DISCUSSION PAPER SERIES

Why Does the U.S. Have the Best Research Universities? Incentives, Resources, and Virtuous Circles

W. Bentley MacLeod
Miguel Urquiola

<div style="border">

DISCUSSION PAPER SERIES

Why Does the U.S. Have the Best Research Universities? Incentives, Resources, and Virtuous Circles

W. Bentley MacLeod
Columbia University, NBER and IZA

Miguel Urquiola
Columbia University and NBER

</div>

ABSTRACT

Why Does the U.S. Have the Best Research Universities? Incentives, Resources, and Virtuous Circles[*]

Around 1870 the U.S. had no research universities of note, while today it accounts for the largest number in the world. Many accounts attribute this transformation to events surrounding World War II. In contrast, this paper traces its origins to reforms that began in the 1870s. We first explain the origins of the American system's weakness at research. We then present an agency theory framework that highlights ingredients necessary for enhanced research performance. These include specialization and meaningful performance metrics. We then discuss reforms that put these ingredients in place. For example: the introduction of specialized and advanced teaching and the ensuing rise of disciplines/departments; the creation of academic journals; the introduction of selective admissions. Throughout, we emphasize the role played by the U.S. university system's free market orientation.

Corresponding author:
W. Bentley MacLeod
Columbia University
Department of Economics
420 West 118th, Mail Code 3308
New York, NY 10027-7296
USA

E-mail: wbmacleod@wbmacleod.net

1. Introduction

Countries' educational performance is not uniform across different dimensions, nor is it static. For example, around 1875 the U.S. had substantially higher levels of schooling than any European country; the gap has since largely disappeared (Goldin and Katz 2008). By contrast, in the same year the U.S. had no leading research universities; today it accounts for the majority of the top-ranked (Cole 2009, Clotfelter 2010). One way to illustrate the latter is by counting the times universities are mentioned in Nobel prize winners' biographies. The idea is that schools that trained or hosted Nobel winners were likely places productive at research (each mention of a university is associated with a year, e.g., that in which the winner graduated from one institution, or became a professor at another). Figure 1 shows that by this measure, around 1875 American universities badly lagged their European counterparts; today they are distinctly ahead.
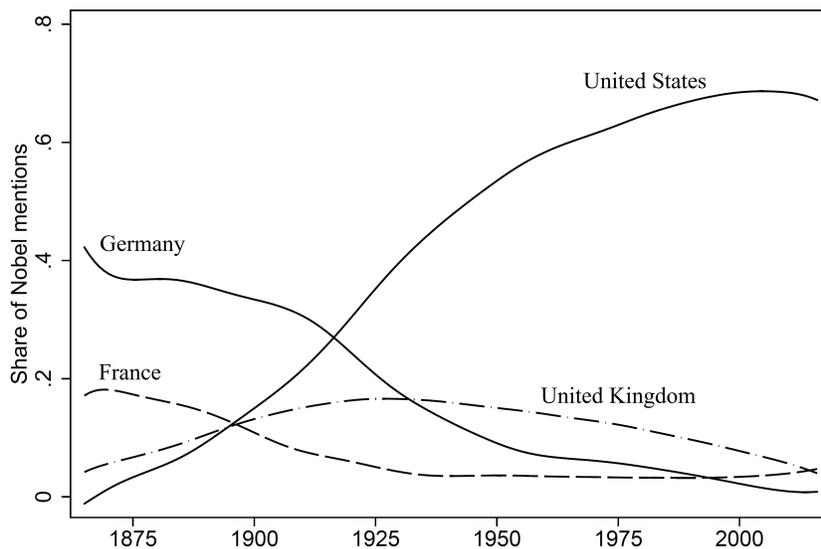
Figure 1.1. University Nobel prize mentions

This figure describes the frequency with which Nobel winners' biographies mention universities in different countries. It plots fitted values of locally weighted regressions of each country's share of total mentions on the year. The data cover the four countries with universities accounting for the most mentions. Source: Urquiola (2020).

Figure 1 displays "smoothed" data, making it hard to observe when the U.S. took the lead. Many observers cite events surrounding World War II as producing the turning point in American performance—for example, they highlight the migration of Jewish academics from Germany to the U.S., and the rise of federal research funding (e.g., Graham and Diamond 1997, Cole 2009, Gruber and Johnson 2019). Emphasizing WWII is analytically attractive: it highlights factors that surely strengthened American universities, and it addresses another salient aspect of Figure 1: the German decline.

However, Table 1 presents the raw data behind Figure 1 by decade, suggesting that U.S. universities' had matched or surpassed most countries' before WWII. Specifically, while American universities were not mentioned in Nobel biographies with associated dates before 1870, they were ahead of all but Germany's by 1901-1910, and ahead of all countries' by 1911-1920. To cite one anecdote illustrating this improvement, in 1901 Theodore Richard—who went on to win a Nobel Prize in Chemistry—became the first American-trained professor to be offered a position at a German university. In short, Table 1 implies that a successful explanation of American universities' research dominance must begin in the 1800s, and likely must emphasize factors other than events in Europe.

	Country	Total number of mentions in:						
		1855-1870	1871-1880	1881-1890	1891-1900	1901-1910	1911-1920	1921-1930
1	United States		2	4	7	24	45	75
2	United Kingdom	1	4	6	11	17	32	38
3	Germany	11	15	18	43	45	36	70
4	France	2	15	7	9	20	7	10
5	Japan							2
6	Switzerland	2	1	3	3	3	11	11
7	Canada					3	2	3
8	Sweden			4	3	4	3	9
9	Netherlands		5	4	1	3	2	3
10	Denmark			3	3	4	4	9

TABLE 1. Country-level university mentions

Notes: The table lists the number of times Nobel winners' biographies mention universities in different countries. It features only the ten countries with the most cumulative mentions. Source: Urquiola (2020).

The explanation we offer centers on reforms that began after the Civil War and enhanced the incentives and resources the American university system directs at research.[1] Understanding the origin of these reforms benefits from two observations. First, the U.S. takes a relatively free market approach to higher education; for instance, it generally allows easy entry by schools seeking to satisfy customer demand. Second, universities do not supply their customers with a single "product"; rather, they provide a complex set of services among which is sorting. When a school produces sorting, it implicitly "sells" peer groups—a chance to be exposed to and associated with certain types of people. For example, certain schools cater to students who are smart, or wealthy, or artistic, or headed to particular careers.

Since colonial times and well into the 1800s, American households demanded denominational sorting. For example, Presbyterians wished to attend college with Presbyterians, Episcopalians with Episcopalians, etc. Further, most preferred to do so close to home. Together, these factors produced massive entry: while in 1776 the U.S. had 9 colleges, about

---

[1] In emphasizing earlier events, our approach is closer to Veysey (1965) and Goldin and Katz (1999). Our account, particularly in sections 2, 4, and 6, draws on Urquiola (2020).

900 more opened before the Civil War. This at a time when most European countries' regulated entry and had a handful of universities.

Mass entry allowed colleges to satisfy the demand for sorting: they catered to students from denominationally-homogeneous regional elites. In some dimensions, however, mass entry also stunted the colleges. Small and underfunded, they offered a basic, narrow, and rigid curriculum delivered by unspecialized, poorly-paid professors. Colleges often hired faculty using criteria other than expertise, and provided them with few *incentives* or *resources* to carry out research. There was almost no way they could compete with European universities on the research front, and as Figure 1 shows, they did not.

This paper discusses how reforms allowed the system to begin providing incentives and resources. We begin with incentives, proceeding in three steps. First, we draw on agency theory to sketch out a framework of how universities provide professors with incentives to do research.[2] It highlights that even if universities are committed to research, producing it requires having reasonably precise research performance measures. We show that precision is enhanced by increasing specialization—roughly, organizing professors into disciplines.

Second, we describe the history of how American universities gained interest in research and increased specialization. To preview, over the 1800s colleges observed a gap grow between the skills they taught and those skills their customers demanded. With industrialization interest grew around areas college curricula essentially ignored (e.g., engineering, business). Incumbent schools attempted, but by and large failed, to lead curriculum reform. After the Civil War two entrants, Cornell and Johns Hopkins, succeeded—they showed that survival could lie in offering specialized and advanced instruction in a range of areas, rather than in supplying denominational sorting. Incumbents like Harvard and Columbia responded forcefully, forming specialized arts and sciences departments and creating professional schools. Soon Chicago, MIT, Stanford, and a couple dozen had joined the fray. All these schools soon saw the need for professors who were specialists at the frontier of a field, i.e., people who could do research. The academic system responded by providing measures of research quantity and quality, e.g., specialized journals.

