

NBER WORKING PAPER SERIES

WHY DOES THE U.S. HAVE THE BEST RESEARCH UNIVERSITIES?
INCENTIVES, RESOURCES, AND VIRTUOUS CIRCLES

W. Bentley MacLeod
Miguel Urquiola

Working Paper 28279
<http://www.nber.org/papers/w28279>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2020

This paper was prepared for the Journal of Economic Perspectives. For useful comments and input we thank Elías Cueto, David Figlio, Patricia Graham, Gordon Hanson, Enrico Moretti, and Timothy Taylor. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2020 by W. Bentley MacLeod and Miguel Urquiola. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

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

W. Bentley MacLeod and Miguel Urquiola

NBER Working Paper No. 28279

December 2020

JEL No. I2,I23,N0

ABSTRACT

Around 1875 the U.S. had none of the world's leading research universities; today, it accounts for the majority of the top-ranked. Many observers cite events surrounding World War II as the source of this reversal. We present evidence that U.S. research universities had surpassed most countries' decades before WWII. An explanation of their dominance must therefore begin earlier. The one we offer highlights reforms that began after the Civil War and enhanced the incentives and resources the system directs at research. Our story is not one of success by design, but rather of competition leading American colleges to begin to care about research. We draw on agency theory to argue that this led to increasing academic specialization, and in turn, to more precise measures of professors' research output. Combined with sorting dynamics that concentrated talent and resources at some schools—and the emergence of tenure—this enhanced research performance.

W. Bentley MacLeod
Department of Economics
Columbia University
420 West 118th Street, MC 3308
New York, NY 10027
and NBER
wbmacleod@wbmacleod.net

Miguel Urquiola
Economics Department
Columbia University
420 W 118TH ST #1022
New York, NY 10027
and NBER
msu2101@columbia.edu

The educational performance of nations is not uniform across different dimensions, nor is it static. Around 1875, the United States had substantially higher levels of schooling than any European country, but that gap has largely disappeared (Goldin and Katz 2008). By contrast, circa 1875 the United States had none of the world’s leading research universities; today, it accounts for the majority of the top-ranked (Cole 2009; Clotfelter 2010).

One way to illustrate the rise of US research universities is by counting the times universities are mentioned in the biographies of Nobel prize winners—the idea being that universities that train or host Nobel winners are likely places active in research. These mentions are associated with years, for example, the year a winner graduated from one institution or became a professor at another. One can use this to create a time series reaching back into the 1800s. Figure 1 shows that by this measure, around 1875 American universities badly lagged their European counterparts; today they are distinctly ahead.

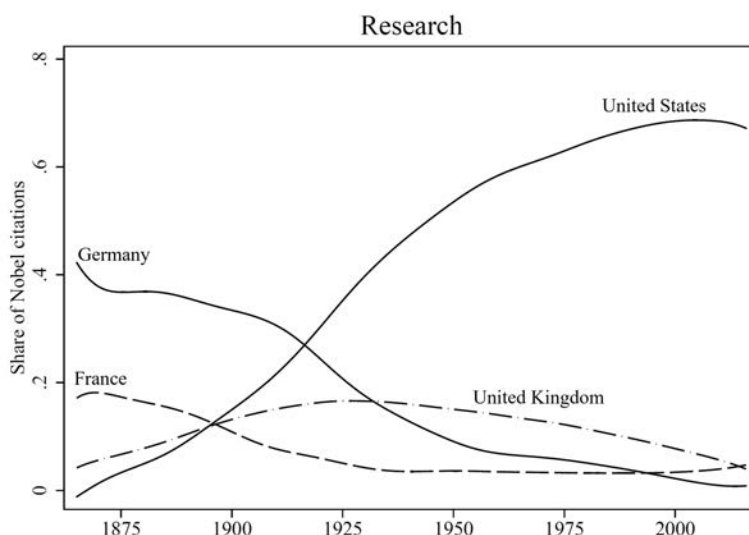


Figure 1: **University Nobel Prize Mentions, By Country**

This figure describes the frequency with which Nobel winners’ biographies mention universities in different countries. Specifically, the dates in the figure refer to the years in which laureates graduated from a given school if they were students there (this accounts for why there are dates for the late 1800s, before the prize was granted). If they were professors, the dates refer to the year of award for the university they were affiliated with at that point, and the year of joining for other schools. The figure 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. They do not adjust for income or population because that does not change the two central facts the paper addresses: American universities’ initial weakness and later leadership. For such adjustments, see Urquiola (2020).

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 United States, and the rise of federal research funding (Graham and Diamond 1997; Cole 2009; Gruber and Johnson 2019). The emphasis on World War II is analytically attractive: it highlights

factors that surely strengthened American universities, and it addresses the relative German decline apparent in Figure 1.¹

However, the raw data behind the figure indicate that US universities had matched or surpassed most countries' well before World War II. By the measure used they were ahead of all but Germany's by 1910, and of all countries' by 1920. To cite one anecdote illustrating this timing, 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, a successful explanation of American universities' research dominance must begin in the late 1800s, and must involve factors other than mid-20th century events in Europe.

The explanation we offer highlights reforms that began after the Civil War and enhanced the incentives and resources the American university system directs at research.² Understanding these reforms benefits from two observations. First, the U.S. takes a relatively free market approach to higher education; for instance, it allows easy entry by schools seeking to satisfy customer demand. Second, universities do not supply their customers with a single “product”; 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, some schools cater to students who are smart, or wealthy, or artistic.

We begin to review the history of US university research by noting that since colonial times and into the 1800s, American households demanded denominational sorting. For example, Presbyterians wished to attend college with Presbyterians, Episcopalians with Episcopalians, etc. Further, most preferred colleges close to home. Together, these factors produced massive entry: while in 1776 the United States had 9 colleges, about 900 more opened before the Civil War.

Mass entry meant that early U.S. colleges tended to be small and underfunded. They offered a basic, narrow, and rigid curriculum delivered by unspecialized, often poorly paid professors. Colleges routinely hired faculty using criteria other than expertise, and provided them with few incentives or resources to do research. It is not surprising that, on average, the colleges did not excel in this dimension.

We then discuss how reforms allowed the system to begin providing incentives and resources. First, we sketch an agency model that asks how universities can incentivize professors to do research. It highlights that even if an institution is committed to research, producing it benefits from reasonably precise research performance measures. Precision is enhanced by academic specialization—in this case, by organizing professors into disciplines.

We then describe how American schools gained an interest in research and increased specialization. To preview, during the last few decades of the 1800s, U.S. colleges let a gap grow between the

¹ Analyses of recent university performance often rely on bibliographic databases. Such data have advantages, but are of limited use here because databases like the Web of Science only have author affiliation information going a few decades back; see Urquiola (2020).

² In emphasizing earlier events, our approach is closer to Veysey (1965) and Goldin and Katz (1999). We draw on Urquiola (2020).

skills they taught and those their customers demanded. With industrialization, interest grew around areas that the college curricula essentially ignored, like engineering and business. In the 1870s and 80s, new universities like Cornell and Johns Hopkins showed that one could attract students by offering specialized and advanced instruction in a range of areas, rather than by supplying denominational sorting. Incumbents like Harvard and Columbia responded forcefully, forming specialized arts and sciences departments and creating professional schools.

Free entry allowed other schools to join the fray. This was helped by private donations, which propelled universities like Chicago and Stanford. It was also aided by public resources, which were crucial for entrants like Berkeley and MIT.

