UNIVERSITY OF ALBERTA
FACULTY OF ARTS
Department of Economics

Working Paper No. 2020-03

Competition in Higher Education

Michael Kaganovich
Indiana University

Sinan Sarpca
Koc University

Xuejuan Su
University of Alberta

April 2020

Competition in Higher Education

Michael Kaganovich* Sinan Sarpça† Xuejuan Su‡

April 2020

Abstract

The structure and functioning of the market of higher education in the United States possess distinctive if not puzzling features such as the wide spectrum of institutional arrangements and sources of funding, stark segmentation in levels of selectivity and instructional resources, and high variance in tuition pricing across and within institutions, including price discrimination based on merit and ability to pay. At the same time, many fundamental questions, including what defines the quality of higher education and explains its (growing) cost continue to be debated. The Chapter surveys theoretical analyses addressing this range of issues.

JEL codes: I210, I220, I230, D400, J240.

*Michael Kaganovich, Department of Economics, Indiana University, Bloomington, IN 47405, USA. Email: mkaganov@indiana.edu.

†Sinan Sarpça, Department of Economics, Koç University, Istanbul, Turkey. Email: ssarpca@ku.edu.tr.

‡Xuejuan Su, Department of Economics, University of Alberta, Edmonton, AB, Canada. E-mail: xuejuan1@ualberta.ca.

1 Introduction

Postsecondary education has seen dramatic expansion in the post-World War II era in the United States. According to the National Center for Education Statistics, the overall enrollment in the US degree-granting postsecondary institutions has increased between 1959-60 and 2016-17 academic years from 3.6 to 19.8 million or by 445% (see Table 303.10 in NCES 2020-009). We provide some detailed statistics of the facets of this expansion in Section 2 immediately below.

As documented by Goldin and Katz (1999), this expansion was a continuation of the trend that started half-a-century earlier, along with the development of institutional characteristics of the provision of higher education in the United States. Goldin and Katz (2008) and Acemoglu (1998, 2009) further demonstrate that the sustained expansion of college enrollment can be explained in the context of the intertwined dynamics of demand for and supply of skilled workforce recurrently fueled by and feeding into temporary relative compressions and increases of the skilled wage premium. In a Goldin and Katz expression, the changes in skill supply were the tail wagging the wage-premium dog. Some significant boosts to the expansion were intermittently driven by exogenous factors, including changes in immigration policy and other demographic and social developments and government policies, such as GI Bill of Rights and Vietnam War deferments, and the institutional ones, such as the expansion of public university system. An essential and growing part, especially since mid-XX century, of this story of spiraling growth of demand and supply of skill belongs, according to *op cit.* and extensive related literature, to the emergence of endogenous skill-biased technological change. According to this reasoning, expanding supply of skill and the resulting temporary reductions in its relative price fueled skill-complementing technological developments, which then propelled a wave of growth in skill premium and, in turn, stimulated the next wave of growth of its supply, and so forth.

Goldin and Katz (2008) volume devotes much attention to documenting and analyzing the evolution of the American system of higher education, which responded to the growing demand for skill and enabled the unprecedented expansion of college enrollment. They underscore its unique consumer orientation expressed, particularly, in the variety of institutions and their development into a competitive industry along with the rise, in the first third of the XX century, of the system of state supported public universities, which offered a dramatically more affordable access. This fact was in turn responsible for rapid expansion of public sector, reaching 70% of overall enrollments among 4-year institutions by 1970. They further document that although for the 1950-1980 period of rapid enrollment expansion, tuition growth at public universities was in line with the growth of median family income, the rise of the former decisively sped past the latter from that point on.[1] The rapid rise of tuition and fees out of proportion to incomes

[1] A survey of some data we present in Section 2 helps relate this to the fact that state appropriations, despite their continued growth, no longer kept the pace with enrollment expansion. This resulted in their

since 1980s is a striking and much debated phenomenon of American higher education. See, for instance, Ehrenberg (2000), who also analyzes the expanding mechanism of decoupling of sticker-price tuition and the actual tuition charged to individual students in the form of financial aid, which effectively plays a role of price discrimination based on ability or willingness to pay.[2]

Given a well-developed historical outlook on the evolution of American market of higher education and the economic processes shaping it, this Chapter will focus on the economic theories, which analyze the structure and functioning of this market. This need is well justified by the unusual multitude of puzzles an economist is confronted with in attempting to understand the workings of this one of the largest of industries, whose economics Winston (1999) aptly labeled as "awkward".[3] One of the challenges continually debated in economic thought is defining the nature of the industrys product. As will be seen, the question of defining the value that a university adds to a consumer is of practical consequence for understanding the nature of competition and market structure in the higher education industry. We shall first examine alternative concepts of university value added in Section 3 and then, in subsequent sections, present the corresponding models of the higher education marketplace they give rise to.

One of the most influential such concepts, due to Spence (1973), is that of signaling value of college education under asymmetric information about student aptitudes. It suggests that college may not add intrinsic educational value, but rather play an intermediary role in the job market by certifying the aptitudes of graduates to potential employers. Costrell (1994) and Betts (1998) advanced this idea to develop a theory of educational standards, which serve as sorting devices or pre-college preparation targets, which can incentivize students to study in order to gain admission but have no effect on the intrinsic gains to human capital while in college. The concept of the value of college education as a signal of quality also offered a powerful explanation to the depressing effect of the higher educations expansion on the wages of workers without college degrees. Goldin and Katz (2008) among others underscored this phenomenon as a substantial factor contributing to the growth of college premium, hence creating positive feedback loop toward increased college enrollment.

The signaling paradigm proved further fruitful in a more recent literature to helping explain the differentiation of academic standards across universities and address the

shrinking share in public university budgets, which was being compensated by the growth of tuitions share.

[2]Fillmore (2019) specifically focuses on tuition price discrimination as a factor enhancing competition in the higher education market. He assesses, in particular, that both are substantially amplified by colleges' access to families' reported ability-to-pay information – by means of the financial disclosure FAFSA form required of all applicants for college "financial aid" consideration, the key vehicle of the effective tuition price discrimination.

[3]In this, he particularly referred to a growing predominance of tuition price discrimination in the form of financial aid based on a combination of merit and ability to pay, and observed that exceedingly rapid rise of sticker price tuition at private non-profit universities was a way to expand their ability to engage in price discrimination.

recently observed phenomenon of intensifying divergence of college selectivity in the U.S. (Hoxby, 2009). This implies that universities differentiate their products to serve different segments of student population. Most of this chapter indeed surveys alternative approaches to modeling quality of education and its differentiation across universities. Signaling tradition in particular emphasizes school reputation as the key characteristic of its quality. According to recent work, reputation is gained through *ex post* evidence, i.e., the *signal*, of the average quality of the university's graduates, so the universities compete by means of admission standards. The resulting sorting of students across different colleges leads to signal differentiation according to endogenous creation of the respective student bodies, those more "reputable" and others less so.

An implication of the above signaling approach is that any student can benefit from being admitted to a university of higher quality. For instance, if a student of relatively low ability gains admission to selective school thanks to lucky draw at an admission test, he or she will benefit from the superior signal to the job market. Recent growing empirical literature (see, e.g., its comprehensive survey by Arcidiacono and Lovenheim, 2016) challenges this premise, particularly as applied to the evaluation of the outcomes of affirmative action programs. This evidence suggests that the educational benefit a student derives from attending a university depends on the quality of the match between the two. Attending selective university with high standards can be beneficial or detrimental to a particular student depending on the fit between the student's preparation and the "pitch" at which instruction is delivered at the college, such as the rigor of its curriculum. In short, this argument suggests that a university, which offers high quality to one set of students, is not the best choice for others. Specifically, for a given curricular standard, the returns to education would tend to be low for insufficiently prepared students as well as the "overqualified" ones, as opposed to students whose aptitude is the right match for this curriculum. This approach thus focuses on defining product differentiation in higher education marketplace horizontally, as a menu of curricular standards, which cater to different segments of potential student population, so students can choose the best available match.

In Section 5, we review the recent work stemming from the two approaches outlined above to modeling university competition in terms of differential academic standards, both offering explanations for the evidence of increasing variation in the selectivity of institutions.

A distinct paradigm of the college value added that proved especially fruitful in terms of developing theories of higher education market structure belongs to Rothschild and White (1995). According to this idea, colleges do add human capital for their students; furthermore, cross-student spillover of knowledge is a central factor in this production, in addition to instructional and other inputs. Thus, the key conceptual innovation of Rothschild and White is that students can enter as inputs in education production, not just its output. This implies that the distribution of quality of its students (the "peer group") is an essential measure of quality of an educational institution, and that therefore some students can be more valuable to a university than others. Notably, this paradigm (uniquely among its competitors) offered a consistent theoretical foundation for modeling tuition pricing in higher education, with its pervasive phenomenon of price discrimination mentioned above. More broadly, the literature that emerged on this conceptual basis developed models of market segmentation of the higher education based on students' intellectual ability (consistent with their roles as "inputs" in education production) as well as their ability to pay. Importantly, it was able to bring these models to data in a comprehensive fashion. We review the state of this literature in Section 4.

Recent advances in the economics of higher education go beyond the explorations of schooling attainment and college premium in general and draw attention to students' choices of the fields of study. This is well motivated by the evidence that the choice of college major is becoming a dominant determinant of the variation in career earnings. In other words, the variation in college *major* premia is overtaking the average college premium (see James, 2012, Hershbein and Kearney, 2014). Furthermore, Altonji et al. (2015) and Kirkeboen et al. (2016) affirm the above even controlling for the quality of peers and the higher education institution overall, and find that the effect on earnings from attending a more selective institution is dominated by the payoffs to a field of study. Accordingly, the literature also provides strong evidence that students' expectation of future earnings associated with college majors is a significant positive determinant of their decisions to choose among them: see Berger (1988), Montmarquette et al. (2002), and Arcidiacono (2004) among others. Student decision-making process about choosing a field of study is often characterized, in campus jargon, as that of "shopping" for majors and classes by students, at least implicitly acknowledging the elements of a marketplace in a campus operation. The flip side of this valid view is that major programs, the units of a university that offer classes, are active participants as "vendors" in this marketplace and often act as competitors for students.

Indeed, departments and their faculty are able to affect the selection of students into their programs by setting degree prerequisites and curricular requirements as well as the grading standards, given that the quantity and quality of students pursuing a major are consequential for its faculty, particularly in terms of the departments position in the intra-university allocation of resources. The latter fact is especially pronounced in universities, which adopted the Responsibility Center Management (RCM) system of budgeting (see Strauss and Curry, 2002). Based on the above argument, Achen and Courant (2009) conjecture that departments can and do use their grading policy as an instrument in the intra-university marketplace to manage their enrollments: to reduce congestion in classes offered by a department, or to counteract falling demand for it. In Section 6, we present a model of intra-university competition between major programs and analyze its implications for the trends in academic standards across disciplines. In concluding Section 7, we briefly survey the recent literature looking into potential disruption of the present competitive market model of U.S. higher education stemming from new technologies of online instruction.

2 Expansion of Post-Secondary Education since 1960s: Some Facts and Figures

The following basic statistics focus on the changes in U.S. higher education over the last six decades. Unless otherwise specified, our data source is NCES (2020).[4] Specifically, most of the figures reported below survey the changes that occurred over the period of 1959/602016/17.

2.1 Demand (student enrollment) side of the higher education market

The number of students enrolled in degree-granting postsecondary institutions has increased from 3.6 to 19.8 million, or by 445% (see Table 303.10 in NCES, 2020). Of those, the enrollment growth in public institutions (by 569%) substantially exceeded that in private institutions (361%). Even accounting for population growth, postsecondary aggregate contemporaneous enrollments tripled over the period, increasing from 2% to 6% of the U.S. population. Starting in 1967, enrollment in private institutions was further broken down into private non-profit versus for-profit schools, acknowledging the emergence and growth of the latter sector. From 1967 to 2016, while the fall enrollments in all private institutions grew by 151%, in private non-profits they grew by 97%, while in private for-profits growth was 5,353%. This reflected a significant structural change: private for-profit enrollments accounted for merely 1.0% of all enrollments in private institutions in 1967, but reached 22.5% in 2016. The following Table 1 compiled from the aforementioned source details the evolution of these dynamics and shows, *inter alia*, that the expansion may be showing recent signs of stabilization.

