Measuring implied field preferences from a university's course offerings

James Thomas*

December 6, 2018

Abstract

Universities are important social institutions yet little is known about how they make institutional decisions. This paper develops a new framework for empirically analyzing course offerings at a sample university. The framework is based on the idea that course offerings directly affect student utilities and the probabilities that students choose courses in a given field. As such, administrators deciding which courses to offer are always implicitly trading off the number of students choosing courses in each field and total student utility. By measuring the marginal effects of offering additional courses in each field on field enrollments and total utility, one can quantify the tradeoffs between total utility and field enrollments which are implied by observed course offerings. In my empirical application, I find that a marginal dollar of spending on social science courses produces 2.5 times as much student utility as a marginal dollar of spending on business or occupational courses. From this, I conclude the university is implicitly sacrificing student utility to draw students out of social science courses and into business or occupational courses. Under stronger assumptions, these implied tradeoffs can be treated as parameters of a university's objective function in a two-sided model of a university and students. With this framework, I conduct counterfactual analyses which predict how course offerings and field enrollments will change in various alternative scenarios.

*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, Richard Mansfield, and Kevin Stange and from presentations at Duke University, University of South Carolina, Yale University, the 2016 SEA Annual Meeting, the 2017 SOLE Annual Meeting, the 2018 AEA Annual Meeting, the US Department of Treasury, the Federal Trade Commission, and the Federal Communications Commission. All remaining errors are my own.

1 Introduction

Universities are very important social institutions—they allow students to acquire human capital which is valuable to them individually and to society more broadly. But universities are not passive parties in the production of human capital, they are active entities which choose their inputs to maximize payoffs subject to constraints. As recent literature shows, choices regarding tuition, institutional spending, and instructor characteristics can play important roles in the production of human capital.[1]

Another important choice which has not been previously studied is how universities allocate their budgets for instruction across different academic fields. This choice directly affects students as it determines the set of courses available to students and thus may influence the courses students choose and the utility they derive from these choices. To contribute to our understanding of how universities make decisions and the implications of these decisions for students, this paper develops a new framework for analyzing course offerings at a university.

Universities are complicated entities with many unobserved constraints; moreover, as with other public or non-profit entities, the structure of a university's objective is unclear.[2] For this reason, empirical studies which aim to understand university behaviors typically focus on estimating specific effects using reduced form frameworks which limit assumptions about university constraints and objectives.[3] While these studies provide important insights, there is additional value in developing full empirical models of universities as these would provide a richer understanding of university behaviors and would allow for counterfactual policy analyses which incorporate university responses into predictions.

To advance towards an empirical model of university behavior while still acknowledging the limits of relying on strong functional form assumptions, this paper develops a framework which can be used both to draw interesting conclusions about the implications of university choices under relatively lenient assumptions and to conduct model based counterfactual analyses under stronger assumptions. The central principle of the framework is that offering

[1]For studies of the effects of tuition on student outcomes see Deming and Walters (2017); Hemelt and Marcotte (2011); Kane (1995). For studies of the effects of aggregate institutional spending on student outcomes see Bound and Turner (2007); Bound et al. (2010, 2012); Dynarski (2008); Turner (2004). For studies of the effects of instructor characteristics on student outcomes see Bettinger and Long (2005, 2010); De Vlieger et al. (2017); Figlio et al. (2015).

[2]For profit universities are probably profit maximizing but they represent a relatively small share of the higher education market comprising only 3% of total enrollment (Turner, 2012). For studies on the objectives of non-profit entities, see Glaeser (2003); Sloan (2000).

[3]Notable exceptions include Epple et al. (2006), Epple et al. (2013), and Fu (2014). These papers develop general equilibrium models of the higher education market which explain observed variation in tuition, admission rates, student characteristics, and other measures across schools. My paper uses microdata to analyze the choices of a specific university.

more courses in a field costs the university money but adds variety of choices within that field. This additional variety increases expected student utility and increases expected enrollment in that field by making the field as a whole relatively more attractive. As such, any reallocation of resources across fields changes the expected number of students choosing courses in each field and total expected student utility.

If one can estimate the marginal effects of offering additional courses in each field on field enrollments and total utility as well as the marginal costs of offerings additional courses, then one can construct the tradeoffs between student utility and field enrollments which are implied by observed course offerings. For example, if estimates of marginal effects and costs imply that reallocating resources from STEM to humanities would increase student utility and humanities enrollment but decrease STEM enrollment, then observed course offerings are implicitly sacrificing student utility to draw students out of humanities and arts courses and into STEM courses.

To estimate the crucial marginal effects of offering additional courses on field enrollments and total student utility, I propose using a nested logit course choice model with panel data of course choices under different sets of offered courses. To estimate the marginal costs of offering additional courses, I propose using a simple cost regression framework with data on course costs and characteristics. In a nested logit course choice model with panel data, marginal effects of offering additional courses are identified by the relationship between the relative number of courses offered in a field and the share of students choosing courses in that field across semesters. If semesters with relatively more courses in a field also have a higher share of students in that field, this suggests adding courses has large marginal effects on student utility and field enrollments. Conversely, if there is little relationship between relative course offerings and enrollment shares, then adding courses has small marginal effects on student utility and field enrollments.

The main identifying assumption underlying this conclusion is that the university is not changing its course offerings in response to unobserved student preferences for fields. While there are certainly threats to this assumption, I argue that in my empirical setting, the patterns of course offerings and the availability of rich student observed covariates imply that the assumption is satisfied.

I use my framework to analyze the introductory course offerings of the University of Central Arkansas in Fall and Spring academic semesters of academic years 2004-05 through 2009-10. While my analysis only applies to one university, I argue University of Central Arkansas (UCA) is a particularly interesting subject for two reasons: First, UCA is a large public four year university with a 45% six year graduation rate.[4] Because the median young

[4]Source: National Center for Education Statistics

American completes some college but does not obtain a degree, and because 45.07% of all full-time equivalent higher education enrollment is at public four year institutions, a public four year university with a 45% graduation rate is somewhat representative of the post secondary education experience of a median American.[5] Second, UCA is a teaching focused university which makes an analysis of its course offerings particularly interesting.[6]

In the first stage of my analysis, I find that the return on marginal dollars of spending in terms of total expected student utility varies widely across academic fields. A marginal dollar spent offering more introductory social science courses produces 2.5 times as much student utility as a marginal dollar spent offering additional introductory business or occupational courses. This implies that observed course offerings implicitly sacrifice significant student utility to increase enrollment in introductory business and occupational courses. These conclusions require reliable estimates of marginal effects and marginal costs but do not rely on the full structural model of the university's course offerings and student course choices.

In the second stage of my analysis, I treat these implied tradeoffs as parameters of the university's objective function. In this analysis, I find that a utility maximizing university which is only allowed to reallocate the share of its instruction budget which is paid to instructors on single-semester contracts would offer four times as many introductory social science courses taught by instructors on single-semester contracts.[7] In fact, one would have to increase the cost of hiring single-semester social science instructors by 45.6% and decrease the cost of hiring single-semester business and occupational instructors by 42.0% to induce this utility maximizing university to offer the courses observed in the data.

Finally, to illustrate the value of the two-sided model, and for higher education models which incorporate supply-side responses more generally, I perform a number of simulations which predict course offerings and student outcomes in counterfactual scenarios where the composition of students or the costs of hiring single-semester instructors are changed. Most notably, I find that a 5% reduction in the cost of hiring a single-semester STEM instructor

[5]Estimates from the 2015 Current Population Survey show that of individuals 25-34 years old residing in the United States, 9.5% did not complete high school, 25.5% completed high school only, 18.5% completed some college but did not complete a degree, 10.4% completed an associates degree only, 25.2% completed a bachelor's degree only, and 10.9% completed an advanced degree (Ryan and Bauman, 2016). Full time equivalent enrollment statistics are author's calculation using IPEDS for academic year 2016-2017.

[6]UCA's teaching focus is apparent 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)

[7]Instructors on single-semester contracts are those whose contracts need to be renewed every semester. 11.4% of all courses are taught by instructors on single-semester contracts.

would lead to a 49.1% increase in the number of introductory courses taught by single-semester STEM instructors and a 6.2% increase in overall enrollment in introductory STEM courses.

This paper relates to a growing literature on the supply side of higher education which analyzes the role of universities in education production.[8] One branch of this literature focuses on estimating the effects of university choices and inputs on student outcomes. This includes studies of "cohort crowding" effects which estimate the effects of aggregate institutional spending on student outcomes (Bound and Turner, 2007; Bound et al., 2010, 2012; Dynarski, 2008; Turner, 2004) and complementary work which estimates the effects of university tuition on student outcomes (Deming and Walters, 2017; Hemelt and Marcotte, 2011; Kane, 1995). Other studies in this branch of supply side higher education literature estimate the effects of instructor characteristics on student outcomes (Bettinger and Long, 2005, 2010; De Vlieger et al., 2017; Figlio et al., 2015). A second branch of this literature aims to form a better understanding of how universities make decisions. This includes studies which develop general equilibrium models of competition in the higher education market (Epple et al., 2006, 2013; Fu, 2014) as well as tests of the "Bennett hypothesis" which predicts that universities will respond to increases in government student aid by increasing tuition (Gibbs and Marksteiner, 2016; Cellini and Goldin, 2014; Long, 2004; Singell and Stone, 2007; Turner, 2017).

To my knowledge, the present paper is the first study to analyze the effects of course offerings on students. This contributes to the first branch of literature which aims to understand the role of supply-side inputs on student outcomes. Furthermore, to my knowledge, this is also the first paper to estimate a two-sided model of university and student choices using micro-level data. This contributes to the second branch of literature by advancing our understanding of how universities make decisions and allowing for counterfactual policy analyses which incorporate university responses into predictions.

The remainder of the paper proceeds as follows: Section 2 introduces my framework for analyzing course offerings at a university, Section 3 describes the data and discusses the empirical specifications used for estimation, Section 4 discusses estimates of implied tradeoffs which can be interpreted under more lenient assumptions, Section 5 discusses additional results and counterfactual predictions which rely on stronger assumptions, Section 6 concludes.

[8]Notable contributions not mentioned in the body include but are not limited to: Andrews and Stange (2016); Bhattacharya et al. (2017); Carrell and West (2010); Cellini (2009, 2010); Dinerstein et al. (2014); Hoffmann and Oreopoulos (2009); Hoxby (1997); Jacob et al. (2015); Pope and Pope (2009, 2014); Tabakovic and Wollmann (2016)

2 General Framework

In this section, I introduce a general framework for analyzing course offerings at a university. The framework views a student's decision of which courses to choose as a static discrete choice problem in which the set of available courses influences student choices and utilities. The main idea is then to compare the marginal effects per dollar of offering additional courses in one field on student choices and utilities to the same marginal effects per dollar for other fields. Differences in these marginal effects per dollar across fields reveal an implicit willingness to sacrifice student utility to increase enrollment in certain fields.

I propose estimating these crucial marginal effects using a nested logit course choice model with panel data of course choices under different sets of offered courses. Under the assumption that the university is not changing its course offerings in response to unobserved student preferences for fields, the nested logit model identifies marginal effects from the empirical relationship between the relative number of courses offered in a field and the share of students choosing courses in that field across semesters.

Under relatively lenient assumptions, one can interpret estimates of the implied tradeoffs between student utility and field enrollments as a measure of the misalignment between student preferences and observed course offerings. Alternatively, under stronger assumptions, one can treat these implicit tradeoffs as structural preference parameters in a two-sided model of a university deciding which courses to offer and students choosing courses from the set of available alternatives.

2.1 Student choices

In this framework, the goal of the student choice model is to yield credible estimates of the marginal effects of offering additional introductory courses on student course choices and choice utility. To that end, I include features which allow marginal effects to better reflect empirical variation but abstract from complex strategic behaviors of students.

To begin, let $i \in [1, N]$ index observations of students choosing introductory courses, let $t \in [1, T]$ index academic semesters, let $j \in [1, J]$ index introductory courses, and let $f \in [1, F]$ index academic fields.[9]

Assume that student observation i's stochastic utility from choosing introductory course j belonging to field f can be additively separated into a field-specific deterministic component

[9] For simplicity, I treat choices of multiple courses in the same semester by the same student as independent observations. For a course choice model which allows students to choose multiple courses at once, see Ahn et al. (2017). In the empirical application, fields are STEM, social science, humanities and arts, and business and occupational. See Appendix A for field definitions.

and a course-specific stochastic component as follows:

$$U_{itj} = \upsilon\left(X_{it}, \beta_f\right) + \epsilon_{itj} \tag{1}$$

where X_{it} are observed student characteristics, β_f are utility parameters, $\upsilon\left(\cdot\right)$ is a smooth function, and ϵ_{itj} are stochastic preference shocks.

An important restriction in Equation (1) is that the deterministic component of utility $\upsilon\left(X_{it}, \beta_f\right)$ does not vary within field f. This restriction implies that marginal effects of offering additional introductory courses in field f on expected student outcomes are well-defined without additional assumptions. While one can generalize the framework to one in which deterministic utilities vary within fields by making assumptions about which courses within a field are marginal, I abstract from this for simplicity and transparency.[10]

I assume stochastic preference shocks ϵ_{ijt} are drawn from a Type 1 Extreme Value distribution with a nesting structure in which nests are defined by academic fields. This implies stochastic preference shocks can be additively decomposed into a field specific component ψ_{ift} and an idiosyncratic course specific component η_{ijt} scaled by a constant λ:

$$\epsilon_{ijt} = \psi_{ift} + \lambda\eta_{ijt} \tag{2}$$

where η_{ijt} are iid draws from a Type 1 Extreme Value distribution, ψ_{ift} and η_{ijt} are independent, and ψ_{ift} is drawn from a conjugate distribution derived in Cardell (1997). I will show that in a panel data setting, this nesting structure implies that marginal effects of offering additional introductory courses on total student utility and field enrollments are identified by the empirical relationship between the relative number of courses offered in a field and the share of students choosing courses in that field across semesters.