Third, this began to allow schools to vary compensation with research output by offering more successful professors a higher salary. Further, the early 1900s saw the emergence of "lumpy" rewards for research, the most salient being tenure. Agency theory implies that these can also promote performance, particularly if professors compete for them against individuals of similar ability. The latter was ensured because the availability of increasingly precise measures of research talent set in motion a sorting process that ensured that professors clustered in departments with colleagues of similar ability.

---

[2] For reviews of agency theory see Eisenhardt (1989) and Prendergast (1999).

Taken together, these developments greatly enhanced the provision of incentives. But they leave open the question of how reforms allowed some American universities to acquire tremendous amounts of resources. For example Stanford provides its researchers with more resources than any European university; often these resources are not explicitly connected to performance, such as housing subsidies and well appointed buildings. A final section discusses the origin of this. To preview, in the early 1900s, the schools cited above were pleased to see their enrollments grow rapidly, but this began to threaten the match between the sorting they provided and that key students demanded. Enrollment growth brought in relatively lower-income students and Jewish students, and—in a period of rising antisemitism—alienated the elites the colleges had traditionally served. By the 1920s Columbia and Harvard implemented selective admissions, and several followed soon after. Selectivity set in motion another sorting process, one that concentrated high ability/income students—and eventually high tuition and donations—at such schools. In short, the top American schools came to enjoy a virtuous circle that gave them resources to invest in research, which they could now effectively incentivize. This created inequality at the same time as it helped to produce the top-performing research university sector in the world.

2. The initial American weakness

A demand for denominational sorting drove the creation of the original nine American colonial colleges. For example, Massachusetts Puritans created Harvard to produce what they saw as a theologically-sound education. They took this seriously; e.g., the colony forced out Harvard's first President, Henry Dunster—even as it widely praised his management—because of his stance on infant baptism. Analogous objectives led to the creation of the College of New Jersey (Princeton) by Presbyterians, King's (Columbia) by Anglicans, Queen's (Rutgers) by the Dutch Reformed, Brown by Baptists, etc.[3]

In addition, like school-users to this day, the early founders desired proximity—to have schools close to home.[4] For example, Connecticut-based Puritans created Yale because they perceived that Boston-based Puritans' Harvard was too physically distant (in addition to too religiously liberal). After independence, the desire for sorting and proximity led to massive entry: historians estimate that 900 colleges opened before the Civil War.[5] This happened as denominations competed, including on the expanding western frontier.

Well into the 1800s, therefore, American colleges mainly catered to small local markets. It was common for them to have fewer than 100 students and 5 faculty members.[6]

---

[3] See Herbst (1982).
[4] For evidence of such a preference see Card (1995), Neilson (2017), and Abdulkadiroglu et al. (2019).
[5] See Tewksbury (1932), Hofstadter (1955), Burke (1982), and Urquiola (2020).
[6] Geiger (1986), Rudolph (1962).

The latter were often hired with attention to their religious affiliation. For example, the desire for clarity on this front led colleges to prefer their own graduates, and to resist sending them away for training. In 1879 Princeton President James McCosh wrote to an alumnus studying in Germany:

> You are aware that the Trustees ... are resolute in keeping the college a religious one. You have passed through varied scenes since you left us ... If a man has the root in him he will only be strengthened in the faith by such an experience. It will be profitable to me to find how you have stood all this.[7]

Wealth-related criteria mattered too. Rudolph (1962) writes that

> From 1835 to 1852 chemistry at Williams was taught by a man of independent wealth who spent his token salary on laboratory equipment. The appointment of Henry D. Rogers ... at ... Pennsylvania ... was a response to his offer to serve without salary. Amasa Walker joined Oberlin ... on the same terms.

This reflects that professorships often paid relatively little. Historians state that even in the 1860s, salaries at Harvard and Yale were "below the cost of living."[8]

Given the salience of religion and wealth in hiring, expertise could take a back seat. For example, in 1853 James Renwick, Columbia's sole professor of natural philosophy, announced his retirement. Among the mainly Anglican trustees who began to search for his replacement was Samuel Ruggles, a real estate investor and politician.

The candidate Ruggles favored was Wolcott Gibbs, a chemist with sterling credentials who would later teach at City College and Harvard. Some trustees, however, noticed a blemish on his record: he was a Unitarian. Ruggles tried to preempt the opposition, proposing a motion that:

> Whereas the original charter incorporating King's ... College ... provides that [its] laws ... shall not 'extend to exclude any person of any religious denomination whatever, from equal Liberty and advantage of Education' ... RESOLVED that in filling [this] Professorship, the Trustees are legally and morally bound to select such Professor, with reference solely to his fitness for the place, without regard to his religious opinions.

This was roundly rejected.

Other types of identity could matter too. In 1802 Yale President Timothy Dwight recruited Benjamin Silliman to teach chemistry. That Silliman would become an illustrious teacher would not have been obvious at the time; as Kelley (1974) notes,

---

[7] Leslie (1992).
[8] Flexner (1946).

> Silliman had never studied chemistry and knew almost nothing about it. ...
> Dwight chose Silliman because he knew it would be difficult to find anyone in
> America with knowledge of chemistry and natural history, and he was afraid
> to select a foreigner.

In short, the colleges' did not prioritize trained or specialized personnel.

It is therefore not surprising that they made an associated choice: well into the 1800s they offered a basic curriculum featuring few if any electives. The curriculum emphasized Latin, Greek, logic, rhetoric, mathematics, physical sciences ("natural philosophy"), and ethics and politics ("moral philosophy"). It excluded "practical" fields such as business and engineering. When it involved science, the emphasis was deductive rather than experimental/inductive.[9] In addition, the teaching methods involved rote learning, e.g., "recitations" in which students declaimed memorized passages.

These curricular and pedagogical choices meant that classes could be staffed by few instructors. One faculty member at Williams taught rhetoric, English literature, aesthetics, and political economy. Another at Dartmouth, was tasked with teaching "... English, Latin, Greek, Chaldee and such other languages as he shall have time for." Even more extreme, in the 1700s tutors delivered all instruction to a given cohort (e.g., the class of 1776) into the junior year. Given such a task, modern professors might quickly warm to rote learning!

These choices kept costs down, which was a priority in the face of limited enrollments and revenues. Serious budgetary pressures were common; Boyer (2015) states that in these years the vast majority of colleges hovered between "genial penury and unmitigated fiscal disaster."

The whole setup also reflected schools' inability/unwillingness to branch out. For example, to the extent that instruction in applied science developed, it was mainly outside colleges. A notable venue was the Rensselaer Polytechnic Institute (RPI), which offered engineering, surveying, and applied science, and featured some of the first laboratories in the U.S.

However, schools like RPI also offered narrow curricula and little advanced instruction. Rudolph (1977) states that around 1830 a motivated college graduate could have earned an RPI degree in twenty-four weeks. Further, such schools' existence facilitated colleges' claim that science belonged elsewhere. When Williams hired Ira Remsen to teach chemistry, he asked for a small room in which he might set up a laboratory using his own equipment. The response: "You will please keep in mind that this is a college and not a technical school."[10]

This is not to say that no research went on in American colleges. For example, in the 1830s, Amherst's Edward Hitchcock completed the first geological survey of Massachusetts;

---

[9] See Butts (1939), Storr (1953).
[10] Rudolph (1962).

Harvard's Asa Gray corresponded with Charles Darwin on the Origin of the Species. But such efforts were exceptions rather than the rule.

Meanwhile, European universities, largely under state control, were fewer and much better-funded. Faculty were often well-paid civil servants, assigned by ministries of education to institutes headed by specialized research-focused "chaired" professors.[11]

The contrast this produced is evident in two schools—one American, one German—both chartered in the mid-1700s by George II (King of England and ruler of German lands around Hanover). By the 1850s, the American, Columbia, had 6 professors and 150 students; the German, Göttingen, about 90 and 1,600, respectively.[12] Table 2 returns to the data to list the twenty universities most mentioned in Nobel winners' biographies in roughly forty-year periods. For 1855-1900, Göttingen was among the top 10 (first column). Columbia—like all other American schools—did not make the top 20.

