42 percent of institutions in the sample below 1, suggesting that the prices set by community colleges are generally not consistent with profit maximization.[17] Nonselective four-year colleges have a median elasticity slightly higher than community colleges of 2.21, with 20 percent having an own-price elasticity below 1. Finally, the median own-price elasticity for for-profit colleges is about 3.32, with only one for-profit college in the sample having a value below 1, suggesting that pricing at for-profit institutions is in fact consistent with profit maximization.

Under the equilibrium assumptions, marginal costs for for-profit institutions, and profit weights and net marginal costs for nonprofit institutions can be recovered. The first row of Table 5 display the first quartile, median, and third quartile marginal costs at for-profit colleges. The median marginal cost is $10,172, with an interquartile range of $3,637. Furthermore, the median margin percentage (given by price minus cost divided by price) for for-profit colleges is 30 percent, with a 25th percentile of 25 percent and an 75th percentile of 36 percent. In comparison, the United States Senate Committee on Health, Education, Labor, and Pensions (2012) surveyed 30 for-profit colleges, and found an average profit margin of 19.4 percent. However, these are profit margins that likely include fixed costs that would not be included as incremental.

For nonprofit colleges, the next panel displays estimates of the profit weights, which are broken down by community college, private nonselective four-year, and public nonselective four year. The profit weight for community colleges is relatively low, indicating that tuition prices are set to maximize the value the college provides to consumers. This effectively implies that community colleges are pricing near their net marginal cost. The estimated profit weight for private four-year colleges is 1, its theoretical maximum, implying that these institutions are profit-maximizers. Considering the anecdotal evidence of private colleges converting between nonprofit and for-profit status (Wong, 2015), it should not be a surprising result that they are profit maximizing. Finally, the profit weight for public four-year colleges is 0.29, suggesting that while there is a profit motive, institutions put relatively more weight on their community value.

[17]Table A.1 present the median own- and cross-price elasticities for each sector of higher education.

Table 5: *Supply-side summary*

Marginal cost of for-profits	
25th percentile	\$8,353.56
50th percentile	\$10,172.48
75th percentile	\$11,990.99
Profit weight for nonprofits	
Community college	0.044
Private 4-year nonprofit	1.000
Public 4-year nonprofit	0.285
Net marginal cost of 4-year nonprofits	
25th percentile	\$1,712.54
50th percentile	\$2,872.75
75th percentile	\$3,714.83

NOTE: Full details on supply-side parameter estimates are available in Appendix Table A.2.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), Fall 2011; and 2011-12 National Postsecondary Student Aid Study (NPSAS:12).

The last panel of Table 5 displays the implied net marginal cost for nonselective four-year colleges (omitting community colleges since they price near, or at marginal cost). Median marginal costs are much lower than the for-profit sector, at about \$2,826, with an interquartile range of \$2,005. This suggests that the average margin percentage is 10%. However, for private four-year colleges, median marginal costs are unsurprisingly higher at \$6,112, with a larger interquartile rate of \$6,666 and an average margin percentage of 45%.

6 Counterfactuals

I use the results of the model to examine the effect of free community college. In particular, I look at the impact of free community college on (1) post-policy equilibrium prices, quantities, and inter-sector substitution patterns, (2) consumer welfare as measured by compensating variation, and (3) postsecondary degree completions. I begin by looking at the counterfactual of an entirely free community college system, then compare it to the implementation of more financially practical programs. This includes a last dollar program, where costs are covered after accounting for other federal grant aid, and a need-based program that makes community college free for low-income students.

In order to conduct such counterfactuals, the following assumptions must be made. First, I assume that community colleges are not capacity constrained and can adapt to a large influx of students.[18] Second, the characteristics and objectives of higher education institutions do not change due to the implementation of free community college. For example, it may be possible for for-profit or nonselective four-year colleges to offer a different set of programs, or change its profit weight, in response to the policy. Third, I also assume that there is no entry or exit of institutions in the short run; enrollment loses to community colleges could alter the viability of operation in the long run for some institutions. Finally, I assume that free community college is supplementary to existing aid programs, like Pell Grants, and does not replace them.

6.1 Equilibrium effects

Table 6 summarizes the equilibrium changes after the introduction of each free college program. Panel A displays pre-policy enrollment, the average tuition price paid, and the average amount borrowed in federal loans, by higher education sector, as a baseline reference. Using the aggregate-level IPEDS data, the baseline for the counterfactual analysis consisted of 2.9 million total students enrolled in community colleges, for-profit colleges, and nonselective four-year colleges in 2012. The average tuition price paid among these students was $5,515, and the average federal loan amount borrowed was slightly higher at $5,940. Next, I examine each policy in turn.

6.1.1 Free community college

Under the counterfactual in which every students can access community colleges for free, for-profit and nonselective four-year colleges will set a post-policy equilibrium tuition that satisfies their first order conditions, taking the tuition of other higher education institutions as given (with the price at community colleges set to zero). Formally, I assume this policy ensures that the price of student i when attending college j is given by

$$p_{ij} = (1 - comm_j)p_j,$$

where $comm_j$ is an indicator for whether or not institution j is a community college.

Panel B of Table 6 presents changes in these quantities due to the introduction of free community college; the first row examines that the average total tuition paid by students, not including grants and loans. In aggregate, students would pay 54 percent less in tuition after the introduction of free community college. At the same time, students that remain

[18] While this assumption may appear unreasonable given recent concerns about inadequate infrastructure at some community colleges, we could reasonably expect that any federal legislation to introduce free community college would be accompanied by a strategy to address these capacity issues.