Table 1: Enrollment in postsecondary degree-granting institutions (in thousands)

Year	Total	Public	Private		
			All	Not-for-profit	For-profit
1959	3,640	2,181	1,459	n/a	n/a
1969	8,005	5,897	2,108	2,088	20
1979	11,570	9,037	2,533	2,462	71
1989	13,539	10,578	2,961	2,731	229
1999	14,850	11,376	3,474	3,055	419
2009	20,314	14,811	5,503	3,768	1,735
2016	19,841	14,583	5,258	4,078	1,180

It is important to note that only a part of the demand growth was driven by young

[4]Digest of Education Statistics 2018, National Center for Education Statistics (2019), NCES 2020-009, `https://nces.ed.gov/programs/digest/d18/`. Unless otherwise indicated, other references to data tables made in this Section likewise refer to this source.

first-time students, i.e., predominantly recent highschool graduates. Among recent high school completers (ages 16 to 24), college enrollments increased from 0.8 million in 1960 (the earliest data available, representing 45.1% of the 1.7 million high school completers) to 2.2 million in 2016 (69.8% of 3.1 million high school completers), or by 189% (Table 302.10). At the same time, enrollments of first-time students increased from 0.9 million in 1960 to 2.9 million in 2016, or by 312%, which is also lower than the growth in all enrollments (Table 305.10). This suggests that a significant factor of the expanding enrollments, in addition to secular growth in the share of high-school graduates choosing to attend college, was also given by the expansion of less traditional cohorts of college attendees, such as: (a) older first-time students, outside the regular school age for higher education (over 24), (b) individuals pursuing multiple degrees, and (c) individuals making multiple attempts at higher education.

The following Table 2 (compiled from NCES, 2020, Table 103.20) offers a more detailed picture of this evolution.

Table 2: Enrollment in postsecondary education by age group (in %)

Year	18-19 years old			20-24 years old			25-29	30-34
	Total	In basic education	In higher education	All	20-21	22-24	years old	years old
1959	36.8	n/a	n/a	12.7	18.8	8.6	5.1	2.2
1969	50.2	11.2	39.0	23.0	34.1	15.4	7.9	4.8
1979	45.0	10.3	34.6	21.7	30.2	15.8	9.6	6.4
1989	56.0	14.4	41.6	27.0	38.5	19.9	9.3	5.7
1999	60.6	16.5	44.1	32.8	45.3	24.5	11.1	6.2
2009	68.9	19.1	49.8	38.7	51.7	30.4	13.5	8.1
2016	69.5	19.0	50.5	39.0	55.5	28.8	13.2	6.4

2.2 Supply (or "producer") side of the higher education market

The post-World War II expansion characterized both the intensive (average size) and extensive (number of institutions) margins. It is important to note that according to Goldin and Katz (2008) the entry barriers for new universities rose at the turn of the XX century, as the cost structure increasingly favored larger institutions, implying that most of the expansion was bound to occur on the intensive margin. The second half of the century, however, did feature substantial net growth in the number of institutions. In the period between 1959/60 and 2016/17, the total number of degree-granting postsecondary institutions increased from 2,004 to 4,360, i.e., net growth by 116%.[5] In particular, the

[5]Note that the reported growth is net of exits from the industry. From 1969/70 (the earliest year of this data reported by NCES) through 2016/17, 922 postsecondary institutions have closed their doors.

rate of increase was larger for 2-year institutions (from 582 to 1,528, or 163%) than that for 4-year institutions (from 1,422 to 2,832, or 99%), suggesting that the extensive expansion was stronger in the lower segment of the market (Table 317.10).

The data differentiating private non-profit and private for-profit institutions became available starting in 1976/77. From that point through 2016/17, while the number of public institutions grew modestly from 1,455 to 1,623 and the number of private non-profit institutions grew from 1,536 to 1,682, the number of private for-profits jumped from 55 to 1,055 (Table 317.10). In other words, the for-profit sector featured most of the extensive growth, which was, by contrast, modest among non-profit universities and colleges, both public and private.

The estimated number of total faculty went from 380 to 1,546 thousand between 1959 and 2016, a 306% growth (Table 301.20). Although some of this growth included part-time faculty, the breakdown available starting in 1970 shows growth in full-time faculty by 121% (along with growth in the part-time faculty by 605%) between 1970 and 2016. Further, over this period, growth in the number of faculty is larger for 2-year institutions (by 280%) than that for 4-year institutions (by 214%) (Table 315.10).

In terms of degrees conferred, over the period from 1959/60 to 2016/17 the number of bachelors, masters, and doctoral degrees grew from 1,065 to 2,942 thousand, or by 176% (of those the graduate degrees grew by 261% and bachelors degrees by 147%), whereas the number of associates degrees conferred increased from 206 to 1,006 thousand, i.e., by 388% (Table 318.10). Notably, this dynamics exhibits a noteworthy U-shaped pattern, with lower and higher levels of attainment featured stronger growth than that in the middle (i.e., at the bachelor level).

2.3 Higher education finance

Measured in constant 2017-2018 dollars, total revenue in postsecondary education increased from 49 to 664 billion, or by 1,259% over the period from 1959/60 to 2016/17. Total expenditure grew from 47 to 597 billion, and the market value of endowment funds grew from 45 billion to 612 billion (Table 301.20). For comparison, real GDP (measured in chained 2012 dollars) increased over the same period from \$3.3 trillion to \$18.0 trillion, or by 453%. Thus, as a share of real GDP, the aggregate revenue of postsecondary education sector increased by 146%.

Note that total revenue includes funding from both public and private sources, the latter of course including tuition payments. Although systematic historical data on the dynamics of these components is scarce, the Digest of Education Statistics 2008 (NCES,

Remarkably, a large share of these exits occurred in most recent years: 112 in 2016/17, 66 in 2015/16, and 54 in 2014/15. (table 317.50) Furthermore, 2-year institutions saw more exits (520) than 4-year institutions (402), and private institutions saw far more exits (871) than public institutions (51). Starting in 1990, the data on private institutions was further broken down into private non-profits and for-profit institutions. From 1990/91 to 2016/17 of the 629 private institutions that exited, there were 444 private for-profits and 185 non-profits.

2009) indeed reported the following statistics over the period from 1980/81 to 2000/01, showing an overall decline in federal and state funding as a share of the current-fund revenue of public institutions (Table 349). Most notably, within the short two decades, share of funding from state governments dropped from 45.6% to 35.6%. The difference was made up, in part, by increased role of tuition and fees, whose share grew from 12.9% to 18.1% over this period (Table 333.10). Note that the increasing importance of tuition and fees in public institutions reflects not only the enrollment growth, as discussed before, but also sharp tuition increases. Measured in constant 2017/18 dollars, the average undergraduate tuition and fees for full-time students in public institutions doubled from $1,819 in 1980/81 to $3,631 in 2000/01, and over a longer time period, nearly quadrupled from $1,883 in 1963/64 (earliest data available) to $6,972 in 2016/17 (Table 330.10).

For private non-profits, the share of tuition and fees in total revenue increased from 24.6% in 1999/2000 to 39.5% in 2015/16; of those, the increase at 4-year schools was from 24.4% to 39.3%, whereas at 2-year institutions it went from 43.0% to 80.2%, i.e., to near-complete budgetary reliance on tuition and fees. A noteworthy distinction of private institutions is a substantial role of private gifts in their financial model, which averaged 13.4% over this period. Over the period 2007/08 2015/16, for which the data is available for both private and public institutions, their share averaged 14% of total revenue in all private institutions (Table 333.40) but less than 3% of total revenue in public institutions (Table 333.10).

2.4 Selectivity of Colleges

Hoxby (2009) measures college selectivity by the average standardized pre-college aptitude test (SAT or ACT) score of their students, translated into current national percentiles of entrance exam takers, which thus could be viewed as representing the absolute exam performance on a stable metric. Using data from various college guides from 1962 to 2007, she found that the top 10 percent of 4-year colleges in the U.S. had become substantially more selective since 1962, while at least the bottom 50 percent of colleges have become substantially less selective. To reiterate, this phenomenon is defined, according to the methodology, by the diverging dynamics of the average aptitude of students between more and less selective colleges. Hoxby (2009) attributes this increased stratification in student sorting across colleges to the greater nationalization of what used to be local markets for college education. At the same time, she found that the diverging selectivity across colleges was also accompanied by the growing inequality in educational resources provided at colleges, with students at more selective institutions enjoying disproportionate growth in quality of education, judging by this metric. More specifically, while "student-oriented resources" were similar at colleges in the data set regardless of selectivity levels in the 1960's, their growth was strongly correlated with the level of selectivity: going from the average 7 percent annual growth rate of real resources per student at the least selective colleges to 13 percent at the most selective ones, resulting in a notable right skewness in the distribution. Furthermore, real subsidies per student (i.e., the difference between the cost of resources and tuition charges) exhibit similar pattern with the average annual growth rate between 7 and 10 percent for groups of less selective colleges, and 25 percent for the most selective ones. This particular finding is consistent with the data on the changes in the composition of university funding we discussed above, whereby the share of tuition revenue shrinks in the operating budgets of selective private schools, while growing at public universities.[6]

Bound et al. (2009) who study a more recent time period from 1986/87 to 2003/04 affirm increasing segmentation of the market, with particularly strong rise in selectivity of the most selective universities. When it comes to the changes in the compositions of student body across the spectrum of 4-year colleges, they focus on the dynamics of aptitude test scores of students in college-specific 75th percentiles. They find that these characteristics of student body, while being markedly higher at more selective colleges, have exhibited growth in all categories of colleges (albeit still higher at the more selective end) over the period, suggesting a degree of selectivity growth even among the less selective universities. This finding is not in contradiction with Hoxby (2009), because of its focus on students at the higher end of aptitude, rather than on central tendency for a college. Furthermore, we find it to be important in potential relation to the recent evidence of the growing student sorting across fields *within* colleges, responding to the developing dominance of the field of study as a factor in returns to college (see Altonji et al., 2015, Kirkeboen et al., 2016, and our discussion in the previous section). Indeed, as pointed out by Arcidiacono et al. (2016), moderately able students improve their chances to graduate in a more lucrative major by trading down in the selectivity ranking of the university they attend.

An important question stemming from the evidence of growing selectivity, particularly at the high end of the distribution, of U.S. universities is its implication for the level of competition in the higher education market. It suggests a possibility that the overall growth of competition may co-exist with growth in market power of the more selective universities. This question and further evidence are discussed in detail in Section 4 of the Chapter.

[6]The market structure of U.S. higher education, where public universities perform the role of ensuring greater access and affordability while the elite layer of the system with superior resources and best-prepared students is largely occupied by private institutions, is strikingly distinct from most other countries. Indeed, in internationally prevalent higher education systems, elite universities, superior in instructional resources and quality of students, are predominantly publicly funded while the private universities tend to serve the less selective segment of the market. These features invite distinct lines of inquiry, such as into the regressivity in the distribution of public subsidy (see, e.g., Psacharopoulos, 1977) and its implications for access and selectivity (see Eckwert and Zilcha, 2020, and Del Rey and Estevan, 2020, for recent theoretical analyses).

3 What do Universities Produce and How Do They Do It

In this Section, we shall survey three main paradigms developed in the literature, which help analyze the nature and ingredients of the value added produced by universities (i.e., the "returns to college") and explain the quality differentiation among them. We'll use these introductions as jumping-off ground for discussing the corresponding alternative characterizations of higher education market in Sections 4 and 5.

3.1 A signaling model of educational standard

The "signaling" paradigm originating from the seminal contribution by Spence (1973) explains the value of colleges as labor market intermediaries helping resolve informational asymmetries between job applicants and employers. It posits that colleges can attract students without necessarily engaging in intrinsic creation of human capital. This paradigm led to the emergence of the concept of academic standard (Costrell, 1994, Betts, 1998). Their modeling approach, laid out in some more detail below, is motivated by the following assumptions:

– employers cannot distinguish well between the levels of productivity of individual college graduates; similarly, they are unable to make such distinctions among prospective workers without college degree;

– they do have the information about the average (or an analogous aggregate characteristic of) productivity of college graduates as well as of the pool of those lacking college education;

– these aggregate productivity levels are determined by curricular standards ensured by the college(s) and the distribution of individual characteristics, such as abilities, of those with and without college degree;

– in this environment, since the fact of college graduation (or lack thereof) is all that an employer can tell about the student, students will decide to acquire higher education (or not) depending on those average returns and the individual cost of meeting the college academic standard.

Let student i's human capital attainment be given by a function $\pi_i = h(1L_i, a_i)$, concave and increasing in each of the arguments, namely, the student's effort (time endowment less leisure L) and exogenous ability, which implies that more effort is required of student of lower ability to reach the same level of achievement. Ability is distributed with a given CDF $F(a)$ on an interval $[\underline{a}, \overline{a}]$. Suppose the college sets a passing standard π_s required for graduation. Based on the above assumptions, the wage of those who end up obtaining the degree will be proportionate to the average attainment, so the "skilled" and "unskilled" wage rates are given, respectively, by

$$w_s = \alpha E(\pi_i | \pi_i \geq \pi_s), \quad w_u = \alpha E(\pi_i | \pi_i < \pi_s),$$

where α is an exogenous positive scale coefficient.