Total number of mentions for:							
1855-1900		1901-1940		1941-1980		1981-2016	
Humboldt Univ. of Berlin	19	University of Cambridge	63	Harvard University	##	Harvard University	32
University of Cambridge	11	Harvard University	35	University of Cambridge	83	Stanford University	28
University of Munich	10	Humboldt Univ. of Berlin	31	Columbia University	64	Massachusetts Inst. of Tech.	27
University of Strasbourg	10	Columbia University	29	University of Chicago	64	Univ. of Calif., Berkeley	20
University of Paris	9	University of Gottingen	27	Univ. of Calif., Berkeley	62	University of Chicago	18
Leiden University	8	University of Munich	25	Massachusetts Inst. of Tech.	48	Yale University	17
University of Wurzburg	7	University of Paris	21	Stanford University	46	Princeton University	16
Heidelberg University	6	University of Chicago	20	University of Oxford	33	University of Cambridge	13
University of Copenhagen	6	University of Manchester	18	Princeton University	32	Columbia University	12
University of Gottingen	6	University of Oxford	17	California Inst. of Tech.	29	University of Texas	12
ETH Zurich	5	ETH Zurich	16	Yale University	26	Johns Hopkins University	10
University of Bonn	5	California Inst. of Tech.	15	Cornell University	23	California Inst. of Tech.	9
University of Vienna	5	University of Copenhagen	15	University of Illinois	20	Rockefeller University	9
Ecole Normale Superieure	4	University of Wisconsin	15	University of Paris	20	Cornell University	8
University of Manchester	4	Johns Hopkins University	14	Carnegie Mellon University	19	Nagoya University	8
Leipzig University	3	Univ. of Calif., Berkeley	12	University of Pennsylvania	19	Northwestern University	8
University of Giessen	3	University of Marburg	12	Johns Hopkins University	15	University of Pennsylvania	8
University of Graz	3	Uppsala University	12	New York University	15	Univ. of Calif., San Diego	7
University of Lyons	3	Cornell University	11	Rockefeller University	15	Univ. of Calif., Santa Barbar	7
University of Stockholm	3	Princeton University	11	Technical Univ. of Munich	15	University of Oxford	7

TABLE 2. University-level mentions in given time periods

Notes: The table lists the number of times Nobel winners' biographies mention specific universities in given time periods. It only lists the top 20 schools within each period. Source: Urquiola (2020).

What happened since the 1850s? To start, even more schools entered the market. Today, the U.S. has about 4,700 post-secondary institutions. By the logic above, its universities

---

[11] Paulsen (1906).
[12] Ruggles (1854).

might be even weaker at research. And yet, the country has about 100 "Research I" universities that lead the world in this dimension.[13] For instance, they account for 17 of the top 20 schools in the last column of Table 2 (for 1981-2016). The next sections review how this transformation happened.

3. AGENCY THEORY AND RESEARCH PERFORMANCE

We begin by sketching out a framework that helps to make sense of these reforms (details are in the appendix). Our starting point is the fundamental idea that individuals respond to rewards, and that rewards are shaped by performance measures.

Suppose there are two faculty members, $i \in \{A, B\}$. For simplicity, assume that both work at the same university. Each professor exerts effort $\vec{e}_i = (r_i, o_i)$, which has two components: effort on research, r_i, and effort on "outside" activities, o_i. We think of outside activities broadly; they could include consulting work, but also staying active in alumni networks or, in the 1800s, religious groups.

We suppose that professors are intrinsically motivated to do research, and would choose effort $r_i > 0$ even in the absence of any explicit reward; e.g., in the early 1800s some American faculty engaged in research at their own expense. In general, though, we suppose there is a participation constraint and professors will not work for the university unless they receive a reservation utility.

Suppose the university chooses to place some weight on its research output, as captured by V^r. The institution's welfare is thus:

$$(3.1) \qquad W = V^r \times (r_A + r_B) - I_A - I_B,$$

where I_A and I_B are the compensation it pays professors A and B. This assumes only human capital is needed to produce research; we return to the role of financial inputs later.

Suppose professor i's utility depends upon her effort and compensation:

$$U(\vec{e}_i, I_i) = E\{I_i\} + V^o \times o_i - C_s(\vec{e}_i) - \frac{\kappa}{2} var(I_i).$$

The first term on the right-hand side is expected compensation. The second is the return to outside effort, which depends upon the parameter V^o.[14] The third is the cost of effort, which we assume is such that outside activities increase the marginal cost of research effort.[15]

---

[13] "Research University-I" was a category in the taxonomy of higher education institutions developed by the Carnegie Commission on Higher Education.

[14] We assume that $V^r < V^o$; otherwise no one would work at the university.

[15] Formally,

$$C_s(\vec{e}_A, r^0) = \left[\vec{e}_A - \begin{bmatrix} e_0 \\ 0 \end{bmatrix}\right]^T \begin{bmatrix} 1/2 & d/2 \\ d/2 & 1/2 \end{bmatrix} \left[\vec{e}_A - \begin{bmatrix} e_0 \\ 0 \end{bmatrix}\right]$$

where $1 > d > 0$.

9

The fourth term, $\frac{\kappa}{2}var(I_i)$, is the cost associated with a risky income stream, where κ is the individual's absolute risk aversion.[16]

A key question is whether compensation, I, can be used to reward research performance. This depends crucially on whether the university observes/uses a research performance measure. Suppose first that it does not, and it simply sets compensation equal to a fixed wage $I_i = \bar{w}$. Proposition 1 summarizes the resulting situation:

Proposition 1. *Suppose faculty are paid a fixed wage that ensures participation. The existence of outside activities (with return V^o) leads professors to supply less research effort than the intrinsic level they desire.*

This captures elements of the American colleges into the 1800s. There were few rewards for research, and yet some professors engaged in it. This was likely easier for wealthier faculty members with less need for outside activity.[17] In addition, the return to outside effort, V^o, could be high. For instance, the Gibbs case illustrates that one might more effectively land a professorship by signaling one's denomination than by doing research.

Now suppose that a measure of professors' research performance exists:

$$s_i = r_i + \alpha_i + \epsilon_i,$$

where r_i is research effort, α_i is ability, and ϵ_i is a noise term distributed normally with mean zero and variance σ_ϵ^2. Principal-agent theory (e.g., Holmstrom 1979) implies that given objective (3.1) the university should make compensation depend on this measure. Indeed, modern universities make some forms of compensation—e.g., salaries and raises—vary with research performance.

Such rewards provide incentives, but they entail a cost because they expose professors to risk (measured by $\frac{\kappa}{2}var(I_i)$). In this sense a more precise performance measure is a "win-win"—it reduces risk and allows for better incentives. A key question, therefore, is how one might enhance the precision of s.

This question matters because in any university, comparing professors' research performance is challenging. For example, in the humanities faculty tend to write books, while in the sciences articles are the norm—it is difficult to compare these. Similarly, a chemistry professor's name may appear on all papers produced by her laboratory, while an economic theorist writes alone—it is hard to tell who is working harder. These examples motivate that measurement can be enhanced by comparing people in the same field/discipline.

---

[16] We follow Holmström and Milgrom (1991) in supposing that individuals have constant absolute risk aversion and all error terms are normally distributed.

[17] Even today faculty members in some countries are paid little. As a result, only a limited share of their effort goes towards research, and they tend to engage in activities like consulting.

To see this, assume that the overall variance on the performance score, ϵ_i, remains fixed, but decompose the error into two terms:

$$\epsilon_i = \gamma_f + \gamma_i$$

γ_f captures the variation in output between fields—e.g., professors in the same discipline are subject to similar standards. γ_i is the within-field variance; essentially, luck.

If we compare the performance of two professors, A and B, in the same field we have:

$$s_{AB} = s_A - s_B = r_A + \alpha_A + \gamma_A - (r_B + \alpha_B + \gamma_B).$$

In words, within-field comparisons remove the common term, γ_f, allowing for a more precise measure. Its precision is given by:

$$\rho_s = \frac{1}{var\,(s_{AB})} = \frac{1}{var\,(\gamma_A) + var\,(\gamma_B)}.$$

In short, specialization is a key way to increase the precision of performance measures.

In the appendix we show that when s_{AB} is available, the optimal compensation scheme for professor A takes the form (the equivalent holds for B):

(3.2) $$I_A = \bar{w} + b_{s_A} \times s_A + b_{s_{AB}} \times s_{AB}.$$

In addition to a fixed wage there is a person-specific reward, weighted by b_{s_A}, and a comparative reward, weighted by $b_{s_{AB}}$, that pays for *relative* performance in the same field. As specialization increases (as the precision ρ_s increases) it makes sense to place greater weight on relative performance: $\rho_s \to \infty$, $b_{s_A} \to 0$ and $b_{s_{AB}} \to V^R$. In the limit the weight on relative performance approaches the weight the university places on research in (3.1).