Table 6: *Counterfactual results for equilibrium effects*

	Total	Community College	For-profit College	4-year Nonprofit
A. Pre-policy				
Enrollment	2,943,021	2,019,505	484,369	439,147
Avg. tuition paid	\$5,515	\$2,685	\$16,622	\$6,280
Avg. fed. loan borrowed	\$5,940	\$5,402	\$8,260	\$5,857
B. Free community college				
Δ tuition paid	-54.77%	-100.00%	6.73%	5.45%
Δ enrollment	26.53%	47.71%	-18.58%	-21.15%
Substitution to comm. coll.	-	-	-15.96%	-19.07%
Income $<$ \$25k	-	-	-24.00%	-32.10%
\$25k $\leq$ Income $<$ \$75k	-	-	-16.56%	-16.85%
Income $\geq$ \$75k	-	-	-6.53%	-7.50%
Entry	27.28%	39.69%	0.03%	0.23%
Exit	-0.74%	0.00%	-2.54%	-2.14%
C. Last dollar program				
Δ tuition paid	-20.98%	-41.70%	-0.14%	0.14%
Δ enrollment	9.77%	16.27%	-2.79%	-6.29%
Substitution to comm. coll.	-	-	-3.02%	-5.38%
Income $<$ \$25k	-	-	-0.87%	-5.02%
\$25k $\leq$ Income $<$ \$75k	-	-	-4.79%	-5.14%
Income $\geq$ \$75k	-	-	-4.38%	-6.20%
Entry	9.90%	14.31%	0.28%	0.24%
Exit	-0.12%	0.00%	-0.09%	-0.73%
D. Need-based program				
Δ tuition paid	-27.30%	-51.69%	6.02%	8.95%
Δ enrollment	11.62%	23.36%	-13.66%	-14.50%
Substitution to comm. coll.	-	-	-10.97%	-11.56%
Income $<$ \$25k	-	-	-23.78%	-32.44%
\$25k $\leq$ Income $<$ \$75k	-	-	-3.27%	-3.09%
Income $\geq$ \$75k	-	-	-1.12%	-1.01%
Entry	12.47%	18.16%	0.04%	0.01%
Exit	-0.83%	0.00%	-2.63%	-2.69%

NOTE: The change in enrollment calculates the percent change due to each policy relative to the baseline. Substitution to community colleges measures the percent of students who chose a for-profit or 4-year nonprofit college without each policy, but would switch to a community college under the counterfactual. Similarly, entry (exit) measures the percent of students who chose the outside option (a higher education option) without each policy, but would choose a higher education option (the outside option) in the counterfactual.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), Fall 2011; and 2011-12 National Postsecondary Student Aid Study (NPSAS:12).

at for-profit colleges would pay 6.7 percent more, which represents an approximately \$1048 increase, while tuition at nonprofits students on averaged saw a 5.5 percent increase, or \$324.

At first glance, the increase in for-profit and 4-year nonprofit tuition may be surprising given the negative shock to demand. However, this can be explained by *which* students are leaving those institutions.

Next, I examine changes to enrollment; the model suggests that enrollment would increase significantly, by 26 percent. Furthermore there would be an almost 47 percent increase at community colleges, with decreases of 18 percent and 21 percent at for-profit and nonselective four-year colleges, respectively. This implies that the policy would result in 963,520 more students enrolled at community colleges, or 1,927 additional students per college. Furthermore, 89,974 less students would enroll in for-profit colleges, or 165 per college, and 92,889 less students would enroll in nonselective nonprofits, or 761 per college. The next row of Panel B displays substitution patterns caused by the introduction of free community college, as measured by the percent of students within each sector to switch to community college because of the policy. In particular, 16 percent of students that would choose to attend a for-profit college without the policy would substitute to a community college. Likewise, 19 percent of students that would choose to attend a nonselective four-year university would switch to a community college.

The next three rows illustrate why tuition at for-profit and nonselective four-year colleges increase in the post-policy equilibrium. Specifically, they display the same substitution metric broken down by income group. There is a clear, negative relationship between income and substitution to community colleges; 24 percent of low-income students (defined by income below $25,000) who would choose to attend a for-profit college absent the policy would switch to a community college, while only 6.5 percent of high-income students (defined by income above $75,000) would do the same. This implies that low-income students, who tend to be more price sensitive, are more likely to change their higher education decision toward a community college. Thus, a for-profit college's price elasticity will decrease at the margin; demand for the marginal customer will be more inelastic, which will incentivize an increase in tuition. The same intuition applies for nonselective four-year colleges, with 32 percent of low income students and 7.5 percent of high income students switching to community colleges when it is free.

As a result, the increase in community college enrollment would not only come from increased access, but also inter-sector substitution. The last two rows of Panel B display the percent of the enrollment change coming from entry or exit from higher education. The model suggests enrollment at community colleges increased by 40 percent due to effects on the extensive margin (students who would have not attended higher education if it were not free). In other words, 83 percent (39.7%/47.7%) of the community college enrollment increase can be attributed to the entry of new students. In addition, because of the equilibrium changes in price, for-profit colleges saw negligible entry of 0.03 percent, and more substantial exits of 2.5 percent. This implies that about 13 percent of the decreased for-profit enrollment can be attributed to higher education exit. Likewise, nonselective four-year colleges also saw a negligible increase in enrollment due to entry, and a 2.1 percent decrease due to exit.

Finally, after accounting for equilibrium effects, the model suggests a completely free community college program would cost about \$8.9 billion annually. In comparison, the Office of Management and Budget estimated in 2016 that free community college in the vision of President Obama's America's College Promise would cost \$9.4 billion in 2023, and \$13.2 billion annually after 2024.[19]

6.1.2 Last dollar programs

While free community college has been the goal of many advocates, practical implementation has been focused on last dollar programs. Specifically, many of the statewide promise programs administer last dollar programs. The model estimated above allows for a comparison between first-dollar and last dollar programs. In addition, an analysis of a last dollar program will allow for a benchmark comparison to the existing literature that evaluates the implementation of statewide promise programs.

To simulate the effect of a last dollar program, I assume students receive a "last dollar" grant if they choose to attend a community college. This last dollar grant is given by the maximum between the difference in tuition and federal grants, and zero.[20] Formally, the price paid by student i when attending college j is given by:

$$p_{ij} = p_j - \max\{p_j comm_j - g_{ij}, 0\}.$$

Panel C of Table 6 presents the equilibrium results when a last dollar promise program is implemented. Overall, last dollar programs result in no change in the average price paid at both for-profit and nonselective four-year colleges. Furthermore, enrollment increases by 9.8 percent, with a 16 percent increase at community colleges, and a 2.8 and 6.3 percent decrease at for-profit and nonselective 4-year colleges, respectively.

[19] https://obamawhitehouse.archives.gov/sites/default/files/omb/budget/fy2016/assets/tables.pdf (Table S-9)

[20] In practice, I assume community colleges do not respond by changing its price in equilibrium. First, the profit weight estimated for community colleges is sufficiently low in which the firm's price is set predominately to maximize the value of the school to its community; this implies that the school will set prices as low as possible, i.e. very close to marginal cost, meaning it will have very little response to its competitors price. Second, computing the new equilibrium price when allowing community colleges to respond does not change the qualitative interpretation of the results, but will often run into convergence problems since its objective function may not have a maximum (since it is near linear).