All these schools began to seek professors who were specialists and, therefore, researchers at the frontier of a field. The academic system responded by providing tools to measure research output, like specialized journals. Schools thus gained the ability to identify, bid for, and begin to recruit successful researchers.

These years also saw some American universities acquire tremendous amounts of resources. In the early 1900s, the schools implementing reforms sought and achieved rapid enrollment growth. However, they found that this threatened the match between the sorting they provided and that which key students demanded. In particular, enrollment growth brought in lower-income students and Jewish students and—in a period of rising antisemitism—alienated the Protestant elites the colleges had traditionally served. By the early 1920s Columbia implemented selective admissions, with several schools following soon after. Selectivity set in motion another sorting process, concentrating high ability/income students—and eventually high tuitions and donations—at such schools. This bolstered their ability to recruit desirable professors.

In addition, the early decades of the 1900s saw the emergence of “lumpy” rewards for research, the most salient being tenure. Agency theory implies that an up-or-out incentive system can promote performance, particularly if professors compete against individuals of similar ability. In concert with increasingly precise research performance measures, tenure furthered a sorting process that led to professors clustering in departments with colleagues of similar ability.

In short, top American schools came to enjoy a virtuous circle that gave them resources to invest in research, which they could now effectively incentivize; this helped attract strong students and funding, which could go into further reforms and enhancements. At the same time, other schools struggled and lost ground, creating the between-school inequality evident in the United States.

Finally, throughout this period, such dynamics operated much more weakly in Europe, particularly on the continent. European states controlled their higher education systems and did not allow easy entry. This prevented the emergence of hundreds of struggling schools but also slowed the rise of extremely well-heeled magnets for global talent, like MIT or Stanford. Overall, sorting of all types was not as extensive in Europe. Concerning incentives, European professorial arrangements—whose origins lie in the high middle ages—use up-or-out schemes like tenure to a much lesser extent.

1 The Initial American Weakness in Higher Education

A demand for denominational sorting drove the creation of the nine American colonial colleges. For example, Massachusetts Puritans created Harvard to produce what they saw as a theologically sound education. They took this task seriously; for instance, the colony forced out Harvard's first President, Henry Dunster—even as it 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, and so on (Herbst 1982).

In addition, like school-users to this day, the early founders desired proximity—to have schools close to home (for modern evidence see Card 1995 and Neilson 2017). For example, Connecticut-based Puritans created Yale because they perceived that the Harvard of Massachusetts-based Puritans was too physically distant (in addition to too religiously liberal). After independence, the desire for sorting and proximity led to massive entry: historians estimate 900 colleges opened before the Civil War (Tewksbury 1932; Hofstadter 1955; Burke 1982). Religious denominations competed to open colleges, including on the changing western frontier (Xiong and Zhao 2020).

Well into the 1800s, therefore, American colleges mainly catered to small local markets. It was common for a college to have fewer than 100 students and, say, 5 faculty members (Rudolph 1962; Geiger 1986). 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 avoid sending them away for training. In 1879, Princeton President James McCosh wrote to an alumnus studying in Europe: “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” (Leslie 1992).

Wealth-related criteria mattered too. Rudolph (1962) writes: “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” (Flexner 1946).

While religion clearly influenced hiring, expertise could matter less. For example, in 1853 James Renwick, Columbia's sole professor of “natural philosophy,” announced his retirement. Among the mainly Anglican trustees who searched for his replacement was Samuel Ruggles, a real estate investor. Ruggles favored hiring 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:

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 rejected (Thompson 1946).

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: "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 early U.S. colleges did not prioritize specialized or even trained personnel. It is not surprising that they made an associated choice: they offered a basic curriculum featuring few, if any, electives. The course of study 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 the curriculum involved science, the emphasis was deductive rather than experimental/inductive (Butts 1939; Storr 1953). The teaching methods involved rote learning in the form of "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 hired to teach "English, Latin, Greek, Chaldee and such other languages as he shall have time for" (Rudolph 1962). Even more extreme, in the 1700s, individual 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 warm to rote learning!

These practices kept costs down, which was necessary due to limited enrollments and revenues. Serious budgetary pressures were common; Boyer (2015) states that most colleges hovered between "genial penury and unmitigated fiscal disaster."

The whole setup also reflected the inability and/or unwillingness of schools to branch out. For example, to the extent that instruction in applied science developed, it was mainly outside colleges. A notable player was the Rensselaer Polytechnic Institute (RPI), which offered engineering, surveying, applied science, and some of the first laboratories in the country. However, schools like RPI also offered narrow curricula and little advanced instruction. Rudolph (1977) states that around 1830, a motivated college graduate could earn an RPI degree in 24 weeks. Further, the existence of such "institutes" facilitated the claim by colleges that science belonged elsewhere. When Williams hired Ira Remsen to teach chemistry, he asked for a small room to set up a self-funded laboratory. The response: "You will please keep in mind that this is a college and not a technical school" (Rudolph 1962).

Table 1: University-Level Mentions in Nobel Biographies During Given Time Periods

Total number of mentions for:							
1855-1900		1901-1940		1941-1980		1981-2016	
Humboldt Univ. of Berlin	19	University of Cambridge	63	Harvard University	101	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 Supérieure	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 Barbara	7
University of Stockholm	3	Princeton University	11	Technical Univ. of Munich	15	University of Oxford	7

The table lists the number of times Nobel winners' biographies mention specific universities in given time periods. Source: Urquiola (2020).

All of the above is not to say that no research went on in the early American colleges. In the 1830s, Amherst's Edward Hitchcock completed the first geological survey of Massachusetts; 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 were in a different situation. Since the Protestant Reformation, European states—particularly on the continent—had tended to control and eventually relatively generously fund their universities. Faculty were often well-paid civil servants, assigned by ministries of education to specialized “chaired” professorships (Paulsen 1906). They gradually began to absorb the research activity previously located in academies founded mainly in the 1600s (for example, the Académie Royale des Sciences, the Royal Society of London, and the Berlin Academy of Sciences).

This created a contrast 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). By the 1850s, the American school, Columbia, had 6 professors and 150 students; the German, Göttingen, about 90 professors and 1,600 students (Ruggles 1854). Table 1 returns to the data to list the 20 universities most mentioned in Nobel winners' biographies in roughly 40-year periods. For 1855-1900, Göttingen was among the top 10; Columbia—like all other American schools—did not make the top 20.

What happened since the 1850s? To start, many more schools entered the market. Today, the

U.S. has about 4,700 post-secondary institutions. By the logic above, its universities might be even weaker. And yet, the country has about 100 “Research I” universities.³ Thanks to these schools, the United States accounts for 17 of the top 20 universities in the last column of Table 1 (for 1981-2016). The next sections review how this transformation happened.

2 Agency Theory and Research Performance

We begin by sketching an agency theory framework that helps make sense of what occurred (a mathematical version of this section is in the appendix). Our starting point is the fundamental idea that individuals respond to rewards, and that performance measures shape those rewards. As Merton (1968) argued, this applies to academic activities.

Suppose professors work at a university, and exert effort on research and “outside activities.” The outside activities may include consulting work and staying active in alumni networks or, in the 1800s, religious groups. We suppose the professors are intrinsically motivated to do research, and so they will engage in this activity to some extent even if they are not compensated for it. We also assume that all else equal, individuals and universities prefer a stable rather than a risky income stream.⁴

A key question is whether compensation can be used to reward research performance, which depends on whether the university observes meaningful measures of research output. Suppose first that the university does not, and so faculty salary cannot vary with research effort.