With this structure, the probability that student i chooses one specific introductory course in field f in semester t is given by:

$$P_{itf} = \frac{\exp\left(\frac{\upsilon\left(X_{it},\beta_f\right)}{\rho_f}\right)\left[\sum_{j\in f}\exp\left(\frac{\upsilon\left(X_{it},\beta_f\right)}{\rho_f}\right)\right]^{\rho-1}}{\sum_{f'=1}^{F}\left[\sum_{j\in f'}\exp\left(\frac{\upsilon\left(X_{it},\beta_{f'}\right)}{\rho_{f'}}\right)\right]^{\rho_{f'}}} \tag{3}$$

$\rho_f \in (0,1]$ is a nesting parameter which measures the degree of independence in unobserved preferences ϵ_{ijt} for courses within field f. When $\rho_f = 1$, variance in ψ_{ift} is zero so that

[10] One important way in which deterministic utilities may differ within fields is in observable instructor characteristics. See Appendix B for a model that treats instructor salary—the most salient instructor characteristic— as a university choice variable that also influences deterministic utility and for empirical evidence that instructor salary has minor effects on student utility at the university I study.

unobserved preferences are iid draws from a Type 1 Extreme Value distribution and choice probabilities are equivalent to those in multinomial logit. When $\rho_f \to 0$, the scalar λ approaches zero so that unobserved preferences are equal for all courses within field f (Train, 2009). Allowing ρ_f to vary across fields implies that the extent to which courses are similar or dissimilar within fields is allowed to vary across fields. This is an important mechanism for capturing heterogeneous marginal effects of offering additional courses across fields.

With this structure, choice probabilities simplify to:

$$P_{itf} = \frac{d_{tf}^{\rho_f - 1} \exp\left(\upsilon\left(X_{it}, \beta_f\right)\right)}{\sum_{f'=1}^{F} d_{tf'}^{\rho_f} \exp\left(\upsilon\left(X_{it}, \beta_{f'}\right)\right)} \tag{4}$$

where d_{tf} represents the number of introductory courses in field f which are offered in semester t.

Before proceeding, note that the general additively separable structure in Equation (1) nests dynamic discrete choice structures which are commonly used in models of student choice as long as the future value term depends only on the field of a course and does not vary across semesters.[11] Specifically, one could parameterize deterministic utility as:

$$\upsilon\left(X_{it}, \beta_f\right) = u\left(X_{it}, \beta_f^1\right) + \delta\mathbb{E}\left[V' \mid X_{it}, \beta_f^2\right] \tag{5}$$

where $u\left(X_{it}, \beta_f^1\right)$ represents the flow utility associated with an introductory course in field f and $\delta\mathbb{E}\left[V' \mid X_{it}, \beta_f^2\right]$ represents the discounted expected next period value associated with choosing an introductory course in field f this period. The expected next period value of choosing an introductory course in field f could reflect the option value of being able to take advanced courses in field f in the future, the future labor market value of coursework in field f, or any other future return associated with introductory coursework in field f. Since my goal is only to obtain estimates of the marginal effects of offering additional introductory courses on field enrollments and total student utility, my empirical specification will be a simple static structure which identifies these marginal effects from empirical variation in a clear and transparent manner. However, if desired, the general framework can accommodate richer models of student choice.

2.2 Student outcomes

The goal of this framework is to estimate the tradeoffs between total expected student utility and expected field enrollments implied by introductory course offerings. To construct these

[11] For papers using dynamic discrete choice models of student choice, see Arcidiacono (2004, 2005); Bordon and Fu (2015); Stinebrickner and Stinebrickner (2014a).

important outcomes as a function of course offerings, let $\mathbf{d}_t = \begin{bmatrix} d_{t1} & \cdots & d_{tF} \end{bmatrix}$ represent a vector containing the offerings in each field in semester t. Total expected student utility in semester t as a function of offered courses $\mathbf{d}_t$ is then given by:

$$\begin{aligned} V_t\left(\mathbf{d}_t\right) &= \sum_{i=1}^{N} \mathbb{E}\left[\max\left\{U_{ijt}\right\} \mid \mathbf{d}_t\right] \\ &= \sum_{i=1}^{N} \left\{\log\left(\sum_{f=1}^{F} d_{tf}^{\rho_f} \exp\left(\upsilon\left(X_{it}, \beta_f\right)\right)\right) + c\right\} \end{aligned} \tag{6}$$

where $c \approx 0.5772$ is the Euler-Mascheroni constant.

Furthermore, the expected number of students choosing courses in field f in semester t as a function of offered courses $\mathbf{d}_t$ is given by:

$$\begin{aligned} n_{tf}\left(\mathbf{d}_t\right) &= \sum_{i=1}^{N} d_{tf} P_{itf} \\ &= \sum_{i=1}^{N} \left[\frac{d_{tf}^{\rho_f} \exp\left(\upsilon\left(X_{it}, \beta_f\right)\right)}{\sum_{f'=1}^{F} d_{tf'}^{\rho_{f'}} \exp\left(\upsilon\left(X_{it}, \beta_{f'}\right)\right)}\right] \end{aligned} \tag{7}$$

These expressions illustrate the close relationship between the outcomes V_t and n_{tf} and course offerings $\mathbf{d}_t$.

As I will show formally in Subsection 2.6, it is not these outcome formulas *per se* which are useful for measuring implied tradeoffs; rather, it is the marginal effects of offering additional introductory courses in each field on these outcomes. These marginal effects are given by:[12]

$$\frac{\partial V_t\left(\mathbf{d}_t\right)}{\partial d_{tf}} = \sum_{i=1}^{N} \rho_f P_{itf} \tag{8}$$

$$\frac{\partial n_{tf'}\left(\mathbf{d}_t\right)}{\partial d_{tf}} = \begin{cases} \sum_{i=1}^{N} \rho_f P_{itf}\left(1 - d_{tf} P_{itf}\right) & f' = f \\ -\sum_{i=1}^{N} \rho_f d_{tf'} P_{itf} P_{itf'} & f' \neq f \end{cases} \tag{9}$$

These formulas illustrate the important roles of the nesting parameters ρ_f in determining the marginal effects of offering additional introductory courses on outcomes. Equation (8) shows that marginal effects of offering additional courses in field f on total expected utility are increasing in ρ_f. This makes sense because larger values for ρ_f imply more independence in unobserved preferences for courses within field f. This greater independence

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

means that additional courses provide more valuable variety.

Similarly, Equation (9) shows that larger values for ρ_f yield more positive own-field marginal effects on enrollment and more negative cross-field effects on enrollment.[13] Once again, this makes sense because greater independence implies that additional courses are less similar to other courses within the same field and thus will induce more students to switch fields in expectation.

2.3 Identification of nesting parameters ρ_f

In Subsection 2.6, I will show that the marginal effects of offering additional introductory courses in each field defined in the previous subsection play a crucial role in measuring the tradeoffs between student utility and field enrollments implied by offered courses. Given the central role of these marginal effects in driving the main conclusions of this paper, it is important to understand how these effects are identified from the data.

Equations (8) and (9) show that in this framework, marginal effects depend on choice probabilities P_{itf}, course offerings d_{tf}, and nesting parameters ρ_f. Choice probabilities are conditional moments and are thus non-parametrically identified from the data. Furthermore, course offerings are directly observed. In this subsection, I show that nesting parameters are identified by the empirical relationship between the relative number of courses offered in a field and the share of students choosing courses in that field across semesters.

To show identification of ρ_f, choose a sub-population of students with observed characteristics $X_{it} = X$ and restrict to two academic semesters t_1 and t_2 and two academic fields f_1 and f_2 which are chosen so that $d_{t_1 f_2} = d_{t_2 f_2}$ but $d_{t_1 f_1} \neq d_{t_2 f_1}$.

Let Φ_{tf} denote the probability that one of these students chooses any course in field f in semester t. These probabilities are given by:

$$\Phi_{tf} = \frac{d_{tf}^{\rho_f} \exp\left(X\beta_f\right)}{\sum_{f'=1}^{F} d_{tf'}^{\rho_{f'}} \exp\left(X\beta_{f'}\right)} \tag{10}$$

and the natural logarithms of these probabilities are:

$$\ln\left(\Phi_{tf}\right) = \rho_f \ln\left(d_{tf}\right) + X\beta_f - \ln\left[\sum_{f'=1}^{F} d_{tf'}^{\rho_{f'}} \exp\left(X\beta_{f'}\right)\right] \tag{11}$$

The difference in log probabilities across the two academic fields within semester t is

[13] Because $d_{tf}P_{itf}$ is the probability that student i chooses any course in field f and is thus less than one, the own-field effect on enrollment is always positive.

then given by:

$$\ln\left(\frac{\Phi_{tf_1}}{\Phi_{tf_2}}\right) = \rho_{f_1}\ln\left(d_{tf_1}\right) - \rho_{f_2}\ln\left(d_{tf_2}\right) + \left(X\beta_{f_1} - X\beta_{f_2}\right) \tag{12}$$

Furthermore, the difference in this difference across the two academic semesters is given by:

$$\begin{aligned}
\ln\left(\frac{\Phi_{t_1f_1}}{\Phi_{t_1f_2}}\right) - \ln\left(\frac{\Phi_{t_2f_1}}{\Phi_{t_2f_2}}\right) &= \rho_{f_1}\left[\ln\left(d_{t_1f_1}\right) - \ln\left(d_{t_2f_1}\right)\right] - \rho_{f_2}\left[\ln\left(d_{t_1f_2}\right) - \ln\left(d_{t_2f_2}\right)\right] && (13)\\
&= \rho_{f_1}\left[\ln\left(d_{t_1f_1}\right) - \ln\left(d_{t_2f_1}\right)\right] && (14)
\end{aligned}$$

where the second equality holds because $d_{t_1f_2} = d_{t_2f_2}$.

Rearranging yields:

$$\rho_{f_1} = \frac{\ln\left(\frac{\Phi_{t_1f_1}}{\Phi_{t_2f_1}}\right) - \ln\left(\frac{\Phi_{t_1f_2}}{\Phi_{t_2f_2}}\right)}{\ln\left(d_{t_1f_1}\right) - \ln\left(d_{t_2f_1}\right)} \tag{15}$$

This illustrates that ρ_f is identified by the empirical relationship between the relative number of courses offered in each field and the relative probability of choosing any course in that field. For example, if $d_{t_1f_1} > d_{t_2f_1}$ then ρ_{f_1} will be close to one if field choice probabilities in field f_1 increase significantly more than field choice probabilities in field f_2 and will be close to zero if field choice probabilities in field f_1 do not increase relative to field choice probabilities in field f_2. This makes sense because larger values for ρ_{f_1} imply more independence in unobserved preferences within field f_1. If there is more independence in unobserved preferences, then offering additional courses in field f_1 provides attractive variety which induces more students to choose courses in this field. Conversely, if unobserved preferences are largely determined by field then offering additional courses in field f_1 does not add variety and will not induce more students to choose courses in this field.

Another way to see how ρ_f is identified from empirical variation is to note that this nested logit choice model yields the same choice probabilities—and is thus equivalent to—an Ackerberg and Rysman (2005) crowding framework. Specifically, assume utility is defined as in Equation (1) but that stochastic preference shocks are given by:

$$\epsilon_{ijt}^{AR} = \delta_f \log\left(d_{tf}\right) + \eta_{ijt} \tag{16}$$

where η_{ijt} are independent draws from a Type 1 Extreme Value distribution and δ_f are parameters to be estimated.

In this setting, δ_f measures field specific "crowding" of the unobserved characteristic space. If δ_f is zero, then the number of options available in field f does not change the unobserved desirability of new courses. However, if δ_f is significantly negative, then new courses in field f will provide less option value when there are already many options available in that field. In estimation, $\log(d_{tf})$ is simply included as a time-varying field characteristic implying that δ_f is identified from the relationship between course offerings and field choice probabilities across semesters (Ackerberg and Rysman, 2005). With this structure, choice probabilities are given by:

$$P_{itf}^{AR} = \frac{\exp\left(\upsilon\left(X_{it}, \beta_f\right) + \delta_f \log\left(d_{tf}\right)\right)}{\sum_{f'=1}^{F} d_{tf'} \exp\left(\upsilon\left(X_{it}, \beta_{f'}\right) + \delta_{f'} \log\left(d_{tf'}\right)\right)} \tag{17}$$

It is straightforward to show that this is equivalent to the expression in Equation (4) when $\delta_f = \rho_f - 1$ implying that the variation which identifies the crowding parameters δ_f in an Ackerberg and Rysman (2005) framework is equivalent to the variation which identifies the (shifted) nesting parameters $\rho_f - 1$ in my nested logit framework.

The Ackerberg and Rysman (2005) representation of this framework also makes it easier to see that the main identifying assumption necessary to recover δ_f (or equivalently, $\rho_f - 1$) is that course offerings d_{tf} must be independent of preference shocks η_{ijt}. In words, this assumption means the university cannot consider unobserved student preferences when deciding how many 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.[14] 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_{ijt} thus any response of the university to these trends will cause misspecification.[15]

Both of these scenarios suggest there could be positive correlation between course offerings d_{tf} and preference shocks η_{ijt}. Because this is a non-linear model, one cannot rigorously sign biases in parameter estimates by signing correlation between unobserved shocks and endogenous variables. With this caveat, the intuition of bias signing in linear models may still be useful when interpreted with sufficient caution. In an analogous linear

[14] 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).