Because of the smaller price effects, substitution to community college is driving a larger proportion of the enrollment drop at for-profit and nonselective four-year colleges: 3.0 percent of students who would have enrolled in a for-profit college, and 5.4 percent of students who would have enrolled in a nonselective four-year college, switched to a community college due to the last dollar program. As a result, the increase in community college enrollment can be attributed to the entry of students who would have otherwise not attended college, with 88 percent of the enrollment increase coming from new entrants. For-profit colleges would see an overall increase in new entry enrollment due to equilibrium price changes, since some institutions would decrease tuition in response to the last dollar program. Nonselective four-year colleges would see a small net decrease on the extensive margin. In terms of overall cost, the model predicts that a last dollar program would cost \$3.0 billion annually.

The findings for last dollar programs are consistent with, and somewhat smaller than, the simulation study of Avery et al. (2019), which suggested that the percent of high school graduates that enrolled into a two-year college increased by 6.5 percentage points, or a 23 percent increase, and enrolled into a four-year college decreased by 3.3 percentage points, or a 6 percent decrease. On the other hand, my findings were larger than Gurantz (2020)'s study of the Oregon Promise Program, which suggested enrollment effects of 4-5 percent at two-year colleges and decreases of 3 percent at four-year colleges in the program's first year of implementation, and no effect on the second. However, the Oregon Promise Program was only available for recent graduates who met certain academic requirements, thus its restrictions on eligibility may contribute to the difference in magnitudes.

6.1.3 Need-based programs

The last counterfactual I examine is a need-based policy in which only low-income students are eligible to enroll into community college for free. Formally, the price paid by consumer i when attending college j is given by:

$$p_{ij} = (1 - comm_j eligible_i) p_j,$$

where $eligible_i$ is an indicator for whether or not student i has an income less than \$25,000 i.e. is eligible for the program. Panel D of Table 6 presents the results. A program in which community college is free for low-income students would increase enrollment overall by 11.6 percent, with increases at community college of 23 percent, and decreases at for-profit and nonselective four-year colleges at around 11-12 percent. Because a high proportion of low-income student substitute away from for-profit and four-year colleges to community colleges, the average tuition prices paid increases by 6-9 percent, which results in the exit of 3 percent

Table 7: *Counterfactual results for welfare*

	IPEDS	NPSAS	
	Mean	Mean	Median
A. Free community college			
Average across all potential students	\$333.53	-	-
Average for enrolled students	\$710.39	\$497.79	\$274.36
By income group:			
Income $<$ \$25k	\$407.27	\$240.48	\$131.21
\$25k $\leq$ Income $<$ \$75k	\$1,046.04	\$708.41	\$461.83
Income $\geq$ \$75k	\$699.08	\$589.34	\$307.78
B. Last dollar program			
Average across all potential students	\$156.09	-	-
Average for enrolled students	\$271.39	\$198.63	\$2.30
By income group:			
Income $<$ \$25k	\$23.11	\$8.17	\$0.37
\$25k $\leq$ Income $<$ \$75k	\$285.04	\$184.93	\$2.46
Income $\geq$ \$75k	\$734.94	\$629.42	\$364.09
C. Need-based program			
Average across all potential students	\$55.71	-	-
Average for enrolled students	\$132.53	\$63.55	\$-0.60
By income group:			
Income $<$ \$25k	\$408.51	\$243.60	\$133.47
\$25k $\leq$ Income $<$ \$75k	\$-46.03	\$-47.87	\$-21.75
Income $\geq$ \$75k	\$-87.06	\$-77.46	\$-36.04

NOTE: Estimates for the IPEDS sample calculates the weighted average of market-level average compensating variations. The NPSAS sample calculates the mean and median compensating variation for the random sample of students.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), Fall 2011; and 2011-12 National Postsecondary Student Aid Study (NPSAS:12).

of students within each sector. Overall, a need-based free community college program would be slightly more successful at increasing access to higher education compared to last dollar programs, with overall increases in enrollment due to new entry of 12.5 percent. The cost of a federal need-based program would be \$3.9 billion.

6.2 Consumer welfare

I use compensating variation to measure the changes in consumer welfare due to the introduction of free community college. Specifically, I examine the change in students' income required to leave them indifferent between being offered free community college and having to pay in full, as in the status quo. This encompasses three aspects of the policy: (1) the decrease in sticker price at community colleges, (2) the equilibrium response from for-profit and nonselective 4-year alternatives, and (3) the change in the burden faced from student loan debt. Panel A of Table 7 summarizes the change to consumer welfare from free community college using both the IPEDS sample and the NPSAS sample. The first column displays the average compensating variation using the IPEDS sample, with the first row displaying the average across all potential students.[21] The model suggests an average compensating variation of \$334, meaning that it would require about \$334 to leave the average potential student indifferent between having and not having free access community college.

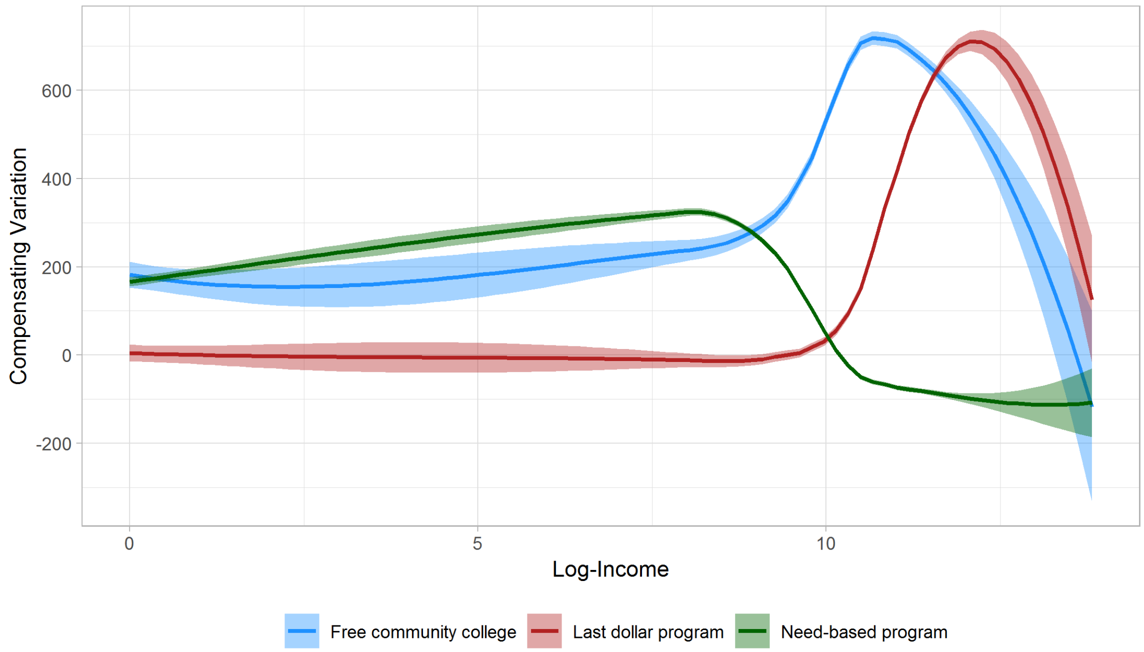
Among students who chose to attend higher education without the policy, the average compensating variation is \$710 in the IPEDS sample. For the NPSAS sample, compensating variation can be calculated for each individual student, yielding a distribution. The mean compensating variation from this sample is lower than the IPEDS sample at \$498, with a median of \$274. By income, both the IPEDS and NPSAS samples both suggest that low income students have the lowest average compensating variation, while middle income students have the highest. The reason for the disparity by income is due to the fact that low income students are the most likely to receive other forms of federal financial aid. The model suggests that a reduction in tuition should result in a reduction in student loan borrowing, which is more prevalent among low-income students. As a result, the compensating variation for low-income student more prominently reflects the dollar value students would pay to increase (discounted) consumption tomorrow. At the same time, middle-income students are less likely to borrow student loans, and thus are able to directly benefit from the policy today. For high-income students, community college is already affordable in the status quo, resulting in a compensating variation between low- and middle-income students.

