What do course offerings imply about university preferences?

James Thomas*

April 15, 2022

Abstract

This paper empirically analyzes how universities decide which courses to offer and the implications of these decisions for students. At a sample university, course offerings significantly impact student course choices and implicitly sacrifice student utility to increase enrollment in STEM and business and occupational courses. This is because new STEM and business and occupational course sections have slightly smaller effects on student utility and cost substantially more than new offerings in other fields. The university changes its course offerings in counterfactual scenarios and ignoring these responses leads to understating the effects of interventions.

1 Introduction

There is a large literature devoted to analyzing how students choose college majors.[1] However, an important contributor to these decisions has thus far been overlooked: the courses offered by universities. Course offerings determine the variety of alternatives available in different fields and greater variety makes fields more attractive. As such, course offerings directly affect student specialization decisions.

*Federal Trade Commission, e-mail: jthomas2@ftc.gov. The views expressed in this article are those of the author and do not necessarily reflect those of the Federal Trade Commission. I am especially thankful to Joseph Altonji, Peter Arcidiacono, V. Joseph Hotz, Robert Garlick, Hugh Macartney, and Arnaud Maurel for ongoing advice on this project and to the Arkansas Research Center for access to and assistance with administrative data. This project has also benefited from comments from Benjamin Cowan, Lisa Kahn, Miriam Larson-Koester, Lance Lochner, Richard Mansfield, Ryne Marksteiner, Benjamin Smith, Kevin Stange, and William Violette. All remaining errors are my own.

[1]See Altonji et al. (2012) and Altonji et al. (2016) for recent reviews.

But course offerings are not fixed inputs, they are chosen by universities to maximize objective functions subject to constraints. Furthermore, universities may adjust their offerings in response to changes in student preferences, changes in costs of hiring instructors, or other counterfactual scenarios. These responses affect student choices implying that ignoring these responses may lead to misstating the effects of various interventions on student decisions.[2]

To advance our understanding of these issues, this paper analyzes how course offerings affect student course choices, how universities decide which courses to offer, how course offerings would change in counterfactual scenarios, and the implications of these changes for student course choices. For this analysis, I build a two-sided model of a university offering courses and students choosing courses from the set of available alternatives.

The university in my model values total student utility and the number of students choosing courses in each field. This allows me to assess whether observed course offerings are aligned with student preferences, and if they are not, to quantify how much student utility the university is implicitly sacrificing to increase enrollment in certain fields.

Students in my framework choose courses according to a nested logit model with unobserved heterogeneity in preferences for fields. The nested logit structure ensures that effects of course offerings on student choices are determined by empirical variation rather than arbitrary functional form assumptions.

An advantage of my framework is that the university's "preference" parameters can be directly interpreted even if the university model is misspecified. For example, if course offerings are determined through complicated bureaucratic processes that largely ignore student utility or field enrollments, university parameters will still measure the implicit tradeoffs between student utility and field enrollments that result from this bureaucratic process. Therefore, even though a misspecified university model cannot be used for counterfactual analyses, it can still be used to assess whether course offerings are aligned with students preferences.

I use my framework to analyze the introductory courses offered by the University of Central Arkansas (UCA) from 2004-05 through 2012-13. UCA is a particularly interesting subject for two reasons: First, UCA is a large public four year university with a 45% graduation rate making it somewhat representative of the post secondary education experience of a median American.[3] Second, UCA is a teaching focused university, which makes an analysis

[2]For example, various papers examine the effects of expected earnings on major choices (e.g. Arcidiacono et al. (2012); Stinebrickner and Stinebrickner (2014); Wiswall and Zafar (2014)); however, these papers do not consider how course offerings might respond to changes in expected earnings and the implications of these responses for major choices.

[3]The median young American completes some college but does not obtain a degree and 45% of all full-

of its course offerings especially pertinent.[4]

My first set of results show that introductory course offerings have important effects on student course choices at UCA. For example, I find that offering an additional STEM course section increases total STEM enrollment by 13.47 students, which is 53.3% of the expected enrollment in that new section. This implies that new STEM offerings attract over half their students from other fields and do not merely siphon enrollment from existing STEM courses.

Next, I show that course offerings at UCA implicitly sacrifice student utility to increase enrollment in STEM and business and occupational classes. If UCA reallocated its budget for adjunct instructors to maximize student utility, it would reduce adjunct instructed STEM offerings by 88% and roughly double adjunct instructed humanities and arts course sections.[5] This reflects the findings that marginal STEM course sections produce slightly less student utility, and are substantially more expensive, than marginal humanities and arts sections. These conclusions hold even if the university's model is misspecified.

Finally, I examine how course offerings would respond to changes in student preferences or the costs of hiring instructors and show that these responses would have important implications for student course choices. One simulation examines a scenario in which all students' observed measures of baseline preparation are increased by one-tenth of a standard deviation. This is meant to simulate a policy such as the Georgia HOPE scholarship program that increased the in-state retention of better prepared students (Cornwell et al., 2006). Because better prepared students are generally more interested in STEM, this would increase introductory STEM enrollment by 0.8% even if course offerings were held fixed. However, the university responds to the increase in STEM interest by offering more introductory STEM course sections. This makes STEM even more attractive resulting in a 5.1% total increase in introductory STEM enrollment. In other words, ignoring the university's response leads to understating effects on STEM enrollment by 4.3 percentage points. Another analysis shows that a 5% reduction in the cost of hiring a STEM adjunct instructor would lead to a 12.6% increase in total sections of introductory STEM courses and a 8.9% increase in overall enrollment in introductory STEM courses.

---

time equivalent higher education enrollment is at public four-year institutions (Ryan and Bauman (2016) and author's calculation using IPEDS for academic year 2016-17).

[4]One can see UCA's teaching focus directly in the data because 82% of student hours of instruction are provided by instructors who receive at least 95% of their compensation for teaching. One can also see UCA's teaching focus in their vision statement:

> The University of Central Arkansas aspires to be a premier learner-focused public university, a nationally recognized leader for its continuous record of excellence in undergraduate and graduate education, scholarly and creative endeavors, and engagement with local, national, and global communities. (Board, 2011)

[5]44.7% of all course sections are taught by adjunct instructors.

This paper relates to two broad strands of higher education literature: First, it relates to a long line of literature that analyzes how students choose courses and majors.[6] Within this literature, the most relevant papers are those that focus on the role of institutional inputs such as grading policies (Ahn et al., 2022; Bar et al., 2009; Butcher et al., 2014) and instructor characteristics (Bettinger and Long, 2005; Carrell et al., 2010; Griffith, 2014) on student decisions. This paper contributes to this literature by providing the first analysis of the effects of course offerings on student course choices.

Moreover, with a few exceptions, papers in this literature estimate the effects of institutional inputs but do not consider how these inputs are chosen by universities.[7] While these estimates illustrate how directly adjusting inputs would affect students, they do not reveal how inputs would evolve over time or change in response to interventions. This paper analyzes how course offerings are chosen to provide a richer understanding of how students are affected by these inputs in various scenarios.

As such, this paper also relates to a more nascent strand of literature that analyzes how universities make decisions.[8] This includes studies that develop general equilibrium models of competition in the higher education market (Cook, 2021; Epple et al., 2006, 2013; Fu, 2014) as well as tests of the "Bennett hypothesis", which predicts that universities will respond to federal tuition subsidies by increasing their tuition (Cellini and Goldin, 2014; Gibbs and Marksteiner, 2016; Long, 2004; Singell and Stone, 2007; Turner, 2017). This paper contributes to this literature by providing the first analysis of how a university decides which courses to offer. Furthermore, it also contributes to this literature by providing the first estimates of a model of university choices using micro-level data. Together, these contributions deepen our understanding of how universities make decisions and the implications of these decisions for students.

The remainder of this paper proceeds as follows: Section 2 introduces a framework for analyzing how universities choose course offerings, Section 3 presents a framework for analyzing how student choices are influenced by course offerings, Section 4 describes the data and discusses estimation, Section 5 presents results, and Section 6 concludes.

[6]Notable contributions not mentioned in the body include, but are not limited to: Altonji (1993); Arcidiacono (2004); Arcidiacono et al. (2012); Beffy et al. (2012); Berger (1988); Bordon and Fu (2015); Eide and Waehrer (1998); Gemici and Wiswall (2014); Stinebrickner and Stinebrickner (2014); Turner and Bowen (1999); Wiswall and Zafar (2014). See Altonji et al. (2012) and Altonji et al. (2016) for recent reviews.

[7]A notable exception is Ahn et al. (2022), which considers the effects of grading policies on student course choices and analyzes how instructors choose grading policies.

[8]Notable contributions not mentioned in the body include, but are not limited to: Andrews and Stange (2019); Bhattacharya et al. (2017); Cellini (2009, 2010); Dinerstein et al. (2014); Hoxby (1997); Jacob et al. (2018); Russell (2021); Thomas (2019).

2 Theoretical Framework: University

In this section, I introduce a general framework for analyzing introductory course offerings at a university. The main idea is to use estimates of the marginal effects of offering additional course sections in each field on student utility and the marginal costs of offering additional sections in each field to assess whether the marginal return on instruction spending differs across fields. Differences in these marginal effects per dollar across fields reveal an implicit willingness to sacrifice student utility to increase enrollment in certain fields.

2.1 University's course offerings

To begin, let $t \in [1, T]$ index academic semesters and let $f \in [1, F]$ index academic fields.[9] Let d_{tf} represent the number of sections of introductory field f courses offered in semester t and collect these offerings into a single vector $\mathbf{d}_t = \begin{bmatrix} d_{t1} & \cdots & d_{tF} \end{bmatrix}$.

Now suppose one has a model for student demand for introductory courses in which the number of course sections offered in each field $\mathbf{d}_t$ affects the expected number of students choosing courses in each field and the total expected utility students derive from their choices. Let $n_{tf}(\mathbf{d}_t)$ represent the university's expectation for total enrollment in introductory courses in field f in semester t, let $V_t(\mathbf{d}_t)$ represent the university's expectation for total student utility from introductory course choices in semester t, and assume both $n_{tf}(\mathbf{d}_t)$ and $V_t(\mathbf{d}_t)$ are continuously differentiable in $\mathbf{d}_t$.[10] In Section 3, I specify a nested logit course choice model in which enrollments and utilities depend on course offerings as desired; however, a wide class of demand models will provide these relationships.

Now suppose the university's payoff from offering courses $\mathbf{d}_t$ is a linear combination of total student utility $V_t(\mathbf{d}_t)$ and field enrollments $n_{tf}(\mathbf{d}_t)$ as follows:

$$\Pi_t(\mathbf{d}_t) = \theta V_t(\mathbf{d}_t) + \sum_{f=1}^{F} \gamma_f n_{tf}(\mathbf{d}_t) \tag{1}$$

Without loss of generality, I normalize $\theta = 1$ and $\gamma_F = 0$.[11] With this structure and normalizations, γ_f measures the amount of student utility that the university is implicitly

[9] In the empirical application, fields are STEM, social science, humanities and arts, and business and occupational. See Appendix A for field definitions.

[10] Note that $\mathbf{d}_t$ is a vector of discrete variables and thus derivatives with respect $\mathbf{d}_t$ are not defined; however, at large universities such as the one I study, the number of introductory course sections in each field is large enough that approximating course offerings as a continuous variable is reasonable.

[11] Normalizing $\gamma_F = 0$ is without loss of generality as long as total enrollment $\sum_{f=1}^{F} n_{tf}(\mathbf{d_t})$ is the same for all $\mathbf{d_t}$. This implies $\mathbf{d_t}$ cannot affect the number of students enrolled at the school or the share choosing introductory courses. I discuss this limitation and others in Section 2.3. Normalizing $\theta = 1$ is without loss of generality because the scale of the university's payoff is not determined.

willing to sacrifice to draw one student out of a field F course and into a field f course in expectation. Because student utility does not have economically meaningful units, the magnitudes of γ_f do not have a direct interpretation; however, one can interpret the ordinal rankings of γ_f as the university's implied ordinal preferences for enrollment in different fields.