Proposition 2 summarizes actions that enhance a university system's research performance:

Proposition 2. *The following actions enhance research performance: i) increasing specialization and hence the precision of the performance measure ρ_s, ii) increasing the weight the university places on research, V^r; iii) reducing the return to outside activities, V^o.*

We now discuss how reforms that began in the 1800s promoted these changes

4. Reforms

As industrialization advanced, the colleges increasingly failed to satisfy some facets of educational demand. Some students were interested in areas the colleges did not cover, such as business and mining/engineering. Others, wanted specialized, modern training in areas like chemistry and economics. Over the first part of the 1800s, reformers at several schools

(e.g., Brown, Harvard, Union) attempted to address these shortcomings, but encountered resistance and failed to sustain their initiatives.[18]

Lasting change arrived with two entrants: Cornell (1865) and Johns Hopkins (1876). These leveraged private and public financial support to expand the curriculum and offer advanced and specialized instruction.[19] Cornell, in particular, saw rapid growth in enrollment, showing that survival could lie in responding to curricular demand rather than in providing denominational sorting. For its part, Hopkins emulated German universities' focus on advanced/graduate instruction. Its first president, Daniel Gilman, promoted hiring with an eye to expertise and research ability, rather than aspects like denominational affiliation:

> The institution we are about to organize would not be worthy the name of a University, if it were to be devoted to any other purpose than the discovery and promulgation of the truth; and it would be ignoble ... if the resources ... given by the Founder ... should be limited to the maintenance of ecclesiastical differences or perverted to the promotion of political strife. ... sectarian and partisan preferences should have no control in the selection of teachers ... .[20]

In our framework, Gilman was increasing the weight on research, V^r, and decreasing that on outside activities, V^o.

Incumbent schools had to decide how to respond. Harvard and Columbia aggressively invested to surpass Cornell and Hopkins in dimensions like offering electives, funding research, and creating graduate schools of arts and sciences and professional schools (business, engineering, law, medicine). Other well-heeled incumbents—notably Princeton and Yale—generally followed, but opted to retain a greater focus on their undergraduate colleges. Yet others like Amherst, Swarthmore, and Williams remained fully committed to undergraduate curricula.

Even this last set, however, began to offer majors and electives. Even these schools began to split academic areas into specialized fields; e.g., natural philosophy into astronomy, biology, chemistry, and physics; political economy into economics, political science, and sociology, and so on. Specialization—measured by $\rho_{s_{AB}}$ in our framework—was on the rise.

On all these fronts, the reformers soon faced additional competition from well-funded public schools—e.g., Berkeley, Michigan, and Wisconsin—and from other well-heeled private competitors—e.g., Chicago, M.I.T., and Stanford. Some of these innovated in ways of their own. For example, Chicago hired faculty unusually aggressively; Berkeley emphasized hiring

---

[18] See Butts (1939), Rudolph (1962), and Urquiola (2020).

[19] These schools received large donations from Ezra Cornell and Johns Hopkins, respectively. In addition, Cornell received federal funds via the Morrill Act, and Hopkins eventually from the state of Maryland (Hawkins 1960, Curti and Nash 1965).

[20] Hakins (1960, 22).

junior faculty and avoided "inbreeding" (another way of reducing V^o); M.I.T. and Stanford moved quickly to collaborate with industry.

Competing to deliver advanced and specialized instruction, these universities needed professors at the cutting edge of their fields. Initially such academics were in short supply and/or hard to identify. For example, early in his long presidency of Harvard (1869-1909), Charles Eliot observed,

> There is in this country a very considerable body of teachers who know how to teach Latin and Greek, and the elements of language; but if you are in search of teachers to teach botany, chemistry, physics and so on, you cannot find them. They do not exist.[21]

This changed with two developments which took place in a decentralized manner. First, there was human capital investment as the supply of Ph.D.-trained specialists grew. Second, professors' expertise began to be ascertained by their research ability, with the academic system providing ways of measuring performance.

Faculty members founded professional associations, e.g., the American Chemical Society (1877) and the American Historical Association (1884). These began to publish journals (e.g., the American Economic Review) which, via the peer-review process, certified research quality. These were supplemented by university department journals; e.g., Harvard's Quarterly Journal of Economics (1886) and Chicago's American Journal of Sociology (1895). In time journals within each field differentiated in quality. As the better ones received more submissions, their editors made them more selective and prestigious, thus attracting more papers—a virtuous circle with corresponding losses for others. The result was a hierarchy in each field.

Charles Eliot reflected on the impact this had over the course of his presidency:

> The chief difficulty that I encountered was the procuring of teachers competent to give advanced instruction. There were really no guides to the discovery and invitation of the persons needed. Then none of the societies organized for the ... mutual support of learned ... men existed. By 1885 I could get some assistance ... from the proceedings of the ... scientific societies. At the beginning there was no such aid.[22]

Again, there appeared more precise measures of research performance—higher $\rho_{s_{AB}}$.

The demand for research talent improved professors' negotiating position. One exponent of the behavior leading to this was William Harper, Chicago's first president. Harper raided campuses including that of Clark University, which had made a significant but ultimately

---

[21] (Flexner 1946, 41). This quote is from 1872.
[22] (James 1930, 12).

underfunded attempt to compete with Johns Hopkins at graduate teaching; Harper hired away fifteen of its professors. To fend off such raids, schools were soon offering higher salaries, reduced teaching loads, and sabbaticals.

At Swarthmore, President Joseph Swain informed an astronomer he was recruiting that he would accommodate his demand for a 24-inch telescope, adding a caveat: "remember, this is a Friend's College and thee should give up thy smoking." In the event, Swain had to accommodate the professor's tobacco habit too.

5. TENURE

Higher salaries and smoking privileges illustrate that compensation began to respond to research performance, including in subtle ways. This is very much in the spirit of compensation as given in (3.2) above. But at the same time, this period saw "lumpy" rewards for research emerge, the most notable being tenure. Unlike salary, tenure does not allow for small enhancements in compensation given to a professor because she performs slightly better. Rather, tenure is a significant reward given for reaching some threshold level of performance, and it cannot be taken away. This section describes the emergence of tenure and then discusses how agency theory suggests it can also incentivize research performance.

5.1. **The emergence of tenure.** One can see tenure as arising from two interrelated sets of demands. The first was a desire on the part of professors for academic freedom and protection from arbitrary dismissal. Several events that illustrate this involves economists, in part because as their field emerged, many acted as reformers in addition to researchers, creating tension with university presidents and donors. A famous case was that of Richard Ely, who Andrew Gilman hired at Johns Hopkins. Ely was active in the progressive movement and supported labor unions.

He later moved to the University of Wisconsin at a time of labor tensions in that state. Soon President Charles Adams heard public complaints that Ely supported strikes, taught elements of socialism, and had "entertained a union organizer in his home." The university's board appointed a committee to investigate. The board supported Ely and made a broader statement:

> As Regents of a university with over a hundred instructors supported by nearly two million of people who hold a vast diversity of views ... we could not for a moment think of recommending the dismissal ... of a teacher even if some of his opinions should, in some quarters, be regarded as visionary. Such a course would be equivalent to saying that no professor should teach anything which is not accepted by everybody as true. ... We cannot for a moment believe that knowledge has reached its final goal, or that the present condition of society is

perfect. We must therefore welcome from our teachers such discussions as shall suggest the means and prepare the way by which knowledge may be extended, present evils be removed and others prevented.

That the outcome could differ, however, is evident in the case of Edward Ross, one of Ely's students. Ross taught economics and sociology at Stanford, where he publicly defended Eugene Debs and opposed Asian immigration. This irritated Jane Stanford, who had remained as the school's main trustee after the death of Leland Stanford. She accused Ross of aiding the "vilest elements of socialism," and pressured President David Jordan to terminate him. Jordan obtained Ross's resignation, but at the cost of much bad press and eight additional departures. The case contributed to the creation of the American Association of University Professors (AAUP), which stated a desire to formulate principles "... respecting the tenure of the professional office ... ."