This basic setup captures elements of the American college system well into the 1800s. There were few rewards for research, and yet some professors engaged in it. This was likely easier for wealthier individuals with less need for outside activity. In addition, in the 1800s the return to outside effort could have been substantial even if one were focused on an academic career. The cases discussed above illustrate 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, and that the university places value on producing research. Principal-agent theory (as in Holmstrom and Milgrom 1991) implies that the university should make faculty compensation depend on this measure. Indeed, at modern universities, rewards like salaries and raises vary with research performance. Such compensation provides incentives, but it entails a cost because it exposes professors to risk. Thus, a more precise research 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 such a measure.

This question matters because it is usually difficult for non-experts to assess professors’ research

³ “Research University-I” was a category in the taxonomy of higher education institutions developed by the Carnegie Commission on Higher Education

⁴ In the case of individuals, this follows from risk aversion; in the case of universities, stable and egalitarian pay reduces organizational conflict (Lazear 1989; MacLeod 2003).

performance. This is particularly true in making comparisons across fields. For example, in the humanities faculty tend to write books, while articles are the norm in the sciences. Similarly, a chemistry professor may be listed as a co-author on all papers produced by her laboratory while an economic theorist writes alone—it is hard to tell who is being more productive.

To obtain more precise measures one useful step is to organize professors into fields, and to compare their performance *within fields* (Lazear and Rosen 1981). With access to a relative performance measure, it makes sense for the university to increase the rewards for research.

In short, our framework highlights factors that improve research performance: i) increasing specialization that organizes faculty into fields; ii) reducing the return to outside activities that compete with research for faculty effort; and iii) increasing the extent to which universities value research. We now discuss how reforms beginning in the 1800s promoted these changes.

3 U.S. Reforms of Higher Education

As industrialization advanced in the second half of the 19th century, American colleges increasingly failed to satisfy some areas of educational demand. Some students were interested in subjects the colleges did not cover, such as business, engineering, law, and mining. Others wanted specialized, modern training in areas like chemistry and economics. Over the first part of the 1800s, reformers at several schools (like Brown, Harvard, Union) attempted to address these shortcomings but failed to sustain reform (Butts 1939; Rudolph 1962).

Lasting change arrived with two entrants: Cornell (1865) and Johns Hopkins (1876). Each received substantial private support, the salient donations being those of Ezra Cornell and Johns Hopkins. But each also received major public support: Cornell got federal funds via the Morrill Act, and Hopkins eventually received support from Maryland (Curti and Nash 1965).⁵

These schools expanded the curriculum and began to offer specialized instruction. Cornell saw rapid enrollment growth, showing that a school could thrive by responding to curricular demand rather than by providing denominational sorting. For its part, Hopkins emulated the focus of German universities on advanced/graduate instruction. Its first president, Andrew Gilman, promoted hiring based on expertise and research ability, rather than aspects like denominational affiliation. He stated: “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” (Hawkins 1960). In our

⁵ The Morrill Act (1862) transferred shares of mostly Western federal land to states. The proceeds of their sale were to benefit higher educational institutions whose objective was “without excluding other scientific or classical studies, to teach such branches of learning as are related to agriculture and the mechanic arts.” States retained wide latitude in the use of funds; for example, they could be awarded to private universities.

framework, Gilman was increasing the weight on research and decreasing that on outside activities.

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

Even these last schools began to offer majors and electives and to split academic areas into specialized fields. For example, natural philosophy into astronomy, biology, chemistry, physics; political economy into economics, political science, sociology, etc. Academic specialization was on the rise.

Reformers soon faced additional competition from other entrants aiming to be research universities. These included Stanford (1885), and the University of Chicago (1890), which relied on private donations, and schools that benefited from Morrill Act funding and/or state support—such as the universities of California, Illinois, Michigan, Minnesota, and Wisconsin, along with Michigan State, Penn State, and Purdue. Yet others, like MIT, enjoyed both private and public funding.

Many of these schools innovated in ways of their own. For example, Chicago was unusually aggressive in raiding other schools' faculty; Berkeley emphasized hiring junior professors and avoided “inbreeding” (another way of reducing the return to outside activities); MIT and Stanford moved to collaborate with the private sector in ways that generated funding, setting a pattern for years to come.

It is worth noting that the scale of entry stood in stark contrast to the situation in Europe. There, states controlled and mostly restricted the creation of universities. For example, around the time of the American Civil War, the U.S. had hundreds of separate colleges, while England still had four universities. Further, many continental universities required state approval to venture into new fields, particularly where states funded faculty salaries.

Competing to deliver advanced and specialized instruction, American universities began to seek 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 . . . but if you are in search of teachers to teach botany, chemistry, physics and so on, you cannot find them. They do not exist” (Flexner 1946).

As the competition for research talent heightened, the supply of Ph.D.-trained specialists grew. The academic system also began to emphasize specialization in ways that provided measures of professors' research performance.

Faculty members founded professional associations, like the American Chemical Society (1877) and the American Historical Association (1884). These began to publish journals: for example,

the American Economic Association (1885) introduced the *Publications of the American Economic Association*, which evolved into the *American Economic Review*. University department journals supplemented such periodicals; for example, Harvard's *Quarterly Journal of Economics* (1886) and Chicago's *American Journal of Sociology* (1895).

Here the United States was catching up to Europe, where academic societies and specialized journals had existed much before. For example, consider mathematics, where European periodicals like the *Journal de l'Ecole Polytechnique* appeared in the 1700s (Bartle 1995). At Johns Hopkins, President Daniel Gilman hired an English Mathematician, James Sylvester, who led the founding of the *American Journal of Mathematics* in 1878. Years later, in his retirement address, Sylvester disputed this role:

You have spoken of our Mathematical Journal. . . . Mr. Gilman is continually telling people that I founded it . . . I assert that he is the founder. Almost the first day I landed in Baltimore, . . . he began to plague me to found a Mathematical Journal on this side of the water . . . I said it was useless, there were no materials for it. Again and again he returned to the charge and again and again I threw the cold water I could on the scheme; nothing but obstinate persistence . . . brought his views to prevail (Flexner 1946).

As American periodicals formalized their procedures and made use of peer review, they differentiated in quality. Those that became better-regarded received more submissions, and their editors made them more selective, thus attracting more papers—a hierarchy evolved in each field.⁶

Charles Eliot reflected on the impact this had over 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” (James 1930).

The demand for research talent improved the negotiating position for professors. To illustrate, Chicago's first president, William Harper, raided campuses for faculty. One of his targets was Clark University, which had made an underfunded attempt to compete with Johns Hopkins at graduate teaching; Harper hired away 15 of its professors. To compete, 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 satisfy his demand for a 24-inch telescope. Still, he added a caveat: “remember, this is a Friend's College and thee should give up thy smoking” (Leslie, 1992). In the event, Swain had to accept the professor's tobacco habit too.

⁶ It is possible that a more thorough hierarchy appeared in the United States than in Europe. Geiger (1986) suggests that Europe displays a greater tendency for academics to submit work to periodicals in their own institutes, universities, or countries; language differences may factor into the latter too.

4 Resources and Admissions-Related Reform

How did universities manage to finance these changes? How were top departments able to pay talented researchers well, offer them sabbaticals and Ph.D. students? How did the American system transition from one in which essentially all colleges experienced “genial penury” to one in which a (small) minority of universities spend lavishly? To elaborate, at present U.S. university spending is high—the country is an outlier in regressions of per-student expenditure on GDP per capita.⁷ Beyond this, it displays distinct inequality. Hoxby (2016) estimates that the most selective universities spend about \$150,000 per student—about six times the U.S. average and about fifteen times as much as their less well-heeled counterparts. How did this situation develop?