Students' choices are determined by individual utility function $U(L_i, w_i)$, increasing and concave in both arguments. A student who is able and willing to pass the standard will have no incentives to exert efforts to achieve beyond it. However, students of high enough ability may exceed the standard effortlessly. This will be the case for those whose individual ability exceeds the level a^{**} defined by the zero effort corner solution $\pi_s = h(1, a^{**})$. For the rest of the college attendees, the decision will be an interior (in terms of effort) optimum $\pi_s = h(1-L_i, a_i)$. It can be shown that there is a cut-off ability level a^* such that students of ability at or above it will decide to attend college while those with ability below it will not, and hence apply zero effort.

What happens if colleges raise educational standard? Since $\frac{dw_s}{s\pi_s}|_{a>a^*} = \frac{\alpha(F(a^{**})-F(a^*))}{1-F(a^*)} + \frac{\partial w_s}{\partial a^*}\frac{\partial a^*}{\partial \pi_s}$, it is obvious that $\frac{\partial w_s}{\partial a^*} > 0$ and it can also be shown that in any stable equilibrium $\frac{\partial a^*}{\partial \pi_s} > 0$. Thus, educational achievement will increase among college graduates, and so will their wage: $\frac{dw_s}{s\pi_s}|_{a>a^*} > 0$. The wage of unskilled workers will also rise. Indeed, since $\frac{\partial a^*}{\partial \pi_s} > 0$, the average ability and hence the average effortless human capital attainment among the unskilled will rise due to the change in the group's composition (as the ability cutoff for going to college increases, fewer people will enroll). The only workers who will see their wage drop as a result of the change are those at the margin who would have otherwise chosen to attend college and now choose to no longer meet the degree standard.

So, how should the higher education standard be set? Costrell's and Betts' papers analyze alternative approaches to setting the "optimal" standard, including those of (i) egalitarian social planner, who maximizes a concave aggregate (planner's) utility over workers incomes and (ii) the planner concerned with the aggregate efficiency, hence maximizing the aggregate income. It can be shown that the comparison of preferred standards based on equity and efficiency may go either way. This is due to the fact of a group's composition affects the group's wages. This is indeed a substantive point: when college enrollment expands significantly, this may hurt wages of the unskilled substantially, because of the perceived lower average human capital level of this group. This depressing effect of expanding enrollments in higher education on the wages of workers without college degrees is well known. It is underscored by Goldin and Katz (2008), in particular, as a factor in the growth of college premium, hence creating positive feedback loop for increased college enrollment. The fact that this is also well understood can be credited to the signaling paradigm.

3.2 Standard differentiation across institutions and their diverging selectivity

The idea of academic standard can be taken further as an institution-specific characteristic, which can differ across universities. This point has received new attention in recent literature as a basis for characterizing the industrial organization of higher education through product differentiation to help explain the differences in the levels of selectivity of universities, which as we discussed above have exhibited continued divergence. According to this characterization of the higher education market, universities compete by differentiating their standards to serve different segments of student population.[7]

A novel line of reasoning explaining the role and the workings of quality (hence vertical) differentiation in the signaling framework was recently advanced by MacLeod and Urquiola (2015) who emphasize school reputation as the key characteristic of its quality. Reputation is gained through *ex post* evidence, i.e., the signal of the average quality of the university's graduates, which is in turn ascertained by means of admission standards. In their model, in the Spence-Costrell-Betts tradition, information about individual productivity is imperfect. As a result, employers must rely on the mean productivity of a pool of graduates as their shared characteristic. The focus now, however, is on sorting of students across different colleges, which leads to signal differentiation according to endogenous creation of the respective student bodies, more or less "reputable". Accordingly, this creates incentives for students to gain admission to the best university possible (i.e., with as able peers as possible, despite the absence of substantive learning spillovers from peers in the model), and for colleges to select students via admission standards accordingly. MacLeod and Urquiola term this mechanism of favorable selection the "anti-lemons" principle since, unlike in the case of Akerlof's (1970) "Lemons", it features the exclusion of inferior participants. In their model, individual ability a_i of potential student i is not observed and can only be estimated through a noisy universal college admission test. Students are (or are not) admitted to a university based on observable test outcome $\tau_i = a_i + r_i + \epsilon_i^\tau$ where r_i is individual effort to prepare to the test, which does not contribute to skill, and ϵ_i^τ is the error term. If student i graduates from college s, she will possess skill level $\theta_i = a_i + e_i + v_s$ where e_i is the individual effort in college, which is productive and happens to be uniform in equilibrium for all attending college s, and v_s is the colleges value added. As a result, individual's post-graduation wage is positively affected by the mean ability of her college peers (notably, without any direct knowledge spillovers from them). Therefore, all students in college s have an interest in higher minimum admission test score τ_s, assuming they get admitted.[8]

[7]De Fraja and Iossa (2002) study intercollegiate duopoly competition in terms of quality as well as geographic proximity with school quality determined by its admission standard, whereas any admitted student benefits from a higher standard, as long as he meets it. Eisenkopf and Wohlschlegel (2012) also analyze a duopoly model where colleges offer distinct standards. Likewise in line with the signaling literatures approach, weaker students do prefer lower curricular standards, because the study cost of attaining high standards is excessively high for such students.

[8]The fact that students can misjudge their true ability based on test results is inconsequential in this

Hence the positive self-selection of peer groups into as selective a college as possible per the aforementioned anti-lemons principle. The reputation-building motive driving this mechanism results in stronger ability sorting of students across colleges and leads to over-investment of effort in unproductive admission test preparation while dulling incentives for productive learning while in college.

An alternative (to that of signaling) approach to explain the differentiation of academic standards across higher education advanced in a recent literature focuses on the direct effect of curricular standards on human capital production function, as opposed to their mere sorting role in the context of asymmetric information. The framework discussed above features, due to the signaling mechanism at work, the "Groucho Marx" condition (as termed by MacLeod and Urquiola): a relatively weak student would benefit from attending a more selective institution, to which he would normally be denied admission unless getting in through a lucky draw. Such premise is challenged in recent empirical literature.[9] In contrast, the curricular standards paradigm is motivated by potential failure of such assumption. For instance, Arcidiacono and Lovenheim (2016) distinguish between the "quality effect" of a more selective college (such as better instructional resources), which can benefit any student, from the "match effect" benefiting only students whose adequate prior preparation makes the college a good "match" for them.

Kaganovich and Su (2019) explicitly incorporate the latter feature by defining the value added of a university as student-specific, dependent on the relationship between a students preparation and the level of the universitys curriculum. For a given such level, returns to education will be low for insufficiently prepared students as well as for the "overqualified" ones, as opposed to students whose aptitude is a "good match" for this curriculum. In their model, *curricular standard* is a discretionary characteristic of education technology, strategically chosen by and potentially differing across colleges, thus determining their levels of selectivity. Specifically, they define human capital value added function for a student of aptitude (prior preparation) a as

$$h_s(a) = \begin{cases} B_s(a - c_s) & \text{if} \quad a > c_s, \\ 0 & \text{if} \quad a \leq c_s, \end{cases} \tag{1}$$

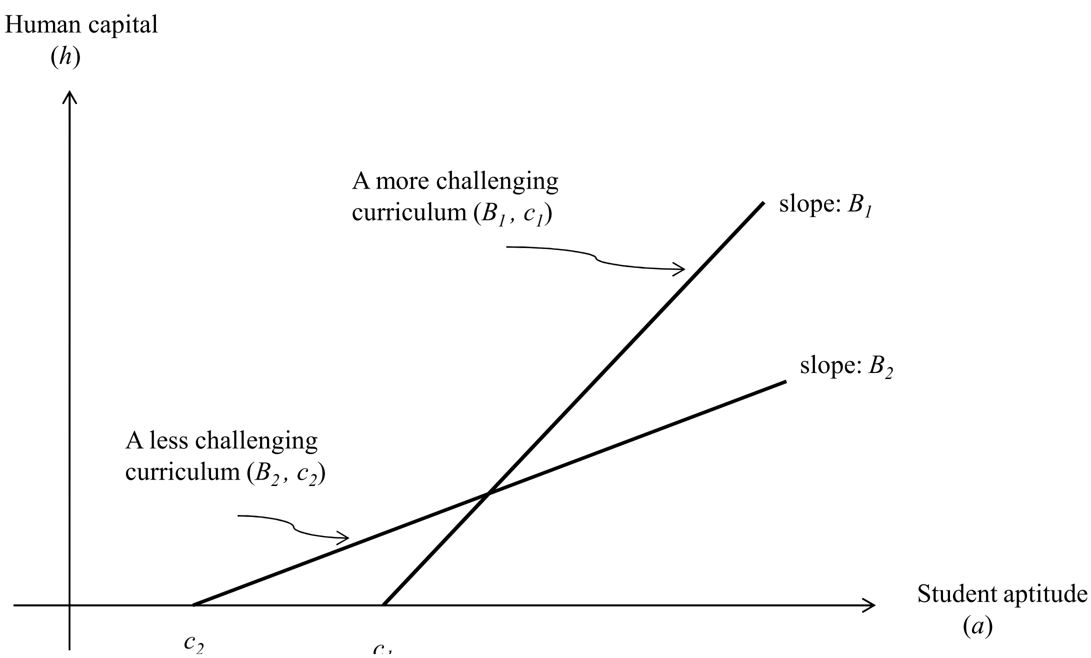
where c_s is the curricular threshold chosen by college s and B_s is the learning progress rate corresponding to and increasing in the chosen curriculum c_s. This implies that students with aptitude below c_s will derive no benefit from college s while those whose aptitude

framework, since higher standards confer benefits without extra costs. If, however, higher selectivity of a college translates into greater curricular challenge, an overly optimistic test outcome may compel the student to make *ex post* adjustments. Manskis (1989) analysis of the widespread college dropout phenomenon was influential for the expanding empirical literature on student responses to performance in college, which is outside the scope of this review.

[9] See, for instance, Light and Strayer (2000), Arcidiacono et al. (2016), and a comprehensive survey by Arcidiacono and Lovenheim (2016), particularly as applied to the evaluation of the outcomes of affirmative action programs.

a is insignificantly above c_s will benefit very little. Thus, a student's human capital gain at a college depends on the relationship between his aptitude and the college's curricular standard, such that cross-college comparison of quality is student-specific. For a given level of student aptitude, there is an individual-specific "Goldilocks" optimal curricular standard for this student's instruction, hence the "match effect". Each student chooses his best match among available colleges according to the curricula they offer. The tradeoff a student of given preparation faces when choosing between two colleges offering different curricular standards can be illustrated by the following Figure 1.

Figure 1: Student trade-off in choosing between two academic curricula



The trade-off depicted in the figure implies that there is a threshold separating the students preferring more selective college 1 from those who are better off attending college 2 (it is obviously determined by the point of intersection between the two lines in the figure). As a result, the colleges are able to compete for students by choosing locations in the space of curricula, i.e., they differentiate their products horizontally. As will be detailed in Section 5, the model conjectures a natural ranking of selectivity among the universities determined by the weights they place on the quantity of students they enroll, besides the quality of students. Their preferences for quantity are in turn derived from the budgeting environments of the universities: the stronger dependence on tuition revenue in the operating budget, the more weight a university will place on the size of enrollments.

3.3 Students as inputs: peer-group effects in college human capital production function

Finally, we lay out the paradigm originating with Rothschild and White (1995), which proved to be one of the most influential for the literature under our review in its ability to reproduce the structure and characteristics of the U.S. higher education market, both qualitatively and quantitatively. According to this theory, which unlike the signaling paradigm, emphasizes the intrinsic value added of college education, the magnitude of value added depends, in addition to the quality of instructional inputs, also on the average quality of students through direct learning spillovers, i.e., the peer group effects in education. This motivated the key innovation of Rothschild and White that students can enter as inputs in education production, not just its output. As a result, some students can be more valuable to a university than others, particularly by enhancing the educational experience for their classmates. This then implies that tuition price discrimination is in order: efficient tuition should deduct the marginal contribution of a student to the university's total human capital production from the marginal cost of the output the student receives, and thus may vary among students in the same classroom.

