

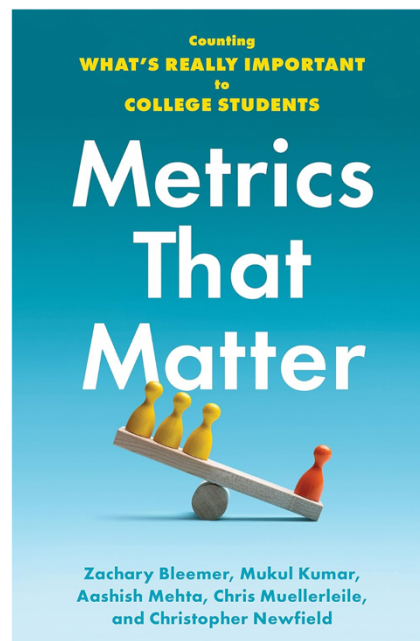
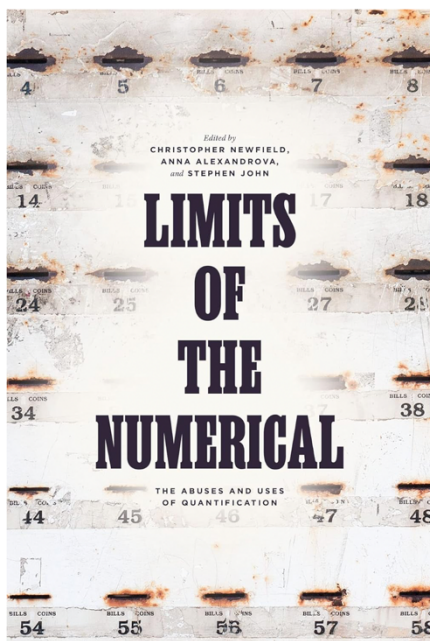
Do Metrics Matter? Metrics Reform and Metrics Abolitionism
Lecture in the School of Education, University of Bristol, 14 November 2023
Christopher Newfield, ISRF, London

I'm going to make a pair of arguments today.

First, I'm going to say it is worthwhile to fix metrics, to replace typical and generally bad metrics with better ones. This is a reform project that advances intellectual and social justice.

Second, I'm going to say that this reform project isn't enough. We also need to subordinate metrics to social and educational missions. We need epistemic equality between quantitative and qualitative methods, and this may well require metrics abolitionism—strategic abolitionism. This project follows from asking how universities got to the point of accepting bad metrics in the first place. I'll point to a conflict or struggle within the professional-managerial class, one that now has to be faced. But more on these connections later.

First, I am co-editor and co-author of a pair of books about metrics that look at them in both their quantitative and qualitative dimensions. One is called *Limits of the Numerical*, and the other *Metrics that Matter*.



The second of these is an example of academic realism. We accepted that metrics “are here to stay” and that the “train has left the station,” to use two annoying phrases. So rather than trying to stop the metrics train, we decided as a diverse interdisciplinary team who disagreed on a lot of things that we’d try to help a specific group, undergraduate students, learn how to take the train to where they actually wanted to go. (Note we posit that subjectivity remains autonomous from metrics—that it isn’t already programmed by them.) We organized the book in the order of metrics that high school students were likely to encounter.

Contents

Preface	vii
Introduction	1
① Return on Investment	11
② University Rankings	20
③ Selectivity	33
④ Tuition Sticker Price	53
⑤ Scorekeeping Student Debt	71
⑥ Average Wages by College Major	85
⑦ Access to Your Preferred Major	113
Conclusion: Metrics and Waste	132

We didn’t try to replace quantitative metrics with qualitative analysis, but to replace each of a set of standard metrics, like wages by course subject, with a better metric. We were aware of the limits of this strategy, and in the second part of this talk I will take up some broader issues with it. But first let me give you an

example of our humble but possibly resonant task of replacing a bad number with a better one.

(1)

One example is the metric of selectivity in admissions. The wider public generally assumes that if Cambridge rejects more applicants than East Anglia—is more selective than East Anglia—then it is a better university. We used different campuses of the University of California because it’s a single university system with widely varying campus acceptance rates—and an informal campus pecking order that has been stable since the 1960s.