[15] 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.

model, positive correlation between d_{tf} and η_{ijt} (and a negative crowding parameter δ_f) would imply that estimates of δ_f are biased towards zero (and estimates of ρ_f are biased towards one). Intuitively, if relatively more courses are offered in field f in semesters where students are unobservably more interested in field f, the model will conclude that the additional offerings attracted the additional enrollment. The model will then infer that the course offerings provided meaningful variety and thus must have largely independent unobserved characteristics (ρ_f close to one / δ_f close to zero).

This suggests one should be concerned that positive correlation between course offerings and preference shocks could be biasing estimates of nesting parameters ρ_f towards one leading to exaggerated estimates of the effects of offering additional courses on field enrollments and total student utility. In my empirical setting, I include detailed observed student characteristics to minimize the possibility that the university is choosing course offerings in response to student characteristics which are observed by the university but not included in my model. Furthermore, I argue that trends in enrollments and courses offerings across semesters at this university do not suggest general trends in preferences which the university is responding to. Finally, because my method for measuring implied tradeoffs between student utility and field enrollments uses the marginal effects of offering additional courses in one field relative to others, bias in estimates of ρ_f will only bias estimates of implied tradeoffs insofar as these biases are stronger in some fields relative to others.

2.4 University's course offerings

In previous subsections, I introduced a framework for measuring the marginal effects of offering additional courses in different academic fields on total expected student utility and expected field enrollments. In this section, I show how these marginal effects can be used to measure the local tradeoffs between student utility and field enrollments which are implied by observed course offerings. Under relatively lenient assumptions, one can interpret these implicit tradeoffs as measures of the misalignment between student preferences and observed course offerings. Under stronger assumptions, one can treat these implicit tradeoffs as structural preference parameters in a two-sided model of a university deciding which courses to offer and students choosing courses from the set of available alternatives.

To begin, I assume 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:

$$C\left(\mathbf{d_t},\, \psi\right) \leq E_t \tag{18}$$

where E_t is a semester specific endowment, $C\left(\cdot\right)$ is a smooth function, and ψ are parameters to be estimated.[16]

To formalize the notion of implied tradeoffs between student utility and field enrollments, consider the following function representing the university's "payoff" to offering courses $\mathbf{d_t}$:

$$\Pi_t(\mathbf{d_t}) = \theta V_t(\mathbf{d}_t) + \sum_{f=1}^{F} \gamma_f n_{tf}(\mathbf{d}_t) \tag{19}$$

Without loss of generality, I normalize $\theta = 1$ and $\gamma_F = 0$.[17] With this structure and normalizations, the university is indifferent between course offerings which yield the following two outcomes:

	1	2
V_t	V	$V - \gamma_f$
n_{tf}	n_1	$n_1 + 1$
n_{tF}	n_2	$n_2 - 1$

As such, γ_f measures the amount of expected student utility which the university is implicitly willing to sacrifice to draw one student out of a field F courses and into a field f course in expectation.

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

$$\mathbf{d}_\mathbf{t}^\star = \text{argmax}_{\mathbf{d_t}} \left\{ V_t(\mathbf{d}_t) + \sum_{f=1}^{F-1} \gamma_f n_{tf}(\mathbf{d}_t) \right\} \quad \text{s.t.} \quad C(\mathbf{d_t}, \psi) \leq E_t \tag{20}$$

2.5 Illustration of implied tradeoffs

In the following subsection, I will derive the first order conditions characterizing the solution to Equation (20) and demonstrate how these can be used to recover implied preference parameters γ_f. In this subsection, I will illustrate the strategy for recovering γ_f graphically in a simplified setting with only two fields ($F = 2$) and one academic semester ($T = 1$).

Figure 1 graphs the set of feasible outcomes which can be achieved given the university's budget constraint, indifference curves for several hypothetical values of γ_1, and optimal course offerings given the set of feasible outcomes and values of γ_1. The horizontal axis measures the expected number of students choosing courses in field 1 and the vertical axis measures

[16] I assume endowments E_t are set exogenously through a process which 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.

[17] Normalizing $\gamma_F = 0$ is without loss of generality because the student model implies total enrollment $\sum_{f=1}^{F} n_{tf}(\mathbf{d_t})$ is preserved. Normalizing $\theta = 1$ is without loss of generality because the scale of the university's payoff is not determined.

total expected student utility.[18] The solid semi-circle represents a production possibilities frontier (PPF) of all possible (n_1, V) outcomes which could be achieved given the university's budget constraint.[19] Dashed line segments represent potential university indifference curves with payoffs increasing in the direction of the arrows.

In this illustration, University A has horizontal indifference curves implying it is not willing to sacrifice any student utility to change field enrollments ($\gamma_1^A = 0$). Given the PPF representing all feasible outcomes, University A chooses to operate at point A—unsurprisingly, this is the feasible outcome which yields the most total expected student utility. Comparatively, University B (C) has downward (upward) sloping indifference curves implying it is willing to sacrifice some student utility to increase (decrease) the expected number of students choosing courses in field 1. Given the PPF, University B (C) chooses to operate at point B (C) which yields less student utility but more (fewer) students choosing courses in field 1 relative to point A.

Suppose the observed university is offering courses which produce outcome B: The goal of this paper is to determine what value of γ_1^B best characterizes the implied tradeoff between student utility and field enrollments at outcome B. This is equivalent to computing the derivative of the PPF—or marginal rate of transformation (MRT)—at point B. Figure 2 zooms in on the choice of University B to illustrate this derivative. Conceptually, the marginal rate of transformation at point B is given by the instantaneous change in total expected student utility relative to the instantaneous change in the expected number of students choosing courses in field 1 as the university marginally reallocates funds from field 1 to field 2. Denote the instantaneous increase in total expected student utility at point B by dV_B. This is given by the marginal gain in utility from spending more in field 2 minus the utility lost my spending less in field 1. In notation:

$$dV_B = \frac{\left(\frac{\partial V}{\partial d_2}\Big|_B \right)}{\left(\frac{\partial C}{\partial d_2}\Big|_B \right)} - \frac{\left(\frac{\partial V}{\partial d_1}\Big|_B \right)}{\left(\frac{\partial C}{\partial d_1}\Big|_B \right)} \tag{21}$$

[18]Since there are only two fields in this example, the expected number of students choosing courses in field 2 is the complement $n_2 = N - n_1$ and thus can be ignored without loss of generality.

[19]The inverted U-shape of the PPF is generated by the nested logit structure and because student characteristics affect relative preferences for fields. If unobserved preferences followed a multinomial logit structure and deterministic utility was the same across all students ($u_{itf} = u_f$) then one field would strictly dominate the other and the utility maximizing bundle might contain either all field 1 courses or all field 2 courses. The nested logit structure allows for crowding in the unobserved characteristic space within fields which makes students value variety across fields. Heterogeneity in deterministic preferences for fields across students implies that some students value variety in courses within field 1 more than variety in courses within field 2 and vice versa. Both features imply that utility maximizing course offerings will generally include courses in both fields.

dV_B is positive since point B has more field 1 courses than the utility maximizing bundle implying that replacing some of these field 1 courses with field 2 courses will increase total student utility.

Next, denote the instantaneous change in the expected number of students choosing courses in field 1 by dn_{1B}. This combines both the marginal effect of making field 1 less attractive by offering fewer field 1 courses and the effect of making field 2 more attractive by offering more field 2 courses. In notation:

$$dn_{1B} = \frac{\left(\frac{\partial n_1}{\partial d_2}\Big|_B\right)}{\left(\frac{\partial C}{\partial d_2}\Big|_B\right)} - \frac{\left(\frac{\partial n_1}{\partial d_1}\Big|_B\right)}{\left(\frac{\partial C}{\partial d_1}\Big|_B\right)} \tag{22}$$

dn_{1B} is always negative since replacing field 1 courses with field 2 courses always makes field 1 relatively less attractive.

Combining both shows that the marginal rate of transformation at point B is given by:

$$\begin{aligned} MRT_B &= \frac{dV_B}{dn_{1B}} \\ &= \frac{\left(\frac{\partial C}{\partial d_2}\Big|_B\right)^{-1}\left(\frac{\partial V}{\partial d_2}\Big|_B\right) - \left(\frac{\partial C}{\partial d_1}\Big|_B\right)^{-1}\left(\frac{\partial V}{\partial d_1}\Big|_B\right)}{\left(\frac{\partial C}{\partial d_2}\Big|_B\right)^{-1}\left(\frac{\partial n_1}{\partial d_2}\Big|_B\right) - \left(\frac{\partial C}{\partial d_1}\Big|_B\right)^{-1}\left(\frac{\partial n_1}{\partial d_1}\Big|_B\right)} \end{aligned} \tag{23}$$

Therefore, the tradeoff between total expected student utility and expected field enrollments implied by observed course offerings is given by $\gamma_1^B = MRT_B$. This illustrates how marginal effects of offering additional courses and marginal costs of offering additional courses can be used to solve for implicit tradeoffs between student utility and field enrollments in a simplified setting with only two fields ($F = 2$) and one academic semester ($T = 1$).

2.6 Formal derivation of implied tradeoffs

To extend the analysis to F academic fields and T semesters, I first derive the first order conditions which characterize an interior solution to the university's problem stated in Equation (20). These first order conditions are:

$$\left(\frac{1}{c_{f_1}}\right)\left[\frac{\partial V_t\left(\mathbf{d}_t^\star\right)}{\partial d_{tf_1}} + \sum_{f'=1}^{F-1}\gamma_{f'}\left(\frac{\partial n_{tf'}\left(\mathbf{d}_t^\star\right)}{\partial d_{tf_1}}\right)\right] = \left(\frac{1}{c_{f_2}}\right)\left[\frac{\partial V_t\left(\mathbf{d}_t^\star\right)}{\partial d_{tf_2}} + \sum_{f'=1}^{F-1}\gamma_{f'}\left(\frac{\partial n_{tf'}\left(\mathbf{d}_t^\star\right)}{\partial d_{tf_2}}\right)\right] \quad \forall f_1, f_2 \tag{24}$$

where

$$c_f = \frac{\partial C\left(\mathbf{d}_t^{\star},\, \psi\right)}{\partial d_{tf}} \tag{25}$$

is the marginal cost of offering additional courses 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 relative to the cost of offering this course 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 expected student utility and net benefit (cost) from drawing students into more (less) implicitly favored fields.

Rearranging and stacking fields and semesters yields:

$$\mathbf{dn}^{\star} \times \Gamma = \mathbf{dV}^{\star} \tag{26}$$

where

$$\begin{aligned}
\underset{(F,\,F-1)}{\mathbf{dn_t}^{\star}}\left(f_1, f_2\right) &= \left(\frac{1}{c_{f_1}}\right)\left(\frac{\partial n_{tf_2}\left(\mathbf{d_t^{\star}}\right)}{\partial d_{tf_1}}\right) - \left(\frac{1}{c_F}\right)\left(\frac{\partial n_{tf_2}\left(\mathbf{d_t^{\star}}\right)}{\partial d_{tF}}\right) \\
\underset{(F\times T,\,F-1)}{\mathbf{dn}^{\star}} &= \begin{bmatrix} \mathbf{dn_1^{\star}} \\ \vdots \\ \mathbf{dn_T^{\star}} \end{bmatrix} \\
\underset{(F,\,1)}{\mathbf{dV_t^{\star}}}\left(f\right) &= \left(\frac{1}{c_F}\right)\left(\frac{\partial V_t\left(\mathbf{d_t^{\star}}\right)}{\partial d_{tF}}\right) - \left(\frac{1}{c_f}\right)\left(\frac{\partial V_t\left(\mathbf{d_t^{\star}}\right)}{\partial d_{tf}}\right) \\
\underset{(F\times T,\,1)}{\mathbf{dV}^{\star}} &= \begin{bmatrix} \mathbf{dV_1^{\star}} \\ \vdots \\ \mathbf{dV_T^{\star}} \end{bmatrix} \\
\underset{(F-1,\,1)}{\Gamma\left(f\right)} &= \gamma_f
\end{aligned}$$

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:

$$\Gamma = \left(\mathbf{dn}^{\star}\right)^{+}\left(\mathbf{dV}^{\star}\right) \tag{27}$$

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

This illustrates how marginal effects and costs of offering additional courses at observed course offerings can be used to measure the tradeoffs between student utility and field enrollments implied by observed course offerings. These tradeoffs can either be interpreted directly under relatively lenient assumptions as measures of the misalignment between student preferences and observed course offerings; or alternatively, they can be treated as structural preference parameters in a two-sided model of university course offerings and the implications for students under stronger assumptions.

2.7 Discussion

Subsection 2.6 shows that implied preference parameters γ_f can be obtained from marginal effects of offering additional courses on field enrollments, marginal effects of offering additional courses on total student utility, and marginal costs of offering additional courses. In this subsection, I discuss how to use and interpret estimates of γ_f under various assumptions and extensions of this framework which may be pursued in future research.

First, I argue that estimates of γ_f are interesting measures of the misalignment between student preferences and observed course offerings and that one can interpret them as such even if the university's problem in Equation (20) is misspecified. If one has credible estimates of the local marginal costs $\frac{\partial C(\mathbf{d}_t^\star, \psi)}{\partial d_{tf}}$ for all fields then one can measure how marginally reallocating dollars across any pair of fields would change observed course offerings. Furthermore, if one has credible estimates of the local marginal effects $\frac{\partial n_{tf'}(\mathbf{d}_t^\star)}{\partial d_{tf}}$ and $\frac{\partial V_t(\mathbf{d}_t^\star)}{\partial d_{tf}}$ for all fields then one can predict how these changes in observed course offerings would affect student utility and field enrollments. As such, if marginally reallocating dollars from field f to field f' would increase student utility and draw students out of field f and into field f' in expectation, then one can conclude that observed course offerings are implicitly sacrificing student utility to prevent students from moving out of field f and into field f'. Therefore, without asserting why observed course offerings were chosen, one can still produce an interesting measure of the misalignment between student preferences and the course offerings that were chosen.