Compensating variation for last dollar programs is displayed in Panel B of Table 7, showing that students are less well-off compared to a fully-free community college scheme, with the average across all potential students of \$156, and between \$2 and \$271 for students who are currently enrolled without the program. Consistent with the intuition of critics, last dollar programs also largely benefit higher-income students, and result in almost negligible benefits in terms of compensating variation for low-income students. This suggests that last dollar programs, while a more affordable method for financing free community college, will largely ignore those with the most need for financial aid, since it provides little to no support for recipients of need-based grant aid, such as Pell Grants.

Finally, Panel C of Table 7 shows that the overall welfare benefit of need-based programs in terms of compensating variation is low relative to a fully free community college program.

[21] Note that the average across all potential students can only be calculated using the IPEDS sample, since the NPSAS sample is restricted to those who chose a higher education option.

Figure 2: *Compensating variation by income for each free community college program.*



SOURCE: U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), Fall 2011; and 2011-12 National Postsecondary Student Aid Study (NPSAS:12).

However, low-income students are willing to pay as much for this program, since they are effectively receiving the same amount of aid. On the other hand, middle- and high-income students are worse off, and actually harmed, under this policy–they face higher prices at for-profit and nonselective 4-year colleges, with no assistance at community colleges.

To visualize these effects by income, Figure 2 presents the nonparametric regression of log-income on compensating variation from the NPSAS sample for each free community college program. A completely free community college program produces moderate welfare improvements up until income is about $10,000. It then begins to increase as students are becoming less reliant on student loans, and starts to decline as higher income students face diminishing marginal utility. For last dollar programs, compensating variation follows a similar pattern, but is mean zero for lower-income students and the benefits are shifted towards higher income students. Lastly, for need-based programs, lower-income students face a moderate welfare benefit, which quickly plummets into harm for middle- and high-income students.

6.3 Higher education outcomes

A common criticism of free community colleges is that the policy may drive students at higher performing four-year universities toward lower performing community colleges, which

Table 8: *Counterfactual results for outcomes*

	Free comm. college	Last dollar program	Need-based program
A. All degree completions			
4 years after enrollment	19.40%	10.58%	2.97%
6 years after enrollment	21.23%	11.34%	3.43%
8 years after enrollment	21.29%	10.81%	4.38%
B. 4-year degree completions			
4 years after enrollment	4.16%	6.19%	-6.00%
6 years after enrollment	16.15%	9.44%	1.44%
8 years after enrollment	19.15%	9.99%	3.58%

NOTE: All degree completions consider the percent increase in either associate or bachelor degrees due to each policy. Four-year degree completions consider the percent increase in bachelor degrees due to each policy.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), Fall 2011; 2011-12 National Postsecondary Student Aid Study (NPSAS:12); and College Scorecard, Fall 1996-2017.

will result in a net negative effect on outcomes. In this subsection, I examine how the composition change of enrollment across institutions affects two sets higher education completion outcomes: (1) the completion of *any* degree four, six, and eight years after enrollment, including those that transferred to other institutions, and (2) the completion of *a 4-year* degree four, six, and eight years after enrollment, including those that transferred to other institutions. These outcome measures are offered through the U.S. Department of Education's online cost and value comparison tool, College Scorecard.

To do this, I run an institution-level regression of an outcome measure $outcome_{jt}$ for institution i in year t on average student characteristics $\bar{\mathbf{d}}_{jt}$ and institution characteristics $\mathbf{X}_{it}$, and include institution and year fixed effects:

$$outcome_{jt} = \boldsymbol{\beta}_j \bar{\mathbf{d}}_{jt} + \boldsymbol{\gamma}\mathbf{X}_{it} + \xi_j + \xi_t + \varepsilon_{jt}. \tag{14}$$

where $\boldsymbol{\beta}_j$ is the coefficient on student characteristics that may vary by features of college j, e.g., its higher education sector. Appendix Table A.3 presents the coefficient estimates of $\boldsymbol{\beta}_j$ from various models of equation (14). To compute the effect of free community college on completion rates, I calculate the change in average demographics for each college at the equilibrium price with and without the policy, and apply the difference to equation (14), yielding the change in the completion rate.

The first panel of Table 8 presents the percent change on total degree completions for each free community college program. If community colleges were completely free, total degree completions would increase by 20 to 22 percent across all three timeframes. This corresponds to about 225 thousand more degree completions. Last dollar programs would cause total completions to increase by half the amount at approximately 11 percent, or about 125 thousand more degree completions. Finally, need-based programs would produce the smallest percent increase, at between 3 and 4 percent, or 40 thousand degree completions.

To explain these results, consider two countervailing effects on degree completions: First, free community college programs induce entry into higher education, meaning there is a larger base of students who can complete a degree. Mountjoy (2019) refers to this as the democratization effect. Second, it changes completion rates through two mechanisms: (1) demographic shifts based on which students the policy incentivizes, and (2) substitution between institutions that differ in quality, referred to as the diversion effect in Mountjoy (2019). The fact that all programs produce positive effects on total degree completions means that the entry effect dominates any shift to completion rates.