Finally, suppose the university faces a semester specific budget constraint which states that the cost of offering $\mathbf{d_t}$ cannot exceed an endowment. Specifically, I assume:

$$C\left(\mathbf{d}_t\right) \leq E_t \tag{2}$$

where E_t is a semester specific endowment and $C\left(\cdot\right)$ is a smooth function.[12]

The university's course offering problem in semester t is then given by:

$$\mathbf{d}_t^{\star} = \text{argmax}_{\mathbf{d_t}} \left\{ V_t\left(\mathbf{d}_t\right) + \sum_{f=1}^{F-1} \gamma_f n_{tf}\left(\mathbf{d}_t\right) \right\} \quad \text{s.t.} \quad C\left(\mathbf{d}_t\right) \leq E_t \tag{3}$$

In this setting, the university payoff parameters γ_f could reflect a variety of underlying mechanisms. They could reflect true university preference parameters rooted in paternalistic beliefs about which courses best serve students' long term interest or social beliefs about which courses produce the most public goods; however, they could also reflect institutional frictions within the university that implicitly favor certain fields as a result of path dependence.

Fundamentally, γ_f represent wedges between the marginal benefits of offering additional courses in terms of student utility and the marginal costs of offering these courses. As I discuss shortly, an advantage of this framework is that these wedges can be directly interpreted as interesting measures of the misalignment between student preferences and observed course offerings even if the misalignment results from institutional frictions or other non-intentional mechanisms.

2.2 Solving for implied preferences

To solve for implied preference parameters γ_f, I first derive the first order conditions that characterize an interior solution to the university's problem stated in Equation (3). These

[12] I assume endowments E_t are set exogenously through a process that is unrelated to course offerings $\mathbf{d_t}$. If offering additional courses in field f has a positive (negative) effect on E_t then I would be ignoring a positive (negative) marginal value to the university of offering additional courses in field f. This would lead to estimates which overstate (understate) implied preferences for enrollment in field f.

first order conditions are:

$$\left(\frac{1}{c_{tf_1}}\right)\left[\frac{\partial V_t\left(\mathbf{d}_t^\star\right)}{\partial d_{tf_1}}+\sum_{f'=1}^{F-1}\gamma_{f'}\left(\frac{\partial n_{tf'}\left(\mathbf{d}_t^\star\right)}{\partial d_{tf_1}}\right)\right]=\left(\frac{1}{c_{tf_2}}\right)\left[\frac{\partial V_t\left(\mathbf{d}_t^\star\right)}{\partial d_{tf_2}}+\sum_{f'=1}^{F-1}\gamma_{f'}\left(\frac{\partial n_{tf'}\left(\mathbf{d}_t^\star\right)}{\partial d_{tf_2}}\right)\right]\quad\forall f_1,f_2 \tag{4}$$

where

$$c_{tf}=\frac{\partial C\left(\mathbf{d}_t^\star\right)}{\partial d_{tf}} \tag{5}$$

is the marginal cost of offering additional course sections in field f at observed course offerings $\mathbf{d}_t^\star$.

Intuitively, these conditions state that the net marginal benefit of offering an additional course section relative to the cost of offering this section must be the same across all academic fields. If this were not the case, the university could improve its payoff by reallocating funds away from fields with low returns to fields with high returns. Net marginal benefit includes both benefit from increasing total student utility and net benefit (cost) from drawing students into more (less) implicitly favored fields.

Rearranging and stacking fields and semesters yields:

$$\mathbf{dn}^\star\times\Gamma=\mathbf{dV}^\star \tag{6}$$

where

$$\begin{aligned}
\underset{(F,F-1)}{\mathbf{dn_t}^\star}\left(f_1,f_2\right)&=\left(\frac{1}{c_{tf_1}}\right)\left(\frac{\partial n_{tf_2}\left(\mathbf{d}_\mathbf{t}^\star\right)}{\partial d_{tf_1}}\right)-\left(\frac{1}{c_{tF}}\right)\left(\frac{\partial n_{tf_2}\left(\mathbf{d}_\mathbf{t}^\star\right)}{\partial d_{tF}}\right)\\
\underset{(F\times T,F-1)}{\mathbf{dn}^\star}&=\begin{bmatrix}\mathbf{dn}_\mathbf{1}^\star & \cdots & \mathbf{dn}_\mathbf{T}^\star\end{bmatrix}'\\
\underset{(F,1)}{\mathbf{dV}_\mathbf{t}^\star}\left(f\right)&=\left(\frac{1}{c_{tF}}\right)\left(\frac{\partial V_t\left(\mathbf{d}_\mathbf{t}^\star\right)}{\partial d_{tF}}\right)-\left(\frac{1}{c_{tf}}\right)\left(\frac{\partial V_t\left(\mathbf{d}_\mathbf{t}^\star\right)}{\partial d_{tf}}\right)\\
\underset{(F\times T,1)}{\mathbf{dV}^\star}&=\begin{bmatrix}\mathbf{dV}_\mathbf{1}^\star & \cdots & \mathbf{dV}_\mathbf{T}^\star\end{bmatrix}'\\
\underset{(F-1,1)}{\Gamma\left(f\right)}&=\gamma_f
\end{aligned}$$

This system of equations can then be inverted to derive the following expression for implied preference parameters $\mathbf{\Gamma}$ as a function of marginal effects and costs:

$$\Gamma=(\mathbf{dn}^\star)^+(\mathbf{dV}^\star) \tag{7}$$

where M^+ denotes the pseudo-inverse of M.

As such, if one can obtain estimates of the marginal effects and costs of offering additional course sections at observed course offerings, one can use Equation (7) to estimate the tradeoffs between total student utility and field enrollments implied by observed course offerings.

2.3 Discussion

To directly interpret estimates of γ_f as measures of the misalignment between student preferences and observed course offerings, one only needs credible estimates of local marginal costs $\frac{\partial C(\mathbf{d}_t^\star)}{\partial d_{tf}}$ and local marginal effects $\frac{\partial n_{tf'}(\mathbf{d}_t^\star)}{\partial d_{tf}}$ and $\frac{\partial V_t(\mathbf{d}_t^\star)}{\partial d_{tf}}$. The student demand model does not need to be globally credible and the university's problem can be misspecified. However, if one wishes to use Equation (3) as a structural model of university decisions in counterfactual analyses, the student demand model must be globally credible and the university's problem must be correctly specified.

One shortcoming of Equation (3) as a structural model of university behavior is that it almost certainly overstates how quickly a university can respond to changes in student demand, costs, or education policies. There is no theoretical mechanism to generate path dependence in the university's offerings—the university can change offerings to solve Equation (3) every semester without frictions.[13] To partially address this concern, I restrict the university in my counterfactual analyses so it can only reallocate its spending on adjunct (non tenure-track) instructors.[14] This provides a coarse measure of institutional friction; however, extensions that handle institutional friction more carefully would likely yield improved predictions.

Another shortcoming of Equation (3) is that it assumes introductory course offerings do not affect the number of students enrolled at the university. Prospective students may have systematically different preferences from enrolled students if other parties, such as parents, have greater influence on enrollment decisions. As such, the effects of course offerings on enrollment may be qualitatively different from the effects on student utility. While it would be interesting to study enrollment responses to course offerings, my data are not well suited for such an analysis. Because of this restriction, a higher value for γ_f could reflect a belief that field f course offerings are more relevant to prospective students than enrolled students.

A final shortcoming of Equation (3) is that this model abstracts from closely related decisions such as how many advanced courses to offer in each field, which introductory courses

[13] In theory, one could use panel variation in student characteristics, instructor costs, and offerings to identify institutional costs of changing offerings; however, this would probably require more than 18 semesters of data. I leave this extension for future work.

[14] 44.7% of all course sections are taught by adjunct instructors.

to offer within fields, and how to match instructors to courses. One direct consequence of abstracting from advanced course offering decisions is that this necessitates assuming introductory course offerings do not affect students' decisions of whether to take advanced or introductory courses. In theory, one could extend this framework to include advanced courses; however, this would require dynamic models of student demand and university course offerings.

While one could use this framework with a less aggregated definition of field if desired, it seems infeasible to model a university's choice of how many sections to offer for every potential course in a semester. As such, some level of abstraction from within field course offerings is probably necessary. Furthermore, although the question of how universities match instructors to courses is interesting, a preliminary analysis found that instructor characteristics have negligible effects on student course choices at UCA. While future work analyzing how instructors are matched to courses would be interesting; I abstract from these choices to focus on the course offering decisions that are more relevant at the university I study.

3 Theoretical Framework: Students

In Section 2, I proposed a model of how universities decide which courses to offer and showed that this model can be estimated with estimates of the marginal effects and marginal costs of offering additional course sections. In this section, I introduce a model of student course choices in which choices are influenced by the courses offered by the university. The student model allows me to analyze the effects of course offerings on student course choices; furthermore, it yields estimates of the marginal effects of offering additional course sections that can be used to estimate the university's implied preference parameters.

3.1 Student choices

As before, let $t \in [1, T]$ index academic semesters, let $f \in [1, F]$ index academic fields, let d_{tf} represent the number of sections of introductory field f courses offered in semester t, and collect these offerings into a single vector $\mathbf{d}_t$.[15] Furthermore, let $j \in [1, J]$ index specific introductory course sections, let $i \in [1, N]$ index students, and let $r \in [1, R_{it}]$ index observations of students choosing courses in a particular semester.[16] Finally, to allow for

[15] In the empirical application, fields are STEM, social science, humanities and arts, and business and occupational. See Appendix A for field definitions.

[16] For simplicity, I treat choices of multiple courses in the same semester by the same student as independent observations. I discuss this limitation and others in Section 3.3.

type specific unobserved heterogeneity in preferences for fields, assume students belong to one of S unobserved types, index types with $s \in [1, S]$, and let τ_{is} indicate whether student i is unobserved type s (Heckman and Singer, 1984).

Assume that the utility student i who is type s receives from choosing introductory course section j belonging to field f depends on observed student characteristics X_{it}, type specific field preferences ϕ_{sf}, and idiosyncratic preferences ϵ_{irtj}:

$$U_{irstj} = X_{it}\beta_f + \phi_{sf} + \epsilon_{irtj} \tag{8}$$

I assume the university knows X_{it}, β_f, ϕ_{sf}, τ_{is}, and the distribution of ϵ_{irtj} but does not observe individual realizations of ϵ_{irtj}. Equation (8) implies that the component of utility that is known by the university $X_{it}\beta_f + \phi_{sf}$ does not vary within field f. This restriction implies that marginal effects of offering additional sections of introductory courses in field f on expected student outcomes are the same regardless of which course within field f receives an additional section. This is central to the methodology because identification of university preference parameters γ_f requires marginal effects of offering additional course sections at the field level. If known utilities vary within fields, either the university model in Section 2 needs to be extended to model course offering decisions within fields or the researcher needs to make a somewhat arbitrary decision about which courses within a field are marginal.

To allow for correlation in idiosyncratic preferences for courses within the same field, I use a nested logit structure with nests defined by academic fields. As I show in Appendix B, this structure is equivalent to a Ackerberg and Rysman (2005) framework where unobserved preferences for field f course sections depend on the number of sections offered in that field and independent Type 1 Extreme Value preference shocks η_{irtj} as follows:

$$\epsilon_{irtj} = (\rho_f - 1)\log(d_{tf}) + \eta_{irtj} \tag{9}$$

The ρ_f nesting parameters, which are typically between 0 and 1, capture the extent to which students have similar unobserved preferences for courses within the same field. Roughly speaking, larger values for ρ_f imply “more independence” in unobserved preferences for courses within field f while smaller values for ρ_f imply “less independence” in unobserved preferences within field f (Train, 2009). As I discuss shortly, this crucial feature ensures that marginal effects reflect empirical variation rather than functional form assumptions.