To be clear, obtaining credible estimates of local marginal effects and local marginal costs still requires assumptions. Strictly speaking, all of the assumptions regarding student choices discussed in Subsection 2.1 are required to identify local marginal effects of offering additional courses on student utility and field enrollments. More loosely, because one only needs credible estimates of marginal effects in the neighborhood of observed course offerings, one might be willing to accept misspecifications in the student model which only affect behavior far away from observed course offerings. Subsection 2.2 shows that in this choice framework, local marginal effects depend on observed courses, choice probabilities at observed courses, and nesting parameters which are identified by the relationship between course offerings and field enrollments across semesters. This suggests real empirical variation is driving estimates of γ_f rather than arbitrary functional form assumptions.

As alluded to in Subsection 2.3, one noteworthy concern is that estimates of nesting parameters could be biased towards one if the university changes its course offerings in response to unobserved student preferences. Here I show that this will only confound estimates of γ_f if the bias is disproportionately large in some fields. Specifically, suppose there is multiplicative bias in estimates of ρ_f so that

$$\hat{\rho}_f = \phi \rho_f \tag{28}$$

It is straightforward to show that this multiplicative bias in estimates of ρ_f leads to multiplicative bias in estimates of marginal effects at observed course offerings as follows:[20]

$$\frac{\partial \widehat{V_t (\mathbf{d}_t^\star)}}{\partial d_{tf}} = \phi \frac{\partial V_t (\mathbf{d}_t^\star)}{\partial d_{tf}} \tag{29}$$

$$\frac{\partial \widehat{n_{tf} (\mathbf{d}_t^\star)}}{\partial d_{tf}} = \phi \frac{\partial n_{tf} (\mathbf{d}_t^\star)}{\partial d_{tf}} \tag{30}$$

This leads to multiplicative bias in estimates of $\mathbf{dn}_\mathbf{T}^\star$ and $\mathbf{dV}_\mathbf{T}^\star$ which then divides out in Equation (26) implying that estimates of γ_f are robust to multiplicative bias in ρ_f which is the same across all fields. This suggests that correlation between course offerings and unobserved preferences will only confound estimation of γ_f if it is stronger in some fields relative to others.

In addition to directly interpreting estimates of γ_f as measures of the misalignment between student preferences and observed course offerings, one can also treat this framework as a two-sided structural model of a university offering courses and students choosing courses from the set of available alternatives. This interpretation requires stronger assumptions: First, the structural interpretation requires assuming that marginal effects and marginal costs can be constructed globally rather than only in a neighborhood around observed course offerings. This implies that the student choice model must be able to predict choice probabilities far away from observed course offerings making the results more sensitive to misspecification of the student choice model. Second, and most importantly, the structural interpretation requires assuming that the objective structure and constraints in the university's problem specified in Equation (20) are correct. There are many potential objective structures which could rationalize observed course offerings; however, these different structures will generally yield differing predictions under counterfactual policies. As such, the predictions of this

[20] Note that this type of misspecification will also affect estimates of choice probabilities P_{itf} which also influence marginal effects; however, because these choice probabilities are conditional moments of observed data, one should expect the estimates to be relatively robust to misspecification.

particular structure will only be correct if this structure is a good approximation of the university's true objective structure. While this is certainly a strong assumption, the structural interpretation gives a deeper understanding into how universities make decisions and allows for counterfactual policy analyses which incorporate university responses in predictions.

Future research may seek to strengthen the credibility of these counterfactual analyses by building richer and more robust models of student and university choices. Two extensions seem particularly interesting: First, the student choice model assumes course sizes are unconstrained so that a student can choose any offered course. In reality, many courses cap enrollments for pedagogical reasons. As such, true demand for courses in which enrollment caps are binding may be substantially higher than censored observed demand which may lead to understating demand for courses in fields where enrollment caps bind more frequently. With data on enrollment caps and methodological advances, one could modify the course choice problem to accommodate supply constraints.

Second, the student choice model assumes class size does not affect course utility. While there are both econometric and theoretical challenges associated with incorporating class size effects, including these effects would yield a richer framework for measuring the marginal effects of offering additional courses. In the current framework, additional courses add value and influence choices by providing particularly good draws of η_{ijt} for some students. In a framework with class size effects, additional courses could also add value by reducing class sizes.

3 Data, descriptive statistics, and empirical specifications

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 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 which is below the national average.[21] 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

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

2011-12 academic years and information on all offered courses and the instructors teaching these courses for all courses offered between the 1994-95 and 2011-12 academic years. I combine these to create a sub-sample of student information and course information from the 2004-05 to the 2009-10 academic years. After excluding required writing courses, required oral communication courses, required health courses, and other special courses, the sample includes 25,056 unique UCA undergraduates and 258,662 observations of students choosing introductory courses.[22]

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 offering additional courses. 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 isolate the costs of marginal courses by identifying which instructors were hired on single-semester contracts and which instructors must be paid to honor pre-existing contracts.

An important empirical decision which must be made to match the administrative data to the theoretical framework is what exactly constitutes a "course" j. In Section 2, I defined a course j as a unit which provides meaningful choice variety to students and presents a marginal cost to the university. Since my primary focus is on the course offerings of the university, I define a course as the unit the university gives teaching credit for (and thus effectively pays for). UCA gives teaching credit for course sections which are defined by a course number, an instructor, and a meeting time. This implies that an instructor who goes from teaching one section of Introductory Economics per week to two sections of Introductory Economics per week receives twice as much teaching credit. As such, I use course sections as the unit of analysis j. For brevity, the remainder of the paper frequently refers to course

[22]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).

sections simply as “courses”.[23]

3.1 Descriptive statistics

For my main empirical analysis, I will be analyzing introductory course offerings and student choices across four academic fields: STEM, social science, humanities and arts, and business and occupational. Before proceeding to the main 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 introductory 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 34.1 students in introductory social science courses, 29.2 students in introductory humanities and arts courses, and 26.6 students in both introductory STEM courses and introductory business and occupational courses. 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 courses 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 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 high GPA but less remarkable ACT scores and students choosing social science or humanities and arts courses are comparable in terms of these measures of baseline preparation. The statistics also show that students choosing introductory business and occupational courses are less

[23] One may argue that offering a new course should provide more choice variety to students than offering an additional section of an existing course. One may also claim that offering a new course should cost the university more than offering an additional section of an existing course. While possible in theory, incorporating these features would require modeling the university’s decision of how many courses (defined by course number) to offer in each field and how many sections of each course to offer. To explain why the number of sections per course number varies widely within fields, the model would also need to allow courses within fields to be heterogeneous further complicating the student choice model. I leave this extension for future research.

likely to be women or freshmen but more likely to be sophomores, juniors, or seniors.

In Subsection 2.3, I showed that the crucial nesting parameters ρ_f are identified by the empirical relationship between introductory course offerings and introductory 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. 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 21% - 25%, and the share of business and occupational courses varies from 12% - 14%. The extent to which enrollment shares move in concert with these 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 as discussed in Subsection 2.3. 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.

3.2 Empirical specifications

In Section 2, I developed a theoretical framework for measuring implied tradeoffs between student utility and field enrollments under a general additively separable course utility function and a general course cost function. In this subsection, I discuss the exact specifications I use in my empirical application.

As discussed previously, these marginal effects are identified by the empirical relationship between course offerings and field enrollments across semesters. As such, for tractability, and to preserve the transparent link between empirical variation and results, I employ the following simple linear structure for the deterministic component of utility:

$$U_{itj} = X_{it}\beta_f + \epsilon_{itj} \tag{31}$$

where X_{it} includes ACT scores, high school GPA, and indicators for gender and year in school.

To estimate the marginal costs of offering additional introductory courses, I assume that the implicit cost of hiring an instructor to teach introductory course j can be additively separated into a field specific effect ψ_f, an instructor rank specific effect ξ_r where r indexes instructor rank, and an idiosyncratic component υ_j as follows:[24]

$$C_j = \psi_f + \xi_r + \upsilon_j \tag{32}$$

I then assume that the marginal cost of adding or removing one introductory course in field f' is given by the expected cost of hiring an instructor on a single-semester contract to teach an introductory course in field f'.[25] In notation,

$$c_f = \mathbb{E}\left[C_j \mid f = f', r = \text{single semester}\right] \tag{33}$$

The idea is that if a university wants to add or subtract a course in field f in semester t, it is generally simpler and more cost effective to do this by hiring or firing an instructor on a single-semester contract.[26] As such, costs of hiring single-semester instructors represent better estimates of marginal costs than average costs within a field.

As discussed in Subsection 2.7, the interpretation of γ_f as implied local tradeoffs only requires estimates of marginal costs which are valid in a neighborhood around observed course offerings. Conversely, treating γ_f as structural parameters requires assuming marginal costs can be estimated globally. Notice that Equation (33) assumes the marginal cost of offering an additional course in field f is independent of the number of courses offered in field f. This is consistent with a framework in which UCA is a wage-taker in the market for single-semester instructors; however, this assumption may still be violated at hypothetical course offerings which are far away from observed offerings. While this does not affect the implied local tradeoff interpretation of γ_f, it may affect counterfactual analyses in which predicted counterfactual offerings 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. To see how ignoring these non-instructor costs affects my analysis, suppose there is general downward multiplicative bias in my estimates of marginal

[24]Possible instructor ranks are: tenured, tenure-track, on a long term contract but ineligible for tenure, and on a single-semester contract. Allowing costs to differ by instructor rank suggests instructor rank should also enter into student utility. Appendix B suggests instructor compensation (which is highly related to rank) has little effect on student utility so I exclude these effects for clarity. 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 course j C_j.

[25]Instructors on single-semester contracts are those whose contracts need to be renewed every semester. 11.4% of all courses are taught by instructors on single-semester contracts.

[26]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.

costs given by:

$$\hat{c}_f = \pi_f c_f \tag{34}$$

where $\pi_f \in (0,1)$. If π_f is equal across fields, then there is multiplicative bias in both dn and dV which divides out in Equation (26). Therefore, estimates of γ_f are robust to ignoring non-instructor costs if these costs are proportional to instructor costs. If $\pi_f < \pi_{f'}$ for all $f' \neq f$ then I am understating the relative marginal cost of offering courses in field f. This leads to downward bias in estimates of the implied local preference for enrollment in field f γ_f.

4 Effects of course offerings and implied preferences

I use the methods described in Section 2 to analyze introductory course offerings at the University of Central Arkansas (UCA) in Fall and Spring academic semesters of academic years 2004-05 through 2009-10. This section reports results of this analysis which can be interpreted under more lenient assumptions about university decisions and student course choices.

I begin by reporting estimates of primitive student preference parameters and cost parameters. I then use these primitive parameters to construct local marginal effects of offering additional introductory courses on total student utility, marginal effects relative to marginal costs, and implied preference parameters γ_f. Results show that an additional dollar of spending offering introductory social science courses produces 2.5 times as much student utility as an additional dollar of spending on introductory business or occupational courses. This implies the university is implicitly sacrificing significant student utility to draw students out of social science courses and into business and occupational courses.

4.1 Student preference parameters and cost parameters

As discussed in Section 2, the fundamental elements needed to measure implied preference parameters γ_f are local marginal effects of offering additional courses on total expected student utility, local marginal effects of offering additional courses on field enrollments, and local marginal costs of offering additional courses.

Table 4 reports estimates of the cost regression described in Equation 32 which will be used to compute marginal costs. Results show that conditional on instructor rank, costs are highest for introductory STEM courses followed by business and occupational, humanities and arts, and social science. Results also show that conditional on field, single semester instructors cost \$5,595 per course less than tenured instructors and \$3,132 per course less than instructors who are on long term contracts but are not eligible for tenure. Because single-semester is the omitted rank category—and because the regression does not include a constant—coefficients on field indicators measure the expected cost of hiring a single semester instructor to teach an introductory course in each field. As discussed previously, I use these single-semester instructor costs as my estimate of the marginal cost of adding or subtracting a course in a given field.

Table 5 reports estimates of student preference parameters from the nested logit course choice model which will be used to measure the marginal effects of offering additional courses. The estimates imply a first year male student with average ACT scores and HS GPA is most attracted to introductory social sciences courses followed by humanities or arts, STEM, and business or occupational. First year female students with average scores and grades have the same relative preferences over fields for introductory courses; however, the magnitudes suggest first year female students are relatively more attracted to social science courses and less interested in business or occupational courses than their male counterparts. While introductory business courses are quite unpopular with freshmen, they are relatively more popular with advanced students. In fact, male sophomores, juniors, and seniors with average scores and grades prefer introductory business courses to introductory courses in all other fields.

The estimates also imply students with conditionally higher ACT scores are relatively more likely to enroll in STEM or humanities or arts courses while students with conditionally higher high school GPA are relatively more likely to enroll in STEM or business or occupational courses. The finding that students with higher ACT scores and high school GPA are relatively more likely to enroll in STEM courses is consistent with existing literature which shows initial preparation is an important determinant of whether a student pursues a STEM education (Arcidiacono, 2004; Stinebrickner and Stinebrickner, 2014b).