RANK	SCHOOL NAME	NICHE OVERALL GRADE	ACCEPTANCE RATE	SAT RANGE	ACADEMICS GRADE	NET PRICE	TOTAL STUDENTS
1	UCLA	A+	17.3%	1180-1470	A+	\$13,816	29,004
2	UC Berkeley	A+	16.9%	1250-1510	A+	\$16,601	26,622
3	UC Santa Barbara	A+	32.6%	1090-1360	A	\$16,211	20,243
4	UC Davis	A	38%	1070-1340	A	\$16,841	27,453
5	UC San Diego	A	38%	1070-1340	A+	\$14,933	25,922
6	UC Irvine	A	38.7%	1030-1290	A	\$14,662	24,854
7	UC Riverside	A-	55.6%	980-1210	A-	\$13,365	18,279
8	UC Santa Cruz	A-	50.3%	1050-1300	A-	\$17,550	15,823
9	UC Merced	B+	60.7%	900-1120	B+	\$13,350	6,164

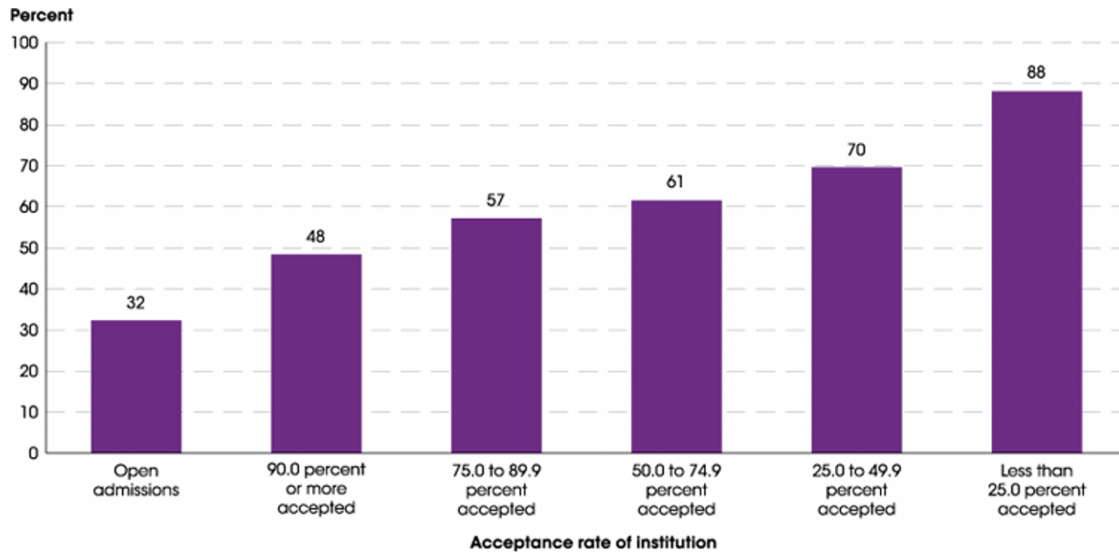
For example, in 2018, UC Berkeley accepted only 16.9% of applicants, which was about half the share of applicants that UC San Diego accepted. The common intuition is this: take the top 1/3rd of students who apply both to Berkeley and to UC San Diego. It appears that Berkeley rejected the weaker half of their top third of applicants, while all of these students were admitted to San Diego. This would drag the level of San Diego’s typical student down below Berkeley’s. Most people seem to accept the ranking magazine *U.S. News’s* assertion that “a school’s academic atmosphere is determined in part by students’ abilities and ambitions.”

UC Berkeley could be said to have a more able and ambitious overall student body because they rejected the bottom half of students among their top-third applicants, while that large tranche of their rejects were admitted by San Diego. Thus Berkeley will seem to have students who will be more academically prepared, find better jobs, help their peers with better connections, and the like. This view concludes that Berkeley is a better university than San Diego and that you can use its higher selectivity rate to quantify that.

Actually, this is wrong. I'll mention 2 of the five problems we go through in the chapter. One is that the relationship between selectivity and quality is circular. A college's perceived high quality attracts more applicants, which creates more rejections, which increases selectivity, which increases perceived quality, which attracts more applicants. This circularity means that universities have an interest in increasing their selectivity by increasing their mainstream marketing to get more applications which they can then reject. In the United States, universities do manipulate their selectivity rates with steadfast diligence.

However, selectivity perversely ties the basic indicator of college success-- completing a bachelor's degree--to the ability of a college to reject a high percentage of applicants before they arrive. Selectivity distracts everyone from the fundamental issue of a university's material resources—the funding that allows it to spend a given level of money on each student's instruction. I'll come back to this later on.

A deeper problem appeared when we looked at the relationship between selectivity and a standard measure of academic quality, which is graduation rates. These are variable in the US, and often quite low. Standard data show graduation rates correlating with selectivity.