A key part of the story involves selective admissions. Before the 1920s, essentially all colleges had “open enrollment” policies, reflecting their desire to grow. Harvard’s Charles Eliot (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.”⁸ Such policies, combined with the reforms that rendered schools attractive, produced higher enrollments. 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 (Rudolph 1962). At this time, size was seen as a signal of quality because it provided resources to pursue reforms.

However, this growth began to generate its own challenges. Because universities continued to serve relatively local markets, growth entailed the arrival of less prepared and/or lower-income students. The resulting heterogeneity threatened universities’ ability to deliver the peer groups their traditional “elite” customers (white and usually Protestant men) expected. In other words, at some schools, traditional customers might have felt that their educational spaces were being invaded, or at least losing some of their signaling value.

Economic theory suggests that if universities had low fixed costs, such concerns would lead to the existence of a large number of schools, each catering to individuals of specific types (Epple and Romano 1998; MacLeod and Urquiola 2015). But in reality, fixed costs and coordination challenges prevented elites from easily seceding from the schools they 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 select fraternities and sororities (Synnott 1979; Karabel 2006; Urquiola 2020). Beyond providing exclusivity, these clubs performed tasks that growth had led schools to neglect—like supplying living spaces and dining rooms.

Many students coveted membership in these clubs. Franklin Roosevelt’s experience is illustrative.

⁷ For cross country data see Organization for Economic Cooperation and Development (2013).

⁸ Some eastern schools had admission exams, but routinely enrolled students “with conditions,” that is, despite having failed (Karabel 2006).

Arriving at Harvard in 1900, he aspired to join Porcellian, the most exclusive club. He later described his failure to do so as the greatest disappointment of his life. Eleanor Roosevelt added that this experience led him to identify “with life’s outcasts” (Karabel 2006). Similarly, a student at Princeton stated, on being rejected by a club: “The news came like a thunderbolt. With a cold, sick feeling the bottom dropped out of my college life” (Karabel 2006). Such reactions may seem out of proportion, but they echo the disappointment some students today feel about not getting into this or that elite college. 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 was problematic. In 1892, for example, Charles Eliot justified it: “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” (Veysey 1965).

But in time, university leaders realized that the clubs posed challenges. First, alumni might donate to their clubs rather than to their school, especially if the clubs were the source of 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 (Karabel 2006). Third, the clubs had their own agenda, one that often glorified unruly/unscholarly behavior; in the words of Princeton’s President Woodrow Wilson (1902-1910): “The sideshows are so numerous, so diverting, so important, if you will, that they have swallowed up the circus” (Brubacher and Rudy 1958).

To complicate matters, this happened in a period of rising antisemitism when many of the new students—particularly at urban schools such as Columbia, Harvard, and Pennsylvania—were Jewish. For example, by 1920, the share of Jewish students at Columbia had reached 30 percent (Karabel 2006). These are estimates since initially, colleges had little data on religious affiliation. But they began to work toward this at this time. At Harvard, President Abbott Lowell asked a committee to use extensive personal data to assign students to three groups: J1—“conclusively Jewish,” J2— “indicatively Jewish,” and J3—“possibly, but not probably Jewish” (Synnott 1979).

In 1919 Columbia responded by implementing selective admissions. It capped its entering class size, started requiring more personal data from applicants, and began to reject some without stating a reason. Soon other schools (including Cornell, Harvard, MIT, Stanford, and Yale) implemented similar policies.

Despite its convoluted origins, selectivity gradually made for an academically stronger student body at these schools. Making enrollments predictable also allowed universities to better supply services like dormitories and dining rooms, aiding efforts to rein in student behavior.

In other words, selectivity set the stage for a massive sorting process—in this case involving students. The academically more talented began to congregate in a few schools. Just as happened

with journals, a pecking order developed. Further, like the appearance of journals facilitated the sorting of professors, standardized tests like the SAT (1926) facilitated student sorting—in both cases, observability was central. Among schools, student sorting created winners and losers. Many schools whose students had been quite similar to Harvard’s in ability found that to no longer be the case (as documented by Hoxby 2009). Further, an “open enrollment” sector continued to exist, as predicted by theory.

Selectivity produced a financially virtuous circle for the schools that moved early. These took in bright and/or well-connected students and provided them with good experiences, including teaching by the best professors (at least in terms of research ability). They gave their students networks with potentially valuable labor market and marriage market contacts. Their satisfied alumni donated to them, allowing these schools to make even more investments, and so forth.

Events around World War II added a powerful ingredient that helped raise spending and reinforced inequality: federal public research funding. Gruber and Johnson (2019) note that in 1938 federal and state governments spent a combined 0.08 percent of national income on research. By 1944, the U.S. government was spending nearly 0.5 percent, a figure which would reach 2 percent by the 1960s (thereafter declining to its pre-war level). Much of this funding is allocated to research projects as a function of quality as assessed by panels of experts. To the extent that the top universities account for more than their fair share of research talent, it is not surprising that they receive a substantial share of these funds (Graham and Diamond 1997).⁹

Here again, a contrast emerges with Europe, where many states suppressed the type of inequality and sorting that characterizes the American university sector. For example, German states and countries like Spain—which largely control their universities—allocate resources to promote equality across schools. In a similar spirit, German states often set faculty salaries largely as a function of seniority and rank. Such restrictions limit universities’ ability to differentiate pay and compete on a global hiring market.

5 Tenure

Higher salaries, lower teaching, and enhanced laboratory space illustrate some ways in which professors’ compensation began to reflect research performance. Furthermore, this period saw the emergence of tenure, a salient reward for performance. Unlike salary, tenure is “lumpy”—it does not allow for small enhancements. Rather, tenure rewards reaching a threshold level of achievement, and cannot be taken away. This section describes the emergence of tenure and discusses how agency theory suggests it can incentivize research performance.

The Emergence of Tenure

⁹ See also https://www.nsf.gov/statistics/nsf13325/content.cfm?pub_id=4240&id=2

One can see the rise of tenure as a response to two related sets of demands. The first was a desire on the part of professors for academic freedom and protection from arbitrary dismissal. Several events illustrating this involve economists, in part because during this period many saw themselves as reformers in addition to researchers—this created tension with university presidents and donors. One famous case involved Richard Ely, who was hired (in 1881) by Johns Hopkins president Andrew Gilman. Ely was active in the progressive movement, supported labor unions, and helped found the American Economic Association (Barber 1987). He later moved to the University of Wisconsin at a time of labor tensions in that state. Soon Wisconsin president Charles Adams heard complaints that Ely taught socialism and had “entertained a union organizer in his home” (Schrecker 1986). The university’s board appointed a committee to investigate. The board not only supported Ely, but made a broader statement (Scott, 2018):

As Regents of a university with over a hundred instructors ... 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.