Finally, estimates of the nesting parameters are in the middle range varying from 0.461 to 0.680. This implies unobserved preferences for courses within the same field are neither fully independent nor perfectly identical. This shows that new courses are sufficiently different from existing courses within the same field to provide students with meaningful choice variety; however, it also shows that assuming independence within fields would lead to grossly overstating the effects of offering additional courses on student choices and utilities. Furthermore, the estimates suggest there is substantial heterogeneity in crowding across fields implying that allowing for heterogeneous crowding is important when comparing marginal effects across academic fields.

4.2 Marginal effects, marginal effects per dollar, and implied preferences

Table 6 uses estimates of the student course choice model and the cost regression to construct marginal effects and to measure implied tradeoffs between student utility and field enrollments. To begin, column 1 uses estimates of student preference parameters to construct the local marginal effects of offering additional introductory courses on total expected student utility averaged across semesters for exposition purposes. In notation, column 1 reports:

$$\frac{1}{T}\sum_{t=1}^{T}\frac{\partial V_t\left(\mathbf{d}_t\right)}{\partial d_{tf}} \tag{35}$$

Average effects are reported relative to the average effect of offering an additional introductory business or occupational course which is normalized to one and stars report whether average effects in other fields are significantly greater than one.

Results show that on average, an additional introductory humanities or arts course produces 1.541 times as much student utility as an additional introductory business or occupational course, an additional introductory social science course produces 1.533 times as much student utility as an additional introductory business or occupational course, and an additional introductory STEM course produces 1.484 times as much student utility as an additional introductory business or occupational course.

While the figures in column (1) show significant differences in the marginal benefits of additional introductory courses in terms of student utility, they do not account for differences in the marginal costs of these courses. Results in Table 4 showed that differences in costs are sizable implying that ignoring these differences would lead to incorrect conclusions. To account for these differences, column 2 reports local marginal effects divided by marginal cost once again averaging across semester. In notation, column 2 reports:[27]

$$\frac{1}{T}\sum_{t=1}^{T}\frac{1}{c_f}\frac{\partial V_t\left(\mathbf{d}_t\right)}{\partial d_{tf}} \tag{36}$$

Results show that on average, an additional dollar spent offering social science courses produces 2.533 times as much student utility as an additional dollar spent offering business or occupational courses, an additional dollar spent offering humanities or arts courses produces 2.249 times as much student utility as an additional dollar spent offering business or occupa-

[27] As before, average marginal effects per dollar are reported relative to the average effect per dollar of offering an additional introductory business or occupational course which is normalized to one and stars report whether effects per dollar in other fields are significantly greater than one.

tional courses, and an additional dollar spent offering STEM courses produces 1.365 times as much student utility as an additional dollar spent offering business or occupational courses.

These differences show that marginally reallocating spending from introductory business and occupational courses to introductory social science courses would increase total expected student utility; however, doing so would decrease variety in introductory business and occupational courses and increase variety in introductory social science courses which would decrease business and occupational enrollment and increase social science enrollment in expectation. Since spending was not reallocated from business and occupational courses to social science courses despite the potential for increasing student utility, observed course offerings are implicitly sacrificing some student utility to keep students from switching from business and occupational courses to social science courses in expectation. This reveals an implicit willingness to sacrifice student utility to draw students out of introductory social science courses and into introductory business and occupational courses.

Column 3 reports estimates of γ_f to precisely quantify these implicit tradeoffs. The omitted field is social science; therefore, estimates for field f report how much total expected student utility the university is implicitly sacrificing to move one student out of an introductory social science course and into an introductory course in field f in expectation. Results show that the university is implicitly sacrificing 0.086 units of total student utility to move a student from social science to humanities and arts, 0.611 units of total student utility to move a student from social science to STEM, and 1.114 units of total student utility to move a student from social science to business or occupational. By showing that observed course offerings are implicitly sacrificing significant student utility to change field enrollments, these estimates quantify the extent to which student preferences and observed course offerings are misaligned.

The interpretation of estimates of γ_f as implicit tradeoffs and measures of misalignment holds even when the university's problem is misspecified and/or when the student model holds only in a neighborhood around observed course offerings. However, quantifying these tradeoffs in terms of units of total student utility makes them difficult to interpret and measures of these implicit tradeoffs alone cannot be used for policy analysis. In the following section, I report results which require stronger assumptions but provide additional interpretation and predictions for policy analysis.

5 Interpretation and Policy Counterfactuals

The preceding section used estimates of local marginal costs and local marginal effects of offering additional courses to measure the tradeoffs between student utility and field enrollments which are implied by observed course offerings. Under relatively lenient assumptions, these tradeoffs can be interpreted as a measure of the misalignment between student preferences and observed course offerings. In this section, I treat these implied tradeoffs as the university's structural preference parameters in a two-sided model of a university choosing course offerings and students choosing courses from the set of available alternatives. While this requires assuming that the university's problem and student choice model are both correctly specified, it allows me to provide additional intuitive ways to quantify misalignment between student preferences and observed course offerings and to conduct counterfactual policy analyses which incorporate university responses.

5.1 Utility maximizing course offerings and equivalent costs

To begin, columns (1) - (5) of Table 7 compare average observed course offerings and field enrollments to offerings and enrollments which would have maximized student utility. Columns (1)-(3) report averages across semesters of the number of introductory course sections taught by instructors on long term contracts in each field, the number of introductory course sections taught by instructors on single-semester contracts in each field, and enrollment in introductory courses by field. Columns (4) and (5) then examine how single-semester offerings and enrollments would change if the portion of the budget allocated to pay single-semester instructors were reallocated to maximize student utility holding contracted offerings in column (1) fixed.[28] Stars indicate where columns (4) and (5) are statistically different from columns (2) and (3) respectively.

Results suggest the utility maximizing allocation of the single semester budget contains no introductory STEM courses taught by single semester instructors, no introductory business or occupational courses taught by single semester instructors, approximately the same number of humanities and arts courses taught by single semester instructors, and four times as many social science courses taught by single semester instructors. The large increase in social science courses reflects the finding in Table 6 that marginal spending on introductory social science courses produces student utility more efficiently than spending in other fields. Column (5) predicts that offering the utility maximizing single-semester courses would increase overall introductory social science enrollment by 11.65% and decrease overall

[28] There are several reasons to reallocate the budget for single semester 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 instructors on long term contracts can only be released when those contracts expire; 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.

introductory STEM and business and occupational enrollment by 13.74% and 5.96% respectively.

To provide an additional intuitive way to measure the misalignment between student preferences and observed course offerings, column (6) of Table 7 reports how much costs of single semesters instructors would need to change to induce a utility maximizing university to offer the single-semester courses reported in column (2). I refer to these as the "equivalent costs" of $\hat{\gamma}_f$ since going from observed costs to equivalent costs with a utility maximizing objective would have the same effect on course offerings as going from a utility maximizing objective to an objective characterized by $\hat{\gamma}_f$ holding costs fixed at observed costs. Figure 3 illustrates the idea with one semester and two fields: The observed production possibilities frontier is PPF and the outcomes associated with observed course offerings are given by B. The goal is to solve for counterfactual costs which yield a production possibilities frontier PPF' which makes it so that a utility maximizing university with indifference curves given by Π^{SUM} would offer courses which achieve outcomes B. Intuitively, I infer these equivalent costs by solving for costs which make it so that a utility maximizing university's first order conditions are satisfied at observed course offerings. This means solving for costs which imply that marginal effects per dollar of offering additional courses on total student utility are equal across fields at observed course offerings. Details are reported in Appendix C.

Results suggest that inducing a utility maximizing university to offer observed courses would require a 45.62% increase in the cost of hiring a single-semester social science instructor, a 24.99% increase in the cost of hiring a single-semester humanities or arts instructor, a 17.01% decrease in the cost of hiring a single-semester STEM instructor, and a 41.99% decrease in the cost of hiring a single-semester business and occupational. This shows that the estimated preference parameters $\hat{\gamma}_f$ have the same effects on course offerings as substantial increases in social science and humanities and arts costs and substantial decreases in business and occupational and STEM costs.

5.2 Counterfactual analyses with university responses

As mentioned previously, one of the primary reasons for moving towards empirical models of university behaviors is the capacity to conduct counterfactual policy analyses which incorporate university responses into predictions. In general, universities are not passive parties in the production of human capital but rather active entities which allocate their resources to maximize their payoffs subject to constraints. While predicting university responses requires strong assumptions, counterfactual policy analyses which assume university inputs remain fixed are arguably making even stronger assumptions.

To illustrate the value of my two-sided model, and for higher education models which incorporate supply-side responses more generally, Tables 8 and 9 predict introductory course offerings and introductory field enrollments in several counterfactual scenarios.[29] To enhance the credibility of these predictions, I make two choices: First, I choose counterfactual scenarios which 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 instructors on single semester contracts. This also restricts the counterfactual scenarios to be close to the observed scenario and examines a short run scenario in which inputs which are costly to vary are held fixed. Furthermore, it acknowledges the limitation that the model has no mechanism for explaining why the university hires instructors of different ranks and thus is not well equipped to predict hiring decisions across ranks.

Given the policy interest in increasing specialization in STEM, one interesting counterfactual scenario to consider is one in which the state subsidizes STEM instructors to increase STEM course offerings and enrollments. To evaluate the effectiveness of such a subsidy, my first counterfactual predicts single-semester introductory course offerings and introductory field enrollments under a subsidy which reduces the cost of hiring a single-semester STEM instructor by 5%. Row 2 of Table 8 shows that this subsidy would increase single-semester introductory STEM offerings by 49.15% and reduce offerings in other fields. Furthermore, row 2 of Table 9 shows that this increase in single-semester introductory STEM offerings would lead to a 6.22% increase in overall introductory STEM enrollment with additional students coming mostly from social science and humanities and arts courses. The subsidy would cost $252.87 per single semester course implying a total cost of $13,498.23 or 4.18% of spending on single-semester introductory courses.

Another interesting scenario to consider is one in which UCA begins attracting higher ability students. From Table 5, we know that higher ability students are generally more interested in STEM suggesting that UCA might respond to a higher ability student body by offering more STEM courses and thus making the STEM field even more attractive. To analyze this scenario, my second counterfactual predicts single-semester introductory course offerings and introductory field enrollments if all student ACT scores and high school GPAs were increased by one-third of a standard deviation holding fixed other student characteristics.[30] Row 3 of Table 8 shows that increasing student abilities would increase single-semester

[29] 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 of averages of predictions across all academic semesters.

[30] These predictions assume that student field preferences depend on absolute abilities rather than abilities

introductory STEM offerings by 48.62% and reduce offerings in other fields. To see how field enrollments would change with higher ability students, row 3 of Table 9 first predicts introductory field enrollments in a partial equilibrium where student characteristics are changed but course offerings remain fixed. Results show that attracting higher ability students would increase introductory STEM enrollment by 3.68% without any response in course offerings. Row 4 of Table 9 incorporates the changes in single-semester introductory course offerings and shows that the total effect of attracting higher ability students is a 10.97% 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 approximately a factor of three.

A final scenario to consider is one in which the gender composition of students at UCA changes. Results in Table 5 suggest men and women do not have wildly different field preferences so I choose an extreme counterfactual setting in which all male students are given female field preferences for illustrative purposes.[31] Row 4 of Table 8 shows that if all students had female preferences, UCA would offer 2.4 times as many single-semester introductory social science courses, 54.86% more single-semester introductory humanities and arts courses, 32.69% fewer single-semester introductory STEM courses, and would virtually eliminate single-semester introductory business and occupational courses. 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 introductory field enrollments in partial equilibrium without changes in course offerings and row 6 of Table 9 predicts introductory field enrollments in general equilibrium with university responses. Results show that stronger female preferences for introductory social science courses imply that giving all students female preferences leads to a 4.71% increase in introductory social science enrollment without any change in course offerings. Incorporating the effects of the increase in single-semester introductory social science leads to a predicted increase in introductory social science enrollment of 8.68%. Once again, ignoring the indirect of effects of changes in course offerings leads to significantly understating changes in field enrollments.

relative to the student body. If field preferences depend on relative abilities then increasing the abilities of all students will have no effect on field preferences and thus no effect on course offerings.

[31]This avoids introducing simulation error by randomly choosing some students to switch preferences and avoids the empirically feasible but theoretically awkward alternative of giving male students "partially female" preferences.

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 over four decades 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.

To advance our understanding of the "supply side" of higher education, this paper proposed a new framework for analyzing course offerings at a university. The main idea of the framework is that offering additional courses in a field provides more variety of choices within that field making the field more attractive to students and increasing student utility in expectation. As such, any marginal reallocation of resources across fields will have effects on expected field enrollments and expected student utility.

One can measure the marginal effects of offering additional courses on expected student utility using a course choice model and use these estimates to identify fields where additional courses have the smallest effects on student utility. Furthermore, one can use data on instructor salaries and other information to estimate the costs of offering courses and combine these with utility effects to identify fields where additional dollars produce the least student utility. If a marginal dollar produces significantly less utility in one field relative to others then observed course offerings are implicitly sacrificing student utility in exchange for more courses in that field. Since more courses means more variety of choices and thus more enrollment, this implies observed course offerings are implicitly sacrificing student utility to increase enrollment in that field.

I propose two alternative interpretations of these implied tradeoffs. First, I claim that these implied tradeoffs can be directly interpreted under relatively lenient assumptions as measures of the misalignment between student preferences and offered courses. This interpretation reveals how observed course offerings are implicitly favoring certain fields without imposing that the econometrician understands why those courses were offered. Second, I propose treating these implied tradeoffs as a university's structural preference parameters in a two-sided model of a university offering courses and students choosing courses from the set of available alternatives. This requires stronger assumptions but allows me to provide additional interpretation and to perform counterfactual policy analyses which incorporate university responses into predictions.