With this structure, the probability that student i who is type s chooses one specific introductory course section in field f in semester t can be expressed as:

$$P_{istf} = \frac{d_{tf}^{\rho_f - 1} \exp\left(X_{it}\beta_f + \phi_{sf}\right)}{\sum_{f'=1}^{F} d_{tf'}^{\rho_{f'}} \exp\left(X_{it}\beta_{f'} + \phi_{sf'}\right)} \tag{10}$$

3.2 Student outcomes

With this framework for student demand, I can now define the total student utility and field enrollment outcomes that enter into the university's objective function in Equation (1). First, expected enrollment in introductory field f courses in semester t is given by:

$$\begin{aligned} n_{tf}\left(\mathbf{d}_t\right) &= \sum_{i=1}^{N}\sum_{s=1}^{S} \tau_{is} d_{tf} P_{istf} \\ &= \sum_{i=1}^{N}\sum_{s=1}^{S} \tau_{is} \left[\frac{d_{tf}^{\rho_f} \exp\left(X_{it}\beta_f + \phi_{sf}\right)}{\sum_{f'=1}^{F} d_{tf'}^{\rho_{f'}} \exp\left(X_{it}\beta_{f'} + \phi_{sf'}\right)}\right] \end{aligned} \tag{11}$$

Second, total expected student utility from introductory courses in semester t is given by:

$$\begin{aligned} V_t\left(\mathbf{d}_t\right) &= \sum_{i=1}^{N}\sum_{s=1}^{S} \tau_{is} \mathbb{E}\left[\max\left\{U_{isjt}\right\} \mid \mathbf{d}_t\right] \\ &= \sum_{i=1}^{N}\sum_{s=1}^{S} \tau_{is} \left\{\log\left(\sum_{f=1}^{F} d_{tf}^{\rho_f} \exp\left(X_{it}\beta_f + \phi_{sf}\right)\right) + c\right\} \end{aligned} \tag{12}$$

where $c \approx 0.5772$ is the Euler-Mascheroni constant. As required by the university model, both outcomes depend closely on course offerings $\mathbf{d}_t$.

However, as shown in Section 2, it is not these outcome formulas *per se* that are useful for measuring implied tradeoffs; rather, it is the marginal effects of course offerings on these outcomes. These marginal effects are given by:[17]

$$\frac{\partial n_{tf'}\left(\mathbf{d}_t\right)}{\partial d_{tf}} = \begin{cases} \rho_f \left(\sum_{i=1}^{N}\sum_{s=1}^{S} \tau_{is} P_{istf}\left(1 - d_{tf} P_{istf}\right)\right) & f' = f \\ -\rho_f \left(\sum_{i=1}^{N}\sum_{s=1}^{S} \tau_{is} d_{tf'} P_{istf} P_{istf'}\right) & f' \neq f \end{cases} \tag{13}$$

$$\frac{\partial V_t\left(\mathbf{d}_t\right)}{\partial d_{tf}} = \rho_f \left(\sum_{i=1}^{N}\sum_{s=1}^{S} \tau_{is} P_{istf}\right) \tag{14}$$

These formulas illustrate the important roles of the nesting parameters ρ_f in deter-

[17]Note that d_{tf} is actually a discrete variable and thus these derivatives are not defined; however, the number of introductory course sections in each field is large enough that approximating it as a continuous variable is reasonable.

mining the marginal effects of offering additional introductory course sections on outcomes. Equation (13) shows that the magnitudes of marginal effects on enrollment are increasing in ρ_f. This makes sense because more independence in unobserved preferences implies that additional sections will provide more meaningful variety, which will induce more students to switch fields in expectation. Similarly, Equation (14) shows that marginal effects on total student utility are increasing in ρ_f. Once again, this makes sense because sections that provide more meaningful variety generate more expected utility.

Moreover, these formulas illustrate a useful property for interpreting estimates of ρ_f. Nested logit is equivalent to multinomial logit when $\rho_f = 1$. As such, Equations (13) and (14) imply that ρ_f equals the ratio of marginal effects in field f in a nested logit framework to the marginal effects in field f in a multinomial logit framework. This implies that estimates of ρ_f directly quantify the empirical importance of allowing for nesting on marginal effects.

Because nesting parameters play a crucial role in determining the marginal effects of course offerings, it is important to understand how these parameters are identified from the data. Substituting Equation (9) into Equation (8), one sees that $(\rho_f - 1)$ is the coefficient on the natural logarithm of the number of field f course sections offered in semester t. Then following standard arguments for identification of coefficients in discrete choice models, one can see that $(\rho_f - 1)$ is identified by the relationship between relative course offerings in field f and the share of students choosing courses in field f across semesters.[18] As such, panel variation in course offerings is critical for producing credible estimates of ρ_f (and thus marginal effects and implied university preferences).

It makes intuitive sense that such variation would be important for measuring marginal effects of additional course offerings. If only one semester of data were available, there would be no empirical variation to assess how additional course offerings affect field enrollments. With multiple semesters, one can see how field enrollments vary when course offerings change. If additional offerings draw many students into a field, marginal offerings must be adding substantial utility. Conversely, if additional offerings do not change field enrollments, the new courses must be mostly redundant.

Although panel variation in course offerings is the most intuitive source of variation for identifying ρ_f, I note that this is not the only source of variation used to estimate ρ_f in practice. As discussed in Berry et al. (2004), variation in observed student characteristics and unobserved student types also plays an important role in identifying substitution parameters such as ρ_f. Intuitively, this variation helps identify ρ_f because these characteristics place students at varying distances from points of indifference between alternatives. This implies that the number of additional sections in field f needed to change a student's choice differs

[18]For example, see Chapter 2 of Train (2009)

across students providing useful variation for identifying ρ_f.

3.3 Discussion

The main identifying assumption necessary to recover ρ_f is that introductory course offerings d_{tf} must be independent of preference shocks η_{irtj}. In words, this assumption means the university cannot consider idiosyncratic student preferences when deciding how many sections of introductory courses to offer in each field. Two violations of this assumption seem most plausible: First, the university may use pre-registration information to cancel unpopular courses or offer additional sections of popular ones.[19] Second, the university may forecast trends in field preferences across semesters either by anticipating general trends in student preferences or by noticing which courses in preceding semesters were over- or under-subscribed. Because the structure assumes field preferences β_f are fixed across semesters, these trends will be subsumed into η_{irtj} thus any response of the university to these trends will cause misspecification.[20]

Both of these scenarios suggest there could be positive correlation between introductory course offerings d_{tf} and preference shocks η_{irtj} that may introduce upward bias in estimates of ρ_f. Because only relative marginal effects matter for inferring university preferences, this will only confound estimates of γ_f if the endogeneity is stronger in certain fields relative to others. Furthermore, I will argue that the presence of detailed baseline student characteristics and the lack of trends in introductory course offerings imply that concerns about the exogeneity of $\mathbf{d}_t$ should be limited. Finally, by allowing for type specific unobserved heterogeneity in preferences for fields, I give the university some scope to tailor its course offerings in response to student characteristics that are known to the university but are not included in my data.

In addition to this endogeneity concern, this framework for student demand possesses several important limitations: First, the framework does not have a mechanism for incorporating section capacity constraints. As most universities, UCA places constraints on the number of students who can enroll in particular course sections. This implies that for sections where the capacity constraint is reached, true student demand may be substantially greater

[19] The university I study posts preliminary Fall (Spring) course offerings by March (October) of the preceding Spring (Fall) semester at which point currently enrolled students can pre-register for courses. While the stated justification for pre-registration is to allow students to plan ahead, the university is not precluded from changing course offerings in response to pre-registration information (UCA, 2006).

[20] In theory, one could allow for some degree of time variation in preferences β_f; however, any time variation in field enrollments that is captured by variation in β_f can no longer be explained by variation in course offerings. In the extreme case, if β_f were semester specific, then all variation in n_{tf} across semesters would be captured by semester specific β_f. As such, allowing for time variation in β_f reduces identifying variation for ρ_f. I will show that there are no detectable trends in field preferences suggesting it is better to assume β_f is fixed to preserve variation for identifying ρ_f.

than constrained demand. Unfortunately, data on capacity constraints are not available; however, even with data on constraints, methodological advances would likely be required to incorporate these in demand estimation.[21] Omitting capacity constraints leads to understating demand for certain course sections; if disproportionately many of these sections are in field f, this may lead to understating the marginal effects of offering additional sections in field f and thus overstating the university's implied preference for enrollment in field f.

Another limitation is that the framework implicitly allows students to retake courses they have already passed. In theory, one could make d_{tf} student specific by subtracting sections of courses that students have already passed; however, this may lead to negative correlation between a student specific d_{tf} and persistent unobserved field preferences in η_{irtj}. While allowing for type specific unobserved heterogeneity may address this issue in some settings, various experiments in my setting found that the endogeneity was too strong to overcome with type specific unobserved heterogeneity. As such, I abstract from choice set variation due to previously passed courses. Because my framework relies on panel variation in course offerings and field enrollments for identification, it will still accurately predict the aggregate marginal effects of offering additional course sections on field enrollments. However, it may overstate (understate) these effects for individual students who have (have not) already passed these courses.

A further limitation is that the framework does not incorporate class size externalities. Although most of the literature on class size externalities has focused on primary school, one may suspect that college students also value small class sizes with more instructor interaction.[22] Class size effects are challenging to include in the demand model because class size will be correlated with any unobserved course section attributes by construction. Bayer and Timmins (2007) propose an iterative IV strategy for addressing this issue. I experimented with this method but found that the instruments lacked power in my setting. As such, I have excluded class size effects from my analysis.

Finally, this framework assumes students choose individual course sections independently rather than complementary bundles of sections. Nevo et al. (2005) and Ahn et al.

[21] Conlon and Mortimer (2013) uses vending machine data to estimate demand in a setting where items can be sold out. However, they observe vending machine inventory every four hours yielding substantial observed variation in item availability. To use their methods in a course choice model, one would need enrollment timestamps or other information to identify which students had the option to enroll in a section which eventually became constrained.

[22] For example, see Angrist and Lavy (1999); Hoxby (2000); Krueger (2003). An exception which examines the effects of class size in higher education is Kokkelenberg et al. (2008). Alternatively, the university might promote smaller classes to increase student learning. The university in Section 2 values student choices and utilities but not their learning. As such, a larger γ_f may partially reflect a university's belief that smaller classes have greater pedagogical benefits in field f (or that learning in field f is valued more than in other fields).

(2022) introduce methods for estimating models of bundled choice; however, these methods greatly increase computational burden and complicate the equations for enrollment, total utility, and marginal effects.[23] Moreover, these methods typically yield parameter estimates similar to those of independent choice models when the number of choices is large as in my setting. As such, I prefer the independent choice specification for transparency and tractability.

4 Data, descriptive statistics, and estimation

The framework introduced previously calls for panel data of offered introductory courses, student characteristics, and student course choices as well as data for estimating the marginal costs of offering additional sections of introductory courses. To this end, I employ administrative data from the University of Central Arkansas (UCA). UCA is a large public teaching focused university located in central Arkansas. Table 1 provides background statistics on UCA. The statistics show UCA is a less selective mid-sized university with a six year graduation rate that is below the national average.[24] Furthermore, almost all students at UCA are full-time, 24 and under, and from the state of Arkansas.

These administrative data include demographic information, admissions information, and full academic transcripts for all students who were enrolled between the 2004-05 and 2012-13 academic years. The data also include information on all offered course sections and the instructors teaching these sections during the same time period. After excluding required writing courses, required oral communication courses, required health courses, and other special courses, the data include 32,445 unique UCA undergraduates and 359,659 observations of students choosing introductory courses.[25]

[23] Gentzkow (2007) analyzes news consumption in a bundled choice framework that also allows for complementarities between choices; however, the framework is only feasible when the cardinality of the choice set is small.

[24] The national average six year graduation rate is 59.4% (Ginder et al., 2017).

[25] Required writing, oral communication, and health courses are specific courses which almost all students take during their Freshmen year. I exclude these courses because students are choosing these courses to satisfy a requirement rather than to maximize utility. Including these courses would lead me to overstate the desirability of fields associated with these courses. I also exclude first year seminar courses (which are only available to freshmen and can only be taken once), English as a second language courses, military science courses, and courses worth fewer than three credit hours (which are predominantly labs associated with other courses, music lessons, and exercise classes). In addition to writing, oral communication, and health courses, UCA also has general education requirements in fine arts, American history and government, humanities, mathematics, natural sciences, behavioral and social sciences, and world cultural traditions. These requirements can be satisfied with many different courses and are often completed in later years. Furthermore, many of these courses also satisfy major specific requirements. I include these courses because many students are choosing these courses to maximize utility. For more information, please see the UCA course bulletin (UCA, 2006).