However, for need-based programs, there would be a significant increase in low-income students, which would considerably reduce completion rates due to the strong positive relationship between income and completions.[22] As a result, the entry effect does not dominate the decrease in completion rates as much as it does for other programs. The model suggests that overall completion rates decrease by 2 percent due to the need-based program, whereas it decreases by 1 percent for free community college and increases by 0.6 percent for last dollar programs.

The second panel displays the percent change on total 4-year degree completions for each free community college program. While percent increases are relatively similar to all degree completions 8 years after enrollment, there is a steep increase over time. This is likely due to the fact that the expansion of community college requires many more students to transfer to four-year institutions before receiving a 4-year degree. This suggests that while free community college programs will increase 4-year degree completions in the long-run, many of those students may not graduate within the traditional four years. Overall, the model suggests that eight years after students enroll, there will be an increase in 4-year degree completions of 115 thousand for free community college, 60 thousand for last dollar programs, and 21 thousand for free for low-income programs. Finally, relative to the literature, Mountjoy (2019) finds magnitudes that are consistent with a last dollar community college program. In particular, he finds that increased access to community college increases bachelor completions by 10.5 percentage points.

[22] Furthermore, some students would shift between higher performing nonprofit colleges, which would reduce aggregate completion rates. However, other students would also be shifting away from for-profit colleges, which could mitigate the strength of this effect.

7 Conclusion

While many state governments have recently adopted free community college programs, there has been much debate over its effectiveness and implementation. This paper compares three forms of free community college in terms of access to higher education, consumer welfare, and degree completions. I find that a fully free community college scheme does well across all measures, significantly increasing access and completions in aggregate. Furthermore, it is valued across all income levels, but is most valuable for middle-income students. Last dollar programs, on the other hand, is limited in terms of increased access and mostly benefits high income students, but materially raises degree completions and limits inter-sector substitution. Lastly, need-based programs that make community college free for low-income students, not surprisingly, benefits low-income students the most. However, it harms middle- and high-income students due to equilibrium price changes at for-profit and nonselective 4-year colleges and is restricted in terms increasing degree completions.

Overall, choosing a method for implementing free community college comes with trade-offs, and will depend on what the what the policy planner values: If they care about overall degree completions irrespective of which population they are coming from, last dollar programs may be a cost effective solution. On the other hand, if they care about increasing access to those who are underrepresented, a need-based program may work the best. While free community college is currently a popular solution to addressing the rising cost of higher education, future research should also explore alternative cost-reducing policies, such as increased subsidies to public 4-year colleges, or student loan forgiveness.

References

Amour, M. S. (2019). Study minimizes impact of free community college. Online. https://www.insidehighered.com/news/2019/09/06/study-plays-down-potential-impact-free-community-college-four-year-graduation.

Andrews, R. J., DesJardins, S., and Ranchhod, V. (2010). The effects of the kalamazoo promise on college choice. *Economics of Education Review*, 29(5):722–737.

Avery, C., Howell, J., Pender, M., and Sacerdote, B. (2019). Policies and payoffs to addressing america's college graduation deficit. *Brookings Papers on Economic Activity*.

Bartik, T. J., Hershbein, B., and Lachowska, M. (2019). The effects of the kalamazoo promise scholarship on college enrollment and completion. *Journal of Human Resources*, pages 0416–7824R4.

Bell, E. (2018). Estimating the impact of tennessee promise eligibility on tuition, fees, and enrollment. *Unpublished working paper*.

Berry, S. (1994). Estimating discrete-choice modesl of product differentiation. *The RAND Journal of Economics*, 25(2):242–262.

Berry, S., Levinsohn, J., and Pakes, A. (1995). Autombile prices in market equilibrium. *Econometrica*, 63(4):841–890.

Berry, S., Levinsohn, J., and Pakes, A. (2004). Differentiated products demand systems from a combination of micro and macro data: The new car market. *Journal of Political Economy*, 112(1):68–105.

Bettinger, E., Gurantz, O., Kawano, L., Sacerdote, B., and Stevens, M. (2019). The long-run impacts of financial aid: Evidence from california's cal grant. *American Economic Journal: Economic Policy*, 11(1):64–94.

Bozick, R., Gonzalez, G., and Engberg, J. (2015). Using a merit-based scholarship program to increase rates of college enrollment in an urban school district: The case of the pittsburgh promise. *Journal of Student Financial Aid*, 45(2):2.

Capps, C., Dranove, D., and Satterthwaite, M. (2003). Competition and market power in option demand markets. *RAND Journal of Economics*, 34(4):737–763.

Carruthers, C. K. (2019). 5 things to know about the tennessee promise scholarship. Online. https://www.brookings.edu/blog/brown-center-chalkboard/2019/05/06/five-things-to-know-about-the-tennessee-promise-scholarship.

Carruthers, C. K. and Fox, W. F. (2016). Aid for all: College coaching, financial aid, and post-secondary persistence in tennessee. *Economics of Education review*, 51:97–112.

Castleman, B. L. and Long, B. T. (2016). Looking beyond enrollment: The causal effect of need-based grants on college access, persistence, and graduation. *Journal of Labor Economics*, 34(4):1023–1073.

Cellini, S. R. (2010). Financial aid and for-profit colleges: Does aid encourage entry? *Journal of Policy Analysis and Management*, 29(3):526–552.

Cellini, S. R., Darolia, R., and Turner, L. J. (2016). Where do students go when for-profit colleges lose federal aid? Working paper.

Cellini, S. R., Darolia, R., and Turner, L. J. (2019). Where do students go when for-profit colleges lose federal aid? *American Economic Journal: Economic Policy*.

Cellini, S. R. and Goldin, C. (2014). Does federal student aid raise tuition? new evidence on for-profit colleges. *American Economic Journal: Economic Policy*, 6(4):174–206.

Cohodes, S. R. and Goodman, J. S. (2014). Merit aid, college quality, and college completion: Massachusetts' adams scholarship as an in-kind subsidy. *American Economic Journal: Applied Economics*, 6(4):251–85.

Committee on Health, Education, Labor, and Pensions (2012). For profit higher education: The failure to safeguard the federal investment and ensure student success. Technical Report S. Prt. 112-37, United States Senate.

Darolia, R. (2013). Integrity versus access? the effect of federal financial aid availability on postsecondary enrollment. *Journal of Public Economics*, 106:101–114.

Deming, D. and Walters, C. (2017). The impacts of price and spending subsidies on us postsecondary attainment. *NBER Working Paper*, 23736.

Denning, J. T. (2017). College on the cheap: Consequences of community college tuition reductions. *American Economic Journal: Economic Policy*, 9(2):155–88.