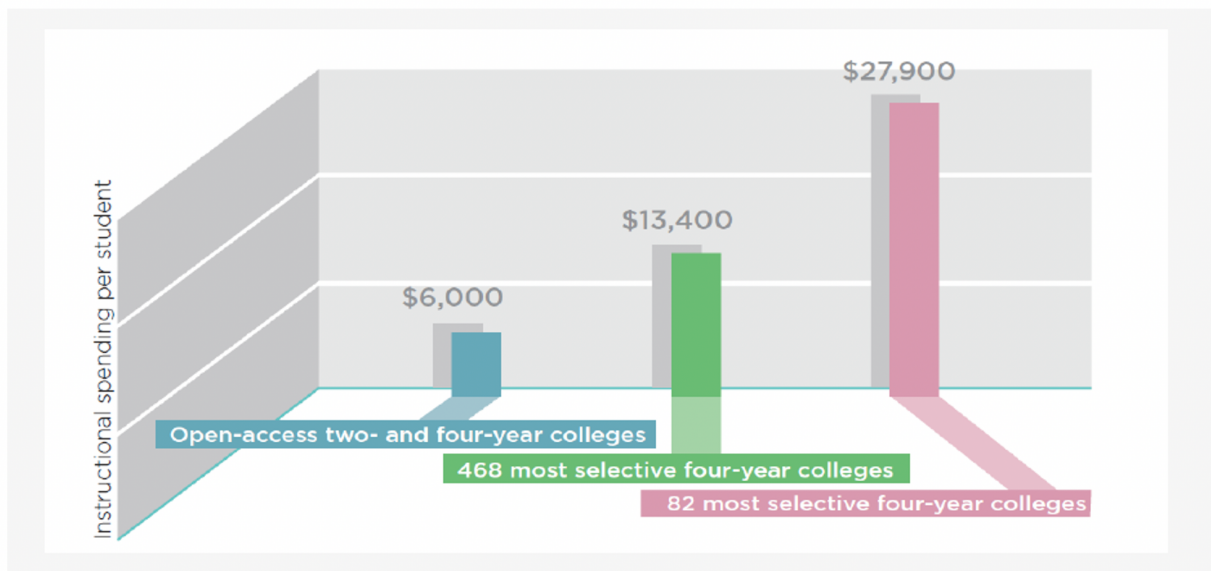
Many people—policymakers as well as parents and students—assume correlation is causality: rejecting weaker students leads to higher average graduation rates. If you go to a selective school, this reasoning goes, you’ll be surrounded by better students who will therefore succeed at graduating, which makes your graduation more likely too. The weaker students will have been rejected and thus aren’t there to drag down the mean, not to mention the level of the typical class.

To repeat, there *is* a correlation between selectivity and graduation rates in the U.S. One particularly detailed analysis, which William G. Bowen, Matthew M. Chingos, and Michael S. McPherson published as *Crossing the Finish Line* (2009), showed that 84% of students graduate in six years from their more selective university group, while 56% do the same for their less selective group. Selectivity *is* a better predictor of graduation rates than some other factors, like the quality of the student’s high school.

But the catch is that they also found *no causal* relation between selectivity and graduation rates. They ran an interesting simulation in which they retroactively “rejected” all students below a chosen high school grade threshold. They then compared graduation rates for their “more qualified” remaining group that was just on the higher side of the threshold. What the researchers found was startling. Most people assume that more selective colleges have better graduation rates for the simple reason I’ve mentioned--they screen out the weaker students from the very beginning. That’s not what this simulation found. In this study, retroactively