These administrative data are ideal for this study for two reasons: First, the data on student choices and characteristics together with information on course offerings allows me to analyze how students make choices given a set of alternatives. Crucially, the panel structure of these data allows me to analyze how choices change when course offerings change providing useful empirical variation for identifying the marginal effects of changing course offerings. Second, the data include information on instructor salaries, teaching loads, and contract characteristics, which allows me to estimate the implied cost of offering course sections with different characteristics and to constrain counterfactuals so that the university can only reallocate its instruction budget for adjunct instructors.

An important empirical decision is whether to use courses (Econ 101), course-instructor pairs (Econ 101 taught by Prof. Smith), or course sections (Econ 101 taught by Prof. Smith at 9AM on Tuesdays and Thursdays) as the unit of analysis j. In Section 2, j represents a unit that presents a marginal cost to the university. In Section 3, j represents a unit that provides meaningful choice variety to students. In this paper, I use course sections—defined by a course number, instructor, and meeting time—as the unit of analysis j. Arguments can certainly be made in favor of alternative choices; however, I feel course sections are the most appropriate unit because they present the most direct cost to the university. When defining full-time instructors and computing each instructor's share of full-time, UCA uses course sections rather than courses as the relevant unit (ADHE, 2011). This choice reflects the fact that although there are fixed preparation costs, instruction and grading time are substantial costs that roughly vary by number of sections. Because this paper studies the decisions of a university, I choose the unit of analysis that presents the most direct cost to the university.

Using course sections as the unit of analysis j implies that variety across j arises from differences in course content, instructor, and meeting time. One may argue that another section of an existing course taught by the same instructor but at a different time provides trivial choice variety to students. However, I would argue that if the university is willing to effectively pay an instructor to teach an additional section, it must be because the university implicitly values this additional section. Moreover, as discussed in Section 3, empirical variation determines whether I will find significant effects of offering additional sections, and results in Section 5 show that these effects are large and significant.

4.1 Descriptive statistics

In my empirical analysis, I analyze introductory course offerings and student choices across four academic fields: STEM, social science, humanities and arts, and business and occupational. Before proceeding to this analysis, Table 2 compares several relevant statistics across introductory courses in these fields. The statistics show that social science is the largest field in terms of courses, sections, and student enrollment. STEM is second in terms of course sections and student enrollment but has relatively fewer courses suggesting offerings in this field may be more homogenous. Humanities and arts is third largest in terms of course sections and student enrollment followed by business and occupational.

Statistics on average introductory enrollment per section show that on average there are 32.6 students in social science sections, 28.2 students in humanities and arts sections, 25.4 students in both STEM sections, and 24.5 students in business and occupational sections. These differences suggest there is substantial variation in the average desirability of introductory courses in different fields. Furthermore, the cost statistics show that social science sections have the lowest implied instruction costs at all quartiles of the cost distributions. The low average costs and large average class sizes in social sciences do not necessarily imply that marginally reallocating resources from STEM to social sciences would increase total student utility; however, they do provide suggestive evidence that there could be some misalignment between student preferences and observed course offerings.

The remaining statistics in Table 2 describe how observed student characteristics affect course choices. The statistics show that students choosing introductory STEM courses have higher ACT scores and high school GPA than students choosing introductory courses in other fields on average. Students choosing business and occupational courses have the second highest ACT scores and high school GPA and students choosing social science and humanities and arts courses have the lowest measures of baseline preparation. The statistics also show that students choosing introductory business and occupational courses are less likely to be women or freshmen but more likely to be sophomores, juniors, or seniors.

In Subsection 3.2, I showed that the crucial nesting parameters ρ_f are identified by the empirical relationship between the share of course sections offered in each field and relative field enrollments across semesters. Table 3 reports the number of introductory courses and sections offered in each field by semester as well as each field's share of total sections and total introductory enrollment by semester to illustrate this identifying variation.[26] The statistics show that the share of STEM sections varies from 28% - 31% across semesters, the share of social science sections varies from 34% - 37% across semesters, the share of humanities and arts sections varies from 20% - 25%, and the share of business and occupational sections varies from 12% - 14%. The extent to which enrollment shares move in concert with these

[26]Note that there are generally fewer introductory sections offered in Spring semesters relative to Fall semesters. This across semester variation in total introductory sections is not a source of identifying variation for ρ_f. Only the share of sections offered in each field across semesters is a source of identifying variation.

fluctuations in section shares helps identify the nesting parameters ρ_f.

As discussed previously, this identification argument relies on the assumption that course offerings are uncorrelated with the unobserved components of student preferences. While this assumption is fundamentally untestable, one can investigate whether there appear to be broad trends in preferences and course offerings that would cause endogeneity. A perusal of Table 3 suggests such trends are not present in these data. Section shares and enrollment shares fluctuate from year to year in a manner that appears random suggesting that estimates of nesting parameters are not confounded by correlated trends in preferences and course offerings.

4.2 Estimation

To estimate the student choice model in the presence of unobserved type specific heterogeneity, I employ the Expectation-Maximization (EM) Algorithm (Dempster et al., 1977). The EM algorithm is an iterative algorithm that uses a student's choices to infer her unobserved type. For instance, suppose unobserved type s has a relative preference for business and occupational classes. If student i chooses many more business and occupational classes over the course of her career at UCA than her observed characteristics would predict, then the algorithm will identify her as more likely to be type s. A full description of the EM algorithm is provided in Appendix B.

In addition to providing estimates of student choice parameters β_f, ϕ_{sf}, and ρ_f, the EM algorithm also yields estimates of the unconditional probability π_s that a random student is type s, and the conditional probability q_{is} that student i is type s given her choice history.[27] Conditional type probabilities can then be used in place of τ_{is} in Equations (14) and (13) to construct the marginal effects of offering additional course sections.

In addition to these marginal effects, inferring university preferences also requires estimates of the marginal costs of offering additional sections. To estimate these costs, I compute the average cost of hiring an adjunct instructor to teach a single course section in each field.[28] Once I have obtained estimates of the marginal effects and marginal costs of offering additional course sections, I estimate implied university preference parameters γ_f using Equation (7).

[27]See Appendix B for formal definitions of π_s and q_{is}. Note that I am assuming unconditional type probabilities are the same for all students in the sample. Variants of the EM algorithm allow these probabilities to depend on observed characteristics (Arcidiacono, 2005).

[28]Using instructor rank to estimate costs suggests rank should also enter into student utility. A preliminary analysis found that instructor rank has negligible effects on student utility; as such, I exclude these effects for power and tractability. See Appendix A for a detailed description of how I use data on instructor salaries, contract details, and teaching histories to construct the implicit cost of hiring an instructor to teach a single course section.

The idea of using adjunct instructor costs is that if a university wants to add or subtract a section, it is generally simpler and more cost effective to do this by hiring or firing an adjunct instructor rather than a tenure-track instructor.[29] As such, costs of hiring adjunct instructors represent better estimates of marginal costs than average costs within a field.

One limitation of this method is that the marginal cost of offering an additional course section in field f is independent of the number of sections offered in that field. This is consistent with a framework in which UCA is a wage-taker in the market for adjunct instructors; however, this assumption may still be violated at hypothetical course offerings that are far away from observed offerings. Another limitation of this cost framework is that it ignores facility costs, material costs, and other non-instructor costs. If these other costs are proportional to adjunct instructor costs, then their omission will not affect the relative marginal costs used to measure implied university preferences. However, if other costs are disproportionately high in field f, this will lead to downward bias in estimates of γ_f.

5 Results

This section reports the results of my analysis of the introductory course offerings at the University of Central Arkansas (UCA). I first report estimates of student preference parameters and show that course offerings have important effects on student course choices. Next, I show that observed course offerings are implicitly sacrificing student utility to draw students out of social science and humanities and arts courses and into STEM and business and occupational courses. Finally, I consider how UCA would change its course offerings in response to changes in student preferences or costs of instruction and show that these responses would have important implications for student course choices.

5.1 Student preference parameters

To begin, Table 4 reports estimates of student preference parameters from the course choice model that will be used to measure the marginal effects of offering additional course sections. Results are for a specification with two unobserved student types and parametric block bootstrapped standard errors are reported in italics.[30] Although student choice parameters are

[29] Research universities may find it optimal to subtract a course by giving a tenured or tenure-track instructor a teaching reduction which allows her to produce more research. This is less likely to be true at a teaching-focused university such as UCA.

[30] Parametric block bootstrapping samples full student panels of observed independent variables and uses point estimates and the course choice model to simulate unobserved types and choices. This is the recommended procedure for estimating standard errors in models with type specific unobserved heterogeneity (McLachlan and Peel, 2004).

not the focus of this paper, several findings are worth highlighting: First, higher ACT scores and high school GPA make students more likely to choose introductory STEM courses. This is consistent with existing literature that shows initial preparation is an important determinant of whether a student pursues a STEM education (Arcidiacono, 2004; Stinebrickner and Stinebrickner, 2014). Second, women are less interested in business and occupational courses on average; conversely, non-white students and students missing race are more interested in business and occupational courses than their white counterparts. Furthermore, Sophomores, Juniors, and Seniors are much more likely to take business and occupational courses than Freshmen.

Finally, estimates of the nesting parameters ρ_f are in the middle range varying from 0.555 to 0.800. Recall from Subsection 3.2 that estimates of ρ_f can be interpreted as the ratio of marginal effects under nested logit to marginal effects under multinomial logit. Estimates of ρ_f thus illustrate that nesting is statistically and economically significant: Allowing for nesting yields marginal effects that are 55.5% to 80.0% of the marginal effects that one would obtain with multinomial logit. Estimates are highest in business and occupational followed closely by STEM and humanities and arts. Estimates are substantially lower in social science implying that new social science sections add less variety than new sections in other fields.

5.2 Marginal effects of course offerings on field enrollments

Estimates of the student course choice model can then be used to compute the marginal effects of offering additional introductory course sections on field enrollments. These effects illustrate the extent to which the university can influence student course choices through its offerings. Because introductory course choices may influence what fields students ultimately choose to specialize in, these effects may have lasting implications for students in the labor market.

The top section of Table 5 reports the marginal effects of offering an additional introductory course section in each field on expected enrollment in all fields. Rows represent the fields where sections are added while columns represent the fields of enrollment. For example, estimates in the first row imply that an additional STEM section increases STEM enrollment by 13.47 students and decreases enrollment in social science, humanities and arts, and business and occupational by 7.61 students, 4.26 students, and 1.6 students, respectively. Diagonal entries show that own field enrollment effects are sizable, ranging from 10.59 students in social science to 15.3 students in humanities and arts.

To contextualize these effects, the bottom section of Table 5 reports the expected enrollment in a new section that is added to each field and the diagonal own-field enrollment effects relative to these expected enrollments. Although technically incorrect, comparing these expected enrollments to the diagonal own-field enrollment effects provides useful intuition for assessing the extent to which new sections draw students from their own field versus other fields.[31]

Results show that new course offerings are not simply cannibalizing their own fields—60.0% of students in a new business and occupational section would have chosen a course in another field if not for that new section. Findings are similar for STEM and humanities and arts where 53.3% and 54.2% of students are attracted from other fields. The share is lowest in social science where only 32.6% of students enrolled in a new section would have chosen a course in another field. This reflects the finding from Table 4 that new social science sections add less variety than new sections in other fields. However, the share is still substantially larger than zero implying that new social science sections still provide meaningful variety.

Altogether, estimates show that course offerings have important effects on student course choices. As such, it is important to understand how course offerings are chosen, how course offerings might change in response to changes in student preferences or instructor salaries, and the implications of these responses for student course choices.