In the Rothschild and White (1995) model, universities' $j = 1, ..., J$ production technologies are given by distinct production functions $F_j(\cdot)$, whereby the differences can be attributed to variation in their exogenously given fixed inputs. The key variable input is a university's student body composed of different student types $(n_1^j, n_2^j, ..., n_T^j)$ where n_t^j is the number of type t students attending university j. Output vector is given by jointly produced aggregate amounts of human capital of each type $(H_1^j, H_2^j, ..., H_T^j)$ with each type t graduate receiving H_t^j/n_t^j units of the corresponding type of human capital (whereas, notably, students of each type may affect the production of human capital of other types). Universities choose type-specific tuition levels to maximize their profits in a competitive market. Under standard assumptions, optimal tuition university j charges a type t student is derived as $p_t^j = H_t^j/n_t^j - \hat{w}_t$, where $\hat{w}_t$ is the value of type t student to a university as an input, which happens to be equal across universities in equilibrium. Thus tuition is differentiated by student type, according to their input value, but also across universities, according to their productivity in a given human capital category. This framework thus played a seminal role in modeling tuition price discrimination, which as noted above, is an essential distinct characteristic of competition in the U.S. higher education market.

In a series of papers Epple, Romano, and Sieg (henceforth, ERS) (2002, 2003, 2006, 2008) build on the Rothschild and White (1995) framework to develop a comprehensive model of the market for higher education and use it for the theoretical and empirical analyses of the markets features. In ERS model, quality-maximizing universities compete for students who are heterogenous in ability, household income, and other characteristics. The quality of education at a university is determined, in line with Rothschild and White, by its students characteristics (e.g., peer quality deriving from student abilities) as well as choices of other inputs such as per-student instructional expenditures. Equilibrium in the higher education market features endogenous product (i.e., quality) differentiation, co-determined by: (i) the optimal admission, pricing, and non-student input decisions by competing schools, (ii) sorting of students (both as consumers and inputs) across universities, where universities optimal (type-specific) prices internalize peer externalities as in Rothschild and White (1995).[10]

ERS modeling paradigm entails two dimensions of tuition price discrimination among students. The first, derived from internalizing the peer-effect externalities in the spirit of Rothschild and White (1995), reflects the common practice of merit-based financial aid in U.S. higher education. This approach can also capture tuition discounts to underrepresented minorities, provided that universities value diversity in the student body (ERS, 2003, 2008). The second channel is tuition differentiation based on family income, hence ability to pay, and corresponds to the universal practice of need-based financial aid and is strongly evidenced in the data sources (see, for instance, ERS, 2003; McPherson and Shapiro, 2006). ERS models are able to capture this income-based price discrimination feature of the market by using pricing above marginal cost, but its magnitude is limited in the presence of close competitors. Indeed, the ability to price discriminate based on income requires that universities possess market power, which may in principle run counter to the evidence of the increased overall levels of competition.[11] Epple, Romano, Sieg, and Sarpça (henceforth ERSS) (2017) develop a novel model of the market for higher education, which incorporates the exercise of market power in the framework of monopolistic competition. As a result, along with the peer-group externalities, the model is able to successfully combine the two aforementioned channels of tuition price discrimination. This theoretical model was given support by the empirical analysis by Epple, Romano, Sieg, Sarpça, and Zaber (ERSSZ, 2019) providing evidence of significant market power, particularly at selective universities, which do in fact feature most stark levels of tuition price discrimination.

Overall, this branch of literature was able to offer a distinct characterization of quality differentiation across universities according to the combination of students academic abilities and their ability to pay, tuition differentiation within and across universities, and in response to changes in public funding policies, as reviewed in more detail in the next section.

[10] Building on ERS (2006), Sarpça (2010) introduces an additional layer of product differentiation: specialization among colleges in types of disciplines. Multi-dimensionality of student abilities and hence of their effects on peers is an essential premise of the model with implications for optimal pricing and allocation of students across colleges.

[11] Hoxby (1997) presents evidence of increased overall competition in the U.S. higher education in the second half of the XX century. However, she argues that this was accompanied by increased vertical product differentiation in terms of quality of education between the universities.

4 Peer Group Effect-Based Models: Market Structure, Competition, and Pricing in Higher Education

In this Section, we first review a basic version of the Epple-Romano-Sieg (ERS) model, introducing the common components of their four referenced papers and their selected findings, extensions and implications. We then discuss innovations introduced in the Epple-Romano-Sarpa-Sieg (ERSS) model to incorporate market power of selective institutions and review the findings of ERSS (2017) and ERSSZ (2019).

4.1 ERS model

There is a continuum of potential students who differ with respect to their household income y and their ability b with joint marginal distribution $f(b, y)$. The utility function of a student $U(\cdot)$ is increasing in its two components: numeraire consumption and educational achievement. Consumption is income net of university tuition p, if one is attended. Educational achievement $h(\cdot)$, is increasing in the quality q of the university attended and the student's ability. A type (b, y) student's utility from attending university j is thus $U(y - p_j, h(q_j, b))$. Ordinary demand for university quality is increasing in income and non-decreasing in ability. Individuals are free to not attend a university, in which case they pay no tuition ($p_0 = 0$) and experience a quality q_0 (same as for those not admitted by any university).

There are J universities competing for students. Universities differ ex-ante in their non-tuition revenues $E_1 < E_2 < ... < E_J$ (such as endowment income, state subsidies, etc.). All universities have the same cost function $C(k_j, I_j) = F + V(k_j) + k_j I_j$ with $V',\ V'' > 0$ where k_j is the number of students admitted to university j and I_j stands for educational expenditures per student there. Quality of education q_j at university j is determined by the average ability of its students θ_j and educational expenditures per student, such that $q_j = q(\theta_j, I_j)$.

A university chooses tuition and admission policy and the levels of expenditure on educational inputs to maximize its quality, subject to a budget constraint, while taking as given the alternative choices available to students in equilibrium. Let $\alpha_j(b, y) \in [0, 1]$ denote the proportion of students with characteristics (b, y) that university j finds optimal to admit. It is optimal for university j to charge a student his/her reservation price $p_j^r(b, y)$ at which the student reaches the level of utility of his/her best alternative, which is given by

$$U(y - p_j^R(b, y), a(q_j, b)) = \max_{i \in \{0,1,2,\ldots,J\},\ i \neq j,\ \alpha_i > 0 \ \text{ is optimal}} U(y - p_i(b, y), a(q_i, b))$$

where university j's optimal admission policy α_j is derived as

$$\alpha_j(b,y)\begin{pmatrix} =1 \\ \in[0,1] \\ =0 \end{pmatrix} \quad \text{if} \quad p_j^R(b,y)\begin{pmatrix} > \\ = \\ < \end{pmatrix} \underbrace{V'(k_j)+I_j+\frac{\partial q/\partial\theta}{\partial q/\partial I}(\theta_j-b)}_{EMC_j(b)}. \tag{2}$$

Thus university j admits all (no) students of type (b,y) if that types reservation price for attending university j exceeds (is below) the types effective marginal cost $EMC_j(b)$. If the two are equal, the university is indifferent what share of the type is admitted. The effective marginal cost is the sum of the cost of admitting any student $V+I$ and the cost of neutralizing the particular students ability externality (the change in the magnitude of peer effect from admitting ability b student multiplied by the resource cost of maintaining the university quality). Note that EMC varies across students within university j only with students ability, and the peer cost enters as a negative term if a students ability is above university mean. Note that if other student characteristics whose averages affect school quality, e.g., racial or income diversity, are incorporated in the model, then corresponding additional terms will appear in the EMC function.

Caps on tuition prices are an essential and realistic attribute of the market pricing featured in ERS model. If a university posts maximum tuition $\overline{p}_j$ so that $p_j(b,y) = max\{p_j^R(b,y),\overline{p}_j\}$, the difference $\overline{p}_j - p_j(b,y)$ between the posted tuition and that charged to a student can be naturally interpreted as the institutional financial aid to a student of the type in question.[12] Similar to (2), the equation between the effective tuition price paid and the effective marginal cost of admitting the student determines student ability threshold for admission to university j, as described in more detail below.

ERS derive the properties of the respective models regarding school characteristics, pricing and admission functions, and the resulting allocation of students across universities. The common properties for the class of models are as follows. Schools vary by quality endogenously in equilibrium such that $q_i < q_2 < ... < q_J$ with quality hierarchy following the endowment ranking. Students are stratified along the income and ability dimensions such that the admission regions of different universities are separated by downward sloping boundary loci in the (b,y) space. Along the boundary loci, tuition charge equals EMC and thus can be expressed as a function of ability only. In the interior of the admission regions, tuition at the attended school exceeds EMC and depends partly on a student's income. Still, the allocation of students across schools is the same that would obtain if they were charged $p = EMC$ at every school. The extent of pricing by income is restricted by competition: tuition charge p_j cannot deviate much from EMC_j in the presence of close substitutes for j. This also implies that the top quality school will have more room for pricing by income than the other schools, as it

[12]The model also allows for the provision non-institutional financial aid, e.g., provided by government based on student characteristics, which can be then incorporated directly into student utility function. Let such aid received by a student be given by $A(\cdot)$, then utility expression becomes $U(yp + A(\cdot), h(\cdot))$.

does not have competition from above.

ERS (2002) and ERS (2008) consider a model, in which students are additionally characterized by race (white or nonwhite). The first of these papers studies alternative specifications of students' and universities' preferences regarding diversity. They then analyze their implications for the distribution of white and nonwhite students within and across universities in equilibrium. ERS (2008) build on this menu of models to characterize optimal university policies for admission and tuition pricing under affirmative action. Affirmative action results in minority students paying lower tuitions and attending higher quality schools compared to non-minority students with similar income and ability. A ban on affirmative action is studied as a policy option and found to have substantial impact on admission of minorities. When the ban runs counter to universities preference for diversity, the universities are able to partly circumvent the ban by exploiting the information on how ability and income are related to race in their admission and pricing so as to achieve desired diversity. This comes at a cost of extending admission, in lieu of some minority students, to previously ineligible categories of non-minority students who fit the statistical profile of the ability-income relationship. Specifically, this group includes some non-minority students with high to moderately high income but relatively low SAT scores. Welfare analysis with a computational counterpart of the model suggests that the loss to minorities resulting from a ban on affirmative action substantially outweighs the gain to non-minorities.

ERS (2003) incorporate non-institutional aid into the model, such as based on policies of and funded by the federal government along with preferences for diversity with multiple dimensions, and go on to investigate whether the model's various predictions are broadly supported by the data. The analysis combines student-level data (including income, ability as measured by standardized test scores, institutional and non-institutional aid) from the National Postsecondary Student Aid Study (NPSAS), and detailed school level data from Peterson's and NSF WebCASPAR.[13] The paper documents correlation of average SAT scores, expenditures per student, endowment per student, and average tuition, along the school quality hierarchy and shows that they are in line with the model's predictions. A series of reduced form regressions also provide some support for stratification of universities by income and ability produced by the model. The empirical analysis in the paper provides evidence of tuition price discrimination by student ability, consistent with schools' preference for peer quality. ERS (2006) advance the empirical analysis by developing a framework to estimate the model, using mostly the data referenced above. An empirical regularity that comes up in many data sources is the presence of price discrimination based on student family income, framed as institutional financial

[13]NPSAS: National Postsecondary Student Aid Study, National Center for Education Statistics, NCES Handbook of Survey Methods, `https://nces.ed.gov/statprog/handbook/npsas.asp`; Peterson's: Petersons CollegeData, `www.petersonsdata.com`; WebCASPAR: National Science Foundation database for Science and Engineering and other fields at U.S. academic instituitions, `https://catalog.data.gov/dataset/webcaspar`.

aid, stronger than predicted by ERS (2003). To capture this empirical phenomenon, ERS (2006) generalize the model to include preferences for income diversity in the student body and, using estimated parameters, study the effects on sorting of students across schools based on counterfactual analyses: (i) a ban on price discrimination by income; (ii) a change in the federal financial aid formula. A more substantial leap in capturing the tuition pricing by income was achieved, however, by introducing significant modeling innovations discussed below.

4.2 ERSS model

ERSS (2017) aims to capture a broad array of qualitative and quantitative characteristics of the U.S. market for higher education, including the institutional differences between universities reflected in their objective functions, funding sources and pricing policies, such as the differences between private and public universities along with the regional variation of pricing in public universities for in-state and out-of-state students. To this end, they develop a model, which is adequately rich in institutional details to produce the requisite variety of outcomes in terms of admission and pricing policies and the allocation of students across universities in equilibrium. Here, we briefly review the main distinctions of this model compared to the baseline models discussed above. The key features of the model include the competition between state and private colleges, which differ in their objectives, and incorporation of federal aid policies reflecting their real world characteristics.