A second set of demands arose because historically university leaders desired mechanisms to incentivize and terminate professors. Harvard's early hiring practices illustrate some associated challenges. Early on Harvard hired tutors using renewable contracts, a practice which likely reflected that immediate indefinite appointments can generate adverse selection. In 1716 Harvard officials stated that permanent contracts raised the risk that bad instructors might be "fixed on the college for life." In addition, renewable contracts can provide incentives; in the words of Harvard's statutes, they could "excite tutors from time to time to greater care and fidelity in their work." Another benefit of finite terms is that if a professor performs badly, the school does not have to state a cause for dismissal—it can simply let the contract end quietly. At the same time, term contracts are not a panacea. As any manager knows, terminating a worker can be costly even if it is contractually feasible; when renewals come up, the temptation to "kick the can" is powerful. Metzger (1973) states that a 1760 Harvard rule limiting appointments within a rank to eight years:

> ... was to prevent incumbencies from being lengthened by re-appointments given out of neglect or sympathy; it was intended to defeat the importance of kindness in the serious business of evaluating personnel.

The two sets of demands—for permanence/freedom and for performance/incentives—could become intertwined. In the 1930s, Harvard president James Conant wished to strengthen his school's social sciences, widely perceived to have fallen behind Chicago's and Columbia's. Two faculty members with somewhat limited publication records, Alan Sweezy and Raymond Walsh, came up for reappointment. For a change, both were economists with labor union ties. When they were terminated, some complained that politics had been the cause. A faculty committee found no evidence of political bias, but questioned the fairness of the timing of their review. Cases like these made clear that it would be difficult to disentangle

the desire for permanence/freedom and the desire to provide incentives. It is possible to see tenure—by creating a predefined trial period and permanence after an "up our out" date—as addressing both sets of demands.

It is not entirely clear when tenure became formalized because, as many features of the American system, it appeared in a decentralized fashion. Salient parts of the package had emerged at Princeton by the 1920s, and at several schools by the 1930s; by 1950 the full package was commonplace.[23]

5.2. **Tenure and agency theory.** Agency theory also provides ways to understand tenure. Even with improvements in the precision of research measures, there always remains an element of subjectivity, which can lead the university to shirk on its compensation commitments.[24] A solution to this problem is for the university to commit in advance to providing rewards for relative performance (Carmichael 1983, Malcomson 1984). For example, it might commit to providing a teaching award or a research honor, leaving in question only who will win it.

There is, however, a wrinkle. If there is substantial variation in ability, then some individuals will have a very low (or high) probability of winning the prize, and for them the scheme will not enhance performance (in the appendix we show that lower variation in ability increases the effectiveness of prizes; see Brown (2011) for related evidence).

This implies that all else equal, the emergence of tenure would have more effectively enhanced incentives for research if it was combined with a process whereby professors were sorted into schools/departments according to their research ability.

In fact, that is exactly what happened. Once measures of professors' research output were available, and once universities began to bid on the most able, stratification by ability/performance was the outcome. In other words, in the U.S. fields often display a relatively clear pecking order of departments/schools in prestige. This sorting process was facilitated by professors, given that they themselves have incentives to segregate by ability. For example, some believe that there are externalities from proximity to talented colleagues. Others might have informational/signaling concerns: individuals who belong to an elite department have their reputation enhanced by association.[25]

Further, tenure itself likely facilitated such sorting. Note that the best individuals to judge faculty in a specialized area are of course other faculty in that area. This can create a problem when hiring. If the employment of incumbents depends upon relative evaluations,

---

[23] It is worth noting that tenure has become less common over the years, it is increasingly available only to research professors at wealthier universities.

[24] The point is that in practice perfect measures do not exist. MacLeod (2003) shows that this necessarily entails conflict; the amount of conflict decreases as the quality of information improves.

[25] MacLeod and Urquiola (2015). See MacLeod et al. 2017 for direct evidence that such informational mechanisms operate.

then they will naturally want to hire individuals less skilled than themselves. Carmichael (1988) observes that the institution of tenure helps solve this problem. If one has an internal hierarchy and faculty with tenure have protected employment, then they have no incentive to misrepresent the quality of new hires.

In short, it is possible to see an incentive-enhancing complementarity in the development of tenure, more precise performance measures, and faculty sorting.

6. Resources and admissions related reform

We have so far not addressed how universities were able to finance so many reforms. How were top departments able to corner the market on talented researchers, pay them high salaries, offer them sabbaticals, etc.? How did the American system transition from one in which all colleges existed in "genial penury" to one in which a (small) minority of its universities are extremely rich? Today, not only does the U.S. spend more on higher education than other wealthy countries, its university system displays substantial inequality. Hoxby (2016) estimates that the most selective universities spend about 150 thousand dollars per student—fifteen times as much as many of their less well-heeled domestic counterparts. How did this wealth and inequality develop?

A crucial part of the story involves selective admissions, implemented by some early reformers cited above. Once these were in place, they produced a financial virtuous circle for the schools that moved early. They took in bright students and provided them with good experiences including teaching by the best professors (at least in terms of their research ability). They gave them outstanding networks that helped produce valuable job contacts and attractive marriage prospects. Their satisfied alumni then turned around and donated, allowing each school to make even more investments, and so forth.

To see the origins of this, note that before the 1920s essentially all colleges had "open enrollment" policies indicating a desire to grow. Harvard's Charles Eliot's (1908) stated:

> The American universities have always and everywhere been desirous of increasing the number of their students; and this is a true instinct of university governors in a democratic country.[26]

Such policies, combined with the curricular reforms that rendered schools more attractive, produced enrollment growth. The historically cash-strapped schools welcomed this development. In the 1910s Columbia and Harvard—perhaps the most aggressive reformers—were the largest American universities by enrollment, larger than any public school. This was still a time in which enrollment was seen as a signal of quality, in part because it provided resources to pursue further reforms.

---

[26] Some of the older eastern schools had admission exams, but even they allowed large percentages of students to enter "with conditions," i.e., despite having failed (Karabel 2006).

And yet enrollment growth began to create its own challenges. Because schools continued to serve relatively local markets, growth entailed the arrival of less prepared/lower income students. This resulted in greater heterogeneity in the student body, which in turn threatened schools' ability to deliver the type of arenas their "elite" users expected. In other words, schools' traditional customers might have felt that their educational spaces were being invaded, or that the old colleges were losing some of their signaling value.

In MacLeod and Urquiola (2015) we show that if schools had low fixed costs, then in the presence of such concerns there would be a large number, each catering to individuals of specific types. But in reality fixed costs and coordination challenges prevented elites from easily seceding from the schools they had historically used.

Instead, they began to form exclusive clubs within their schools. At Harvard, Princeton, and Yale, wealthier students sorted into private dormitories, "eating clubs," and secret societies; at Cornell, Michigan, and Wisconsin, into certain fraternities and sororities.[27] Beyond providing exclusivity, these clubs performed tasks that due to rapid growth, the schools had neglected, e.g., supplying dormitories and dining rooms.

Many students coveted membership in these clubs intensely. Franklin Roosevelt's experience is illustrative; from the moment he arrived at Harvard, he wished to join Porcellian, the most exclusive "final" club. He later described his failure to do so as the greatest disappointment of his life. Eleanor Roosevelt added that it was this experience helped FDR "identify with life's outcasts."[28] Similarly, a student at Princeton remarked on being rejected by an eating club: "The news came like a thunderbolt. With a cold, sick feeling the bottom dropped out of my college life."[29]

While such reactions might feel completely out of proportion, they are reminiscent of the disappointment students today feel about not getting into some elite colleges. This illustrates that in this period the clubs—rather than the colleges themselves—began to provide the sorting students coveted.

Initially many university leaders did not see that this created problems. In 1892, for example, Charles Eliot wrote that the clubs' existence simply reflected that:

> Rich people cannot be made to associate comfortably with poor people, or poor with rich. They live, necessarily, in different ways, and each set will be uncomfortable in the habitual presence of the other. Their common interests are unlike, and their pleasures are as different as their more serious occupations.[30]

---

[27] See Synnott (1979), Karabel (2006), Urquiola (2020).
[28] (Karabel 2006).
[29] (Karabel 2006, 59).
[30] (Veysey 1965, 91).