I use my framework to analyze introductory course offerings at the University of Central Arkansas (UCA) and find that UCA is implicitly sacrificing student utility to draw students out of social science courses and into STEM and business and occupational courses. The misalignment is so large that if one were to reallocate the portion of the introductory course budget paid to instructors on single-semester contracts to maximize student utility, one would eliminate all introductory STEM and business and occupational courses taught by single-semester instructors and offer four times as many introductory social science courses taught by single-semester instructors. Another way to quantify the misalignment is to see that the implied field preferences yield the same over-supply of business and occupational courses and under-supply of social science courses as increasing the cost of offering a social science course by 45.62% and decreasing the cost of offering a business or occupational course by 41.99%.

Finally, to illustrate the value of the two-sided model, and for higher education models which incorporate supply-side responses more generally, I perform a number of simulations which predict course offerings and student outcomes in counterfactual scenarios. Most notably, I find that a 5% reduction in the cost of hiring a single-semester STEM instructor would lead to a 49.1% increase in the number of introductory courses taught by single-semester STEM instructors and a 6.2% increase in overall enrollment in introductory STEM courses. I also show that ignoring changes in course offerings in response to changes in student composition leads one to significantly understate the effects of student composition on field enrollments in general equilibrium.

To my knowledge, this is the first analysis of course offerings at a university and the first attempt to estimate a two-sided model of a university and students with microdata. As such, there is substantial room for future research to extend this analysis: First, the model could be extended to accommodate course capacity constraints or class size externalities. Second, the analysis could be broadened to understand how universities decide how many instructors to hire of different ranks. Finally, the analysis could be repeated for a set of schools or over a longer time period to analyze how implicit field preferences vary by institutional characteristics or over time. These extensions and others will broaden our understanding of the higher education market and may lead to more informed policies which benefit students, families, and taxpayers.

Appendix A: Data Appendix

Definitions

STEM: Biology; Chemistry; Computer Science; Mathematics; Physics and Astronomy.

Social Science: Family and Consumer Sciences; Geography; History; Political Science; Psychology and Counseling; Sociology; World Languages, Literatures, and Cultures.

Humanities and Arts: Art; Communication; English; Mass Communication and Theater; Music; Philosophy and Religion; Writing.

Business and Occupational: Accounting; Economics, Finance, Insurance, and Risk; Education; Elementary, Literacy, and Special Education; Health Sciences; Kinesiology and Physical Education; Management Information Systems; Marketing and Management; Nursing; Occupational Therapy

Long term contracts: Tenured instructors, tenure-track instructors, and instructors who teach on a recurring contractual basis but are ineligible for tenure. See ADHE (2011) for further information.

Short term contracts: Instructors with a non-recurring appointment where funding is temporary and there is no guarantee of a continuing appointment and graduate student instructors. See ADHE (2011) for further information.

Average costs of offering a non-contract course

To compute the implicit cost of hiring an instructor to teach course 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 course. Generally speaking, this method uses credit hours to allocate an instructor's total salary to specific courses. 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.

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 yields the instructor salary paid for each course.

Importantly, this method ensures that faculty members who are paid for activities other than research 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: Intensive margin of instruction spending

The body of this article assumes instruction spending only affects students through the number of courses offered. However, if higher paid instructors are more attractive to students, universities could also influence student choices and utility by spending more on instructors. There can be both budget allocation decisions on the extensive margin—how many courses to offer in each field—and budget allocation decisions on the intensive margin—how much to pay instructors in each course—which are made by the university and directly affect student outcomes.

In this appendix, I modify the model presented in Section 2 to include both intensive and extensive margin spending decisions and discuss alternative methods for recovering university preference parameters in this setting.[32] Following this, I present evidence which suggests intensive margin spending has minor effects on student choices at UCA and justify my decision to abstract from intensive margin spending decisions in the analysis.

Theoretical model with intensive margin spending decisions

To incorporate intensive margin spending, in this Appendix only, let c_{jt} represent spending on instruction in course j in semester t, let m_f represent the minimum cost of offering a course in field f, and let e_{jt} represent spending in excess of this minimum which may affect

[32] An earlier draft of this paper (available upon request) contains a more detailed discussion of this model and these estimation methods.

the desirability of course j. For courses taught by instructors on long term contracts, c_{jt} must be paid to honor these contracts. For the share of the budget that remains after all existing contracts are honored, a course in field f is offered if and only if $c_{jt} \geq m_f$.

To allow for the possibility that excess spending affects the desirability of a course, modify student utility in Equation (31) to be:[33]

$$U_{ijt} = X_{it}\beta_f + \theta \log\left(e_{jt} + 1\right) + \epsilon_{ijt} \tag{37}$$

The parameter θ measures the extent to which higher paid instructors make courses more attractive to students.

For simplicity, assume ϵ_{ijt} are iid draws from a type 1 extreme value distribution. Similar to Section 2.2, total expected student utility in semester t, the expected number of students choosing courses in field f in semester t, and the effects of both extensive margin spending and intensive margin spending on both of these outcomes can be defined as a function of model parameters and observed data. The effects of intensive margin spending on total expected student utility in semester t and on the expected number of students choosing courses in field f in semester t are given by:[34]

$$\frac{\partial V_t\left(\mathbf{d_t}, \mathbf{e_t}\right)}{\partial e_{jt}} = \sum_{i=1}^{N} P_{itj}\left(\frac{\theta}{e_{jt}+1}\right) \tag{38}$$

$$\frac{\partial n_{tf}\left(\mathbf{d_t}, \mathbf{e_t}\right)}{\partial e_{jt}} = \begin{cases} \sum_{i=1}^{N}\left\{\left(\frac{\theta}{e_{jt}+1}\right) P_{itj}\left(1 - P_{itj}\right) - \sum_{j' \in f \setminus j}\left(\frac{\theta}{e_{jt}+1}\right) P_{itj}P_{itj'}\right\} & j \in f \\ -\sum_{i=1}^{N}\sum_{j' \in f}\left(\frac{\theta}{e_{jt}+1}\right) P_{itj}P_{itj'} & j \notin f \end{cases} \tag{39}$$

where $\mathbf{e_t}$ is a vector containing all excess spending decisions, $\mathbf{d_t}$ is a vector containing all offered courses, and P_{itj} is the probability that student i chooses course j in semester t. As one might expect, these equations illustrate the crucial role of the parameter θ in determining the effects of intensive margin spending on student outcomes.

With these marginal effects, one can construct the set of intensive margin university first order conditions analogous to the extensive margin conditions given by Equation (24):

$$\frac{\partial V_t\left(\mathbf{d_t}, \mathbf{e_t}\right)}{\partial e_{j_1 t}} + \sum_{f'=1}^{F-1}\gamma_{f'}\left(\frac{\partial n_{tf'}\left(\mathbf{d_t}, \mathbf{e_t}\right)}{\partial e_{j_1 t}}\right) = \frac{\partial V_t\left(\mathbf{d_t}, \mathbf{e_t}\right)}{\partial e_{j_2 t}} + \sum_{f'=1}^{F-1}\gamma_{f'}\left(\frac{\partial n_{tf'}\left(\mathbf{d_t}, \mathbf{e_t}\right)}{\partial e_{j_2 t}}\right) \quad \forall f_1, f_2 \tag{40}$$

[33]The log function is used to ensure diminishing marginal returns to avoid a corner solution in which the university spends its entire discretionary instruction budget on a single course. I add 1 to ensure marginal effects of excess spending are finite over the entire support of excess spending.

[34]Other equations are straightforward to derive and are omitted for brevity.

As in Section 2.6, this system can be rearranged to solve for the university preference parameters which best explain why observed intensive margin spending decisions were preferred to all feasible alternative decisions.

The intuition underlying this method is analogous to the intuition behind the extensive margin methods discussed in the body: If the university were purely trying to maximize total expected student utility, it would choose excess spending levels so that the marginal effect of increasing excess spending on total expected student utility is the same across all courses. If the university is consistently overpaying instructors in a certain field relative to the allocation which maximizes student utility, it must be that the university is trying to draw more students into this field thus revealing an institutional preference to increase the number of students in this field.

Effects of intensive and extensive margin spending

I chose to abstract from intensive margin spending decisions in my analysis because empirical evidence suggests intensive margin spending has much smaller effects on student choices than extensive margin spending. Panel A of Table A1 reports estimates of the elasticity of enrollment with respect to spending on instructors estimated with several specifications of the regression:

$$\log\left(S_{jt}\right) = \tilde{\theta}\log\left(c_{jt}\right) + \xi_k + \eta_{jt} \tag{41}$$

where S_{jt} is the number of students enrolled in course j in semester t and ξ_k is a course number fixed effect (e.g. ECON 101). Specification 2 suggests the elasticity of enrollment with respect to instructor salary could be as large as 0.162 for courses taught by single-semester instructors. This would imply that doubling spending on instruction for all single-semester field f courses but keeping other course characteristics fixed would increase single-semester field f enrollment by 16.2%. However, specification 4 suggests this moderately large estimate is driven by a small number of very small courses. When I exclude 45 course observations with five or fewer students, the elasticity drops to 0.0534. This suggests doubling spending on single-semester field f instruction but keeping other course characteristics fixed only increases single-semester field f enrollment by 5.34%. Elasticities for all instructor contract types (columns 1 and 3) suggest similarly small effects.

While it is not the focus of this paper, I should note that this finding is in line with existing literature which finds that higher paid instructors have small or zero effects on student outcomes at universities (Bettinger and Long, 2010; Figlio et al., 2015).

Comparatively, Panel B of Table A1 reports estimates of elasticities of enrollment with respect to spending on course offerings computed using estimates of the nested logit course choice model, observed non-contract course offerings, and estimates of costs of offering non-contract courses.[35] Estimates of these elasticities range from 0.3229 in social science to 0.4999 in humanities and arts. This suggests that doubling the number of single-semester field f courses offered to students increases single-semester field f enrollment by 32.29 - 49.99%.

The large differences between intensive margin elasticities and extensive margin elasticities suggest UCA can increase student utility more and attract more students into desirable fields by spending marginal dollars offering additional courses rather than increasing spending on instruction. This implies that no values for γ_f can rationalize both observed intensive and observed extensive margin spending decisions at UCA. Furthermore, the small effects of intensive margin spending suggest variation in spending on instruction at UCA exists for some reason other than influencing student choices and utility. For this reason, I focus on extensive margin decisions which have significant effects on student choices and utility at UCA. Future research may seek to better explain variation in spending on instruction.

Appendix C: Solving for equivalent costs

In this appendix, I describe my method for estimating the equivalent costs reported in Column 6 of Table 7. The goal of this exercise is to solve for counterfactual costs of hiring single-semester instructors which come closest to inducing a utility maximizing university to offer observed single-semester courses.

A utility maximizing university's problem is given by:

$$\mathbf{d_t^{SUM}} = \text{argmax}_{\mathbf{d_t}} \left\{ V_t\left(\mathbf{d}_t\right) \right\} \quad \text{s.t.} \quad \sum_{f=1}^{F} d_{tf}^N c_f \leq E_t^N \tag{43}$$

where c_f is the cost of hiring a single-semester instructor to teach a field f course, d_{tf}^N is the number of single-semester field f courses offered in semester t, and E_t^N is the residual share of the semester t instruction budget which is paid to instructors on single-semester contracts. This equation is similar to Equation 20 except that it excludes the implied preference terms $\gamma_f n_{tf}$, it uses the empirical linear budget constraint, and it imposes the counterfactual restriction that the university can only reallocate the portion of its budget paid to instructors

[35]Specifically, the formula is:

$$\epsilon_{tf} = \frac{\partial n_{tf}\left(\mathbf{d}_t\right)}{\partial d_{tf}} \times \frac{d_{tf}^N}{n_{tf}^N\left(\mathbf{d}_t\right)} \tag{42}$$

where d_{tf}^N is the number of field f single-semester courses offered in semester t and $n_{tf}^N\left(\mathbf{d}_t\right)$ is observed enrollment in single-semester field f courses in semester t. Figures in Panel B of Table A1 are field specific averages of elasticities across academic semesters.

on single-semester contracts. The goal of the equivalent cost exercise is then to solve for equivalent costs $\tilde{c}_f$ which imply that the solutions to Equation (43) are as close as possible to observed course offerings.

To solve for equivalent costs $\tilde{c}_f$, note that the system of first order conditions characterizing a solution to 43 if single-semester costs are given by $\tilde{c}_f$ is:

$$\left(\frac{1}{\tilde{c}_{f_1}}\right)\left[\frac{\partial V_t\left(\mathbf{d}_t^\star\right)}{\partial d_{tf_1}}\right]=\left(\frac{1}{\tilde{c}_{f_2}}\right)\left[\frac{\partial V_t\left(\mathbf{d}_t^\star\right)}{\partial d_{tf_2}}\right]\quad \forall f_1, f_2 \tag{44}$$

$$\sum_{f=1}^{F} d_{tf}^N c_f = E_t^N \tag{45}$$

Because this university's objective is to maximize student utility, optimal course offerings simply equate marginal utility per dollar of additional course offerings across all academic fields.