Dubé, J.-P., Fox, J. T., and Su, C.-L. (2012). Improving the numerical performance of static and dynamic aggregate discrete choice random coefficients demand estimation. *Econometrica*, 80(5):2231–2267.

Dynarski, M. (1994). Who defaults on student loans? findings from the national postsecondary student aid study. *Economics of Education Review*, 13(1):55–68.

English, J. (2019). Free college is the wrong answer. Online. https://www.politico.com/agenda/story/2019/10/17/free-college-is-the-wrong-answer-001289.

Flint, T. A. (1997). Predicting student loan defaults. *Journal of Higher Education*, 68(3):322–354.

Fu, C. (2014). Equilibrium tuition, applications, admissions and enrollment in the college market. *Journal of Political Economy*, 122(2):225–281.

Gibbs, C. and Marksteiner, R. (2016). Effects of removing state postsecondary grant aid eligibility. *Unpublished manuscript*.

Goodman, J., Hurwitz, M., and Smith, J. (2017). Access to 4-year public colleges and degree completion. *Journal of Labor Economics*, 35(3):829–867.

Government Accountability Office (2011). Postsecondary education: Student outcomes vary at for-profit, nonprofit, and public schools. Technical Report GAO-12-143, Government Accountability Office.

Government Accountability Office (2018). Actions needed to improve oversight of schools' default rates. Technical Report GAO-18-163, Government Accountability Office.

Gowrisankaran, G., Nevo, A., and Town, R. (2015). Mergers when prices are negotiated: Evidence from the hospital industry. *American Economic Review*, 105(1):172–203.

Grosz, M. (2020). The returns to a large community college program: Evidence from admissions lotteries. *American Economic Journal: Economic Policy*, 12(1):226–254.

Gurantz, O. (2020). What doest free community college buy? early impacts from the oregon promise. *Journal of Policy Analysis and Managment*, 39(1):11–35.

Mountjoy, J. (2019). Community colleges and upward mobility. *Working Paper*.

Murphy, K. M. and Topel, R. H. (2002). Estimation and inference in two-step econometric models. *Journal of Business & Economic Statistics*, 20(1):88–97.

Nevo, A. (2000). A practitioner's guide to estimation of random-coefficients logit models of demand. *Jounral of Economics and Management Strategy*, 9(4):513–548.

Page, L. C., Iriti, J. E., Lowry, D. J., and Anthony, A. M. (2019). The promise of place-based investment in postsecondary access and success: Investigating the impact of the pittsburgh promise. *Education Finance and Policy*, 14(4):572–600.

Petrin, A. (2002). Quantifying the benefits of new products: The case of the minivan. *Journal of Political Economy*, 110(4):705–729.

Pingel, S., Parker, E., and Sisneros, L. (2016). Free community college: An approach to increase adult student success in postsecondary education. Technical report, Education Commission of the States.

Quilantan, B. (2019). The potential disaster of free community college. Online. https://www.politico.com/news/2019/12/17/2020-free-community-college-086315.

Singell Jr, L. D. and Stone, J. A. (2007). For whom the pell tolls: The response of university tuition to federal grants-in-aid. *Economics of Education Review*, 26(3):285–295.

Town, R. and Vistnes, G. (2001). Hospital competition in hmo networks. *Journal of Health Economics*, 20(5).

Trammell, K. (2019). California joins a handful of other states that provide community college tuition-free. Online. https://www.cnn.com/2019/08/28/us/california-free-community-college-tuition-trnd/index.html.

Turner, L. J. (2017). The economic incidence of federal student grant aid. *Unpublished manuscript*.

Wermund, B. (2019). The red state that loves free college. Online. https://www.politico.com/agenda/story/2019/01/16/tennessee-free-college-000867.

Wong, A. (2015). Why for-profit colleges become nonprofits. Online. https://www.theatlantic.com/education/archive/2015/10/the-covert-for-profit/409477/.

Zimmerman, S. D. (2014). The returns to college admission for academically marginal students. *Journal of Labor Economics*, 32(4):711–754.

A Estimation Details

In this appendix section, I provide additional details on estimation that were omitted from the main text.

A.1 Maximum likelihood

In the individual-level NPSAS data, three outcome quantities are observed: Define d_i as the institution chosen by student i, L_i^f as the observed amount of federal loans borrowed by student i, and L_i^p as the observed amount borrowed of private loans. The probability of observing the data for individual i, conditional on the parameters $\boldsymbol{\theta}$, is given by:

$$P_{\boldsymbol{\theta}}(d_i, L_i^f, L_i^p) = P_{\boldsymbol{\theta}}(d_i) P_{\boldsymbol{\theta}}(L_i^f, L_i^p | d_i).$$

The log-likelihood function is then given by:

$$\log \mathcal{L}(d_i, L_i^f, L_i^p; \boldsymbol{\theta}) = \sum_i \log P_{\boldsymbol{\theta}}(d_i) + \log P_{\boldsymbol{\theta}}(L_i^f, L_i^p | d_i),$$

where $P_{\boldsymbol{\theta}}(d_i)$ is the probability that student i selects college d_i, and $P_{\boldsymbol{\theta}}(L_i^f, L_i^p | d_i)$ is the probability of observing student i borrow L_i^f and L_i^p in federal and private loans, respectively, conditional on student i choosing to attend college d_i. Estimation is done in two steps: The first step maximizes the partial likelihood of the observed loan amounts; the log-likelihood is given by

$$\log \mathcal{L}_1 = \sum_i \log P_{\boldsymbol{\theta}_1}(L_i^f, L_i^p | d_i),$$

where $\boldsymbol{\theta}_1 = (\boldsymbol{\tau}_f, \sigma_f, \boldsymbol{\tau}_p, \sigma_p)$. Given the estimates of the loan parameters, accrued interest and counterfactual loan amounts can be calculated. Using these quantities, the second step maximizes the partial likelihood of observed college choices; the log-likelihood is given by:

$$\log \mathcal{L}_2 = \sum_i \log P_{\boldsymbol{\theta}_2}(d_i),$$

where $\boldsymbol{\theta}_2 = (\alpha, \boldsymbol{\gamma}, \boldsymbol{\xi})$. I discuss details of each step in turn.