However, this outcome was not preordained, as is evident in the case of Edward Ross, one of Ely’s students. Ross obtained a Ph.D. in political economy at Johns Hopkins and taught economics and sociology at Stanford, where he publicly defended the socialist Eugene Debs and opposed Asian immigration. This irritated Jane Stanford, whose fortune partially derived from her late husband’s investment into railroads built using Chinese labor. She accused Ross of aiding the “vilest elements of socialism,” and pressured President David Jordan to terminate Ross (Dorn, 2017). Jordan obtained Ross’s resignation (1900), but at the cost of bad press and eight additional faculty departures. The case contributed to creating the American Association of University Professors (AAUP), which in 1915 stated a desire to “safeguard freedom of inquiry ... against both covert and overt attacks.”¹⁰

The second set of demands reflected universities’ desire to incentivize and occasionally terminate professors. Harvard’s early hiring practices illustrate some associated challenges. In the 1700s, Harvard hired tutors using renewable contracts. Its officials explained this stating that permanent contracts raised the risk that bad instructors might be “fixed on the college for life” (Metzger, 1973). Further, renewable contracts could provide incentives; as put by Harvard’s statutes, they could “excite tutors ... to greater care and fidelity in their work” (Metzger, 1973). A further benefit

¹⁰ American Association of University Professors, “Declaration of principles on academic freedom and academic tenure.”

of finite terms is that a school does not have to cite a cause to dismiss a professor who performs poorly—it can simply let his contract end. Yet as any manager knows, terminating a worker is difficult even if it is feasible; when renewals come up, it is always tempting to “kick the can down the road.” Metzger (1973) states that a 1760 Harvard rule limiting appointments within a rank to eight years “was to prevent incumbencies from being lengthened by reappointments given out of neglect or sympathy; it was intended to defeat the importance of kindness in the serious business of evaluating personnel.”

These two sets of demands—for permanence/freedom and performance/incentives—could become intertwined. For example, in the 1930s, Harvard president James Conant wished to strengthen his school’s social sciences, widely perceived to have fallen behind those at Chicago and Columbia. Two economists with somewhat limited publication records and labor union ties came up for reappointment. When they were terminated, some complained that politics had been the cause. A faculty committee found no evidence of this but questioned the timing of their review.

Cases like this illustrate that it is difficult to disentangle the desire for permanence/freedom and the desire to provide incentives. It is possible to see tenure—by creating a pre-defined trial period leading to an “up-or-out” date—addressing both sets of demands.¹¹ It is not entirely clear when tenure became formalized because, as happened with many features of the U.S. 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.¹²

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 allow the university to shirk on its compensation commitments. A solution to this problem is for the school to commit in advance to providing rewards for relative performance (Lazear and Rosen 1981; Carmichael 1983; Malcomson 1984). For example, it might create a teaching award, leaving in question only who will win it.¹³

However, 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. This implies that all else equal, tenure will more effectively promote research if combined with a process whereby professors sort into schools/departments according to their research ability.

Indeed, once measures of professors’ research output were available, and once universities began

¹¹ As a matter of law, tenure does not guarantee permanence of employment, but rather sets the bar for dismissal high. Faculty who violate behavioral policies or fail to teach, for example, can be dismissed.

¹² In recent decades, tenure has become less common in the United States, and it is increasingly available only to research-focused professors at wealthier universities (Figlio, Schapiro, and Sotter 2015).

¹³ In practice, perfect performance measures do not exist. MacLeod (2003) shows that this necessarily leaves space for conflict, which decreases as the quality of information improves. In the appendix we show that lower variation in ability increases the effectiveness of prizes; see Brown (2011) for evidence.

to bid on faculty, stratification by ability/performance resulted. In other words, academic fields came to display a pecking order of departments/schools. This sorting process was facilitated by professors, given that they themselves have incentives to segregate by ability. For example, some believe there are externalities from proximity to talented colleagues, or that belonging to an elite department enhances their reputation (MacLeod et al. 2017).

Tenure likely facilitated such sorting. The individuals most qualified to judge faculty in a specialized area are other academics in that area. This can create a problem when hiring—if incumbents’ employment depends upon relative evaluations, they will naturally want to hire individuals less skilled than themselves. Carmichael (1988) observes that tenure helps solve this problem. If faculty with tenure have protected employment, they have no incentive to misrepresent the quality of new hires.

Agency theory provides additional insights into tenure. First, the length of the tenure clock may allow for a more precise measure of performance. Second, labor economists often treat labor as a standard commodity; for instance, workers produce a certain output in an hour and are paid for that hour. Work on research can be quite different; for example, a researcher might work for years to make one important discovery—MacLeod (2007) calls this an *innovative* commodity. In such cases, optimal compensation can feature delayed rewards and bonuses (Levin 2003; Fuchs 2007; Manso 2011), which tenure helps introduce. Third, universities wish to hire professors with an intrinsic preference for research (Stern 2004). A long tenure clock and relatively low pay can result in the self-selection of faculty who will keep exerting research effort even after receiving tenure. Fourth, with tenure, faculty come to have a vested interest in their institution, and will be more willing to contribute towards its advancement (Carmichael 1983, 1988).

Note that tenure mostly did not emerge in Europe. To be clear, European universities certainly have a set of senior professors who enjoy job security and other privileges. For example, chaired German professors enjoy arguably greater power in their schools than American full professors—they are the heirs of the “masters” that controlled ancient universities like Paris and Heidelberg.

However, younger academics in Europe are much less likely to face an internal “up-or-out” research-based evaluation. At some schools (for instance in France and the U.K.) faculty essentially receive immediate tenure. Their progression through the ranks may indeed depend upon their performance, but their permanence is assured. In other settings (for example in Germany), universities hold open contests for all senior professorships, in some cases even precluding applications by junior academics at the same school. This differs from the American arrangement where the junior “tenure-track” professor takes part in an internal evaluation and the only question is whether she has attained an absolute level of attainment.

Further to the extent that tenure has positive impacts, these are likely reinforced by the fact that it interacts with faculty sorting by research talent. Sorting implies that each assistant professor is matched to a department at which she is not far, in expected output, from her colleagues; combined

with a tenure goal, this will raise effort through a wide range of the ability distribution.

6 Conclusion

We have discussed why the United States accounts for a high proportion of the world’s leading research universities. Our story is not one of success by design, but rather of competition helping to create a confluence of incentive mechanisms that help explain current performance.

The rise of American higher education began in the late 1800s, when industrialization increased universities’ interest in advanced/specialized instruction, hence on professors’ ability to do research. The emergence of specialized fields (with accessories like journals) produced improved measures of performance. Armed with these, schools began to bid for and reward good researchers. This began to concentrate talent at a few schools, a process reinforced by the emergence of tenure that selected dedicated individuals and created incentives for them. Schools that aggressively pursued these reforms began to experience growth and responded to associated challenges by implementing 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-talented professors and strong students—increasingly from all over the world—at a few schools.

In thinking about the emergence and persistence of US leadership in university research, it is worth contrasting this outcome with the situation in K-12 education. A country’s performance in K-12—say as measured by international tests—depends on the state of hundreds if not thousands of schools. In addition, the existence of the “teacher value-added” literature illustrates the difficulty of measuring performance in K-12. By contrast, when it comes to university research, it is easier to measure performance, sorting may be less of a concern, and top quality output is disproportionately important. Thus national performance can depend on what happens at a few dozen universities.

The American system is well suited to producing top schools, although at the cost of inequality. To illustrate, del Corral (2020) compares the performance of the Spain and the United States in a recent “Shanghai” university ranking. The United States accounts for 40 of the top 100 universities; Spain for 0. On the other hand, 83 percent of public Spanish universities (delivering in-person instruction) appear somewhere in the ranking; only 23 percent of their American counterparts do.

In closing, we note that agency theory highlights that performance depends crucially on high-quality performance metrics. Today the academy is rife with complaints that individuals focus excessively, and with calls for “inter-disciplinary.” Our results suggest that a move in that direction is likely to entail trade-offs because it may bring on noisier evaluation systems.

7 Appendix

One goal of our paper is to highlight the role of faculty incentives in explaining universities' research output.¹⁴ Specifically, we are interested in the impact that specialization has on performance. To explore this, this appendix sets out a simple agency model building on the multi-tasking framework developed by Holmstrom and Milroam (1991). To explore the effects of competition, we augment this with a tournament model along the lines of Lazer and Rosen (1981) and Green and Stokey (1983).