But in time leaders realized the clubs were problematic. First, they could impact the student experience and alumni generosity. Graduates might be grateful to their club rather than to their school, especially if the clubs were the ones providing the valuable networks. Second, a school that loses control of the student experience loses control over its admissions and image. For instance, in the early 1900s some students avoided Yale because they perceived getting into its clubs as too difficult.[31] Third, the clubs could promote unruly/unscholarly behavior; in the words of Princeton's Woodrow Wilson: "The side shows are so numerous, so diverting, so important, if you will, that they have swallowed up the circus."[32]

To complicate matters, this happened in a period of rising anti-semitism when many of the new students—particularly at urban schools such as Columbia, Harvard, and Pennsylvania— were Jewish. While in the early 1800s Jewish students barely accounted for any of these schools' enrollments, by 1920 their share was 30 percent at Columbia. These are estimates, since initially colleges had little data on religious affiliation. At Harvard—in an exercise dreamed up for today's machine learning—President Abbott Lowell asked a committee to use personal data to assign students to three groups: J1—"conclusively Jewish", J2— "indicatively Jewish", and J3—"possibly, but not probably Jewish".[33]

In-demand schools, led by Columbia, responded to these perceived challenges by implementing selective admissions. They capped their class sizes, started requiring extensive data from applicants, and began to reject some without stating a reason. Selectivity made for an academically stronger student body, and the fact that its size became predictable made it easier to supply services like dormitories and dining rooms. While selective admissions began at few schools (e.g., Columbia, Harvard, Stanford) they were soon adopted by many competitors.

Selectivity set the stage for a massive sorting process—in this case involving students. The academically more talented began to congregate in the schools that had reformed their curricula early and become selective. Just as happened with journals, a pecking order of selectivity developed. In addition, just as the appearance of journals facilitated the sorting of professors, the introduction of the SAT facilitated student sorting—in both cases observability was key. Just as with journals, the process created winners and losers. Many schools whose students had been quite similar to Harvard's in ability, found that was no longer the case, as documented by Hoxby (2009). Finally, an "open enrollment" sector continued to exist, as predicted by theory (Epple and Romano 1998, MacLeod and Urquiola 2015).

It is also relevant that federal research funding—which became salient around WWII— by and large reinforced rather than negated this pattern of inequality. Such funding is

---

[31] Karabel (2006).
[32] Brubacher and Rudy (1958)
[33] (Synnott 1979).

distributed according to proposal quality and assessments by panels of experts. To the extent that the top schools have cornered a lot of research talent, it is not surprising that schools like Berkeley and Hopkins receive substantial shares of NSF and NIH funding.

7. Conclusion

We have addressed why the American university system performs well at research—why a proportion of its institutions (e.g., "R1" universities) excel in this dimension. Our claim has been that this originates in decentralized reforms that improved the *incentives* and *resources* directed at research.

In the late 1800s, industrialization increased universities interest in advanced/specialized instruction, and hence on professors' ability to do research. The emergence of specialized fields (with accessories like academic journals), allowed for improved measures of performance. Armed with these, schools began to bid for and reward good researchers. This concentrated talent at a few schools, and the emergence of tenure reinforced these processes. The schools that most aggressively pursued these reforms also began to experience growth, and responded to associated challenges by putting in place selective admissions. This unleashed a further sorting process, allowing them to attract the most talented students—along with high tuition payments and donations.

In short, complementary and self-reinforcing dynamics concentrated research talent (increasingly recruited from all over the world) at a few schools, along with money to provide laboratories, students, etc.

It is useful to contrast the result with another educational setting: K-12. A country's performance in international K-12 tests like PISA necessarily depends on the state of thousands of schools. By contrast, the nature of research is such (e.g., top quality output is disproportionately important) that national performance can depend on what happens at a few dozen universities. In addition, the very existence of the "teacher value added" literature illustrates the difficulty of measuring performance in K-12. We have made the case that—when it comes to research—the university sector is rather different. Of course, performance measures at the university level are not costless—witness the many hours invested by faculty into drafting referee reports and tenure letters.

In closing we note that agency theory highlights that performance depends crucially on having high quality performance metrics. Today the academy is rife with complaints that individuals focus too narrowly, and with calls for "interdisclipinary" work or academic units. Our results show that a move in that direction comes with costs too, since it may entail noisier evaluation systems and lower research effort and output.

References

Abdulkadiroglu, A., P. Pathak, J. Schellenberg, and C. Walters (2019). Do parents value school effectiveness? Technical report, National Bureau of Economic Research Working Paper No. 23912.

Brown, J. (2011). Quitters never win: The (adverse) incentive effects of competing with superstars. *Journal of Political Economy 119*(5), 982–1013.

Brubacher, J. S. and W. Rudy (1958). *Higher Education in Transition.* New York: Harper & Row Publishers, Inc.

Burke, C. B. (1982). *American Collegiate Populations: A Test of the Traditional View.* New York and London: New York University Press.

Butts, R. F. (1939). *The College Charts its Course: Historical Conceptions and Current Proposals.* New York and London: McGraw-Hill Book Company, Inc.

Card, D. (1995). Using geographic variation in college proximity to estimate the return to schooling. In E. Christofides and R. Swidinsky (Eds.), *Aspects of labor market behaviour: Essays in Honour of John Vanderkamp.* Toronto: University of Toronto Press.

Carmichael, H. L. (1983). The agents-agents problem: Payment by relative output. *Journal of Labor Economics 1*, 50–65.

Carmichael, H. L. (1988). Incentives in academia: Why is there tenure? *Journal of Political Economy 96*(3), 453–472.

Clotfelter, C. T. (2010). Introduction. In C. T. Clotfelter (Ed.), *American Universities in a Global Market.* Chicago and London: The University of Chicago Press.

Cole, J. R. (2009). *The Great American University: Its Rise to Preeminence, its Indispensable National Role, Why it Must be Protected.* New York: Public Affairs.

Curti, M. and R. Nash (1965). *Philanthropy in the Shaping of American Higher Education.* New Brunswick: Rutgers University Press.

Eisenhardt, K. M. (1989). Agency theory: An assessment and review. *The Academy of Management Review 14*(1), 57–74.

Eliot, C. W. (1908). *University Administration.* Boston and New York: Houghton Mifflin Company.

Epple, D. and R. E. Romano (1998). Competition between private and public schools, vouchers, and peer-group effects. *American Economic Review 88*(1), 33–62.

Flexner, A. (1946). *Daniel Coit Gilman: Creator of the American Type of University.* New York: Harcourt, Brace and Company.

Geiger, R. L. (1986). *To Advance Knowledge: The Growth of American Research Universities, 1900-1940.* New Brunswick and London: Transaction Publishers.

Goldin, C. and L. F. Katz (1999). The shaping of higher education: The formative years in the United States, 1890-1940. *Journal of Economic Perspectives 13*(1), 37–62.

Goldin, C. and L. F. Katz (2008). *The Race Between Education and Technology.* Cambridge and London: The Belknap Press of Harvard University Press.

Graham, H. D. and N. Diamond (1997). *The Rise of American Research Universities: Elites and Challenges in the Postwar Era.* Baltimore and London: The Johns Hopkins University Press.

Gruber, J. and S. Johnson (2019). *Jump-Starting America: How Breakthrough Science Can Revive Economic Growth and the American Dream.* New York: PublicAffairs.

Hawkins, H. (1960). *Pioneer: A History of the Johns Hopkins University, 1874-1889.* Baltimore and London: The Johns Hopkins University Press, 2002.

Herbst, J. (1982). *From Crisis to Crisis: American College Government, 1636-1819.* Cambridge and London: Harvard University Press.

Hofstadter, R. (1955). *Academic Freedom in the Age of the College.* New Brunswick and London: Transaction Publishers, 1996.

Holmstrom, B. (1979). Moral hazard and observability. *Bell Journal of Economics 10*(1), 74–91.

Holmström, B. and P. Milgrom (1987). Aggregation and linearity in the provision of intertemporal incentives. *Econometrica 55*, 303–328.

Holmström, B. and P. Milgrom (1991). Multi-task principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, and Organization 7*, 24–52.

Hoxby, C. (2009). The changing selectivity of American colleges. *Journal of Economic Perspectives 23*(4), 95–118.

Hoxby, C. M. (2016). The productivity of U.S. postsecondary institutions. Technical report, Stanford University.

James, H. (1930). *Charles W. Eliot: President of Harvard University, 1869-1909, Volume I.* Boston and New York: Houghton Mifflin Company.

Karabel, J. (2006). *The Chosen: The Hidden History of Admission and Exclusion at Harvard, Yale, and Princeton.* Boston and New York: Mariner Books, Houghton Mifflin Company.

Kelley, B. M. (1974). *Yale: A History.* New Haven and London: Yale University Press.