Rearranging and stacking fields and semesters yields:

$$\left(\mathbf{M}^1+\mathbf{M}^2\right)\tilde{\mathbf{c}}=\mathbf{ME} \tag{46}$$

where

$$\underset{(F-1,\,F-1)}{\mathbf{M}_t^1}\left(f_1,\ f_2\right)=\left(\frac{\partial V_t\left(\mathbf{d_t}\right)}{\partial d_{tf_1}}\right)\left(\frac{d_{tf_2}}{d_{tF}}\right)$$

$$\underset{((F-1)\times T,\,F-1)}{\mathbf{M}^1}=\begin{bmatrix}\mathbf{M}_1^1\\ \vdots\\ \mathbf{M}_T^1\end{bmatrix}$$

$$\underset{(F-1,\,F-1)}{\mathbf{M}_t^2}\left(f_1,\ f_2\right)=\begin{cases}\frac{\partial V_t(\mathbf{d_t})}{\partial d_{tF}} & f_1=f_2\\ 0 & f_1\neq f_2\end{cases}$$

$$\underset{((F-1)\times T,\,F-1)}{\mathbf{M}^2}=\begin{bmatrix}\mathbf{M}_1^2\\ \vdots\\ \mathbf{M}_T^2\end{bmatrix}$$

$$\underset{(F-1,1)}{\tilde{\mathbf{c}}\left(f\right)}=\tilde{c}_f$$

$$\underset{(F-1,1)}{\mathbf{ME}_t\left(f\right)}=\left(\frac{\partial V_t\left(\mathbf{d_t}\right)}{\partial d_{tf}}\right)\left(\frac{E_t^N}{d_{tF}}\right)$$

$$\underset{((F-1)\times T,\,1)}{\mathbf{ME}} = \begin{bmatrix} \mathbf{ME}_1 \\ \vdots \\ \mathbf{ME}_T \end{bmatrix}$$

This system of equations can then be inverted to derive the following expression for equivalent costs:

$$\tilde{\mathbf{c}} = \left(\mathbf{M}^1 + \mathbf{M}^2\right)^+ \mathbf{ME} \tag{47}$$

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, "Grade Inflation in General Equilibrium with Implications for Female Interest in STEM Majors," 2017.

Andrews, Rodney and Kevin Stange, "Price Regulation, Price Discrimination, and Equality of Opportunity in Higher Education: Evidence from Texas," 2016.

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.

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

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.

_ **and** _ , "Does cheaper mean better? The impact of using adjunct instructors on student outcomes," *Review of Economics and Statistics*, 2010, *92* (3), 598–613.

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.

Bound, John and Sarah Turner, "Cohort crowding: How resources affect collegiate attainment," *Journal of Public Economics*, 2007, *91* (5), 877–899.

_ , Michael F Lovenheim, and Sarah Turner, "Why have college completion rates declined? An analysis of changing student preparation and collegiate resources," *American Economic Journal: Applied Economics*, 2010, *2* (3), 129–157.

_ , _ , and _ , "Increasing time to baccalaureate degree in the United States," *Education*, 2012, *7* (4), 375–424.

Cardell, N Scott, "Variance components structures for the extreme-value and logistic distributions with application to models of heterogeneity," *Econometric Theory*, 1997, *13* (2), 185–213.

Carrell, Scott E and James E West, "Does professor quality matter? Evidence from random assignment of students to professors," *Journal of Political Economy*, 2010, *118* (3), 409–432.

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.

Deming, David and Chris Walters, "The Impacts of Price and Spending Subsidies on US Postsecondary Attainment," 2017.

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.

Dynarski, Susan, "Building the stock of college-educated labor," *Journal of Human Resources*, 2008, *43* (3), 576–610.

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.

Figlio, David N, Morton O Schapiro, and Kevin B Soter, "Are tenure track professors better teachers?," *Review of Economics and Statistics*, 2015, *97* (4), 715–724.

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

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.

Glaeser, Edward L, *The governance of not-for-profit organizations*, National Bureau of Economic Research, 2003.

Hemelt, Steven W and Dave E Marcotte, "The impact of tuition increases on enrollment at public colleges and universities," *Educational Evaluation and Policy Analysis*, 2011, *33* (4), 435–457.

Hoffmann, Florian and Philip Oreopoulos, "Professor qualities and student achievement," *Review of Economics and Statistics*, 2009, *91* (1), 83–92.

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

Jacob, Brian, Brian McCall, and Kevin M Stange, "College as Country Club: Do Colleges Cater to Students' Preferences for Consumption?," *Journal of Labor Economics*, 2015, *forthcoming.*

Kane, Thomas J, "Rising public college tuition and college entry: How well do public subsidies promote access to college?," 1995.

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.

Pope, Devin G and Jaren C Pope, "The impact of college sports success on the quantity and quality of student applications," *Southern Economic Journal*, 2009, pp. 750–780.

_ **and** _ , "Understanding college application decisions: Why college sports success matters," *Journal of Sports Economics*, 2014, *15* (2), 107–131.

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.

Sloan, Frank A, "Not-for-profit ownership and hospital behavior," *Handbook of health economics*, 2000, *1*, 1141–1174.

Stinebrickner, Ralph and Todd Stinebrickner, "Academic Performance and College Dropout: Using Longitudinal Expectations Data to Estimate a Learning Model," *Journal of Labor Economics*, 2014, *32* (3), 601–644.

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

Tabakovic, Haris and Thomas Wollmann, "The Impact of Money on Science: Evidence from Unexpected NCAA Football Outcomes," 2016.

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

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

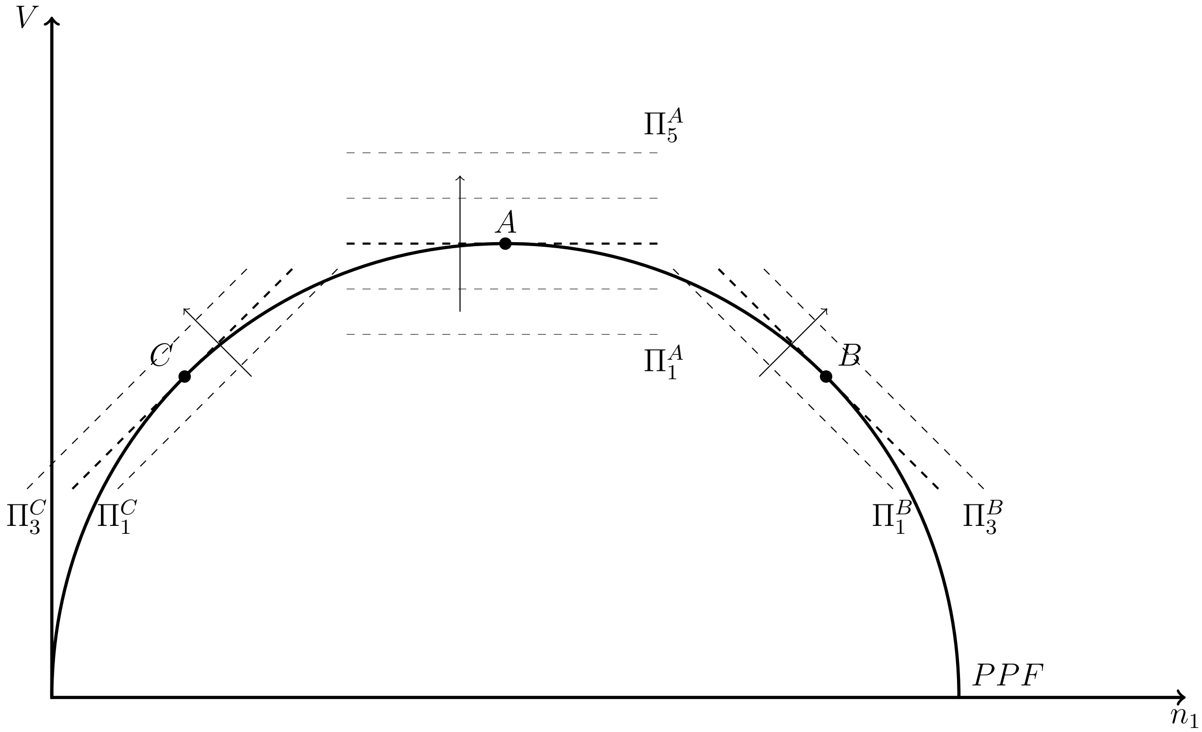
Turner, Sarah, "Going to college and finishing college. Explaining different educational outcomes," in "College choices: The economics of where to go, when to go, and how to pay for it," University of Chicago Press, 2004, pp. 13–62.

Turner, Sarah E, "For-Profit Colleges in the Context of the Market for Higher Education," in David W Breneman, Brian Pusser, and Sarah E Turner, eds., *Earnings from Learning: The Rise of For-Profit Universities*, SUNY Press, 2012.

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

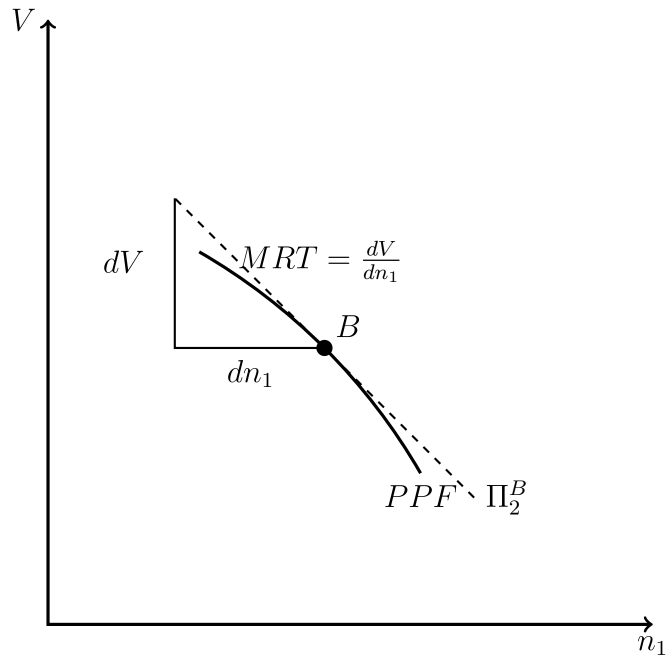
Vlieger, Pieter De, Brian Jacob, and Kevin Stange, "Measuring Instructor Effectiveness in Higher Education," in "Productivity in Higher Education," University of Chicago Press, 2017.

Figure 1: Optimal Course Offerings



The vertical axis is total expected student utility. The horizontal axis is expected number of students choosing courses in field 1 (the expected number of students choosing courses in field 2 is the complement). The solid semi-circle is a production possibilities frontier representing the frontier of outcomes which can be achieved given the university's constraints. Dashed line segments represent potential university indifference curves with payoffs increasing in the direction of the arrows. University A only values total expected student utility $\left(\gamma_1^A = 0\right)$ and offers courses to achieve outcome A. University B has institutional preferences to increase the expected number of students choosing courses in field 1 $\left(\gamma_1^B > 0\right)$ and offers courses to achieve outcome B. University C has institutional preferences to decrease the expected number of students choosing courses in field 1 $\left(\gamma_1^C < 0\right)$ and offers courses to achieve outcome C.

Figure 2: Revealed Institutional Preferences



This is Figure 1 zoomed in to focus on the tangency condition of university B. The derivative of the PPF at point B, or marginal rate of transformation (MRT), is given by the instantaneous change in total expected student utility relative to the instantaneous change in the expected number of students choosing courses in field 1 as the university marginally reallocates funds from field 1 to field 2. The instantaneous change in total expected student utility is given by the marginal effect per dollar of offering an addition field 2 course on total expected student utility minus the marginal effect per dollar of offering an addition field 1 course:

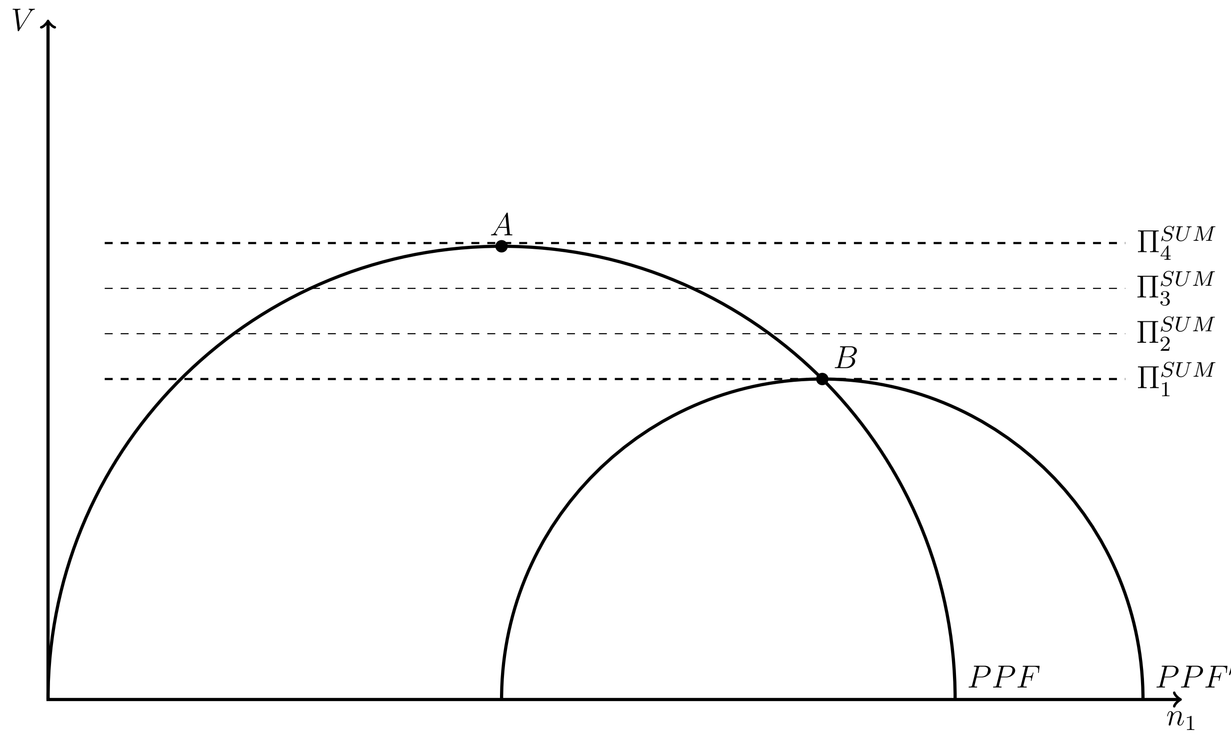
$$dV = \left(\frac{1}{c_2^N}\right)\left(\frac{\partial V}{\partial d_2}\right) - \left(\frac{1}{c_1^N}\right)\left(\frac{\partial V}{\partial d_1}\right)$$

The instantaneous change in the expected number of students choosing courses in field 1 is given by the marginal effect per dollar of offering an addition field 2 course on the expected number of students choosing courses in field 1 minus the marginal effect per dollar of offering an addition field 1 course on the expected number of students choosing courses in field 1.

$$dn_1 = \left(\frac{1}{c_2^N}\right)\left(\frac{\partial n_1}{\partial d_2}\right) - \left(\frac{1}{c_1^N}\right)\left(\frac{\partial n_1}{\partial d_1}\right)$$

This graphically demonstrates how marginal effects of spending can be used to solve for the slope of the indifference curves which rationalize why point B was optimal for this university.