Suppose, as discussed in the main text, that universities can enhance the demand for their services by improving their reputation for research production. To keep matters simple, suppose that a university hires a single professor, i , who can exert research effort, denoted r_i . This produces the following payoff for the school:

$$\Pi = V_r \times r_i - w_i, \tag{1}$$

where V_r is the value the university assigns to research and w_i is the compensation paid to the professor. Part of the latter may be in amenities such as lab space, reduced teaching assignments, etc.

7.1 Faculty performance and preferences

Suppose that when evaluating a candidate, the university observes a measure of individual research performance. Specifically, in considering an individual i who works in field f , the university observes:

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

where r_i is research effort, α_i is ability (assumed to be non-random for simplicity), and $\gamma_f \sim N(0, \eta_f)$ is field-specific noise. This last component reflects the fact that non-experts—say a university hiring or tenure committee—have trouble assessing professors' research performance. This is particularly true in making comparisons across fields. For example, in the humanities faculty tend to write books, while articles are the norm in the sciences. Similarly, a chemistry professor may be listed as a co-author on all papers produced by her laboratory, while an economic theorist writes alone—it is hard to tell who has higher performance.

One way to tackle this problem is by comparing a candidate to other professors in her field. To see this, suppose the university observes a second signal that consists of the average performance in a candidate's field:

$$s_f = \bar{r}_f + \alpha_f + \gamma_f.$$

The first term \bar{r}_f is the average effort in the field; this will be determined as an equilibrium outcome

¹⁴The seminal work of Merton (1968) emphasizes the importance of reward design to understand the structure and performance of academic research.

below. The random variable α_f represents the average ability of individuals in the field, and we let this be given by $\alpha_f \sim N(\bar{\alpha}_f, \Omega_f)$; i.e, Ω_f captures the variance in ability in field f . As before, $\gamma_f \sim N(0, \eta_f)$.¹⁵

Using both signals, the university can create a signal of *relative* performance—comparing individuals against their field (Lazear and Rosen 1981). In particular, it can compute:

$$\begin{aligned}\delta_{if} &= s_i - s_f \\ &= r_i - \bar{r}_f + \alpha_i - \alpha_f.\end{aligned}$$

A key point is that comparing within field causes the term γ_f to drop out. The variance of this signal is thus Ω_f .

Following Lazer and Rosen (1981) and Holmstrom and Milgrom (1991), compensation can then take the form:

$$w_i = \bar{w}_i + b_i s_i + b_f \delta_{if},$$

which features a base wage and bonuses paid as a function of the individual and the relative performance measures.¹⁶ Let $\vec{p} = \{\bar{w}, b_i, b_f\}$ summarize the contract offered by the university.

Suppose that in addition to doing research, professors can engage in non-research, “outside” activities. Broadly, this could include consulting work, but also staying active in alumni networks or, in the 1800s, religious groups. Formally we assume that the effort exerted by a professor is given by:

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

where r_i is effort on research and o_i is effort on outside activities. This highlights the fact that effort is fundamentally a resource allocation problem—how to allocate one’s time. To keep things tractable, we posit a quadratic effort cost function (Holmstrom and Milgrom 1991):

$$C(\vec{e}_i) = \vec{e}_i^T \begin{bmatrix} 1 & d \\ d & 1 \end{bmatrix} \vec{e}_i / 2.$$

The off-diagonal terms, $d \in (0, 1)$, imply that the two types of effort are substitutes—an increase in outside activity, o_i , increases the cost of supplying research effort and vice versa. Note that the matrix $A = \begin{bmatrix} 1 & d \\ d & 1 \end{bmatrix} / 2$ is invertible and positive definite, and hence $C(\vec{e}_i) \geq 0$ for any effort and

¹⁵It would be possible to add an additional subscript t to γ , exploring over-time variation. We leave that for future work.

¹⁶We have assumed that measures are normally distributed, and hence this equation can be viewed as a measure of log wages.

is zero if and only if:

$$\vec{e}_i = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

The payoff for professor i takes the form:

$$U_i = E \{w_i\} + V_o \times o_i + V_{ri} \times r_i - C(\vec{e}_i) - \frac{\rho}{2} \text{var}(w_i).$$

In words, the parameters that determine faculty performance are the marginal value of outside activities, V_o , the intrinsic value the professor attaches to research, V_{ri} , and relative risk aversion, ρ . In order to have a smooth trade-off between research and outside activities, we assume that the value of both activities is sufficiently similar such that professors always choose a combination of research and outside activities:

$$1 \geq \min \left\{ \frac{V_o}{V_r}, \frac{V_r}{V_o} \right\} > d > 0.$$

Finally, note that this setup does not model the issue of matching faculty to appointments. There is work showing that some individuals have an intrinsic taste for research that leads them to seek university positions (Stern 2004; Agarwal and Ohya 2013). Our focus is on the design of rewards.

7.2 Optimal Compensation

We now derive the optimal contract for the university. The problem facing it is to choose a pay package, $\vec{p}_i = \{\bar{w}_i, b_i, b_f\}$, subject to participation and incentive constraints.¹⁷

Relying on the assumption that the expected value of the noise term γ_f is zero, the professor's payoff given compensation, $\vec{p}_i = \{\bar{w}_i, b_i, b_f\}$, is:

$$U_i(\vec{e}_i, \vec{p}_i) = E \{w_i\} + V_{ri}r_i + V_o o_i - C(\vec{e}_i) - \frac{\rho}{2} \text{var}(w_i) \quad (2)$$

$$= \bar{w}_i + b_i(r_i + \alpha_i) + b_f(r_i - \bar{r}_f + \alpha_i - \bar{\alpha}_f) + V_{ri}r_i + V_o o_i - C(\vec{e}_i)$$

$$- \frac{\rho}{2} (b_i^2 \eta_f + b_f^2 \Omega_f).$$

$$= \bar{w}_i + (b_i + b_f)r_i + V_{ri}r_i + V_o o_i - C(\vec{e}_i) \quad (3)$$

$$+ b_i \alpha_i + b_f(\alpha_i - \bar{\alpha}_f - \bar{r}_i) - \frac{\rho}{2} (b_i^2 \eta_f + b_f^2 \Omega_f).$$

Let $B_i \equiv (b_i + b_f)$ be the total reward for effort provided by the university. Notice that effort is a function of the total reward (B_i), the professor's intrinsic motivation, V_{ri} , and the return to the outside option V_o . We suppose that parameters are set so we have an interior solution with positive

¹⁷ To keep matters simple, we fix the candidate's alternative utility to the average ability of individuals in the field, and hence $U^0 = \bar{\alpha}_f$.

levels of both types of effort, r_i and o_i . Then we can set:

$$\vec{e}_i^*(B_i) = \begin{bmatrix} r_i(B_i) \\ o_i(B_i) \end{bmatrix} \quad (4)$$

$$= \operatorname{argmax}_{\vec{e}} U_i(\vec{e}, \vec{p}), \quad (5)$$

$$= \frac{1}{1-d^2} \begin{bmatrix} B_i + V_{ri} - dV_o \\ V_o - d(B_i + V_{ri}) \end{bmatrix} \quad (6)$$

Thus, as the total reward for research, B_i , increases, the effort devoted to research goes up, while that devoted to outside activities declines. The payoff for the university is given by:

$$\begin{aligned} \Pi(r_i, \vec{p}_i) &= V_r \times r_i - (\bar{w}_i + b_i(r_i + \alpha_i) + b_f(r_i - \bar{r}_f + (\alpha_i - \bar{\alpha}_f))), \\ &= V_r \times r_i - (\bar{w}_i + B_i(r_i + \alpha_i) - b_f(\bar{\alpha}_f + \bar{r}_f)). \end{aligned}$$

Note that variable employment costs thus have two terms. One term is the direct incentive cost that rewards total productivity, $r_i + \alpha_i$. The second term reflects the relative performance evaluation. If the professor is above average, that creates an additional cost. As we shall see, the optimal compensation also depends upon the risk the professor must face, which affects the base wage, \bar{w} .