In addition to ability b and income y, students differ also in state/region of residence $s \in S$ and unobserved idiosyncratic preferences over universities. Utility of student of type (s, b, y) from attending university j is given by

$$U(y - p_{sj}(b, y), h(q_j, b)) + \epsilon_j,$$

where ϵ_j denotes an idiosyncratic preference shock for school j known only to the student. Students choose among their college options to maximize utility, given the prices and qualities $P(s, b, y) = \{p_{sj}(b, y)\}_{i=0}^{J}$ and $Q = \{q_i\}_{i=0}^{J}$ of all J universities, public and private, as well as the outside option. Conditional choice probabilities $r_{sj}(b, y, P(s, b, y), Q)$ are obtained by integrating out the idiosyncratic the taste components for each type. A private university's problem can be formulated in a similar fashion to that in ERS by replacing the admission function $\alpha_j(b, y)$ with the conditional demand $r_{sj}(b, y, P(s, b, y), Q)$ aggregated over all states $s \in S$. For student of type (s, b, y) with $y_{sj} > 0$, the optimal tuition at university j satisfies

$$p_{sj}(b, y) + \frac{r_{sj}(b, y; \cdot)}{\partial r_{sj}(b, y; \cdot)/\partial p_{sj}(b, y)} = \underbrace{V'(k_j) + I_j + \frac{q_\theta}{q_I}(\theta_j - b)}_{EMC_j(b)}.$$

The left hand side of this equation is marginal revenue from admitting student of type (s, b, y), with $r_{sj}(b, y, \cdot)$ representing the type's demand for university j. The right-hand side is the effective marginal cost as discussed in the ERS model, see expression (2).

The following three functional and parametric specifications of the model, which are used in the computational application in ERSS (2017) and estimation purposes in ERSSZ (2019), provide further insight into the workings of the model:

(i) The university quality function is given by for $q_j = \theta_j^{\gamma} I_j^{\omega} \Gamma_j^{\kappa}$ for $\gamma, \omega, \kappa > 0$.

(ii) The utility of a student from attending university j is given by function $U_j(y - p_{sj}, a(q_j, b)) = \alpha \ln[(y - p_{sj})q_j b^{\beta}] + \epsilon_j$ where $\beta, \alpha > 0$ with α representing the weight student places on the systematic component of utility.

(iii) The idiosyncratic disturbances ϵ_j are independent and identically distributed with Type I Extreme Value Distribution.

Let $J_a(s, b)$ denote the choice set of students of ability b in region s, i.e., the set of schools, which will find it optimal to admit the type in equilibrium. The conditional choice probability for type (s, b, y) student for university $j \in J_a(s, b)$ is then given by

$$r_{sj}(b, y) = \frac{[(y - p_{sj}(b, y))\, q_j]^{\alpha}}{\sum_{k \in J_a(s,b)} [(y - p_{sk}(b, y))\, q_k]^{\alpha}}. \tag{3}$$

With a continuum of students of each type, these expressions represent conditional market shares. This illustrates a major distinction compared to a baseline ERS model: Admission spaces overlap, such that students of the same type appear in different schools, though in different proportions. Furthermore, the optimal tuition price is now obtained as a weighted average of the effective marginal cost of admitting the student and student's income:

$$p_{sj}(b, y) = \frac{(1 - r_{sj})\alpha}{1 + (1 - r_{sj})\alpha} EMC_j(b) + \frac{1}{1 + (1 - r_{sj})\alpha}\, y \tag{4}$$

With the incorporation of idiosyncratic preferences, pricing by income arises naturally as an equilibrium outcome within a framework of monopolistic competition. The first term captures merit (i.e., ability) based financial aid, or that based on any other student characteristic whose average directly contributes to the school quality. The second term reflects the exercise of market power. The markup is monotonically increasing in students family income. The weight on income increases with (type-specific) demand $r_{sj}(b, y)$ and decreases with α. The university captures more revenue from higher income students with stronger demand. Note that pricing by income persists even if individual colleges have negligible market shares, since the markup for a student depends not on the overall market share of the university, but on the market share conditional on observed student characteristics.

Another important feature of this model is the inclusion of the public sector in the competition. A theoretical challenge has been to capture the different objectives of private and public universities and the different constraints they face within a general equilibrium model. Public universities face state mandates and incentives to provide affordable education to in-state students. In the model, tuition levels for in-state and out-of-state students are legislatively fixed. The public universities then choose admission policies and expenditures to maximize aggregate achievement of their in-state students subject to the requirement that tuition revenues plus state subsidies cover costs. This objective results in setting a minimum ability threshold for in-state students, and (for realistic parameter values) a higher one for out-of-state students. It is optimal for state universities to admit out-of-state students for two reasons: (i) they pay higher tuition than in-state students, (ii) the higher ability threshold for out-of-state results in admission of out-of-state students who enhance the average peer quality.

ERSS (2017) then use a quantitative version of this model to examine the effects of changes in public funding policies on university enrollment. It is often conjectured that increased demand caused by increases in federal financial aid may induce universities to increase their tuitions, i.e., is only partly passed on to students in terms of reducing their financial need. Adopting a realistic approximation of the provision of federal financial aid in the U.S., the papers analysis suggests that a 25% increase in maximum federal aid increases total enrollment in universities by 6%, mostly among relatively poor students and mainly at public universities. Private universities indeed reduce their institutional aid, increase expenditure on educational inputs, and substitute some high-ability but lower-income students for some higher-income but somewhat less-able students. The policy change leads average private university student tuition costs to rise, but the effects are uneven across student types, and poorer students in fact experience a cost saving. Although the average student costs among all public students fall by a modest amount, the effects also vary by student characteristics.

ERSSZ (2019) use data from the NPSAS to estimate the market equilibrium for the model developed in ERSS (2017). The estimated model includes minority status as a student characteristic, which affects quality of education as a positive externality (e.g., by better preparing all students for a diverse workplace), and studies the implications for tuition pricing. The estimation strategy does not require solving for the equilibrium, and involves the estimation of conditional market shares given in (3), then using them in the estimation of the pricing equation (4). When estimating conditional market shares, the analysis takes into account the fact that a student will be admitted to a subset of universities in equilibrium and predicts the set of alternatives a student is choosing from. The key results of this analysis suggest that pricing by income is prevalent among private universities in the U.S. Specifically, the paper estimates that $10,000 increase in family income increases tuition by $210 to $510 on average. Average university markups, computed as the difference between price and effective marginal cost, range between $750 and $13,200, with higher levels and greater variation in markups characterizing a subset of colleges, particularly the highly selective ones.

5 Product Differentiation in Higher Education Market via Academic Standards

A distinguishing characteristic of the class of models discussed in the previous section is that education technology does not vary across universities. Differential quality of outcomes obtains because the qualities of inputs, teaching resources and average student ability, do vary. As we could see, these models turn out effective in replicating many stylized features of the US higher education market. In particular, they are able to produce segmentation of the market where universities are ranked by their quality and students are sorted into them according to the academic ability and ability or willingness to pay. The key mechanism, as highlighted by Rothschild and White (1995), is that of tuition price discrimination. If there are no caps on tuition, then this mechanism is in fact sufficient for excluding academically undesirable students out. However, mediocre students with high ability to pay are still eligible and desirable for admission at an elite university thanks to their financial contribution to enhancing teaching resources. Reciprocally, such mediocre students, in this modeling framework, clearly do gain in human capital accumulation from studying at an elite university thanks to the spillovers they would enjoy from superior ability peers.

In contrast, two modeling approaches discussed in this section feature academic mechanisms where excluding students falling below a colleges set standards plays a key role in determining the quality of the product students receive there. These approaches, employing their respective paradigms of academic standards, offer alternative theories of market segmentation in higher education. Although these recent contributions have not been developed to the point of explicitly incorporating tuition policies and other resource related characteristics of university operation and competition, they do offer a proof of concept, particularly in demonstrating the emergence of school reputation and curricular standards as sorting mechanisms resulting in segmentation along the selectivity axis. They also help gain insights into the observed trends of diverging selectivity of colleges in the U.S.

MacLeod and Urquiola (2015) develop a signaling model based on *reputations* of individual colleges as indicators of quality. This means that a graduates job market valuation is determined by the perceived quality of the college she attended. College reputation is in turn based on the average quality ("skills") of its graduates, which it manages by means of admission standards. The average skill level of a specific colleges graduates is also the main basis for wage determination by employers, as they lack information about individual skill levels. This implies the following incentive structure: any student wishes to gain admission to as selective a school as possible, whereas each school, defined by its pool of admitted students, has an *ex post* interest in more stringent admission standards to raise the pool's average quality.

In the model, where all variables are presented in log terms, true individual ability of potential student i, which remains unobserved, follows $a_i\ N(,\sigma^2)$. Potential students take a universal college admission test whose results are observable and given by

$$\tau_i = a_i + r_i + \epsilon_i^\tau, \tag{5}$$

where r_i is individual effort to prepare to the test, which does not contribute to skill, and $\epsilon_i^\tau\ N(,\sigma_\tau^2)$ is the error term with variance σ_τ^2 characterizing the test's precision. Since the test is taken without the knowledge of true ability, students are *ex ante* identical and will exert test prep effort at identical level r. Thus, test prep effort is rational on students part but unproductive and constitutes pure efficiency loss, as it does not enhance revelation of information.

The admission test outcome produces an estimate of an applicant's true ability. Given test result τ and the distribution of individual abilities a_i, one can infer the posterior distribution of true abilities $\hat{a}_i(\tau)$ of students who obtained this score. Accordingly, by setting a minimum acceptable test score τ_s as its admission standard, each college s truncates the distribution of true ability of eligible students. Colleges can thus be ranked in terms of their selectivity.

If student i graduates from college s, she will possess true skill level

$$\theta_i = a_i + e_i + v_s, \tag{6}$$

where e_i is the individual effort in college and v_s is observable college-specific value added determined by its exogenous characteristics. Since the true skill of individual college graduate i is not observed by employers, her wage is determined through signaling, in this case two signals: (i) the fact of having graduated from college s, whose reputation is known in equilibrium, and (ii) a noisy outcome $t_i(a_i, e_i)$ of a college graduation test. The latter gives students (the only) incentive to make study effort e_i while in college, which, however, also productively contributes to the actual skill acquisition according to (6).[14] An essential feature of the model (with students preferring to receive higher wages and make less effort in preparing to admission and graduation tests) is that, in equilibrium, marginal return to effort within a given college, i.e., the incentive to improve the graduation test performance, does not vary with ability. Therefore, the effort e_{si} of student i is determined by the college s_i she attends, hence is uniform across students

[14] A case of such testing is documented in a companion paper by MacLeod et al. (2017), as systematically introduced in Columbia starting in 2004 for the purposes of certifying college quality. They take the predictions of MacLeod and Urquiola (2015) to Columbian data, which also makes it possible to derive the measures of college reputations based on the available distributions of admission scores of their students, and to tie them to career outcomes based on comprehensive wage data for graduates. The results confirm, in particular, the effect of college reputation on career earnings as well as the trade-off between it and the effect of the information provided by the college graduation test: the availability of the latter lessens the weight of the former. It is worth noting that the concept of graduation test in the model has broader relevance, because even in the absence of such formal testing, student grade performance in college plays a similar informational role.

within the college. As a result, the students future wage is given by

$$w(s_i, t_i) = E(\theta_i|s_i, t_i) = E(a_i|s_i, t_i) + e_{si} + v_s,$$

which is thus likewise fundamentally determined by the college attended. True individual ability factors in wage determination with noise, through the graduation test outcome along with what can be inferred, as posterior estimate of true ability, based on the fact of this students admission to the college the information employers can use for deriving conditional expectation of true skill given the distribution of the error term in test outcomes.

Based on the structure of decision-making by the parties described above, a perfectly competitive equilibrium allocation of students to colleges is given by the following conditions: (i) Students optimally choose: the effort level in the admission test prep, the most selective college available given the test result, and the level of effort there to prepare to graduation test. (ii) Each college sets minimum admission standard and makes students who meet it eligible for admission subject to capacity constraints; moreover, no individual college with available capacity can affect students choices by changing its admission standards.