Leslie, W. B. (1992). *Gentlemen and Scholars: College and Community in the "Age of the University," 1865-1917.* University Park: Pennsylvania State University Press.

MacLeod, W. B. (2003). Optimal contracting with subjective evaluation. *American Economic Review 92*(1), 216–240.

MacLeod, W. B. (2020). *An Introduction to Microeconomics for Contract and Organizational Economics.* M.I.T. Press, under contract.

MacLeod, W. B., E. Riehl, J. E. Saavedra, and M. Urquiola (2017). The big sort: College reputation and labor market outcomes. *American Economic Journal: Applied Economics 9*(3), 223–261.

MacLeod, W. B. and M. Urquiola (2015). Reputation and school competition. *American Economic Review 105*(11), 3471–3488.

Malcomson, J. M. (1984). Work incentives, hierarchy, and internal labor markets. *Journal of Political Economy 92*(3), 486–507.

Neilson, C. (2017). Targeted vouchers, competition among schools, and the academic achievement of poor students. Unpublished manuscript, Princeton University.

Paulsen, F. (1906). *The German Universities and University Study.* New York: Charles Scribner's Sons.

Prendergast, C. (1999). The provision of incentives in firms. *Journal of Economic Literature 37*(1), 7–63.

Rudolph, F. (1962). *The American College and University: A History.* Athens and London: The University of Georgia Press, 1990.

Rudolph, F. (1977). *Curriculum: A History of the American Undegraduate Course of Study Since 1636.* San Francisco: Josey-Bass Publishers.

Ruggles, S. B. (1854). *The Duty of Columbia College to the Community, and its Right to Exclude Unitarians from its Professorships of Physical Science.* New York: John F. Trow, Printer.

Storr, R. J. (1953). *The Beginnings of Graduate Education in America.* Chicago: The University of Chicago Press.

Synnott, M. G. (1979). *The Half Opened Door: Discrimination and Admissions at Harvard, Yale, and Princeton, 1900-1970.* New Brunswick and London: Transaction Publishers, 2010.

Tewksbury, D. G. (1932). *The Founding of American Colleges and Universities Before the Civil War: With Particular Reference to the Religious Influences Bearing Upon the College Movement.* Mansfield Centre: Martino Publishing, 2011.

Urquiola, M. (2020). *Markets, Minds, and Money: Why America Leads the World in University Research.* Cambridge, MA: Harvard University Press.

Veysey, L. R. (1965). *The Emergence of the American University.* Chicago and London: The University of Chicago Press.

8. Mathematical Appendix

This appendix introduces a multi-tasking model that illustrates the points made in the text. We assume that faculty have some intrinsic preference for research, given by $\bar{r} > 0$. The amount of research carried out is given by effort r_i. In addition to their research, faculty engage in outside activities, denoted by o_i. They derive no intrinsic pleasure from these, but they can increase compensation. These could include consulting, outreach activities to make their research better known, or explicit gaming of the reward system to divert resources towards themselves. Here we are not concerned with the normative question of the appropriate level of such activities, but rather have variation in reward leads to variation in the level of such activities. In short, total effort by a professor is given by:

$$\vec{e}_i = [r_i, o_i]^T.$$

We suppose the university cares only about research output (and hence the competence of individuals to be knowledgeable in the field they are instructing). Hence, the university's payoff is:

$$W(r_A, r_B, I_A, I_B) = V^r \times (r_A + r_B) - I_A - I_B,$$

where V^r is the value of research effort, and $I_i, i \in \{A, B\}$ is compensation. We derive the optimal compensation system for Professor A. Given the symmetry of our setup, the same system applies to Professor B.

The university has two performance signals, absolute performance and relative performance:

(8.1)
$$s_A = r_A + \alpha_i + \gamma_A + \gamma,$$
$$s_{AB} = (r_A - r_B) + \alpha_A - \alpha_B + \gamma_A - \gamma_B,$$

where the random noise is i.i.d. and satisfies $\gamma_i \sim N(0, \bar{\sigma}^2)$ and $\gamma \sim N(0, \sigma^2)$. Hence, $var(s_A) = \bar{\sigma}^2 + \sigma^2$, while $var(s_{AB}) = 2\bar{\sigma}^2$. The signals for Professor B are similarly defined with A replaced by B. The variances play a role in explaining both effort and overall system performance. The variance of s_A (or s_B) is the overall uncertainty associated with measured output given effort for the full population of professors. We assume this is exogenous and let $\rho_0 = \frac{1}{\bar{\sigma}^2 + \sigma^2}$ be the corresponding precision (that will play a role in the analysis below). Specialization is measured by $\rho_s = \frac{1}{2\bar{\sigma}^2}$, with higher values of ρ_s corresponding to more specialization. When professors A and B are in the same specialty, then one can more precisely measure their effort by bench marking one against the other, which corresponds to a larger ρ_s.

Consider first the case in which both faculty have the same ability and thus $\alpha_A = \alpha_B = 0$. We follow Holmstrom and Milgrom (1987, 1991) and use the standard multi-tasking framework that supposes compensation is an affine function of the performance measures:

$$I_A = \bar{w} + b_{s_A} \times s_A + b_{s_{AB}} \times s_{AB},$$

where b_{s_A} is the reward based upon individual performance, and $b_{s_{AB}}$ is the reward based upon relative performance to other Professor in the same field.

We suppose professors have constant absolute risk aversion κ and hence the utility of professor A is:

$$U_A = E\{I_A\} + V^o \times o_A - C_s(\vec{e}_A, \bar{r}) - \frac{\kappa}{2} var(I_A).$$

The term effort terms o_A corresponds to outside activities with a return $V^o < V^R$. The cost of effort satisfies:

$$C_s(\vec{e}_A, \bar{r}_A) = \left[\vec{e}_A - \begin{bmatrix} \bar{r}_A \\ 0 \end{bmatrix}\right]^T \begin{bmatrix} 1/2 & d/2 \\ d/2 & 1/2 \end{bmatrix} \left[\vec{e}_A - \begin{bmatrix} \bar{r}_A \\ 0 \end{bmatrix}\right],$$

where $1 > d > 0$ is a parameter that measure the degree of substitution between research and outside effort. Notice that the matrix $A = \begin{bmatrix} 1/2 & d/2 \\ d/2 & 1/2 \end{bmatrix}$ is invertible and positive definite, and hence $C_s(\vec{e}_A, \bar{r}_A) \geq 0$ for any effort and is zero if and only if:

$$\vec{e}_A = \begin{bmatrix} \bar{r}_A \\ 0 \end{bmatrix}.$$

In the absence of any financial reward, the professor will produce research output $\bar{r}_A > 0$. The off diagonal terms, d, imply that efforts are substitutes—an increase in rent seeking activity, o_A, increases the cost of supply research effort and vice-verso.

The problem facing the University is to choose a pay package, $\vec{p} = \{\bar{w}, b_{s_A}, b_{s_{AB}}\}$, to maximize its payoff given that the professor must earn at least U^0 and she chooses effort in response to incentives. The professor's payoff given compensation is:

$$U_A(\vec{e}_A, \vec{e}_B, \vec{p}) = \bar{w} + b_{s_A} \times (r_A - \bar{r}_A) + b_{s_{AB}} \times (r_A - r_B) + V^o \times o_A$$
$$- \left((r_A - \bar{r}_A)^2 + o_A^2\right)/2 - d \times (r_A - \bar{r}_A) \times o_A$$
$$- \frac{\kappa}{2}\left(b_{s_A}^2\left(\bar{\sigma}^2 + \sigma^2\right) + b_{s_{AB}}^2 \times 2\bar{\sigma}^2\right).$$

Notice that the return on effort is $(b_{s_A} + b_{s_{AB}}) \times r_A$, and hence $B = b_{s_A} + b_{s_{AB}}$ is the total reward for effort that we will derive below. We suppose that parameters are set so we have an interior solution. Then we can set:

$$\vec{e}_A(\vec{p}) = argmax_{\vec{e}_A} U_A(\vec{e}_A, \vec{e}_B, \vec{p}).$$

25

This, combined with the condition that $\vec{e}_A = \vec{e}_B$ and setting $B = b_{s_A} + b_{s_{AB}}$ we get:

$$(8.2) \qquad r_A(\vec{p}) = \bar{r}_A + \frac{b_{s_A} + b_{s_{AB}} - d \times V^o}{1 - d^2} = \bar{r}_A + \frac{B - d \times V^o}{1 - d^2},$$

$$(8.3) \qquad o_A(\vec{p}) = \frac{V^o - d \times B}{1 - d^2}.$$

Observe that due to the complementarities, if the university does not reward research, then effort is less than the professor's desired research level due to the existence of the outside rent seeking activities. The more remunerative these activities, the less is the level of research. If there are no rewards to research, and d is sufficiently large, then no research will be done.