The determination of bonus pay is assumed to maximize university welfare subject to the professor's participation constraint:

$$\begin{aligned} \max_{\vec{p}_i} \Pi(\vec{p}_i) & \quad (7) \\ \vec{e}_i & \in \operatorname{argmax}_{\vec{e}_i} U_i(\vec{e}_i, \vec{p}_i), \\ U_i(\vec{e}_i, \vec{p}_i) & \geq U^0 = \bar{\alpha}. \end{aligned}$$

The solution to this problem is relatively straightforward. We can substitute the closed form solution for effort from (4). The Lagrangian for the optimal contract is:

$$L(\vec{p}_i, \lambda) = \Pi(r_i(\vec{p}_i), \vec{p}_i) + \lambda \{U_i(\vec{e}_i^*(\vec{p}_i), \vec{p}_i) - \bar{\alpha}_f\}.$$

In order to have an optimal solution for the fixed payment \bar{w} it must be the case that $\lambda = 1$. Using (2) and (4) this becomes:

$$\begin{aligned} L(\vec{p}_f, 1) &= (V_r + V_{ri}) r_i(B_i) + V_o \times o_i(B_i) - C(\vec{e}_i^*(B_i)) - \bar{\alpha}_f \\ &\quad - \frac{\rho}{2} (b_i^2 \eta_f + b_f^2 \Omega_f). \end{aligned}$$

Note that the first line depends only upon total performance-based compensation, B_i . The second line measures the trade-off between the two forms of compensation, based upon the relative precision

of the performance measures. Given B_i we follow Grossman and Hart (1983) and define the risk premium that must be paid to faculty as a function of performance pay:

$$RPC(B_i) = \min_{B_i \leq b_i + b_f} \frac{\rho}{2} (b_i^2 \eta_f + b_f^2 \Omega_f).$$

This is increasing in B_i with each component given by:

$$b_i(B_i) = \frac{\Omega_f}{(\Omega_f + \eta_f)} B_i, \quad (8)$$

$$b_f(B_i) = \frac{\eta_f}{(\Omega_f + \eta_f)} B_i. \quad (9)$$

Thus the ratio of relative performance pay (b_{if}) to performance pay is:

$$\frac{b_f}{b_i} = \frac{\eta_f}{\Omega_f}. \quad (10)$$

In other words, the use of relative performance pay increases with η_f , i.e., with the difficulty non-experts have in assessing performance in a field f . From this we can compute the overall costs of performance pay:

$$\begin{aligned} RPC(B_i) &= \frac{\rho}{2} \frac{\Omega_f \eta_f}{(\Omega_f + \eta_f)} B_i. \\ &\equiv \kappa B_i \end{aligned} \quad (11)$$

where κ measures the marginal cost of performance pay upon faculty payoffs. Thus, the cost of performance pay falls with more precise signals. If η_f increases due to the increase in the complexity of a field, then costs can be reduced if the variance of individual performance, Ω_f , can be reduced. This can be achieved with hierarchical sorting by skill. From this we can compute both the optimal level of performance pay and the total expected compensation to faculty as the solution to $\frac{dL(B_i)}{dB_i} = 0$, where:

$$\begin{aligned} L(B_i) &= (V_r + V_{ri}) \times r_i^*(B) + V_o \times o_i^*(B_i) - C(\bar{e}_i^*(B_i)) \\ &\quad - \bar{\alpha}_f - RP(B_i). \end{aligned} \quad (12)$$

Taking the first order condition we get:

$$\begin{aligned}\frac{dL(B_i)}{dB_i} &= \left(V_r + V_{ri} - \frac{\partial C}{\partial r_i} \right) \frac{dr_i}{dB_i} \\ &+ \left(V_o - \frac{\partial C}{\partial o_i} \right) \frac{do_i}{dB_i} \\ &- \kappa.\end{aligned}$$

Using the closed form solution for effort we get:

$$B_i^* = V_r - \kappa(1 - d^2), \quad (13)$$

$$= V_r - \frac{\rho}{2} \frac{\Omega_f \eta_f}{(\Omega_f + \eta_f)} (1 - d^2). \quad (14)$$

By construction, faculty utility is fixed at $\bar{\alpha}_f$, the average quality of individuals in the peer pool. Rewards have to compensate individuals choosing effort above their normal level given by the intrinsic return to research, V_{ri} . Total expected compensation is given by:

$$w_i = \bar{\alpha}_f + U^0(B_i^*) + \kappa B_i^*.$$

Thus, the faculty member receives additional pay to compensate her for additional effort,

$$U^0(B_i^*) = V_{ri} r_i^*(B_i^*) + V_o o_i^*(B_i^*) - C(\bar{e}_i^*(B_i^*)), \quad (15)$$

and the costs associated with relative performance evaluation, κB_i^* .

7.3 Implications

This yields two propositions. Consider first the situation in which there is no bonus pay ($B_i = 0$). In that case result (6) implies:

Proposition 1. *Suppose faculty are paid a fixed wage that ensures participation (with no bonus pay for research performance). Then research effort is given by:*

$$r_i = \frac{V_{ri} - dV_o}{1 - d^2}.$$

Hence research is increasing in the intrinsic preference for research, and decreasing in the return to outside activities.

It is optimal for colleges to reward faculty as a function of their research performance. The amount of reward is given by (8), (9) and (13). Hence, we have:

Proposition 2. *Optimal performance pay is decreasing with the variance of the ability of the candidates. The use of relative performance evaluation increases rewards for research, and the level of research activities.*

Notice that the level of performance pay is determined by the value of κ :

$$\kappa = \frac{\rho}{2} \frac{\Omega_f \eta_f}{(\Omega_f + \eta_f)}.$$

A smaller κ leads to lower costs, and this can only be achieved with universities becoming increasingly selective in their pool of applicants which determines the value of Ω_f . Thus we have:

Proposition 3. *If departments can draw from a more homogeneous pool of candidates (smaller Ω_f), this decreases κ , and hence increases the optimal bonus (from 13), and in turn increases research output (from 6).*

References

- [1] Agarwal, Rajshree, and Atsushi Ohyama. (2013) Industry or Academia, Basic or Applied? Career Choices and Earnings Trajectories of Scientists. *Management Science*, 59(4): 950–970.
- [2] Barber, W. J. (1987) Should the American Economic Association have toasted Simon Newcomb at its 100th birthday party? *Journal of Economic Perspectives*, 1(1), 179-183.
- [3] Bartle, R. G. (1995) A brief history of the mathematical literature. *Publishing Research Quarterly*, 11, 3-13.
- [4] Boyer, John W. 2015. *The University of Chicago: A History*. Chicago and London: The University of Chicago Press.
- [5] Brown, J. (2011). Quitters never win: The (adverse) incentive effects of competing with superstars. *Journal of Political Economy* 119 (5), 982–1013.
- [6] Brubacher, John S. and Willis Rudy. 1958. *Higher Education in Transition*. New York: Harper & Row.
- [7] Burke, C. B. (1982). *American Collegiate Populations: A Test of the Traditional View*. New York and London: New York University Press.
- [8] Butts, R. F. (1939). *The College Charts its Course: Historical Conceptions and Current Proposals*. New York and London: McGraw-Hill Book Company, Inc.