When colleges are atomistic and can perfectly differentiate themselves by admission requirements, students who surpass minimum standard in a school are motivated to move up the selectivity ladder. Moreover, students beliefs about composition of a colleges student body are assumed pessimistic, namely that no students with test outcome above the minimum admission standard will enroll. This leads to the "*free choice*" equilibrium regime, with colleges unregulated in their admission policies, where the pessimistic belief is fulfilled. Specifically, the "free choice" equilibrium features perfect segregation by admission score, i.e., the pool of students in college s consists of students whose admission test score is at its minimum admission threshold τ_s. Since strong sorting discourages study effort, underinvestment in productive study while in college is a costly implication of the extreme stratification of college selectivity in this perfectly competitive regime.[15]

An alternative industry structure with partial segregation by admission score, is given by the allocations where each college s admits students with test scores in an interval $\tau \in [\underline{\tau}_s, \overline{\tau}_s]$. According to (5), this is equivalent to composition of student body of a college s being defined by a certain range $\hat{a} \in [\underline{a}_s, \overline{a}_s]$ of posterior expected true abilities (i.e., conditional on the test outcomes falling into $[\underline{\tau}_s, \overline{\tau}_s]$. This leads to the following comparative statics result: If college s marginally raises its minimum admission standard $\underline{\tau}_s$ and by implication increases the lower admission threshold $\underline{a}_s$ in terms of posterior expected ability, then all remaining students (those not excluded by the increased level of selectivity) will enjoy higher payoffs. Therefore, all students in college

[15] This offers an important insight into the consequences of the phenomenon of diverging selectivity of colleges, which is consistent with the findings by Babcock and Marx (2011) that study effort has fallen over time in higher education throughout its selectivity spectrum.

s whose admission scores are marginally above its minimum admission standard prefer this standard to be raised (to the exclusion of their marginally less fortunate peers), hence the positive self-selection of peer groups into as selective a college as possible, i.e., as per aforementioned *anti-lemons principle*. The reputational mechanism at work here motivates the exclusion of inferior participants, contrary to the case of adverse selection, which repels high value participants. The above result also shows that because schools enhance reputation through exclusion, more selective schools tend to be smaller in equilibrium under this mechanism, which is consistent with empirical facts.

We now turn to Kaganovich and Su (2019) model whose education technology features college specific curricular standards, as defined by expression (1) in Section 3, where the model was briefly referenced. It is assumed that a higher curricular standard c_s of college s in (1) is associated with its higher progress rate B_s , so the relationship $B_s = Bc_s$ is assumed for simplicity, where B is a given constant. This education technology makes returns to education at a college dependent on a match between students aptitude (prior preparation) and the standard of the curriculum. According to the model, curricular standards c_s are discretionary policy variables strategically chosen by each college s (an existing one or a potential entrant to the market). Given the distribution of aptitude in the population of potential students, curricular choices made by colleges will partition the population into non-overlapping segments of those who will find it optimal to enroll at a college (or choose the "no college" option) offering them the best match, similar to the two-college situation illustrated in Figure 1 in Section 3. Thus, distinctly from the other paradigms of education value added we discussed above, colleges differentiate their products (the curricula) horizontally in this model. Indeed, cross-college comparison of quality is student aptitude-dependent.

Let there be N colleges ranked in accordance with their selectivity. Some of these colleges may be present in the industry, while others can be seen as potential entrants. College objectives are assumed to represent a combination of concerns about the quality and quantity of their students:

$$O_s = H_s + \gamma_s N_s, \tag{7}$$

where γ_s is an exogenously given coefficient, which represents the degree of selectivity of college s, H_s is the aggregate human capital gain by students enrolling in this college, which can be seen as a measure of its quality, while N_s is the quantity of its students.

The above thus presumes that the degree of selectivity of each college, characterized by the relative weight it places on quantity of students (so higher γ_s corresponds to lower level of selectivity), is given, which can be viewed as a historical characteristic of a university. Indeed, it is assumed that such historical rankings are given:

$$\gamma_1 < \gamma_2 < ... < \gamma_N. \tag{8}$$

Thus colleges differ in their exogenously (historically) established relative priorities over the quality, in terms of human capital outcomes, vs. the quantity of their graduates. These exogenous priorities determine relative positions of colleges in the selectivity rankings. A high priority that less selective colleges give to the number of students implies that such colleges can also be more exposed to additional incentives to expand enrollments further, stemming for instance from political pressure from state governments to ensure greater access to higher education, expressed directly or through financial incentives.[16] Although, as will be seen, weight coefficients γ_s are subject to marginal shocks, those are assumed to not disturb the above ranking. Along the same lines, it is assumed that potential new entrants can only emerge at the bottom of the rankings, proceeding from an understanding that establishing a selective university anew is prohibitively expensive. Note that for a highly selective college $\gamma_s < 0$ is a meaningful situation because the aggregate human capital gain H_s combines the quality of students and their quantity with equal weights, so $\gamma_s = 0$ implies such a balance while $\gamma_s < 0$ will shift the balance in favor of quality. Likewise, a large $\gamma_s > 0$ could reflect greater dependence of less selective colleges on tuition revenues or the dependence of state appropriations for the college being tied to the measure of access by in-state students to it (neither of which our model explicitly incorporates).

Students aptitude is assumed, for the sake of tractability, to follow triangular distribution on $[0, A]$, such that its density for its declining portion where $a > A/2$ is given by $f(a) = 1/Aa/A^2$. This realistically implies that medium ability students constitute

[16]This interpretation fits the case of less selective public colleges most subject to such government policies, imposed by statutory means or financial channels. More broadly, although the financial dimension of colleges operation is not explicitly present in the model for the sake of analytical tractability, allusions to it offer realistic motivation of the models assumptions. In particular, a stronger preference of less selective colleges, public or private, in favor of quantity of their students has much to do with the colleges increased reliance on tuition revenues. Indeed, the public policies to expand access to higher education are often expressed in the U.S. via tuition subsidies, either through direct appropriation for public colleges, or through financial aid to students.

The case of for-profit colleges, which substantially rely on students receiving public subsidies (see, for example, Cellini 2010), can be viewed as occupying the extreme end of this spectrum in terms of their exclusive interest in the quantity of students, to the point, within this highly stylized framework, of not imposing any curricular requirements and accordingly, not delivering educational value added to their students. However, one must draw a conceptual distinction between such case and that of the so-called open-enrollment colleges (e.g., some community colleges), which do not impose explicit admission requirements. Despite relying on tuition revenue and hence placing high weight on the quantity of students, such community colleges do have a mandate to deliver educational value added to students and therefore set certain curricular standards, which serve as effective barriers to entry and/or graduation, the nominally open-enrollment feature notwithstanding.

On the other hand, the business model of elite private colleges (whose formalization is offered by Hoxby, 2014a) is based in part on operating a private endowment, which allows a college to balance its budgets intertemporally while banking on future contributions by graduates commensurate with their career earnings, whose expected levels can be deemed proportionate to the attained human capital. This allows selective colleges to play a "long game", such as focusing on the quality of students more than their quantity, a policy that may entail running budget deficits, with cost of instruction exceeding tuition revenues, but bringing rewards in the form of alumni contributions in the long run, with endowments playing a self-fulfilling role in this business model. In contrast, less selective colleges have little or no reserves and thereby must meet short-run budget constraints. Therefore, tuition revenues, including government subsidies tied to the quantity of students, play a more dominant role in their business model.

a large share of the population, while high ability ones are less numerous. Colleges engage in strategic competition for students in pursuit of their objectives (7) where their strategies are given by the choices of curricular thresholds c_s, which determine colleges locations along the axis of student ability. Given the ranking of colleges by selectivity according to (8), Nash equilibrium is characterized by curricular choices such that $c_1 > c_2 > ...c_N$, i.e., less selective colleges choose lower curricular standards, and, accordingly, a partition of the college-bound population where the student body of college s is given by the ability segment $[a_s, a_{s-1}]$ where $a_0 = A$.

The two-college version of the model captures the gist of the results. When one of the colleges is selective and the other is not, i.e., γ_1 is negative and γ_2 is positive and both obey certain bounds, a unique locally stable equilibrium in terms of curricular standards $c_1^* > c_2^* > 0$ exists. They, in turn, define the equilibrium college attendance: students with aptitudes falling into the intervals $[c_2^*, c_1^* + c_2^*]$ and $[c_1^* + c_2^*, A]$ will attend colleges 2 and 1, respectively.

The following comparative statics results carries the main message. Suppose college 2 experiences a shock to γ_2, which increases its bias in favor of the quantity of students. (This may be caused, for instance, by a reduction in state support of a public college increasing its incentive to increase enrollments.) This will have the following effect on the Nash equilibrium: $\frac{\partial c_2^*}{\partial \gamma_2} < 0$ while $\frac{\partial c_1^*}{\partial \gamma_2} > 0$. In other words, while college 2 further reduces its selectivity, college 1 will move in the opposite direction, i.e., becoming more selective, which results in the overall diverging selectivity of the colleges. The intuition is straightforward. As college 2 reaches more students in the lower ability segment of the population by adjusting its curricular standard downward, this comes at the expense of human capital attainment of the top ability students who were bound for college 2. For a subset of these students, this will shift the trade-off between the colleges toward college 1. Thus, the competition faced by college 1 will become somewhat weaker, so it will be able to afford giving less attention to human capital gains in its lower marginal cohort and yet remain more attractive to them than college 2 with its watered down curriculum. Therefore, college 1 will be able to raise its curricular standard to the benefit of its better students. A noteworthy implication of the diverging selectivity outcome is that the quality of the closest match provided by the available college curricula will worsen for students in the middle of the college-bound population. Indeed, the less selective college increases its appeal to the less able students to compel them to enroll, while the selective college will gear its curriculum to its more able students.

The diverging selectivity result carries over to the general, multi-college case. When a relatively low selectivity college among those in the market receives a marginal positive shock to its preference for the quantity of students, then the existing set of colleges partitions into two groups. All those below a certain selectivity level (not just the college receiving the shock) will become less selective, while the group of originally more selective universities will further increase its selectivity. Recall that according to this model, the level of a universitys selectivity does not translate into superior quality for each individual student, which is instead a function of a match between student aptitude and college curriculum. This then implies that diverging selectivity of colleges will tend to have a negative impact on the quality of education available to the medium-aptitude segment of student population. The model obtains this outcome as a consequence of the overall expansion of higher education. An analysis of such negative side effect of the expansion of higher education is a distinction of this curricular standard-based model of education technology and its approach to this market as horizontally differentiated.

6 Intra-University Competition

In the Introduction, we discussed the growing differences in college premia by students major concentration along with the evidence that choice of a major can be a stronger determinant of students career earnings than the selectivity of a university he/she attends. There is also strong evidence that students expectations of future career earnings (to which well refer as *lucrativity* of a major) have a significant effect on students choices of majors, although the strength of this factor tends to differ across demographic groups. It is known that grades also vary strongly across disciplines. In fact, this variation appears to be in inverse relationship with that in the lucrativity of the majors: e.g., grades are markedly higher in humanities than in math and sciences. See, for instance, Achen and Courant (2009) who also document that expected grades in courses have an effect on students decisions about taking them, hence on course enrollments, controlling for institutional and curricular characteristics of the courses.[17] This leads Aachen and Courant to conjecture that departments (major programs) may determine grading policy strategically as a tool to manage enrollments.

The above reasoning implies that struggling departments whose majors are less lucrative, such as in some social sciences and humanities disciplines, could be conducting a more generous grading policy strategically. Specifically, they could be offering higher grades as a compensatory instrument to counteract the loss of students, which would endanger the departments lot in the allocation of university resources, especially in the Responsibility Center Management (RCM) system prevalent in the American higher education. On the contrary, the departments attractive in the financial reward dimension do not have the same incentive to extend generosity in grading. Moreover, some may be compelled to uphold tighter grading standards either to prevent congestion in their classes or to control their composition to maintain a reputation of exclusivity, similar to the mechanism of maintaining reputation of a selective university discussed in the previous section.

Kaganovich and Su (2018) develop a model of strategic interaction between two majors, which provides support for the conjecture that the differences in grading standards

[17] Sabot and Wakeman-Linn (1991) and Bar et al. (2009) find that receiving better grades appears to affect students' utility directly (apart from their indirect effect as signals of higher achievement).

across fields of study have a compensatory role in managing student demand and are used by units of a university in their competition for students.

Potential students are characterized by exogenously given aptitude a assumed, for simplicity, to be uniformly distributed on $[0,1]$. Each major $m = 1,2$ is characterized by an exogenously given wage rate w_m per unit of human capital a student attains in the major, which determines the graduates earnings over the course of their careers. The first major is presumed to be more lucrative than the other, i.e., $w_1 > w_2$. Student career earnings are thus given by $w_m h_m(i)$ where $h_m(i)$ is the human capital attained by student i in major m he/she chooses.

The learning effort it will take a student to attain a desired level of human capital is inversely related to aptitude. Specifically, $e(i) = h_m(i)/a(i)$. As will be detailed shortly, both career income and the learning effort level enter students preferences, along with the psychic benefit of a grade earned in college.

Each major evaluates its students performance on a categorical (letter) grading scale and sets human capital attainment standards corresponding to each grade. For simplicity, there are only two grades: "C", for low achievement, which is the minimum graduation requirement, and "A", for high achievement. Each major m sets its own grading standards, i.e., the minimum human capital attainment level $h_{m,g}$ for obtaining grade $g = A, C$. Thus, $h_{m,C}$ is the minimum standard for graduating in major m. We denote $g(h_m(i))$ the grade received by student i in major m based on his human capital attainment relative to the grading standards in the major.