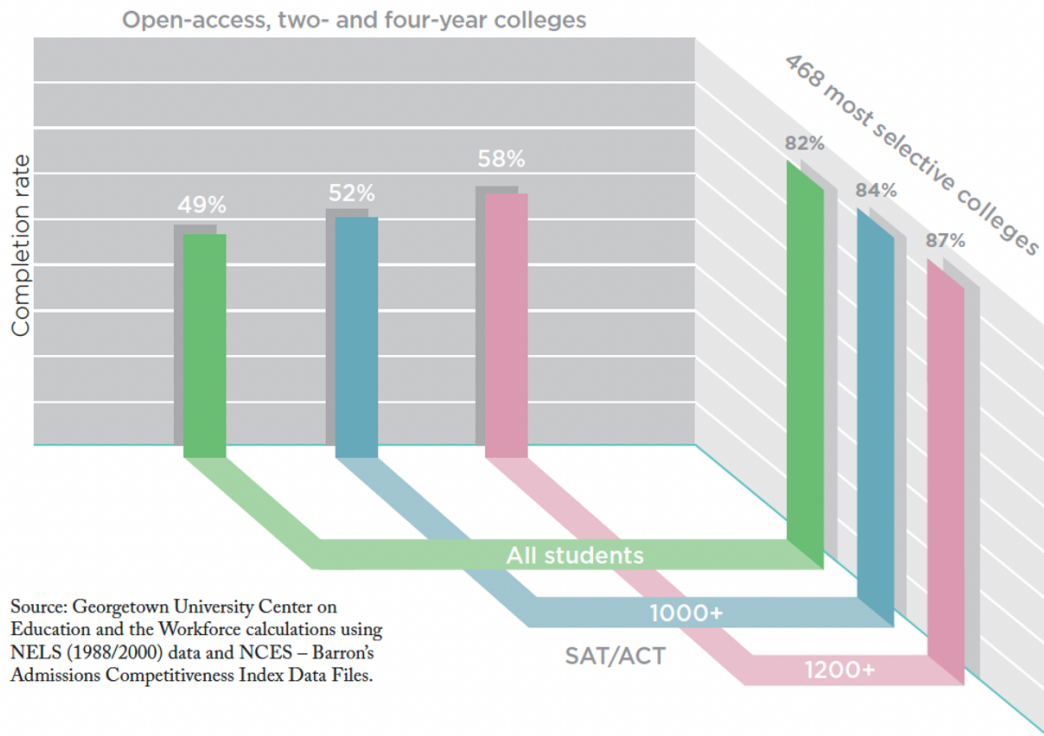
“rejecting” weaker students did not change the simulated graduation rates at the most selective universities. It produced “only a tiny gain in the overall graduation rate” for the next group of schools, and it increased the graduation rate by just six points at the least selective schools (Bowen, Chingos, and McPherson 2009, 197). In fact, the surprising “reality is that graduation rates vary dramatically across universities even when we look [only] at students with good high school grades and impressive test scores” (Bowen, Chingos, and McPherson 2009, 198). In short, graduation rates are higher at selective colleges not because they weeded out weaker students ahead of time, but because of something about the universities that affects students while they are there.

What is that something about the universities, if it’s not having a higher share of smarter students? When all is said and done, the key thing about the better universities is that they are *richer* universities. Here’s a figure from a landmark Georgetown study (“Separate and Unequal,” 2013).



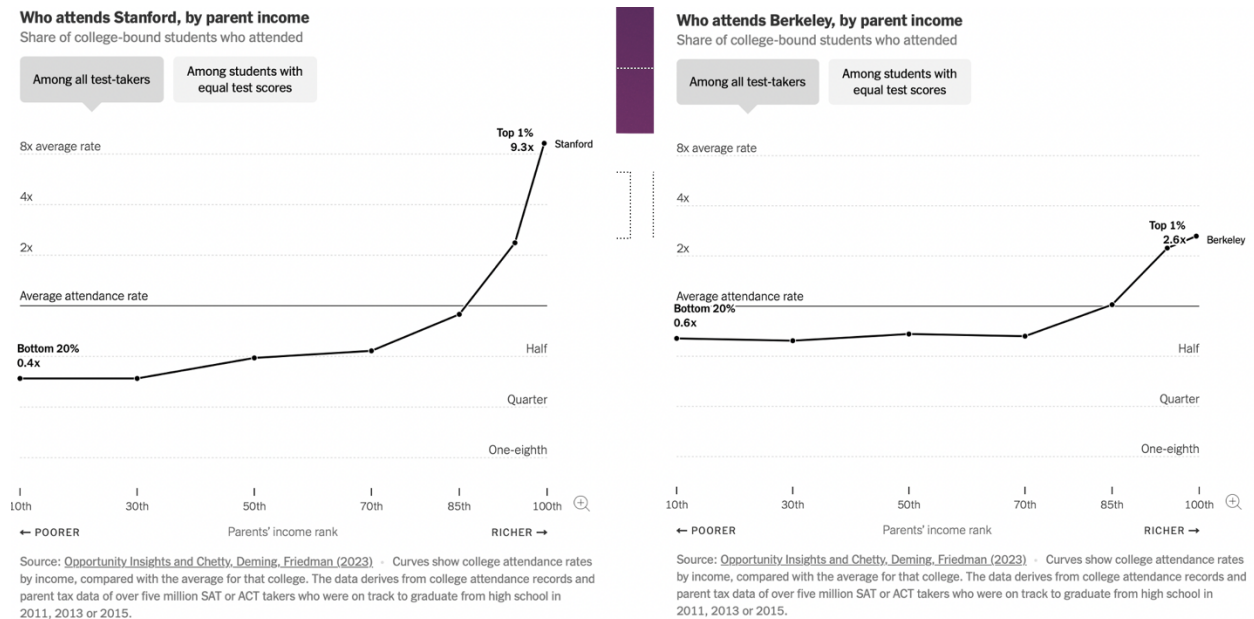
The simple pattern is that the more selective universities in the United States spend more money per student, often much more money. (These averages conceal the *really* big spending at places like Harvard and Stanford.) Wealthy universities spend more money per student on the range of things that multiple studies have found do help students persist in university, succeed at learning, and then graduate. These things include well-staffed living arrangements, individual advising, tutoring staff, psychological and peer counseling services, support