A.2 Step 1: Simulated maximum likelihood for loan choice

The main text describes a partial likelihood that depends on cases, such that the probability of observing case k is given by P_{ik}. In this appendix subsection, I present the formulations of P_{ik} for each case in terms of the observables and parameters. As a reminder, the latent variables for federal and private loans are given by:

$$\ell_{ij}^f = \frac{1}{\lambda_i^f} + (p_j - y_i - g_{ij})$$
$$\ell_{ij}^p = \frac{1}{\lambda_i^p} + (p_j - y_i - g_{ij}),$$

such that $\lambda_i^f = \exp(\mathbf{d}^T\boldsymbol{\pi}^f + \sigma_f \eta)$ and $\lambda_i^p = \lambda_i^f \exp(\mathbf{d}^T\boldsymbol{\pi}^p + \sigma_p \eta^p)$. The random variables η and η^p both follow a standard normal distribution and are assumed to be independent. Finally, to simplify notation, let $\tilde{n}_{ij} = p_j - y_i - g_{ij}$.

Case 1: Student i does not borrow any money from either federal or private sources. The probability of this case can be written as:

$$P_1 = P(\lambda^f < \lambda^p)P(\tilde{\ell}^f \leq 0|\lambda^f < \lambda^p) + P(\lambda^f \geq \lambda^p)P(\tilde{\ell}^p \leq 0|\lambda^f \geq \lambda^p),$$

where

$$\begin{aligned}
P(\lambda^f \geq \lambda^p) &= \Phi\left(-\frac{\mathbf{d}^T\boldsymbol{\pi}^p}{\sigma_p}\right) \\
P(\ell^f \leq 0|\lambda^f < \lambda^p) &= 1 - \Phi\left(\frac{-\log(-\tilde{n}) - \mathbf{d}^T\boldsymbol{\pi}^f}{\sigma_f}\right) \\
P(\ell^p \leq 0|\lambda^f \geq \lambda^p) &= \int 1 - \Phi\left(\frac{-\log(-\tilde{n}) - \mathbf{d}^T(\boldsymbol{\pi}^f + \boldsymbol{\pi}^p) - \sigma_p\eta^p}{\sigma_f}\right) \\
&\qquad \times P\left(\eta^p \middle| \eta^p \leq -\frac{\mathbf{d}^T\boldsymbol{\pi}^p}{\sigma_p}\right) d\eta^p.
\end{aligned}$$

The last equation is an expectation under a truncated normal distribution, and can be simulated using standard numerical methods.

Case 2: Student i only borrows with federal loans. This case will only consist of students with private interest rates that exceed federal interest rates. The probability of this case can be written as:

$$P_2 = P(\lambda^f < \lambda^p)P(\ell^f = f|\lambda^f < \lambda^p),$$

where

$$P(\ell^f = f|\lambda^f < \lambda^p) = \frac{1}{\sigma_f}\phi\left(\frac{-\log(f - \tilde{n}) - \mathbf{d}^T\boldsymbol{\pi}^f}{\sigma_f}\right)$$

Case 3: Student i is borrowing exactly the federal loan maximum, but is not taking out any private loans. In this case, the student's marginal utility of consumption is high enough that they require more than the loan maximum, but their private interest rates are too high to incentivize them to borrow from a private lender. The probability of this case can be written as:

$$P_3 = P(\lambda^f < \lambda^p)P(\ell^p < \bar{L} \leq \ell^f|\lambda^f < \lambda^p),$$

where

$$P(\ell^p < \bar{L} \leq \ell^f|\lambda^f < \lambda^p) = \int \Bigg[\Phi\left(\frac{-\log(\bar{L}-\tilde{n}) - \mathbf{d}^T\boldsymbol{\pi}^f}{\sigma_f}\right) - \Phi\left(\frac{-\log(\bar{L}-\tilde{n}) - \mathbf{d}^T(\boldsymbol{\pi}^f + \boldsymbol{\pi}^p) - \sigma_p\eta^p}{\sigma_f}\right)\Bigg] \times P\left(\eta^p \middle| \eta^p > -\frac{\mathbf{d}^T\boldsymbol{\pi}^p}{\sigma_p}\right) d\eta^p$$

Case 4: Student i is borrowing exactly the federal loan maximum, and is taking also taking out private loans. In this case, if the amount of private loans the student borrows is z, then the latent variable ℓ^p will take the value $z + \bar{L}$, since the amount of private loans a student borrows will already take into account the amount of federal loans the student has already borrowed. This condition encompasses the fact that $\ell^f > \bar{L}$, since $\ell^f > \ell^p$ when $\lambda^f < \lambda^p$. The probability of this case can be written as:

$$P_4 = P(\lambda^f < \lambda^p)P(\ell^p = z + \bar{L}|\lambda^f < \lambda^p),$$

where

$$P(\ell^p = z + \bar{L}|\lambda^f < \lambda^p) = \int \frac{1}{\sigma_f}\phi\left(\frac{-\log(z + \bar{L}-\tilde{n}) - \mathbf{d}^T(\boldsymbol{\pi}^f + \boldsymbol{\pi}^p) - \sigma_p\eta^p}{\sigma_f}\right) \times P\left(\eta^p \middle| \eta^p > -\frac{\mathbf{d}^T\boldsymbol{\pi}^p}{\sigma_p}\right) d\eta^p$$

Case 5: Student i only borrows from private lenders because her private interest rate is lower than her federal interest rate: The probability of this case can be written as:

$$P_5 = P(\lambda^f \geq \lambda^p)P(\ell^p = z|\lambda^f \geq \lambda^p)$$

where

$$P(\ell^p = z|\lambda^f \geq \lambda^p) = \int \frac{1}{\sigma_f}\phi\left(\frac{-\log(z-\tilde{n}) - \mathbf{d}^T(\boldsymbol{\pi}^f + \boldsymbol{\pi}^p) - \sigma_p\eta^p}{\sigma_f}\right) \times P\left(\eta^p \middle| \eta^p \leq -\frac{\mathbf{d}^T\boldsymbol{\pi}^p}{\sigma_p}\right) d\eta^p$$

A.3 Step 2: Simulated maximum likelihood for enrollment choice

In the second step, I maximize the partial likelihood based on the college choice probabilities $P_{\boldsymbol{\theta}_2}(d_i)$. These choice probabilities are integrated over the unobservable portion of federal and private loans, η_i and η_i^p. I use Monte Carlo simulations with 200 draws to simulate this integral. As addressed in the main text, moment conditions that set predicted demand equal to enrollment are required to identify the model. Following the demand estimation literature, I set predicted market share equal to observed market share, which is mathematically equivalent to setting demand equal to observed enrollment, given market size. While the main text writes the problem out as constrained maximum likelihood, such as the formulation suggested by Dubé et al. (2012), in practice, I use the equivalent nested fixed point procedure outlined by Berry et al. (1995). This is mainly because the nested fixed point procedure was computationally faster for this application.