Following the signaling framework, it is assumed that the information observable by future employers is limited to students major and grade received in it. This implies that no student set on pursuing major m with grade g has an incentive to make any effort to attain human capital level beyond the minimum needed to this end.

Student preferences are defined by maximizing the utility function

$$\max_{m,g} \quad w_m h_m(i) - \left(\frac{h_m(i)}{a(i)}\right)^2 + s(g(h_m(i))) \quad s.t. \quad h_m(i) \geq h_{m,g} \tag{9}$$

where the psychic benefit of high grade $s(A) = \delta$, a given constant, while the benefit of passing with low grade is normalized to $s(C) = 0$. Since, according to the signaling assumption, students have no incentive to exceed the minimum requirement for the chosen grade in the chosen major, the above optimization reduces for each student to discrete choice over $m = 1,2$ and $g = A, C$ possibilities, given the grade standards $h_{m,g}$.

Thus, when grade standards $h_{m,g}$ for each major/grade combination are given, students sort themselves across these possibilities, plus the option of not choosing either, i.e., not enrolling at all. This implies that student allocation across majors are determined by the departments grading policies, which are thereby effectively strategic. Suppose each major programs objective is to maximize the aggregate human capital it produces.

For major m, it can be expressed as

$$H_m = H_{m,C}(h_{1,C}, h_{1,A}, h_{2,C}, h_{2,A}) + H_{m,A}(h_{1,C}, h_{1,A}, h_{2,C}, h_{2,A}) \tag{10}$$

where the first component $H_{m,C}(h_{1,C}, h_{1,A}, h_{2,C}, h_{2,A})$ represents the aggregate human capital of the pool of students who choose to pursue grade C in major m. As stated, this expression is a function of grading standards in the major as well as those in the alternative major, because this entire set of standards determines the choice of each individual student according to (9). This means, that when a department sets its grading standards, it needs to factor in such decisions by its counterpart in order to determine the resulting enrollments and grade level pursuits by its students. The meaning of the second component in (10) analogously applies to the pool of students pursuing grade A in major m. Note that each of these human capital aggregates is given by a product of human capital standard for the grade, times the number of students who choose this particular option. The assumed departments objective to maximize the aggregate human capital attained by its students therefore implies that the program gives equal weights to student quantity and quality. This objective could be, in principle, modified in either direction.

The strategic interaction between the major programs described above is a Cournot-Nash game, if each program takes the others grading policies as given when setting ones own. It can be shown that in equilibrium students sort themselves into non-overlapping segments in the aptitude distribution, each segment corresponding to specific major-grade combination, plus the lowest aptitude segment of students choosing not to enroll. Two types of equilibria can arise. In the separated majors equilibrium, all students pursuing more lucrative major 1 are of superior aptitude than those choosing major 2; the essential part being that students earning C in major 1 have higher aptitude than those earning A in major 2. Kaganovich and Su (2018) show that when the wage differential between the majors is sufficiently large, then only the separated majors regime emerges.[18]

They define *grade inflation* as a policy by a major to relax either or both of its grade standards. In this framework, this leads to an important distinction in the menu of available grading policy options: (i) to pursue "C-inflation", which only relaxes the minimum standards for qualifying to graduate in the major, while maintaining high standards for its elite group of students (i.e., no "A-inflation"); (ii) to engage in "A-inflation" only; or (iii) to engage in both. It is not hard to envision potential rationales for departments to pursue either of the possibilities on this policy menu. For instance, engaging in C-inflation while resisting A-inflation can help boost the size of the major

[18] An alternative possibility, which under some parametric conditions can co-exist with the separated majors regime, is equilibrium with interlaced majors, where "A" students in the less lucrative major 2 are of superior aptitude than "C" students in major 1. Given the space constraint, we shall limit the focus here to just the case where only separated majors equilibrium exists, i.e., under the assumption of sufficiently large wage rate differential.

while maintaining the reputation for the elite group of its graduates.

Under the assumption that wage rate differential between the majors is large enough to produce only the separated majors equilibrium (see the last footnote), the following results are obtained for comparative statics effects of increasing exogenous wage rate differential between the majors. A marginal increase in w_1 will cause major 2 to lower standards for both of its grades further, now focusing more on its lower ability students to maintain enrollments. In contrast, major 1 elevates grade A standards but expands its total enrollment engaging in "C-inflation". Thus, as the absolute rewards of the lucrative major increase, it can compel its better students to exert more effort. By the same token, it becomes more competitive vs. major 2, and can steal some of its better students by lowering its lower bar, and major 2 is then forced to respond by engaging in grade inflation.

An alternative cause of increasing cross-major wage differential is an absolute marginal decline in w_2, wage rate of the less lucrative major. This means that major 2 becomes less competitive against the "no college" option, as well as against major 1. This will cause major 2 to compensate enrollment losses by further reducing all its grade standards to attract lower ability students who previously opted against matriculating and to minimize losses of its better students to major 1. The rise of the wage rate differential will allow major 1 to maintain its elite group of students and attract more students to its C cohort without changing its grade standards.

Both results characterize the erosion of standards in a less lucrative major as an optimal response to its further loss of relative earning capacity in the job market. Kaganovich and Su then extend the model to argue that this erosion can be self-reinforcing as a longer-term negative external effect. To this end, they posit in the spirit of Costrell (1994) and Betts (1998) that average human capital attainment levels in a major have, in the long run, an effect on the wage rate in the corresponding occupation. This implies that a decline of human capital standards in a major can generate a further decline in the associated wage rate in the long run. This effect is external, resulting from the evolution of standards in the major, which are assumed similar across the higher education system, such that a program alone is unable to internalize it. The result of this is a positive feedback loop where the initial relatively small exogenous increase in the wage rate differential between majors triggers a cycle of endogenous adjustments by the competing majors within each university, which have a negative effect on the standards of a less-paying major magnifying the initial negative impact on its relative value and on its attraction to students. It is interesting to note that such endogenous magnification mechanism can occur even if the intrinsic value of education in the underlying academic discipline is high. Indeed, the mechanism underscores that the erosion of standards as the culprit. We conjecture that such analysis may offer a potential insight into the mechanics of the much-discussed "crisis of liberal arts".

7 Concluding Remarks

Our review focused on three paradigms of value creation in the higher education and the features of its market structure they help analyze. The models based on *peer group effects* in the production of human capital, which we reviewed in Section 4, proved remarkably successful in reproducing many notable characteristics of this market such as: (a) vertical differentiation of universities in the quality of education measured in per student resources; (b) sorting of students based on ability to study and to pay, and other characteristics of student diversity, (c) tuition price discrimination in the form of merit- and need-based financial aid as the mechanism of said sorting. This success is all the more remarkable given that the empirical assessment of the actual strength and mechanics of peer group spillovers has not been conclusive and compete against the evidence of benefits of educational tracking.[19]

The two paradigms reviewed in Section 5 offer distinctly different mechanisms for differentiating education quality and segmenting the market accordingly, both based on the variation of academic standards across universities. The first of these approaches underscores the *reputational* mechanism of student sorting, i.e., that signaling their underlying differential quality based on meeting admission standards of more or less selective colleges, without necessarily having educational value added while in college. The second approach, while also focusing on the variation of standards across colleges, emphasizes the role of *curricular standards* as factors in the technology of delivering added educational value. Most importantly, this approach underscores that quality of education measured by the magnitude of value added depends on a students prior preparation (such that a less demanding curriculum may deliver higher quality for a less prepared student). This leads to a novel view of the competition in the higher education marketplace as that based on horizontal differentiation of curricular products segmenting student population according to their preparedness. These recent approaches offer important insights into the observed increasing spread in the spectrum of selectivity of colleges with its escalating exclusivity at the top and virtually open access at the bottom. The curricular standards model, in particular, captures the negative impact of the expanding access to higher education on its quality available to students in the mid-range of the aptitude distribution.

Finally, in Section 6, we outlined a new direction of research focusing on a universitys intrafirm marketplace and the impact of the inter-unit competition for academic standards across disciplines. This is well justified by the macro-level phenomenon whereby the field of study has become a key factor in the variation of wage premium. It also draws attention to the fact that the industry, excepting for-profit and niche markets, continues to be dominated by multi-product conglomerates rather than gravitating toward specialization. This in turn raises questions about the joint production nature of

[19]The latter suggests that positive spillovers from superior quality peers may depend on the proximity of the donor and recipient aptitudes – see, e.g., Booij et al. (2017).

the education technology as its essential feature along with the roles of institutional protectionary policies of bundling educational products, which too deserve theoretical and empirical investigation.

Our review of the economic theories analyzing the structure and competition in the current U.S. higher education system over would be remiss without mentioning the potential impact of developing technological shifts in the industry. For over a century of its growth, higher education has been spared the effects of a major technological disruption. This is in fact a remarkable distinctive feature of higher education, beyond many discussed in this Chapter. A potential challenge in the form of online education is well recognized, particularly with broadband Internet offering increasingly viable alternatives to face-to-face instruction. Whether and how such technology will disrupt the current structure of higher education market and its product remains to be seen. Emerging literature on the subject attempts to conceptualize what can be gleaned from the recent developments. We shall conclude our review on this open-ended note.

From the standpoint of potential effects of online education on the market, it is essential to distinguish between the growing presence of online courses offered for credit at traditional postsecondary institutions vs. the emergence of outlets such as Massive Open Online courses (MOOCs) which operate outside the traditional providers. When it comes to the former, available evidence thus far seems to suggest that students fare worse in terms of academic outcomes when taking a course online compared to taking it in person, be it in community colleges (Xu and Jaggars, 2013), research universities (Figlio et al., 2013), or for-profit universities (Bettinger et al., 2017). It is not unreasonable to expect similar results for MOOCs, where lack of academic success can be similarly attributed to the need for self-discipline and focus without the benefit of peer groups or structured study time (Banerjee and Duflo, 2014). On the other hand, for students who face obstacles to attending traditional classes, the access to online instruction expands educational opportunities, albeit with yet uncertain educational benefits. For instance, Hoxby (2018), finds little evidence of students shifting into high productivity industries after enrolling in online education.

Given the potentially large fixed costs of online content development but low marginal costs of serving additional students, it is conceivable that online education could lead to increasing market concentration. Comen and Tabarrok (2014) suggest that a subscription model (similar to video games) could make online education financially viable, and that the economies of scale may result in both high market concentration in the upper tier and the proliferation of fringe market segments. In contrast, Acemoglu et al. (2014) highlight the complementarity between two types of teaching services: a "global" non-rival online lecture component, and the other with "local" face-to-face and hands-on instruction. They reason that if such complementarity is sufficiently strong in human capital production, the development of pervasive online education may not lead to increased concentration in the education sector overall. More specifically, a division of labor may develop between "superstar" teachers providing online lectures and enjoying the economies of scale, and *in situ* rank-and-file teachers specializing in the face-toface supervised and hands-on instruction. Hoxby (2014b) further distinguishes the potential impact of the technological change on non-selective postsecondary education (NSPE) institutions from that on highly selective (HSPE) ones. She envisions a development of industry structure with the former selling educational services for contemporaneous tuition payment, and the latter acting as venture capitalists investing in their students in return for a share of their future returns, the role that alumni donations can be seen as playing in the present model of operation of elite universities. She argues that the MOOC model may be partly compatible with NSPE institutions, assuming they can harness the revenue stream from MOOCs; on the other hand, the MOOC model is fundamentally incompatible with HSPE institutions, whose selectivity is at the core of their business model.

References

[1] Acemoglu, D. (1998). "Why do New Technologies Complement Skills? Directed Technical Change and Wage Inequality." *Quarterly Journal of Economics* 113, 1055–90.

[2] Acemoglu, D. (2009). *Introduction to Modern Economic Growth*. Princeton University Press, Princeton.

[3] Acemoglu, D., D. Laibson and J.A. List (2014). "Equalizing Superstars: The Internet and The Democratization of Education." *American Economic Review* 104, 523–27.

[4] Achen, A.C. and P.N. Courant (2009). "What Are Grades Made of?" *Journal of Economic Perspectives* 23, 77–92.

[5] Akerlof, G.A. (1970). "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism." *Quarterly Journal of Economics* 84, 488–500.

[6] Altonji, J.G., P. Arcidiacono, and A. Maurel (2015). "The Analysis of Field Choice in College and Graduate School: Determinants and Wage Effects." NBER Working Paper 21655.

[7] Arcidiacono, P. (2004). "Ability Sorting and The Returns to College Major." *Journal of Econometrics* 121, 343-75.

[8] Arcidiacono, P., E.M. Aucejo, and V.J. Hotz (2016). "University Differences in the Graduation of Minorities in STEM Fields: Evidence from California." *American Economic Review* 106, 525–62.