- [9] 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.
- [10] Carmichael, H. L. (1983) Firm-Specific Human Capital and Promotion Ladders, *Bell Journal of Economics*, 14, 251-258
- [11] Carmichael, H. L. (1988). Incentives in academia: Why is there tenure? *Journal of Political Economy* 96 (3), 453–472.
- [12] Clotfelter, C. T. (2010). Introduction. In C. T. Clotfelter (Ed.), *American Universities in a Global Market*. Chicago and London: The University of Chicago Press.
- [13] 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.
- [14] Curti, M. and R. Nash (1965). *Philanthropy in the Shaping of American Higher Education*. New Brunswick: Rutgers University Press.
- [15] del Corral, J. (2020) Ranking de Shanghai: ¿comparando España con Estados Unidos? *mimeo*, Studia XXI.
- [16] Dorn, Charles. 2017. *For the common good: a new history of higher education in America*. Ithaca, NY: Cornell University Press.
- [17] Eliot, C. W. (1908). *University Administration*. Boston and New York: Houghton Mifflin Company.
- [18] 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.
- [19] Figlio, David N., Morton O. Schapiro, and Kevin B. Soter. 2015. Are tenure track professors better teachers? *The Review of Economics and Statistics* 97 (4): 715-724.
- [20] Flexner, A. (1946). *Daniel Coit Gilman: Creator of the American Type of University*. New York: Harcourt, Brace and Company.
- [21] Fuchs, W. (2007) Contracting with repeated moral hazard and private evaluations, *American Economic Review* 97(4), 1432-1448.
- [22] Geiger, R. L. (1986). *To Advance Knowledge: The Growth of American Research Universities, 1900-1940*. New Brunswick and London: Transaction Publishers.
- [23] Ghosh, S. and M Waldman (2010) Standard promotion practices versus up-or-out contracts. *The RAND Journal of Economics*, 41(2), 301-325.

- [24] 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.
- [25] Goldin, C. and L. F. Katz (2008). *The Race Between Education and Technology*. Cambridge and London: The Belknap Press of Harvard University Press.
- [26] 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.
- [27] Green, Jerry R., and Nancy L. Stokey. (1983) A Comparison of Tournaments and Contracts. *Journal of Political Economy*, 91(3): 349–364.
- [28] Grossman, Sanford J., and Oliver D. Hart. (1983) An Analysis of the Principal-Agent Problem. *Econometrica*, 51(1): 7–45.
- [29] Gruber, J. and S. Johnson (2019). *Jump-Starting America: How Breakthrough Science Can Revive Economic Growth and the American Dream*. New York: PublicAffairs.
- [30] Hawkins, H. (1960). *Pioneer: A History of the Johns Hopkins University, 1874-1889*. Baltimore and London: The Johns Hopkins University Press, 2002.
- [31] Herbst, J. (1982). *From Crisis to Crisis: American College Government, 1636-1819*. Cambridge and London: Harvard University Press.
- [32] Hofstadter, R. (1955). *Academic Freedom in the Age of the College*. New Brunswick and London: Transaction Publishers, 1996.
- [33] Holmstrom, 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.
- [34] Hoxby, C. (2009). The changing selectivity of American colleges. *Journal of Economic Perspectives* 23 (4), 95–118.
- [35] Hoxby, C. M. (2016). The productivity of U.S. postsecondary institutions. Technical report, Stanford University.
- [36] James, H. (1930). *Charles W. Eliot: President of Harvard University, 1869-1909, Volume I*. Boston and New York: Houghton Mifflin Company.
- [37] 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.
- [38] Kelley, B. M. (1974). *Yale: A History*. New Haven and London: Yale University Press.

- [39] Lazear, E. P. and Rosen, S. (1981) Rank-Order Tournaments as Optimal Labor Contracts *Journal of Political Economy*, 89, 841-864.
- [40] Lazear, E. P. (1989) Pay Equality and Industrial Politics *Journal of Political Economy*, 97, 561-580. Leslie, W. B. (1992). *Gentlemen and Scholars: College and Community in the "Age of the University," 1865-1917*. University Park: Pennsylvania State University Press.
- [41] Levin, J. (2003) Relational Incentive Contracts. *American Economic Review*, 93, 835-857
- [42] MacLeod, W. B. (2003). Optimal contracting with subjective evaluation. *American Economic Review* 92 (1), 216–240.
- [43] MacLeod, W. B. (2007). Reputations, relationships, and contract enforcement. *Journal of Economic Literature* 45 (3), 597-630.
- [44] 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.
- [45] MacLeod, W. B. and M. Urquiola (2015). Reputation and school competition. *American Economic Review* 105 (11), 3471–3488.
- [46] Malcomson, J. M. (1984). Work incentives, hierarchy, and internal labor markets. *Journal of Political Economy* 92 (3), 486–507.
- [47] Manso, G. Motivating Innovation, *The Journal of Finance*, 2011, 66(5), pp. 1823-1860.
- [48] Merton, RK, The Matthew Effect in Science, *Science*, 1968, 159, 56-63.
- [49] Metzger, W. P. 1973. Academic tenure in America: a historical essay. *Faculty Tenure: A Report and Recommendations* by the Commission on Academic Tenure in Higher Education. San Francisco: Josey-Bass, Inc.
- [50] Neilson, C. (2017). Targeted vouchers, competition among schools, and the academic achievement of poor students. Unpublished manuscript, Princeton University.
- [51] Organization for Economic Cooperation and Development (2013) Education expenditures by country. Mimeo, https://nces.ed.gov/programs/coe/pdf/Indicator_CMD/coe_cmd_2013_03.pdf.
- [52] Paulsen, F. (1906). *The German Universities and University Study*. New York: Charles Scribner's Sons.
- [53] Rudolph, F. (1962). *The American College and University: A History*. Athens and London: The University of Georgia Press, 1990.

- [54] Rudolph, F. (1977). *Curriculum: A History of the American Undergraduate Course of Study Since 1636*. San Francisco: Josey-Bass Publishers.
- [55] 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.
- [56] Schrecker, E. W. 1986. *No Ivory Tower: McCarthyism and the Universities*. New York and Oxford: Oxford University Press.
- [57] Scott, Joan Wallach. (2018). The tension between the university and the state. In *Academic Freedom: The Global Challenge*. Edited by Michael Ignatieff and Stefan Roch. Budapest: Central European University Press.
- [58] Stern, S. (2004) Do scientists pay to be scientists? *Management Science*, 50, 835-853.
- [59] Storr, R. J. (1953). *The Beginnings of Graduate Education in America*. Chicago: The University of Chicago Press.
- [60] 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.
- [61] 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.
- [62] Thompson, D. G. Brinton. 1946. *Ruggles of New York: A Life of Samuel B. Ruggles*. New York: Columbia University Press.
- [63] Urquiola, M. (2020). *Markets, Minds, and Money: Why America Leads the World in University Research*. Cambridge, MA: Harvard University Press.
- [64] Veysey, L. R. (1965). *The Emergence of the American University*. Chicago and London: The University of Chicago Press.
- [65] Xiong, Heyu and Yiling Zhao (2020) Sectarian Competition and the Market Provision of Human Capital. Mimeo, Case Western Reserve University.