Given the contract, the cost to the university of Professor A is given by:

$$I_i(\vec{p}) = \bar{w} + b_{s_i} r_i(\vec{p}), i \in \{A, B\}.$$

Notice that since the expected effort of each professor is the same, then relative performance term does not have a direct cost, though the outside option ensures that total compensation must provide the professor with payoff U^0 and hence, there is an implicit reward in the fixed wage to compensate for higher effort. The University's optimal contract is the solution to:

$$(8.4) \qquad max_{\vec{p}} v(r_A(\vec{p}) + r_B(\vec{p})) - \sum_{i \in \{A, B\}} I_i(\vec{p})$$

subject to:

$$(8.5) \qquad U_i(\vec{e}_A(\vec{p}), \vec{e}_B(\vec{p}), \vec{p}) \geq U^0, i \in \{A, B\}.$$

The solution to this problem is a straightforward computation (see MacLeod (2020)). It follows from the fact that the participation constrain is binding and hence the solution to:

$$max_{\vec{p}} \sum_{i \in \{A, B\}} V^r \times r_i(\vec{p}) - \left(U^0 - V^o \times o_i(\vec{p}) + C_s(\vec{e}_i(\vec{p}), \bar{r}_i) + \frac{\kappa}{2}\left(b_{s_i}^2(\bar{\sigma}^2 + \sigma^2) + b_{s_{ij}}^2 \times 2\bar{\sigma}^2\right)\right).$$

using (8.2-8.3) and the fact that $\partial C_s(\vec{e}_A(\vec{p}), \bar{r}_A)/\partial e_A = B^*$ implies that the optimal performance pay satisfies:

$$b_{s_A}^* = \rho_0\left(\frac{V^r - B^*}{\kappa(1 - d^2)}\right),$$

$$b_{s_{AB}}^* = \rho_s\left(\frac{V^r - B^*}{\kappa(1 - d^2)}\right).$$

This in turn implies:

$$(8.6) \qquad B^* = \frac{(\rho_0 + \rho_s) V^r}{\kappa(1 - d^2) + \rho_0 + \rho_s},$$

while:

$$b^*_{s_A} = \frac{\rho_0}{\rho_0 + \rho_s} B^* = \frac{\rho_0 V^r}{\kappa(1 - d^2) + \rho_0 + \rho_s},$$

$$b^*_{s_{AB}} = \frac{\rho_s}{\rho_0 + \rho_s} B^* = \frac{\rho_s V^r}{\kappa(1 - d^2) + \rho_0 + \rho_s}.$$

These expressions assume that $\bar{r}_A = \bar{r}_B$ and $o_A = o_B > 0$. If the return to research is sufficiently high, then $o_A = o_B = 0$ and one simply sets the "d" terms equal to zero to get the optimal level of effort and bonus pay. However, when there is some rent seeking activity notice that larger d increases B^* because rent seeking does not benefit the University, and hence increasing the reward to research reduces the level of outside activities via the substitution effect.

Notice that increasing specialization, ρ_s, also increases both B^* and effort, $r_i, i \in \{A, B\}$. In the limit as $\rho_s \to \infty$ we get $B^* \to V^r$, the marginal return of research effort to the university. Also, as specialization increases the weight on personal effort, b^*, goes to zero, while $b^*_{s_{AB}} \to B^*$, in other words pay for performance depends only upon relative and not absolute performance. Moreover, since $V^r > V^o > d \times V^o$, from (8.3) we get that it must be the case that $o_A = o_B = 0$, in other words specialization leads to a complete shutdown of rent seeking activities.

Finally, the fixed wage does not affect the performance pay slope, it is set to make the Professor indifferent between research and leaving the university. Hence, we have:

$$\bar{w}^* = U^0 - (b^* \times e^*_A + V^o \times o^*_A) + C_s(\bar{e}^*_A, \bar{r}_A) + \frac{\kappa}{2}\left(\frac{V^r}{\kappa(1 - d^2) + \rho_0 + \rho_s}\right)^2 (\rho_0 + \rho_s).$$

This expression simply points out that the fixed wage rises with effort and risk. Notice that as specialization increases, $\rho_s \to \infty$, we have $b^* \to 0$, $o^*_A \to 0$ and hence:

$$(8.7) \qquad\qquad \bar{w}^* \to U^0 + a \times \frac{V^{r^2}}{2}.$$

while University payoff satisfies:

$$(8.8) \qquad\qquad W^* \to (1 - a) V^{r^2} + 2V^r e_0 - 2U^0.$$

8.1. **Tournaments.** This section derives the relationship between dispersion in faculty ability and the incentive effect of prizes upon effort and now allows ability to vary, and let $\alpha_A \geq \alpha_B$, so that it is always Professor A with the higher skill level. Let P denote a prize or reward that is given to the faculty member for whom $s_i > s_j$, for $i, j \in \{A, B\}$. The issue of outside effort and risk aversion is put aside, and we let $a = 1$, this it is assumed payoffs

are given by:

$$U_A = \bar{w} + P \times Pr\left[s_A > s_B | e_A, e_B\right] - r_A^2/2,$$
$$U_B = \bar{w} + P \times Pr\left[s_B > s_A | e_A, e_B\right] - r_B^2/2,$$

where s_A is defined by (8.1), and s_B is defined similarly. Given that the common component is factored out, the probability of winning is given by:

$$Pr\left[s_i > s_j\right] = Pr\left(r_i + \alpha_i + \gamma_i + \gamma_s - (r_j + \alpha_j + \gamma_j + \gamma_s) > 0\right),$$
$$= F\left(\sqrt{\rho_s}\left(r_i + \alpha_i - r_j - \alpha_j\right)\right),$$

where $F\left(\right)$ is the cumulative Normal distribution and ρ_s is the degree of specialization defined above. Given that the Normal density, $f\left(\right)$, is symmetric around zero we have that the first order condition for effort is given by:

$$r_A^* = r_B^* = \sqrt{\rho_s}Pf\left(\sqrt{\rho_s}\left(r_i + \alpha_i - r_j - \alpha_j\right)\right) > 0,$$

and hence:

(8.9)
$$r_i^* = \sqrt{\rho_s}Pf\left(\sqrt{\rho_s}\Delta\right),$$

where $\Delta = \alpha_A - \alpha_B \geq 0$. This implies that increasing the prize increases effort $\left(\frac{\partial r_i^*}{\partial P} > 0,\right)$. This also implies:

$$\frac{\partial r_i^*}{\partial \Delta} = \rho_s Pf'\left(\sqrt{\rho_s}\Delta\right) < 0,$$

hence increasing the variance in ability in a department reduces effort. If everybody has the same skill $\left(\Delta = 0\right)$, then

$$r_i^* = \sqrt{\rho_s}Pf\left(0\right) = \sqrt{\rho_s/2\pi}P.$$

This implies that holding the prize fixed, increasing specialization $\left(\rho_s \to \infty\right)$ results in $\left(r_i^* \to \infty\right)$ without increasing the size of the prize. However, there is a limit. The expected payoff with equal skill is each wins the prize P with equal probability and thus:

(8.10)
$$U_i^* = \bar{w} + P/2 - \rho_s P^2/4\pi.$$

Eventually for large ρ_s we have $U_i^* < \bar{w}$, in which case Professor i gains by putting in no effort. When this happens, the only possible equilibrium is a mixed strategy. When there is asymmetric ability we have:

$$U_A^* = \bar{w} + P \times F\left(\sqrt{\rho_s}\Delta\right) - \rho_s\left(P \times f\left(\sqrt{\rho_s}\Delta\right)\right)^2,$$
$$U_B^* = \bar{w} + P \times F\left(-\sqrt{\rho_s}\Delta\right) - \rho_s\left(P \times f\left(\sqrt{\rho_s}\Delta\right)\right)^2.$$

In this formulation, the marginal benefit of effort by the Professor is not affected by ability, and hence the prize system makes both work hard at the same level. The interesting observation is that if the skill difference is too large, then the low probability of winning can cause Professor B to drop out, and we get a discontinuous drop in performance when either the prize is too large, or the area is very specialized.