5.3 Marginal effects on student utility and implied university preferences

To analyze how course offerings are chosen, Table 6 reports the marginal effects of offering additional course sections on total student utility and the tradeoffs between total student utility and field enrollments that are implied by observed course offerings. To begin, column 1 reports the local marginal effects of offering additional sections of introductory courses on total student utility. For expositional purposes, I report relative marginal effects averaged across semesters. In notation, column 1 reports:

$$\frac{1}{T}\sum_{t=1}^{T}\frac{\left({}^{\partial V_t(\mathbf{d}_t)}\big/_{\partial d_{tf}}\right)}{\left({}^{\partial V_t(\mathbf{d}_t)}\big/_{\partial d_{tHum}}\right)} \tag{15}$$

Results indicate that marginal effects of offering additional course sections are fairly similar across fields. Point estimates suggest effects in social science are 10% smaller than

[31]This is technically incorrect because the model predicts enrollment probabilities rather than particular course choices. As such, it is impossible to identify students who choose courses in different fields in different scenarios. To make this technically correct, one would need to assume that idiosyncratic preferences η_{irtj} are fixed in observed and counterfactual scenarios. This would imply that all students would either remain in the same section or switch to the newly added section. While theoretically reasonable, this is empirically infeasible because η_{irtj} are not point identified.

effects in humanities in arts; however, this difference is not statistically significant. However, although marginal effects on utility are similar, column 2 shows that marginal costs vary significantly across fields. Adjunct instructors for business and occupational courses cost 38.5% more than adjunct instructors in social science.

Column 3 combines marginal effects and marginal costs to form the marginal effects per dollar that will be used to infer university preferences. Once again, I report relative marginal effects per dollar averaged across semesters. Marginal effects per dollar are very similar in humanities and arts and social sciences; however, effects per dollar in STEM and business and occupational are 22% and 23% smaller than effects per dollar in humanities and arts.

These differences show that marginally reallocating spending from STEM and business and occupational sections to social science and humanities and arts would increase total student utility; however, doing so would decrease variety (and thus enrollment) in the former fields and increase variety (and thus enrollment) in the latter fields. Since spending was not reallocated in this manner despite the potential for increasing student utility, observed course offerings are implicitly sacrificing some student utility to keep students from switching from STEM and business and occupational courses to social science and humanities and arts. This reveals an implicit willingness to sacrifice student utility to draw students out of social science and humanities and arts courses and into STEM business and occupational courses.

Column 3 reports estimates of γ_f to precisely quantify these implicit tradeoffs. The omitted field is humanities and arts; therefore, estimates for field f report how much total student utility the university is implicitly sacrificing to move one student out of an introductory humanities and arts course and into an introductory course in field f in expectation. Recall that these estimates can be interpreted as the implicit tradeoffs made by observed course offerings even if the university's problem in Equation (3) is misspecified.

5.4 Student utility maximizing course offerings

To further quantify the misalignment between student preferences and observed course offerings, columns (1) - (5) of Table 7 compare average observed course offerings and field enrollments to cost-equivalent offerings and enrollments that would have maximized total student utility. Columns (1) - (3) report averages across semesters of the number of introductory course sections taught by tenure-track instructors in each field, the number of introductory course sections taught by adjunct instructors in each field, and enrollment in introductory courses by field. Columns (4) and (5) then examine how adjunct instructed offerings and enrollments would change if the portion of the budget allocated to pay adjunct instructors were reallocated to maximize total student utility holding tenure-track offerings in column (1) fixed.[32] Stars indicate where columns (4) and (5) are statistically different from columns (2) and (3) respectively.

Results suggest the student utility maximizing (SUM) allocation of the adjunct instructor budget for introductory courses would contain much fewer STEM and business and occupational courses sections and much more social science and humanities and arts sections. These differences reflect the finding in Table 6 that marginal humanities and arts and social sciences sections yield more student utility per dollar than marginal STEM and business and occupational sections. Column (5) predicts that SUM adjunct offerings would increase humanities and arts enrollment and social science enrollment by 26% and 10% and decrease STEM and business and occupational enrollment by 28% and 27%.

To provide an additional intuitive way to measure the misalignment between student preferences and observed course offerings, column (6) of Table 7 illustrates a revenue neutral tax/subsidy policy that would induce UCA to offer the SUM sections reported in column (4). This illustrates how much costs of adjunct instructors would need to change to "price-out" UCA's implied preferences for STEM and business and occupational enrollment. Intuitively, I infer these by solving for costs that make it so the university's first order conditions are satisfied at SUM course offerings. Details are reported in Appendix B. I stress that these are provided to assist in interpreting γ_f and are not intended as a specific policy recommendation—because social welfare may differ substantially from total student utility, it is not clear whether pricing out UCA's implied preferences for STEM and business and occupational enrollment would benefit society as a whole.

Results suggest it would take a 20.3% increase in the cost of hiring a business or occupational adjunct, a 15.6% increase in the cost of hiring a STEM adjunct, and modest subsidies for social science and humanities and arts instructors to induce UCA to offer the student utility maximizing sections reported in column (4). This shows that the estimated preference parameters $\hat{\gamma}_f$ are large enough that it would take substantial changes in costs to price out these implied preferences.

[32] There are several reasons to reallocate the budget for adjunct instructors only: First, this mechanically restricts counterfactual course offerings to remain relatively close to observed offerings where I am more confident in the predictive power of the estimated student choice model; second, this represents a realistic picture of what could be achieved in the short run since tenure-track instructors are generally more difficult to fire; third, the model provides no mechanism for explaining why the university hires instructors of different ranks and thus is not well equipped to predict hiring decisions across ranks.

5.5 Counterfactual analyses with university responses

Previous results showed course offerings have important implications for student course choices. Previous results also provided insights into how course offerings are chosen by the university—revealing that course offerings implicitly sacrifice student utility to increase enrollment in STEM and business and occupational courses. This subsection combines both insights to predict how UCA would change its course offerings in response to changes in student preferences or costs of instruction and to analyze the implications of those responses for student course choices.

To enhance the credibility of predictions, I make two choices:[33] First, I choose counterfactual scenarios that are relatively close to the observed scenario to increase confidence in the ability of my model to predict university and student choices. Second, as in Table 7, I only allow the university to reallocate the portion of its budget for introductory courses paid to adjunct instructors. This also restricts the counterfactual scenarios to be close to the observed scenario and examines a short run scenario in which inputs that are costly to vary are held fixed.

Given the policy interest in increasing specialization in STEM, I first consider a scenario where the state introduces a 5% subsidy for STEM instructors to increase STEM course offerings and enrollments. Row 2 of Table 8 shows that this subsidy would increase the number of adjunct instructed STEM sections by 32.0%; furthermore, row 2 of Table 9 shows that this increase in STEM offerings would lead to a 8.9% increase in overall STEM enrollment.[34] The total cost of the subsidy would be $34,526 or $74.33 per additional STEM course choice.

A subsidy for STEM instructors would have both income and substitution effects making it unclear *a priori* whether the subsidy would increase or decrease offerings in other fields. Estimates show that offerings in other fields would decrease, which would reinforce student migration towards STEM. This illustrates the importance of treating the university as a strategic agent—a forecast that assumes offerings in other fields remain fixed would understate the effect of the subsidy on STEM offerings and enrollment.

Next, I consider a scenario in which UCA begins attracting better prepared students.

[33]To predict course offerings in counterfactual scenarios, I solve the university's problem in Equation (3) using numerical constrained maximization methods. I use numerical methods because the first order conditions characterizing the solution to the university's problem are complicated non-linear functions of course offerings. As such, it is unclear whether a closed form expression for a solution to the university's problem exists.

[34]Tables 8 and 9 also report predicted introductory course offerings and introductory field enrollments in the observed state in row 1. Stars in rows 2 and beyond indicate whether predictions in counterfactual scenarios are statistically different from predictions in the observed state. Reported figures are averages of predictions across all academic semesters.

Cornwell et al. (2006) show evidence that the Georgia HOPE scholarship program increased the SAT scores of students attending public universities in Georgia relative to the national student population. From Table 4, we know that students with higher scores are generally more interested in STEM suggesting that universities might respond to such a change in student composition by offering more STEM courses and thus making the STEM field even more attractive.

To analyze this scenario, my second counterfactual predicts adjunct instructed introductory course offerings and introductory field enrollments if all student ACT scores and high school GPAs were increased by one-tenth of a standard deviation. These increases are calibrated to approximately match the effects of the Georgia HOPE scholarship documented in Cornwell et al. (2006).[35] Row 3 of Table 8 shows that increasing student preparation would increase the number of adjunct instructed STEM sections by 11.5% and reduce offerings in other fields. To see the effects of this intervention on field enrollments, row 3 of Table 9 first predicts field enrollments in a partial equilibrium where student characteristics are changed but course offerings remain fixed. Results show that attracting better prepared students would increase introductory STEM enrollment by only 0.8% without any response in course offerings. Row 4 of Table 9 incorporates the changes in adjunct instructed course offerings and shows that the total effect of attracting better prepared students is a 5.1% increase in introductory STEM enrollment. This illustrates the importance of incorporating university responses into counterfactual policy analyses; ignoring changes in course offerings leads to understating increases in STEM enrollment by 4.3 percentage points.

My final analysis considers a scenario in which the gender composition of students changes, as happened in the US over the course of the 20th century. According to the National Center for Education Statistics, female share of college enrollment increased from 40.4% in 1968 to 57.0% in 2018.[36] Table 4 showed that female students have relatively stronger preferences for social science and humanities and arts courses, suggesting that an increase in female attendance may lead to more course offerings in those fields. To consider how greater female representation might have affected course offerings and field enrollments, I compare UCA in the observed state (where the gender share of 59.0% closely matches the current national share) to a counterfactual state where the gender share matches the 1968 share of 40.4%.

[35] Specifically, I use the data in Figure 3 of Cornwell et al. (2006) to conclude that the SAT scores of Freshmen at public universities in Georgia increased 25 points after 1993 versus 8.3 points for high school seniors nationwide. This implies an effect of 16.75 SAT points. This is equivalent to 0.396 ACT points, which represents 0.091 standard deviations in my data.

[36] Source: NCES Table 303.10.: Total fall enrollment in degree-granting postsecondary institutions, by attendance status, sex of student, and control of institution: selected years, 1947 through 2029.

Row 4 of Table 8 shows that with a 1968 gender ratio, UCA would offer 11.8% more adjunct instructed business and occupational sections, 4.3% more adjunct instructed STEM sections, 5.9% fewer adjunct instructed humanities and arts sections, and 8.5% fewer social science sections. Once again, to separate out the direct effects of changes in students characteristics and the indirect effects of changes in course offerings, row 5 of Table 9 predicts field enrollments in partial equilibrium without changes in course offerings and row 6 of Table 9 predicts field enrollments in general equilibrium with university responses. Results show that a 1968 gender ratio would increase business and occupational enrollment by 3.5% without any change in course offerings. Incorporating the effects of the increase in adjunct instructed introductory business and occupational sections leads to a total predicted increase in introductory business and occupational enrollment of 8.7%. Once again, ignoring the indirect effects of changes in course offerings leads to significantly understating changes in field enrollments.

6 Conclusion

In 1973, Daniel Bell described the university as "the axial institution of post-industrial society" (Bell, 1973). This is more true today than it was nearly half a century ago. Despite this, there is still a great deal we do not know regarding how universities make decisions and the implications of these decisions for students. These knowledge gaps limit our understanding of the "axial institution" and prevent higher education policymakers from choosing policies that best serve students and society as a whole.

To advance our understanding of the "supply side" of higher education, this paper empirically analyzed introductory course offerings at the University of Central Arkansas (UCA). First, I showed that UCA's introductory course offerings influence student course choices. For example, offering an additional STEM course section increases total STEM enrollment by 13.47 students, which is 53.3% of the expected enrollment in that new section. Next, I showed that introductory course offerings at UCA implicitly sacrifice student utility to increase enrollment in STEM and business and occupational classes. If UCA reallocated its budget for adjunct instructors to maximize student utility, it would reduce adjunct instructed STEM offerings by 88% and roughly double adjunct instructed humanities and arts course sections. Finally, I showed that UCA would change its course offerings in counterfactual scenarios and that these responses would have important implications for student course choices. For example, ignoring university responses would lead to understating the effects of attracting better prepared students on STEM enrollment by 4.3 percentage points.

To my knowledge, this is the first analysis of how universities decide which courses to offer and the implications of these decisions for students. As such, the analysis must come with important caveats and there is substantial room for subsequent extensions. First, although the static student demand model provides a transparent and intuitive mapping from empirical variation to effects of course offerings, this model abstracts from many factors that affect student choices and assumes course offerings are orthogonal to idiosyncratic student preferences. Future work may build deeper models of student demand or exploit quasi-experimental variation in course offerings to enhance the credibility of results.