[9] Arcidiacono, P. and M. Lovenheim (2016). "Affirmative Action and the Quality-Fit Trade-off." *Journal of Economic Literature* 54, 3-51.

[10] Babcock, P. and M. Marks (2011). "The Falling Time Cost of College: Evidence from Half a Century of Time Use Data." *Review of Economics and Statistics* 93, 468–78.

[11] Banerjee, A.V. and E. Duflo (2014). "(Dis)Organization and Success in an Economics MOOC." *American Economic Review* 104, 514–18.

[12] Bar, T., V. Kadiyali, and A. Zussman (2009). "Grade Information and Grade Inflation: The Cornell Experiment." *Journal of Economic Perspectives* 23, 93–108.

[13] Berger, M.C. (1988). "Predicted Future Earnings and the Choice of College Major." *Industrial and Labor Relations Review* 41, 418–29.

[14] Bettinger, E.P., L. Fox, S. Loeb and E.S. Taylor (2017). "Virtual Classrooms: How Online College Courses Affect Student Success." *American Economic Review* 107, 2855–75.

[15] Betts, J. (1998). "The Impact of Educational Standards on the Level and Distribution of Earnings." *American Economic Review* 88, 266–75.

[16] Booij A.S., E. Leuven, and H. Oosterbeek (2017). "Ability Peer Effects in University: Evidence from a Randomized Experiment." *Review of Economic Studies* 84, 547–78.

[17] Bound, J., B. Hershbein and B.T. Long (2009). "Playing the Admissions Game: Student Reactions to Increasing College Competition." *Journal of Economic Perspectives* 23, 119–46.

[18] Cellini, S.R. (2010). "Financial Aid and For-Profit Colleges: Does Aid Encourage Entry?" *Journal of Policy Analysis and Management* 29, 526-52.

[19] Costrell, R.M. (1994). "A Simple Model of Educational Standards." *American Economic Review* 84, 956–71.

[20] Cowen, T. and A. Tabarrok (2014). "The Industrial Organization of Online Education." *American Economic Review* 104, 519–22.

[21] De Fraja, G. and E. Iossa (2002). "Competition among Universities and the Emergence of the Elite Institution." *Bulletin of Economic Research* 54, 275–93.

[22] Del Rey, E. and F. Estevan (2020). "Assessing Higher Education Policy in Brazil: A Mixed Oligopoly Approach." *The B.E. Journal of Economic Analysis and Policy* 20, 1–16.

[23] Eckwert B., and I. Zilcha (2020). "The Role of Colleges within the Higher Education Sector." *Economic Theory* 69, 315–36.

[24] Ehrenberg, R.G. (2000). *Tuition Rising: Why College Costs So Much.* Harvard University Press, Cambridge, MA.

[25] Eisenkopf, G., and A. Wohlschlegel (2012). "Regulation in the Market for Education and Optimal Choice of Curriculum." *Journal of Urban Economics* 71, 53–65.

[26] Epple, D., R. Romano, S. Sarpça, and H. Sieg (2017). "A General Equilibrium Analysis of Access to Higher Education in the U.S." *Journal of Public Economics* 55, 164–78.

[27] Epple, D., R. Romano, S. Sarpça, H. Sieg and M. Zaber (2019). "Market Power and Price Discrimination in the U.S. Market for Higher Education." *RAND Journal of Economics* 50, 201–25.

[28] Epple, D., R. Romano, and H. Sieg (2002). "On the Demographic Composition of Colleges and Universities in Market Equilibrium." *American Economic Review* 92, 310–14.

[29] Epple, D., R.E. Romano, and H. Sieg (2003). "Peer Effects, Financial Aid, and Selection of Students into Colleges and Universities: An Empirical Analysis." *Journal of Applied Econometrics* 18, 501–25.

[30] Epple, D., R. Romano, and H. Sieg (2006). "Admission, Tuition, and Financial Aid Policies in the Market for Higher Education." *Econometrica* 74, 885-928.

[31] Epple, D., R. Romano and H. Sieg (2008). "Diversity and Affirmative Action in Higher Education." *Journal of Public Economic Theory* 10, 475–501.

[32] Figlio, D., M. Rush and L. Yin (2013). "Is It Live or Is It Internet? Experimental Estimates of the Effects of Online Instruction on Student Learning." *Journal of Labor Economics* 31, 763–84.

[33] Fillmore, I. (2019). "Price Discrimination and Public Policy in the U.S. College Market." Washington University in St. Louis, Mimeo.

[34] Goldin, G. and L.F. Katz (1999). "The Shaping of Higher Education: The Formative Years in the United States, 1890 to 1940." *Journal of Economic Perspectives* 13, 37–62.

[35] Goldin, G. and L.F. Katz (2008). *The Race between Education and Technology.* The Belknap Press of Harvard University Press. Cambridge, MA.

[36] Hershbein, B. and M.S. Kearney (2014). "Major Decisions: What Graduates Earn over Their Lifetimes." *The Hamilton Project Papers*, The Brookings Institution.

[37] Hoxby, C. (1997). "How the Changing Market Structure of U.S. Higher Education Explains College Tuition." NBER Working Paper 6323.

[38] Hoxby, C.M. (2009). "The Changing Selectivity of American Colleges." *Journal of Economic Perspectives* 23, 95–118.

[39] Hoxby, C. (2014a). "Endowment Management Based on a Positive Model of the University." In J. Brown and C. Hoxby (Eds), *How the Financial Crisis and Great Recession Affected Higher Education*, University of Chicago Press.

[40] Hoxby, C.M. (2014b). "The Economics of Online Postsecondary Education: MOOCs, Nonselective Education, and Highly Selective Education." *American Economic Review* 104, 528–33.

[41] Hoxby, C. (2018). "Online Postsecondary Education and Labor Productivity." In C.R. Hulten (ed.), *Education, skills, and technical change: implications for future US GDP growth* (Vol. 77), University of Chicago Press.

[42] James, J. (2012). "The College Wage Premium." *Economic Commentary*, Cleveland Fed.

[43] Kaganovich, M. and X. Su (2018). "Grade-Compensating Differentials in the Competition between College Majors." Indiana University, mimeo.

[44] Kaganovich, M. and X. Su (2019). "College Curriculum, Diverging Selectivity, and Enrollment Expansion." *Economic Theory* 67, 1019–50.

[45] Kirkeboen, L.J., E. Leuven, and M. Mogstad (2016). "Field of Study, Earnings, and Self-Selection." *Quarterly Journal of Economics* 131, 1057–1111.

[46] McPherson, M.S. and M.O. Schapiro (2006). "U.S. Higher Education Finance." In *Handbook of Education*. Elsevier North Holland.

[47] MacLeod, B. and M. Urquiola (2015). "Reputation and School Competition." *American Economic Review* 105, 3471–88.

[48] MacLeod, B., E. Riehl, J. Saavedra, and M. Urquiola (2017). "The Big Sort: College Reputation and Labor Market Outcomes." *American Economic Journal: Applied Economics* 9, 223–61.

[49] Manski, C.F. (1989). "Schooling as Experimentation: a Reappraisal of the Postsecondary Dropout Phenomenon." *Economics of Education Review* 8, 305–12.

[50] Montmarquette, C., K. Cannings, and S. Mahseredjian (2002). "How do Young People Choose College Majors?" *Economics of Education Review* 21, 543–56.

[51] NCES (2009). *Digest of Education Statistics 2008*. National Center for Education Statistics, March 2009. US Department of Education, NCES 2009-020.

[52] NCES (2019). *Digest of Education Statistics 2018.* National Center for Education Statistics, December 2019. US Department of Education, NCES 2020-009.

[53] NPSAS. National Postsecondary Student Aid Study, National Center for Education Statistics, NCES Handbook of Survey Methods. `https://nces.ed.gov/statprog/handbook/npsas.asp`.

[54] Psacharopoulos G. (1977). "The Perverse Effects of Public Subsidization of Education or How Equitable Is Free Education?" *Comparative Education Review* 21, 69–90.

[55] Rothschild, M. and L.J. White (1995). "The Analytics of Pricing in Higher Education and Other Services in Which Customers are Inputs." *Journal of Political Economy* 103, 573–86.

[56] Sabot, R. and J. Wakeman-Linn (1991). "Grade Inflation and Course Choice." *Journal of Economic Perspectives* 5, 159–70.

[57] Sarpça, S. (2010). "Multi-Dimensional Skills, Specialization, and Oligopolistic Competition in Higher Education." *Journal of Public Economics* 94, 800–11.

[58] Spence, M. (1973). "Job Market Signaling." *Quarterly Journal of Economics* 87, 355–74.

[59] Stinebrickner, T and R. Stinebrickner (2012). "Learning about Academic Ability and the College Dropout Decision." *Journal of Labor Economics* 30, 707–48.

[60] Strauss, J.C. and J.R. Curry (2002). *Responsibility Center Management: Lessons from 25 Years of Decentralized Management.* Washington, D.C.: National Association of College and University Business Officers.

[61] Winston, G.C. (1999). "Subsidies, Hierarchy and Peers: The Awkward Economics of Higher Education." *Journal of Economic Perspectives* 13, 13–36.

[62] Xu, D. and S.S. Jaggars (2013). "The Impact of Online Learning on Students Course Outcomes: Evidence from a Large Community and Technical College System." *Economics of Education Review* 37, 46–57.

Department of Economics, University of Alberta
Working Paper Series

2020-02: Misallocation across Establishment Gender – **Ranasinghe, A.**
2020-01: Does Locational Marginal Pricing Impact Generation Investment Location Decisions? An Analysis of Texas's Wholesale Electricity Market – **Brown, D.,** Zarnikau, J., Woo, C.
2019-16: A Unified Explanation of Trade Liberalization Effects Across Models of Imperfect Competition – **Alfaro, M.**, Lander, D.
2019-15: Analyzing Firm Behaviour in Restructured Electricity Markets: Empirical Challenges with a Residual Demand Analysis – **Brown, D., Eckert, A.**
2019-14: LinkedIn(to) Job Opportunities: Experimental Evidence from Job Readiness Training – **Wheeler, L.**, Garlick, R., Johnson, E., Shaw, P., Gargano, M.
2019-13: Entry Preemption by Domestic Leaders and Home-Bias Patterns: Theory and Empirics – **Alfaro, M.**
2019-12: How Much are Good Schools Worth? Evidence from School Acquisitions in Beijing – **Su, X.,** Yu, H.
2019-11: Enumerating Rights: More is Not Always Better – Ball, S., **Dave, C.,** Dodds, S.
2019-10: Motivating the Optimal Procurement and Deployment of Electric Storage as a Transmission Asset – **Brown, D.,** Sappington, D.
2019-09: Pricing Patterns in Wholesale Electricity Markets: Unilateral Market Power or Coordinated Behavior? – **Brown, D., Eckert, A.**
2019-08: Son Preference and Child Under Nutrition in the Arab Countries: Is There a Gender Bias Against Girls? – **Sharaf, M.,** Rashad, A., Mansour, E.
2019-07: How Local Economic Conditions Affect School Finances, Teacher Quality, and Student Achievement: Evidence from the Texas Shale Boom – **Marchand, J.,** Weber, J.
2019-06: Subjects in the Lab, Activists in the Field: Public Goods and Punishment – **Dave, C.,** Hamre, S., Kephart, C., Reuben, A.
2019-05: Fiscal Austerity in Emerging Market Economies – **Dave, C.,** Ghate, C., Gopalakrishnan, P., Tarafdar, S.
2019-04: Inequality and Trade Policy: Pro-Poor Bias of Contemporary Trade Restrictions – **Ural Marchand, B.**
2019-03: Determinants of Locational Patenting Behavior of Canadian Firms – **Eckert, A., Langinier, C., Zhao, L.**
2019-02: The Microeconomics of New Trade Models – **Alfaro, M.**
2019-01: Universal Childcare for the Youngest and the Maternal Labour Supply – Kunze, A., **Liu, X.**
2018-19: On the Benefits of Behind-the-Meter Rooftop Solar and Energy Storage: The Importance of Retail Rate Design – **Boampong, R., Brown, D.**
2018-18: The Value Premium During Flights – **Galvani, V.**
2018-17: Asymmetric Information, Predictability and Momentum in the Corporate Bond Market – **Galvani, V., Li, L.**
2018-16: The Momentum Effect for Canadian Corporate Bonds – **Galvani, V., Li, L.**
2018-15: Green Technology and Patents in the Presence of Green Consumers – **Langinier, C.,** Ray Chaudhuri, A.
2018-14: Subjective Performance of Patent Examiners, Implicit Contracts and Self-Funded Patent Offices – **Langinier, C.,** Marcoul, P.
2018-13: Ordered Leniency: An Experimental Study of Law Enforcement with Self-Reporting – **Landeo, C.,** Spier, K.