Figure 3: Equivalent Costs to Make a Utility Maximizing University Offer Observed Courses



PPF' is the production possibilities frontier under "equivalent costs" which would induce a student utility maximizing university to offer courses producing observed allocation B. They are equivalent in the sense that going from preferences characterized by Π^A to preferences characterized by Π^B while holding PPF fixed in Figure 1 has the same effect on course offerings as going from PPF to PPF' while maintaining utility maximizing preferences characterized by Π_4^{SUM} in this figure.

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	33.1	64.7	52.6	25.2
Avg. intro sections per semester	210	259	165	88
Avg. intro enrollment per semester	5590	8833	4802	2330
Avg. intro enrollment per section	26.6	34.1	29.2	26.6
Intro section cost (25th pctile)	$5,766	$4,566	$5,184	$5,168
Intro section cost (Median)	$8,684	$6,088	$6,801	$7,012
Intro section cost (75th pctile)	$10,927	$8,781	$9,382	$11,407
Avg. ACT score	24.2	23.7	23.9	23.9
Avg. HS GPA	3.42	3.37	3.36	3.41
Share Female	58.0%	59.8%	57.9%	48.2%
Share Freshmen	43.9%	40.5%	40.2%	11.4%
Share Sophomores	27.9%	31.8%	34.7%	40.6%
Share Juniors	16.9%	18.1%	16.1%	35.5%
Share Seniors	11.2%	9.6%	9.0%	12.5%

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

STEM	Semester											
	F04	S05	F05	S06	F06	S07	F07	S08	F08	S09	F09	S10
Courses	34	30	33	33	34	32	33	33	33	34	33	35
Sections	194	175	221	179	235	208	248	211	222	207	219	201
Sections (%)	30%	28%	31%	28%	30%	29%	30%	28%	29%	29%	28%	29%
Enrollment (%)	27%	24%	28%	24%	26%	25%	27%	25%	26%	26%	26%	26%

Social Science	Semester											
	F04	S05	F05	S06	F06	S07	F07	S08	F08	S09	F09	S10
Courses	59	64	65	61	65	64	65	66	66	67	69	65
Sections	227	220	250	231	284	269	297	271	272	270	279	240
Sections (%)	35%	35%	36%	36%	36%	37%	36%	37%	35%	37%	36%	34%
Enrollment (%)	42%	43%	41%	42%	41%	41%	40%	41%	40%	41%	41%	40%

Humanities and Arts	Semester											
	F04	S05	F05	S06	F06	S07	F07	S08	F08	S09	F09	S10
Courses	47	48	49	52	51	50	55	53	56	54	56	60
Sections	146	146	148	148	169	158	182	172	182	160	187	176
Sections (%)	22%	23%	21%	23%	22%	22%	22%	23%	24%	22%	24%	25%
Enrollment (%)	21%	23%	21%	23%	22%	22%	22%	23%	22%	21%	23%	23%

Business and Occupational	Semester											
	F04	S05	F05	S06	F06	S07	F07	S08	F08	S09	F09	S10
Courses	26	26	26	24	26	24	25	25	26	24	25	25
Sections	90	81	85	77	95	86	98	87	96	85	91	82
Sections (%)	14%	13%	12%	12%	12%	12%	12%	12%	12%	12%	12%	12%
Enrollment (%)	10%	11%	10%	11%	10%	11%	11%	11%	11%	12%	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: Course Section Cost Regression

	Course Section Cost
STEM	5057.4
	124.9
Social Science	2816.9
	127.1
Humanities and Arts	3191.3
	152.2
Business and Occupational	4650.9
	198.1
Tenured	5595.1
	143.4
Tenure-track	5433.6
	161.2
Contracted non-tenure	3132.5
	135.0
Single semester	*omitted*
Course Section Observations	8857

Notes: Block bootstrapped standard errors (1000 iterations, sampling course sections) are in italics. Dependent variable 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. All course sections are categorized into either STEM, social science, humanities and arts, or business and occupational (the regression does not include a constant). As such, coefficient on STEM indicates that the predicted cost of offering a STEM course section with an instructor on a single semester contract is $5,057.40. Instructors on single semester contracts are hired to teach for one semester and have no explicit guarantee that their contracts will be renewed.

Table 5: Student Course Choice Parameters

	STEM	Social Science	Humanities and Arts	Business and Occupational
Intercept	0.473***	1.344***	0.543***	*omitted*
	0.032	*0.045*	*0.023*	
ACT Z-Score	0.155***	0.073***	0.127***	*omitted*
	0.013	*0.012*	*0.014*	
Missing ACT	-0.203***	-0.161***	-0.233***	*omitted*
	0.029	*0.026*	*0.030*	
GPA Z-score	0.003	-0.093***	-0.126***	*omitted*
	0.013	*0.013*	*0.015*	
Missing GPA	0.089***	0.162***	0.180***	*omitted*
	0.031	*0.029*	*0.032*	
Female	0.416***	0.525***	0.457***	*omitted*
	0.024	*0.021*	*0.025*	
Sophomore	-1.737***	-1.495***	-1.395***	*omitted*
	0.028	*0.026*	*0.029*	
Junior	-2.086***	-1.923***	-2.016***	*omitted*
	0.031	*0.030*	*0.034*	
Senior	-1.449***	-1.522***	-1.561***	*omitted*
	0.040	*0.037*	*0.042*	
Nesting Parameter ρ_f	0.680	0.547	0.642	0.461
	0.007	*0.011*	*0.008*	*0.008*

Notes: Block bootstrapped standard errors (1000 iterations, sampling student panels) are in italics. *** indicates significantly different from omitted category (normalized to zero) at 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.

Table 6: Marginal Effects and Implied Preferences

	Average Marginal Effect on Total Utility	Average Marginal Effect on Total Utility per Dollar	Implied Preferences
	(1)	(2)	(3)
STEM	1.484***	1.365***	0.611***
	0.033	*0.059*	*0.041*
Social Science	1.533***	2.533***	0
	0.035	*0.117*	
Humanities and Arts	1.541***	2.249***	0.086***
	0.031	*0.109*	*0.029*
Business and Occupational	1	1	1.114***
			0.068

Notes: Block bootstrapped standard errors (1000 iterations, sampling student panels for student parameters and course sections for costs) are in italics. Column 1 contains marginal effects of offering an additional course section in the specified field on total student utility. These are averaged across academic semesters and reported relative to the effect for a business or occupational course which is normalized to one. Column 2 divides marginal effects by the cost of hiring an instructor on a single semester contract to teach a course section in the specified field once again averaging across semesters and normalizing relative to the effect per dollar for a business or occupational course. Column 3 reports estimates of implied preference parameters γ_j with social science as the omitted field. Estimates quantify how much student utility the university is implicitly willing to sacrifice to move one student from a social science course to a course in the specified field.

Table 7: Utility Maximizing Course Offerings and Equivalent Costs

	Observed			**Utility Maximizing**		
	(1) Average Contracted Courses	(2) Average Single Semester Courses	(3) Average Field Enrollment	(4) Average Single Semester Courses	(5) Average Field Enrollment	(6) Equivalent Cost Differences
STEM	175.08	34.92	5589.75	0.00***	4821.49***	-17.01%***
				0.00	33.45	1.13
Soc Sci	234.58	24.58	8833.33	98.35***	9862.71***	45.62%***
				10.52	178.39	3.57
Hum and Arts	150.17	14.33	4801.83	14.69	4681.97	24.99%***
				8.72	169.01	3.18
Bus and Occ	81.00	6.75	2330.25	0.00***	2191.49***	-41.99%***
				0.00	24.52	2.08

Notes: Block bootstrapped standard errors (1000 iterations, sampling student panels for student parameters and courses for costs) are in italics. Columns 1-3 are the observed number of courses taught by instructors on long term contracts, the observed number of courses taught by instructors on single semester contracts, and observed field enrollments averaged across semesters. Column 4 reallocates the residual budget spent on instructors on single semester contracts to maximize student utility and column 5 reports estimated field enrollments under these utility maximizing offerings. In columns 4 and 5, *** indicates significantly different from observed values at 1% significance. Column 6 reports how much the cost of hiring a single semester instructor would need to change to induce a utility maximizing university to offer the observed single-semester courses reported in column 2. In other words, the implied preferences reported in Table 6 have the same effect on course offerings as changing costs by the percentages reported in column 6. In column 6, *** indicates significantly different from zero at 1% significance.

Table 8: Single Semester Course Offerings in Counterfactual Scenarios

		STEM	Social Science	Humanities and Arts	Business and Occupational
(1)	Observed state	35.79	20.55	12.15	9.73
	(predicted)	*0.15*	*0.36*	*0.24*	*0.29*
(2)	Reduce cost of STEM	53.38***	6.31***	5.00***	7.03***
	instructors by 5%	*0.30*	*0.53*	*0.28*	*0.24*
(3)	Increase all SAT scores and	53.19***	2.71***	4.25***	7.04***
	GPA by 1/3 of a std dev	*0.60*	*0.57*	*0.54*	*0.42*
(4)	Make all students	24.09***	49.91***	18.82***	0.14***
	female	*1.95*	*2.39*	*2.04*	*0.12*

Notes: Block bootstrapped standard errors (1000 iterations, sampling student panels for student parameters and courses for costs) are in italics. Row 1 is the average number of course sections taught by instructors on single semester contracts predicted by the estimated model in the observed state. Rows 2-4 are the average number of course sections taught by instructors on single semester contracts in counterfactual states. In rows 2-4, *** indicates significantly different from row 1 at 1% significance.

Table 9: Field Enrollments in Counterfactual Scenarios

		STEM	Social Science	Humanities and Arts	Business and Occupational
(1)	Observed state	5619.26	8787.54	4779.57	2371.28
	(predicted)	*36.11*	*43.37*	*29.19*	*26.78*
(2)	Reduce cost of STEM	5968.65***	8566.8***	4667.73***	2354.48
	instructor by 5%	*37.93*	*42.52*	*28.86*	*26.23*
(3)	Increase all SAT scores and	5825.93***	8676*	4749.75	2305.98*
	GPA by 1/3 of a std dev (PE)	*39.48*	*44.91*	*30.54*	*26.27*
(4)	Increase all SAT scores and	6235.6***	8369.65***	4615.37***	2337.03
	GPA by 1/3 of a std dev (GE)	*46.08*	*45.5*	*33*	*28.64*
(5)	Make all students	5563.19	9201.64***	4865.66*	1927.17***
	female (PE)	*40.9*	*49.31*	*36.2*	*26.5*
(6)	Make all students	5288.73***	9550.23***	4884.21	1834.49***
	female (GE)	*70.75*	*69.65*	*61.13*	*26.46*

Notes: Block bootstrapped standard errors (1000 iterations, sampling student panels for student parameters and courses for costs) are in italics. Row 1 are average field enrollments predicted by the estimated model in the observed state. Rows 2-6 are average field enrollments in counterfactual states. In rows 2-6, ***/**/* indicates significantly different from row 1 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.

Table A1: Elasticities of enrollment

Panel A: Elasticity with respect to instructor salaries (log-log regression)

	log(Enroll)	log(Enroll)	log(Enroll)	log(Enroll)
log(Instructor Salary)	0.0888***	0.162***	0.0220***	0.0534***
	0.0112	*0.0388*	*0.0064*	*0.0191*
Course fixed effects	Yes	Yes	Yes	Yes
Single-semester only	No	Yes	No	Yes
Enrollment>5	No	No	Yes	Yes
Observations	8,556	873	8,280	819
R-squared	0.499	0.564	0.609	0.589

Panel B: Elasticity with respect to spending on course offerings (model based)

	STEM	Soc Sci	Hum and Arts	Bus and Occ
Elasticity	0.4963***	0.3229***	0.4999***	0.3844***
	0.0055	*0.0064*	*0.0064*	*0.0067*

Notes: Standard errors in italics. *** denotes p-value for test that coefficient is not equal to zero is $p<.01$. Panel A contains estimates of the elasticities of enrollment with respect to spending on instructor salaries which are estimated using the log-log regression specification given in Equation 41. Panel B contains estimates of the elasticities of enrollment with respect to spending on course offerings which are derived from estimates of the nested logit student choice model. In panel A, single-semester means only courses taught by instructors on single semester contracts are included.