Furthermore, while my analysis quantifies how much student utility a university is willing to sacrifice to increase STEM and business and occupational enrollment, my framework cannot identify why the university prefers these fields or whether these preferences benefit society as a whole. STEM and business and occupational courses generally have higher labor market returns (Altonji et al., 2012). As such, the university may favor these fields paternalistically to nudge myopic students into majors with higher returns. Alternatively, if STEM and business and occupational education have larger social externalities, the university may favor those fields to internalize their externalities. Future work that explores these possibilities will broaden our understanding of the "axial institution" and may lead to policies that better serve students and society as a whole.

Appendix A: Data Appendix

Field Definitions

Fields are defined using Classification of Instructional Program (CIP) Codes as follows:

STEM: Biological and Biomedical Sciences (26), Mathematics and Statistics (27), Physical Sciences (40).

Social Science: Area, Ethnic, Cultural and Gender Studies (5), Foreign Languages, Literatures, and Linguistics (16), Family and Consumer Sciences / Human Sciences (19), Psychology (42), Social Sciences (45), History (54).

Humanities and Arts: Communication, Journalism, and Related Programs (9), English Language and Literature / Letters (23), Philosophy and Religious Studies (38), Visual and Performing Arts (50).

Business and Occupational: Computer and Information Sciences and Support Services (11), Education (13), Parks, Recreation, Leisure, and Fitness Studies (31), Health Professions and Related Clinical Sciences (51), Business, Management, Marketing, and Related Support Services (52).

Instructor costs

To compute the implicit cost of hiring an instructor to teach course section j C_j, I use information on instructor salaries, contract details, and teaching histories. Instructor salaries are typically paid for multiple services across multiple semesters so one must make assumptions regarding what share of an instructor's total salary is paid for a specific section. Generally speaking, this method uses credit hours to allocate an instructor's total salary to specific sections. I make use of the following information: how much the instructor is paid for an entire contract, a contract identifier which indicates which semesters are covered by the same contract, the number of credit hours that a full time instructor teaches, a numeric measure of what share of full time each instructor is, and the credit hour value of each course section.

The first step is to calculate the number of credit hours each instructor would be teaching in each semester if they were only paid to teach. This involves multiplying the share of full time measure by the number of credit hours that a full time instructor teaches. For example, if an instructor has a 50% part time contract and a full time instructor teaches 12 credit hours per semester, then this instructor would teach 6 credit hours if she were only paid to teach. The second step is to sum these teaching only credit hours across all semesters covered by the same contract. This represents the total number of credit hours the instructor would teach in each contract if they were only paid to teach. The third step is to divide instructor salary for each contract by this measure of total contract teaching only credit hours. This yields a measure of salary per credit hour for each contract which can be interpreted as an instructor wage. Finally, multiplying this salary per credit hour measure by the credit hour value of each course section yields the instructor salary paid for each course.

Importantly, this method ensures that faculty members who are paid for activities other than teaching are not assigned inflated "wages" despite having high salaries relative to the number of credit hours they teach. To see this, suppose the 50% part time instructor from the previous example only teaches a three credit hour course and receives the rest of her compensation for administrative duties. If she is on a one semester contract with a salary of \$60,000, her salary per credit hour of teaching is:

$$\frac{\$60,000}{6\text{hrs}} = 10,000\frac{\$}{\text{hr}}$$

Dividing by 6—the credit hours she would teach if she were only paid to teach—rather than 3—the credit hours she actually taught—ensures that her pay for administrative activities does not inflate the true cost of hiring her to teach.

Appendix B: Technical Appendix

Equivalence of Nesting and Crowding Models

This subsection shows the equivalence of a nested logit model with nests defined by academic fields and a Ackerberg and Rysman (2005) crowding model with the error structure given in Equation (9). In the nested logit model, probabilities of choosing any course section in field f are given by:

$$P_{istf} = \frac{\exp\left(\frac{X_{it}\beta_f+\phi_{sf}}{\rho_f}\right)\left[\sum_{j'\in f}\exp\left(\frac{X_{it}\beta_f+\phi_{sf}}{\rho_f}\right)\right]^{\rho_f-1}}{\sum_{f'=1}^{F}\left[\sum_{j'\in f'}\exp\left(\frac{X_{it}\beta_{f'}+\phi_{sf'}}{\rho_{f'}}\right)\right]^{\rho_{f'}}}$$

These simplify as follows:

$$\begin{aligned}
P_{istf} &= \frac{\exp\left(\frac{X_{it}\beta_f+\phi_{sf}}{\rho_f}\right)\left[d_{tf}\exp\left(\frac{X_{it}\beta_f+\phi_{sf}}{\rho_f}\right)\right]^{\rho_f-1}}{\sum_{f'=1}^{F}\left[d_{tf'}\exp\left(\frac{X_{it}\beta_{f'}+\phi_{sf'}}{\rho_{f'}}\right)\right]^{\rho_{f'}}} \\
&= \frac{\exp\left(\frac{X_{it}\beta_f+\phi_{sf}}{\rho_f}\right)d_{tf}^{\rho_f-1}\left[\exp\left(\frac{X_{it}\beta_f+\phi_{sf}}{\rho_f}\right)\right]^{\rho_f-1}}{\sum_{f'=1}^{F}d_{tf'}^{\rho_{f'}}\left[\exp\left(\frac{X_{it}\beta_{f'}+\phi_{sf'}}{\rho_{f'}}\right)\right]^{\rho_{f'}}} \\
&= \frac{d_{tf}^{\rho_f-1}\left[\exp\left(\frac{X_{it}\beta_f+\phi_{sf}}{\rho_f}\right)\right]^{\rho_f}}{\sum_{f'=1}^{F}d_{tf'}^{\rho_{f'}}\left[\exp\left(\frac{X_{it}\beta_{f'}+\phi_{sf'}}{\rho_{f'}}\right)\right]^{\rho_{f'}}} \\
&= \frac{d_{tf}^{\rho_f-1}\exp\left(X_{it}\beta_f+\phi_{sf}\right)}{\sum_{f'=1}^{F}d_{tf'}^{\rho_{f'}}\exp\left(X_{it}\beta_{f'}+\phi_{sf'}\right)} \\
&= \frac{\exp\left(\log\left(d_{tf}^{\rho_f-1}\right)\right)\exp\left(X_{it}\beta_f+\phi_{sf}\right)}{\sum_{f'=1}^{F}d_{tf'}\exp\left(\log\left(d_{tf'}^{\rho_{f'}-1}\right)\right)\exp\left(X_{it}\beta_{f'}+\phi_{sf'}\right)} \\
&= \frac{\exp\left(X_{it}\beta_f+\phi_{sf}+\log\left(d_{tf}^{\rho_f-1}\right)\right)}{\sum_{f'=1}^{F}d_{tf'}\exp\left(X_{it}\beta_{f'}+\phi_{sf'}+\log\left(d_{tf'}^{\rho_{f'}-1}\right)\right)} \\
&= \frac{\exp\left(X_{it}\beta_f+\phi_{sf}+(\rho_f-1)\log\left(d_{tf}\right)\right)}{\sum_{f'=1}^{F}d_{tf'}\exp\left(X_{it}\beta_{f'}+\phi_{sf'}+(\rho_{f'}-1)\log\left(d_{tf'}\right)\right)}
\end{aligned}$$

which is equivalent to the choice probabilities one would obtain with the error structure given in Equation (9).

Expectation-Maximization (EM) Algorithm

This subsection describes the Expectation-Maximization (EM) Algorithm used to estimate student choice parameters in the presence of type specific unobserved heterogeneity in student preferences for fields. Let y_{irtf} indicate whether observation r of student i in semester t who is type s chooses a field f course section. Furthermore, let $\Theta = \{\beta_f, \phi_{sf}, \rho_f\}_{f=1}^{F}$ contain all student choice parameters. Then the likelihood of observing student i's panel of choices in the state where she is type s as a function of Θ is given by:

$$l_{is}(\Theta) = \prod_{t=1}^{T}\prod_{r=1}^{R_{it}} P_{istf}^{y_{irtf}}$$

Let π_s represent the unconditional probability that a student is type s. The full log-likelihood as a function of Θ would then be:

$$\mathscr{L}(\Theta) = \sum_{i=1}^{N} \ln\left(\sum_{s=1}^{S} \pi_s \left(\prod_{t=1}^{T}\prod_{r=1}^{R_{it}} P_{istf}^{y_{irtf}}\right)\right)$$

In theory, one can estimate student choice parameters using this full log-likelihood function; however, doing so would be computationally burdensome. Instead, define the augmented log-likelihood function

$$Q(\Theta) = \sum_{i=1}^{N}\sum_{s=1}^{S}\sum_{t=1}^{T}\sum_{r=1}^{R_{it}} q_{is} y_{irtf} \ln P_{istf} \tag{16}$$

where q_{is} represent the probability that student i is type s conditional on her observed choices given by:

$$q_{is}(\Theta) = \frac{\pi_s l_{is}}{\sum_{s'=1}^{S} \pi_{s'} l_{is'}} \tag{17}$$

The EM algorithm proceeds by iteratively maximizing Equation (16) to obtain estimates of Θ and using Equation (17) to infer conditional type probabilities given these estimates. Specifically, the algorithm proceeds as follows:

1. Begin with arbitrary guesses for Θ^0 and π_s^0
2. Evaluate Equation (17) at Θ^0 and π_s^0 to obtain q_{is}^0
3. Estimate Θ^1 by maxizing Equation (16) given q_{is}^0. Estimate π_s^1 with the sample average $\pi_s^1 = \frac{1}{N}\sum_{i=1}^{N} q_{is}^0$
4. Return to step 2 using Θ^1 and π_s^1 and iterate until successive estimates of Θ and π_s become arbitrarily close.

Solving for tax/subsidy costs

In this subsection, I describe my method for estimating the tax/subsidy costs reported in Column 6 of Table 7. The goal of this exercise is to solve for counterfactual costs of hiring adjunct instructors that come closest to inducing the observed university to offer student utility maximizing (SUM) adjunct instructed courses.

Let d_{tf}^{N} represent the number of adjunct instructed field f course sections offered in semester t and let E_t^N represent the residual share of the semester t instruction budget that is paid to adjunct instructors. The goal is then to solve for counterfactual costs $\tilde{c}_f$ that come closest to implying that SUM course offerings $\mathbf{d_t^{SUM}}$ satisfy the observed university's first order conditions given by Equation (4).

Rearranging Equation (4) yields:[37]

$$\mathbf{M}\tilde{\mathbf{c}} = \mathbf{E}$$

where

$$\underset{(F,F)}{\mathbf{M}_t} = \begin{bmatrix} -\frac{\partial V_t}{\partial d_{tF}} - \sum_{f'=1}^{F-1} \gamma_{f'} \left(\frac{\partial n_{tf'}}{\partial d_{tF}}\right) & & & \frac{\partial V_t}{\partial d_{t1}} + \sum_{f'=1}^{F-1} \gamma_{f'} \left(\frac{\partial n_{tf'}}{\partial d_{t1}}\right) \\ & \ddots & & \vdots \\ & & -\frac{\partial V_t}{\partial d_{tF}} - \sum_{f'=1}^{F-1} \gamma_{f'} \left(\frac{\partial n_{tf'}}{\partial d_{tF}}\right) & \frac{\partial V_t}{\partial d_{tF-1}} + \sum_{f'=1}^{F-1} \gamma_{f'} \left(\frac{\partial n_{tf'}}{\partial d_{tF-1}}\right) \\ d_{t1}^N & \cdots & \cdots & d_{tF}^N \end{bmatrix}$$

$$\underset{(F\times T,F)}{\mathbf{M}} = \begin{bmatrix} \mathbf{M}_1 & \cdots & \mathbf{M}_T \end{bmatrix}'$$

$$\underset{(F,1)}{\mathbf{E}_t} = \begin{bmatrix} \mathbf{0}_{1,F-1} & E_t^N \end{bmatrix}'$$

$$\underset{(F\times T,1)}{\mathbf{E}} = \begin{bmatrix} \mathbf{E}_1 & \cdots & \mathbf{E}_T \end{bmatrix}'$$

$$\underset{(F,1)}{\tilde{\mathbf{c}}} = \begin{bmatrix} \tilde{c}_1 & \cdots & \tilde{c}_F \end{bmatrix}'$$

This system of equations can then be inverted to derive the following expression for counterfactual costs of hiring adjunct instructors that come closest to inducing the observed university to offer SUM adjunct instructed courses.

$$\tilde{\mathbf{c}} = \mathbf{M}^{+}\mathbf{E}$$

[37]Note that I am also applying the empirical linear budget constraint and the counterfactual restriction that the university can only reallocate the portion of its budget paid to adjunct instructors.

where M^+ denotes the pseudo-inverse of M.