Following Berry et al. (1995), I first identify the institution-specific parameters ξ_j using the contraction mapping such that:

$$\boldsymbol{\xi}^{t+1} = \boldsymbol{\xi}^t + \log share_j - \log s_j(\alpha, \boldsymbol{\gamma}, \boldsymbol{\xi}^t),$$

where s_j is the college choice probability integrated over the loan unobservables η_i and η_i^p, as well as demographics $\mathbf{d}$, and $share_j$ is the observed share. Monte Carlo simulations using 1000 draws are used to approximate the integrals. The series converges at an iteration T such that $||\boldsymbol{\xi}^{t+1} - \boldsymbol{\xi}^t||$ is smaller than a tolerance level of $1e^{-9}$. The value of $\boldsymbol{\xi}^T(\alpha, \boldsymbol{\gamma})$ is used as an approximation of $\boldsymbol{\xi}$.

The parameters α and $\boldsymbol{\gamma}$ are estimated by maximum likelihood by including the approximation $\boldsymbol{\xi}(\alpha, \boldsymbol{\gamma})$ in place of $\boldsymbol{\xi}$:

$$(\widehat{\alpha}, \widehat{\boldsymbol{\gamma}}) = \max_{(\alpha, \boldsymbol{\gamma})} \quad \log \sum_i P_{(\alpha, \boldsymbol{\gamma}, \boldsymbol{\xi}(\alpha, \boldsymbol{\gamma}))}(d_i).$$

Finally, as mentioned in the main text, the mean utility parameters on school characteristics are not directly identified since they are absorbed into the institution-specific term $\boldsymbol{\xi}$. I use a minimum distance procedure to back them out. In particular, estimates of $\boldsymbol{\beta}$ are given by:

$$\widehat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{V}_{\boldsymbol{\xi}}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{V}_{\boldsymbol{\xi}}^{-1} \boldsymbol{\xi}(\widehat{\alpha}, \widehat{\boldsymbol{\gamma}}),$$

where $\mathbf{V}_{\boldsymbol{\xi}}$ is the variance-covariance matrix of the estimated $\boldsymbol{\xi}$.

A.4 Additional estimation results

Table A.1: *Median implied elasticity estimates*

	Comm. Coll.	For-Profit	4-year Public	4-year Private
Community College	1.26	−0.04	−0.05	−0.15
For-profit College	−0.02	3.32	−0.10	−0.13
4-year Private Nonprofit	−0.03	−0.08	2.21	−0.07
4-year Public Nonprofit	−0.16	−0.06	−0.07	1.34

NOTE: The table presents the median own- and cross-price elasticities implied by the demand model.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), Fall 2011; and 2011-12 National Postsecondary Student Aid Study (NPSAS:12).

Table A.2: *Supply-side estimation results*

	Estimate	Std. Err.
Profit weights		
Community college	0.044	0.040
Private 4-year nonprofit	1.000	0.000
Public 4-year nonprofit	0.285	0.180
Net marginal cost parameters		
Alabama	-925.89	934.04
Arizona	-1466.28	851.41
California	-2743.36	577.99
Connecticut	-1267.29	1920.20
Florida	-1450.50	869.65
Georgia	-863.08	644.76
Illinois	-190.75	841.15
Louisiana	-2775.72	794.96
Maryland	1016.63	1264.58
Massachusetts	2097.23	2199.59
Michigan	36.22	1149.46
Minnesota	828.11	1240.83
Missouri	1100.70	1210.12
New Jersey	515.15	870.30
New Mexico	-1845.85	1312.19
New York	201.23	973.14
North Carolina	-3765.01	1287.22
Ohio	1252.22	665.85
Pennsylvania	843.36	1125.45
South Carolina	992.79	1349.74
Texas	-1350.10	604.95
Virginia	-275.34	641.95
Washington	-1605.45	1140.13
Other states	3451.50	567.44

NOTE: Standard errors are bootstrapped using 250 draws.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), Fall 2011, and 2011-12 National Postsecondary Student Aid Study (NPSAS:12).

Table A.3: *Results for outcome regression*

	All degree completions			4-year degree completions		
	4 years	6 years	8 years	4 years	6 years	8 years
Family income (log)	0.193***	0.197***	0.173***	0.079***	0.096***	0.094***
	(0.007)	(0.008)	(0.008)	(0.005)	(0.005)	(0.006)
Enrollment (log)	−0.002	0.006**	0.006**	−0.0005	0.002	0.004**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Age	0.010***	0.006***	0.002	0.006***	0.001	−0.0005
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
Female	−0.059**	−0.061**	0.043	0.174***	0.119***	0.095***
	(0.025)	(0.027)	(0.030)	(0.018)	(0.019)	(0.021)
Dependent	0.101***	0.098***	0.100***	0.076***	0.037*	0.043*
	(0.030)	(0.030)	(0.032)	(0.022)	(0.021)	(0.022)
Minority	−0.010	−0.027**	−0.039***	−0.005	0.001	−0.001
	(0.012)	(0.013)	(0.014)	(0.009)	(0.009)	(0.010)
For-profit × Income	−0.137***	−0.175***	−0.177***	−0.070***	−0.090***	−0.091***
	(0.009)	(0.010)	(0.010)	(0.006)	(0.007)	(0.007)
For-profit × Enrollment	−0.011***	−0.016***	−0.007**	−0.008***	−0.009***	−0.008***
	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
For-profit × Age	−0.015***	−0.007***	0.001	−0.012***	−0.003**	0.001
	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
For-profit × Female	−0.034	−0.036	−0.131***	−0.188***	−0.148***	−0.118***
	(0.028)	(0.031)	(0.034)	(0.020)	(0.021)	(0.024)
For-profit × Dependent	−0.038	0.034	0.026	−0.059**	0.003	0.0003
	(0.039)	(0.042)	(0.044)	(0.028)	(0.029)	(0.031)
For-profit × Minority	−0.034**	−0.006	0.006	−0.029***	−0.031***	−0.034***
	(0.014)	(0.015)	(0.016)	(0.010)	(0.010)	(0.011)
Observations	15,262	13,133	11,051	15,262	13,133	11,051
R^2	0.920	0.917	0.921	0.927	0.938	0.944
Adjusted R^2	0.914	0.909	0.911	0.921	0.932	0.938

NOTE: *p<0.1; **p<0.05; ***p<0.01. Standard errors are in parentheses. All regressions include institution and year fixed effects.
SOURCE: U.S. Department of Education, National Center for Education Statistics, College Scorecard, Fall 1996-2017.