References

Ackerberg, Daniel A and Marc Rysman, "Unobserved product differentiation in discrete-choice models: Estimating price elasticities and welfare effects," *RAND Journal of Economics*, 2005, *36* (4), 771–789.

ADHE, "AHEIS Reference Manual for the Student Information System," Technical Report 2011.

Ahn, Thomas, Peter Arcidiacono, Amy Hopson, and James Thomas, "Equilibrium grade inflation with Implications for female interest in STEM majors," 2022.

Altonji, Joseph G, "The demand for and return to education when education outcomes are uncertain," *Journal of Labor Economics*, 1993, *11* (1, Part 1), 48–83.

_ **, Erica Blom, and Costas Meghir**, "Heterogeneity in human capital investments: High school curriculum, college major, and careers," *Annual Review of Economics*, 2012, *4*, 185–223.

_ **, Peter Arcidiacono, and Arnaud Maurel**, "The analysis of field choice in college and graduate school: Determinants and wage effects," in "Handbook of the Economics of Education," Vol. 5, Elsevier, 2016, pp. 305–396.

Andrews, Rodney and Kevin Stange, "Price Regulation, Price Discrimination, and Equality of Opportunity in Higher Education: Evidence from Texas," *American Economic Journal: Economic Policy*, 2019, *11* (4), 31–65.

Angrist, Joshua D and Victor Lavy, "Using Maimonides' rule to estimate the effect of class size on scholastic achievement," *The Quarterly journal of economics*, 1999, *114* (2), 533–575.

Arcidiacono, Peter, "Ability sorting and the returns to college major," *Journal of Econometrics*, 2004, *121* (1), 343–375.

_ , "Affirmative action in higher education: How do admission and financial aid rules affect future earnings?," *Econometrica*, 2005, *73* (5), 1477–1524.

— , **V Joseph Hotz, and Songman Kang**, "Modeling college major choices using elicited measures of expectations and counterfactuals," *Journal of Econometrics*, 2012, *166* (1), 3–16.

Bar, Talia, Vrinda Kadiyali, and Asaf Zussman, "Grade information and grade inflation: The Cornell experiment," *The Journal of Economic Perspectives*, 2009, *23* (3), 93–108.

Bayer, Patrick and Christopher Timmins, "Estimating Equilibrium Models Of Sorting Across Locations," *The Economic Journal*, 2007, *117* (518), 353–374.

Beffy, Magali, Denis Fougere, and Arnaud Maurel, "Choosing the field of study in postsecondary education: Do expected earnings matter?," *Review of Economics and Statistics*, 2012, *94* (1), 334–347.

Bell, Daniel, *The Coming of Post-Industrial Society: A Venture in Social Forecasting*, Basic Books, 1973.

Berger, Mark C, "Predicted future earnings and choice of college major," *ILR Review*, 1988, *41* (3), 418–429.

Berry, Steven, James Levinsohn, and Ariel Pakes, "Differentiated products demand systems from a combination of micro and macro data: The new car market," *Journal of political Economy*, 2004, *112* (1), 68–105.

Bettinger, Eric P and Bridget Terry Long, "Do faculty serve as role models? The impact of instructor gender on female students," *American Economic Review*, 2005, *95* (2), 152–157.

Bhattacharya, Debopam, Shin Kanaya, and Margaret Stevens, "Are university admissions academically fair?," *Review of Economics and Statistics*, 2017, *99* (3), 449–464.

Board, UCA Trustee, "Mission, Vision, and Core Values," 2011. Retrieved from: http://uca.edu/about/mission/.

Bordon, Paola and Chao Fu, "College-major choice to college-then-major choice," *The Review of Economic Studies*, 2015, *82* (4), 1247–1288.

Butcher, Kristin F, Patrick J McEwan, and Akila Weerapana, "The Effects of an Anti-Grade-Inflation Policy at Wellesley College," *The Journal of Economic Perspectives*, 2014, pp. 189–204.

Carrell, Scott E, Marianne E Page, and James E West, "Sex and science: How professor gender perpetuates the gender gap," *The Quarterly Journal of Economics*, 2010, *125* (3), 1101–1144.

Cellini, Stephanie Riegg, "Crowded colleges and college crowd-out: The impact of public subsidies on the two-year college market," *American Economic Journal: Economic Policy*, 2009, *1* (2), 1–30.

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

_ **and Claudia Goldin**, "Does federal student aid raise tuition? New evidence on for-profit colleges," *American Economic Journal: Economic Policy*, 2014, *6* (4), 174–206.

Conlon, Christopher T and Julie Holland Mortimer, "Demand estimation under incomplete product availability," *American Economic Journal: Microeconomics*, 2013, *5* (4), 1–30.

Cook, Emily E, "Competing Campuses: Equilibrium Prices, Admissions, and Undergraduate Programs in US Higher Education," 2021.

Cornwell, Christopher, David B Mustard, and Deepa J Sridhar, "The enrollment effects of merit-based financial aid: Evidence from Georgia's HOPE program," *Journal of Labor Economics*, 2006, *24* (4), 761–786.

Dempster, Arthur P, Nan M Laird, and Donald B Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society*, 1977, *Series B (methodological)*, 1–38.

Dinerstein, Michael F, Caroline M Hoxby, Jonathan Meer, and Pablo Villanueva, "Did the Fiscal Stimulus Work for Universities?," in "How the Financial Crisis and Great Recession Affected Higher Education," University of Chicago Press, 2014, pp. 263–320.

Eide, Eric and Geetha Waehrer, "The role of the option value of college attendance in college major choice," *Economics of Education review*, 1998, *17* (1), 73–82.

Epple, Dennis, Richard Romano, and Holger Sieg, "Admission, tuition, and financial aid policies in the market for higher education," *Econometrica*, 2006, *74* (4), 885–928.

_ **,** _ **, Sinan Sarpça, and Holger Sieg**, "The US market for higher education: A general equilibrium analysis of state and private colleges and public funding policies," 2013.

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

Gemici, Ahu and Matthew Wiswall, "Evolution Of Gender Differences In Post-Secondary Human Capital Investments: College Majors," *International Economic Review*, 2014, *55* (1), 23–56.

Gentzkow, Matthew, "Valuing new goods in a model with complementarity: Online newspapers," *American Economic Review*, 2007, *97* (3), 713–744.

Gibbs, Christa and Ryne Marksteiner, "Effects of Removing State Postsecondary Grant Aid Eligibility," 2016.

Ginder, Scott A, Janice E Kelly-Reid, and Farrah B Mann, "Graduation Rates for Selected Cohorts, 2007-12; Outcome Measures for Cohort Year 2007; Student Financial Aid, Academic Year 2014-15; and Admissions in Postsecondary Institutions, Fall 2015," Technical Report, National Center for Education Statistics 2017.

Griffith, Amanda L, "Faculty gender in the college classroom: Does it matter for achievement and major choice?," *Southern Economic Journal*, 2014, *81* (1), 211–231.

Heckman, James and Burton Singer, "A method for minimizing the impact of distributional assumptions in econometric models for duration data," *Econometrica*, 1984, *52*, 271–320.

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

_ , "The effects of class size on student achievement: New evidence from population variation," *The Quarterly Journal of Economics*, 2000, *115* (4), 1239–1285.

Jacob, Brian, Brian McCall, and Kevin M Stange, "College as Country Club: Do Colleges Cater to Students' Preferences for Consumption?," *Journal of Labor Economics*, 2018, *36* (2), 309–348.

Kokkelenberg, Edward C, Michael Dillon, and Sean M Christy, "The effects of class size on student grades at a public university," *Economics of Education Review*, 2008, *27* (2), 221–233.

Krueger, Alan B, "Economic considerations and class size," *The Economic Journal*, 2003, *113* (485), F34–F63.

Long, Bridget Terry, "How do financial aid policies affect colleges? The institutional impact of the Georgia HOPE scholarship," *Journal of Human Resources*, 2004, *39* (4), 1045–1066.

McLachlan, Geoffrey and David Peel, *Finite Mixture Models*, Wiley, 2004.

Nevo, Aviv, Daniel L Rubinfeld, and Mark McCabe, "Academic journal pricing and the demand of libraries," *American Economic Review*, 2005, *95* (2), 447–452.

Russell, Lauren, "Price Effects of Nonprofit College and University Mergers," *Review of Economics and Statistics*, 2021, *103* (1), 88–101.

Ryan, Camille L and Kurt Bauman, "Educational Attainment in the United State: 2015," Technical Report 2016.

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

Stinebrickner, Ralph and Todd Stinebrickner, "A major in science? Initial beliefs and final outcomes for college major and dropout," *Review of Economic Studies*, 2014, *81* (1), 426–472.

Thomas, James, "The signal quality of grades across academic fields," *Journal of Applied Econometrics*, 2019, *34* (4), 566–587.

Train, Kenneth E, *Discrete choice methods with simulation*, Cambridge university press, 2009.

Turner, Lesley J, "The Economic Incidence of Federal Student Grant Aid," 2017.

Turner, Sarah E and William G Bowen, "Choice of major: The changing (unchanging) gender gap," *ILR Review*, 1999, *52* (2), 289–313.

UCA, "Undergraduate Bulletin, 2006-2008," Technical Report 2006.

Wiswall, Matthew and Basit Zafar, "Determinants of college major choice: Identification using an information experiment," *Review of Economic Studies*, 2014, *82* (2), 791–824.

Table 1: University of Central Arkansas

Institutional Characteristics	
Undergraduates	9,887
Full-time faculty	547
Admission Rate	92%
Yield	44%
ACT 25th pctile	20
ACT 75th pctile	26
6 year graduation rate	45%
Student characteristics	
Full-time	84%
24 and under	90%
In-state	89%
Female	59%
White	66%
Black	18%
Hispanic	5%
Other race	11%

Source: National Center for Education Statistics. Fall, 2015. Yield is the percent of students who choose to enroll conditional on being offered admission. ACT scores are composite scores. Graduation rate is for students pursuing a Bachelor's degree.

Table 2: Field Characteristics at UCA

	STEM	Social Science	Humanities and Arts	Business and Occupational
Avg. intro courses per semester	33.7	65.8	49.5	25.2
Avg. intro sections per semester	204	250	160	87
Avg. intro enrollment per semester	5181	8161	4516	2123
Avg. intro enrollment per section	25.4	32.6	28.2	24.5
Intro section cost (25th pctile)	\$6,000	\$4,566	\$5,034	\$5,441
Intro section cost (Median)	\$8,811	\$6,088	\$6,734	\$7,140
Intro section cost (75th pctile)	\$10,880	\$8,366	\$8,865	\$11,708
Avg. ACT score	24.4	23.7	23.9	24.0
Avg. HS GPA	3.43	3.35	3.34	3.39
Share Female	57.5%	59.5%	58.1%	47.5%
Share Freshmen	43.8%	39.7%	40.3%	11.6%
Share Sophomores	27.9%	31.8%	33.8%	40.6%
Share Juniors	17.1%	18.3%	16.5%	35.0%
Share Seniors	11.2%	10.2%	9.4%	12.7%

Notes: Statistics are for introductory courses at the University of Central Arkansas. "Courses" are defined by a course number (e.g. Econ 101). "Sections" are defined by a course number, instructor and meeting time (e.g. Econ 101 taught by Prof. Jane Doe meeting MWF from 9 - 10:30AM). Section cost is the amount an instructor is implicitly paid to teach a course section. This depends on an instructor's salary, teaching load, and other responsibilities. Average student scores and demographic proportions treat every instance of a student choosing an introductory course as an observation and compute statistics conditional on the field of the introductory course.

Table 3: Course Offerings and Enrollment Shares

	Semester																	
STEM	F04	S05	F05	S06	F06	S07	F07	S08	F08	S09	F09	S10	F10	S11	F11	S12	F12	S13
Courses	34	30	33	33	34	32	33	33	33	34	33	35	34	34	34	36	35	37
Sections	194	175	221	179	233	207	248	211	221	201	219	199	212	184	197	181	206	188
Sections (%)	30%	29%	32%	29%	31%	29%	30%	29%	29%	29%	29%	29%	29%	28%	28%	28%	29%	28%
Enrollment (%)	27%	24%	28%	24%	26%	25%	27%	25%	26%	25%	26%	26%	27%	26%	26%	26%	27%	26%

	Semester																	
Social Science	F04	S05	F05	S06	F06	S07	F07	S08	F08	S09	F09	S10	F10	S11	F11	S12	F12	S13
Courses	59	64	65	60	64	64	65	66	66	67	68	65	66	69	69	69	70	68
Sections	227	220	250	230	275	261	294	267	268	261	274	237	249	236	249	227	248	233
Sections (%)	35%	36%	36%	37%	36%	37%	36%	37%	35%	37%	36%	34%	34%	36%	36%	35%	35%	35%
Enrollment (%)	42%	43%	41%	42%	41%	41%	40%	41%	40%	41%	41%	40%	40%	41%	41%	40%	40%	40%

	Semester																	
Humanities and Arts	F04	S05	F05	S06	F06	S07	F07	S08	F08	S09	F09	S10	F10	S11	F11	S12	F12	S13
Courses	41	43	44	45	48	46	50	49	52	50	50	55	53	51	53	52	53	56
Sections	134	135	142	140	165	153	175	166	177	156	180	171	172	156	169	160	168	166
Sections (%)	21%	22%	20%	22%	22%	22%	21%	23%	23%	22%	24%	25%	24%	24%	24%	25%	24%	25%
Enrollment (%)	21%	23%	21%	23%	22%	23%	22%	23%	23%	22%	23%	24%	23%	22%	23%	23%	23%	24%

	Semester																	
Business and Occupational	F04	S05	F05	S06	F06	S07	F07	S08	F08	S09	F09	S10	F10	S11	F11	S12	F12	S13
Courses	26	26	25	24	25	24	25	25	26	24	25	25	26	25	27	26	25	25
Sections	90	81	84	77	90	85	97	87	96	84	91	80	92	87	84	82	87	86
Sections (%)	14%	13%	12%	12%	12%	12%	12%	12%	13%	12%	12%	12%	13%	13%	12%	13%	12%	13%
Enrollment (%)	10%	11%	10%	11%	10%	11%	11%	11%	11%	12%	10%	10%	10%	11%	10%	11%	10%	11%

Notes: Statistics are for the University of Central Arkansas. FXX/SXX indicate fall/spring semester of 20XX. "Courses" are defined by a course number (e.g. Econ 101). "Sections" are defined by a course number, instructor and meeting time (e.g. Econ 101 taught by Prof. Jane Doe meeting MWF from 9 - 10:30AM).

Table 4: Student Course Choice Parameters

	STEM	Social Science	Humanities and Arts	Business and Occupational
Intercept	1.095**	2.186***	0.934*	*omitted*
	0.478	*0.494*	*0.485*	
ACT Z-Score	0.119***	0.019*	0.073***	*omitted*
	0.010	*0.010*	*0.011*	
Missing ACT	-0.261***	-0.212***	-0.273***	*omitted*
	0.017	*0.015*	*0.017*	
GPA Z-score	0.060***	-0.063***	-0.101***	*omitted*
	0.009	*0.009*	*0.009*	
Missing GPA	0.139***	0.162***	0.202***	*omitted*
	0.019	*0.018*	*0.020*	
Female	0.243***	0.366***	0.305***	*omitted*
	0.014	*0.013*	*0.014*	
Non-White	-0.035*	-0.086***	-0.128***	*omitted*
	0.020	*0.019*	*0.020*	
Missing Race	-0.266***	-0.145***	-0.129***	*omitted*
	0.016	*0.015*	*0.016*	
Sophomore	-1.766***	-1.500***	-1.461***	*omitted*
	0.020	*0.019*	*0.020*	
Junior	-2.162***	-1.962***	-2.099***	*omitted*
	0.021	*0.019*	*0.021*	
Senior	-1.649***	-1.630***	-1.754***	*omitted*
	0.026	*0.025*	*0.026*	
Unobs. Type 2	2.231***	2.163***	2.459***	*omitted*
(uncond. prob. = 0.63)	*0.016*	*0.016*	*0.016*	
Nesting Parameters	0.730***	0.555***	0.709***	0.800*
	0.055	*0.058*	*0.063*	*0.110*

Notes: Parametric block bootstrapped standard errors (300 iterations) are in italics. */**/*** indicate significantly different from omitted category (or from one for ρ_f) at 10%/5%/1% significance. ACT/GPA Z-scores are scores that have been rescaled to have mean 0 and standard deviation 1 in my observed sample of students. Unconditional probability of being unobserved type 1 (omitted) is 0.37 and unconditional probability of being unobserved type 2 is 0.63.

Table 5: Marginal Effects of Course Offerings on Field Enrollments

Field of Add. Section		Field of Enrollment			
		STEM	Soc Sci	Hum and Arts	Bus and Occ
	STEM	13.47***	-7.61***	-4.26***	-1.6***
		1.02	*0.58*	*0.32*	*0.12*
	Social Science	-4.72***	10.59***	-4.17***	-1.7***
		0.50	*1.10*	*0.44*	*0.17*
	Humanities and Arts	-5.3***	-8.37***	15.3***	-1.63***
		0.48	*0.75*	*1.37*	*0.14*
	Business and Occupational	-4.14***	-7.08***	-3.39***	14.62***
		0.55	*0.94*	*0.45*	*1.94*
	Exp. Enroll. of New Section	25.27***	32.53***	28.23***	24.37***
		0.15	*0.14*	*0.14*	*0.29*
	Own-Field Effect / Enroll.	53.3%***	32.6%***	54.2%***	60.0%***
		3.8	*3.4*	*4.8*	*8.6*

Notes: Parametric block bootstrapped standard errors (300 iterations) are in italics. Top section contains the marginal effects of offering additional sections in each field on expected enrollment in all fields. Rows represent fields where course sections are added, columns represent fields of enrollment. Bottom section contains the expected enrollment of a new section added to a field and the diagonal own-field enrollment effects relative to new section enrollments. *** indicate significantly different from zero at 1% significance.

Table 6: Relative Marginal Effects and Implied Preferences

	Average Marginal Effect on Total Utility	Cost of Adjunct Instructors	Average Marginal Effect on Total Utility per Dollar	Implied Preferences
	(1)	(2)	(3)	(4)
STEM	0.926	\$6387.60	0.781***	0.266***
	0.077	*92.61*	*0.065*	*0.084*
Social Science	0.904	\$4941.05	0.985	0.024
	0.086	*56.95*	*0.094*	*0.085*
Hum. and Arts	1	\$5387.07	1	0
		99.35		
Business and Occ.	0.979	\$6845.59	0.771**	0.307
	0.126	*126.96*	*0.099*	*0.189*

Notes: Parametric block bootstrapped standard errors (300 iterations) are in italics. Column 1 contains marginal effects of offering an additional course section in the specified field on total expected student utility. These are averages across academic semesters of effects relative to humanities and arts. Column 2 reports average costs of hiring adjunct instructors to teach one course section in each field. Column 3 divides marginal effects by the costs of hiring adjunct instructors. Once again, these are averages across semesters of effects per dollar relative to humanities and arts. Column 4 reports estimates of implied preference parameters γ_j with humanities and arts as the omitted field. Estimates quantify how much student utility the university is implicitly willing to sacrifice to move one student from a humanities and arts course to a course in the specified field.

Table 7: Student Utility Maximizing Course Offerings

	Observed			**Student Utility Maximizing (SUM)**		
	(1)	(2)	(3)	(4)	(5)	(6)
	Average Tenure-Track Courses	Average Adjunct Courses	Average Field Enrollment	Average Adjunct Courses	Average Field Enrollment	Tax/ Subsidy
STEM	125.28	78.94	5181.39	9.26***	3744.78***	15.6%***
				14.00	*320.56*	*5.4*
Soc Sci	132.33	118.00	8161.22	176.54	8991.03	-3.6%
				45.27	*817.78*	*3.0*
Hum and Arts	100.11	60.17	4515.94	128.90	5697.85	-1.9%
				42.26	*831.46*	*3.6*
Bus and Occ	29.94	56.72	2122.50	25.40**	1547.40**	20.3%
				14.29	*266.29*	*16.7*

Notes: Parametric block bootstrapped standard errors (300 iterations) are in italics. Columns 1-3 are the observed number of course sections taught by tenure-track instructors, the observed number of course sections taught by adjunct instructors, and observed field enrollments averaged across semesters. Column 4 reallocates the residual budget spent on adjunct instructors to maximize total student utility and column 5 reports estimated field enrollments under these student utility maximizing (SUM) offerings. In columns 4 and 5, **/*** indicates significantly different from observed values at 5%/1% significance. Column 6 reports how much the costs of hiring an adjunct instructor would need to change to induce the university to offer the SUM course sections reported in column 4. In column 6, *** indicates significantly different from zero at 1% significance.

Table 8: Adjunct Instructed Course Offerings in Counterfactual Scenarios

		STEM	Social Science	Humanities and Arts	Business and Occupational
(1)	Baseline (predicted)	81.91	118.54	56.92	56.12
		0.59	*0.30*	*0.54*	*0.19*
(2)	Reduce cost of STEM	34.8%***	-8.7%***	-18.8%***	-7.5%***
	adjunct by 5%	*6.9*	*2.0*	*6.9*	*3.0*
(3)	Increase all SAT scores	12.7%***	-4.8%***	-8.7%***	-2.0%**
	and GPA by 1/10 of a std dev	*2.6*	*0.9*	*3.2*	*1.0*
(4)	1968 gender ratio	4.7%**	-8.5%***	-6.2%*	13.1%**
		2.2	*1.6*	*3.2*	*5.4*

Notes: Parametric block bootstrapped standard errors (300 iterations) are in italics. Row 1 is the average number of course sections taught by adjunct instructors predicted by the estimated model in the observed state. Rows 2-4 are the percent change in the average number of course sections taught by adjunct instructors in counterfactual states relative to baseline. In rows 2-4, */**/*** indicates significantly different from zero at 10%/5%/1% significance.

Table 9: Field Enrollments in Counterfactual Scenarios

		STEM	Social Science	Humanities and Arts	Business and Occupational
(1)	Baseline (predicted)	5219.30	8169.82	4474.81	2117.12
		30.72	*34.09*	*24.16*	*24.88*
(2)	Reduce cost of STEM	8.9%***	-2.3%***	-4.7%***	-3.1%
	adjunct by 5%	*2.4*	*0.7*	*1.7*	*2.1*
(3)	Increase all SAT scores and	1.2%***	-0.8%***	1.3%*	2.0
	GPA by 1/10 of a std dev (PE)	*0.2*	*0.1*	*0.8*	*4.2*
(4)	Increase all SAT scores and	5.1%***	-1.7%***	-2.4%***	-0.8%
	GPA by 1/10 of a std dev (GE)	*0.9*	*0.3*	*0.7*	*0.7*
(5)	1968 gender ratio (PE)	0.5%**	-1.4%***	1.6%**	5.6%
		0.2	*0.1*	*0.8*	*4.4*
(6)	1968 gender ratio (GE)	2.5%***	-3.2%***	-1.3%	8.9%
		0.8	*0.6*	*0.8*	*5.6*

Notes: Parametric block bootstrapped standard errors (300 iterations) are in italics. Row 1 are average field enrollments predicted by the estimated model in the observed state. Rows 2-6 are the percent changes in the average field enrollments in counterfactual states relative to baseline. In rows 2-6, ***/**/* indicates significantly different from zero at 1%/5%/10% significance. (PE) indicates that student characteristics are changed but course offerings are held fixed. (GE) indicates that course offerings change in response to counterfactual student characteristics as reported in Table 8.