programs for students from underrepresented racial groups and from poorer or disfavored communities, and better-paid teaching assistants. Equally important is financial support that allows students to work less while in university and therefore study more. These student and academic services all cost money, which mostly elite universities have. Here’s another Georgetown slide, showing the correlation between high spending per student and high graduation rates.



To tighten up the *causal* link, my economist co-authors ran a regression to see whether the per-student spending level does a better job of predicting graduation rates than does the selectivity rate. Universities that look good on one of those measures usually look good on the other. We found that level of spending on instruction *does* appear to be a much more useful piece of information about universities if the student also has other information, such as average SAT scores and where the school is located. If you know these other things, selectivity doesn’t add any new valuable information, while knowing how much a university or college spends on instruction does. Students, we conclude, should always look at the money a university will spend actually on their education *rather than* on its selectivity rate and related rankings.

So in this case if we replace a bad metric—selectivity—with a better one—per-student instructional spending—we are asking students and the wider society to confront the political economy of higher education in all of its gross inequity. In the US at least, poor students go to poor universities and get the least help piled onto the lesser help they got from their poor high schools. The reverse is also the case: Stanford spends at least 4 times and perhaps up to 10 x as much as UC Berkeley per student on a similarly qualified student body, while also having much wealthier students.



I have found academic managers and politicians to be generally phobic about discussing the economics of higher education. They have spent the 2010s not having conversations about how underfunding holds back learning, social effects, and racial justice, among other things. Better metrics are a foot in the door to having real debates about the political economy that good learning and research need in order to have their full social effects. Better metrics could help get the US and UK discussions about higher ed funding back on track after years and years of negligence. They can help overcome the silencing of the discussion about economic injustice built into our educational systems. As things stand, they will also need to spark a mini-social movement for educational equality that expenditure data—not rankings—helps define.

A second indicator we critiqued is what the U.S. calls wages by major. We'd call it income by course here. The UK has the Longitudinal Educational Outcomes metric, which should be called Longitudinal Economic Outcomes where the latter is defined very narrowly by individual salary. Whatever we call it, wages by major is another junk metric. We found a range of data problems. There's the sampling problem posed by entry requirements to particular majors or courses. There's the sampling problem rooted in student employment preferences unrelated to education. For example, 18 year olds who are already intent on high future incomes skew towards courses like economics rather than history with reputations for yielding higher incomes, but such preferences and mental habits precede what they learn in those courses and determine what they get out of them. There is also a sampling problem with students choosing courses based on their self-perceived capacity to do maths, again predating arrival at university. There are measurement problems: bad data, missing data, and others. Wages by course is a technically crappy metric, and I want to focus on just one issue, which is the effect of individual *preferences* on wages.

The book's discussion is based on a study of a unique case, Norway, which has complete data on both the courses applicants want to study and the courses their scores require that they attend. Often there is a mismatch, and the study looked at academically nearly-identical students who were forced into different courses by tiny differences in scores. In a typical case in this study, one student who wanted to study social sciences was allowed to, and their academic twin, by getting one or two fewer points on a test somewhere, was required to study the humanities or the natural sciences instead.

Skipping to the conclusion, it turns out that a student who wanted to study social sciences but was forced into the humanities was making \$18,700 less, eight years later, than the social sciences aspirant who did indeed study the social sciences. But before you conclude that the humanities-related occupations just pay less, note they found that the reverse is also true: a student who wanted to study the humanities but was forced into the social sciences made \$21,400 less than the student who could stay in their desired course, the humanities. Similarly, the humanities aspirant forced into the sciences made \$5,000 less as a scientist than their counterpart who stayed in the humanities. In other words, the wage isn't attached strictly to the field (with some qualifications); it's attached to *one's*

power to satisfy an intellectual preference. I think this should blow our minds. Let me quote our understated conclusion:

It appears that the wages of humanities, social science, and science majors importantly depend on students' preferences. Whichever field the student prefers, completing that major will lead them to higher wages. . . . students' preferences align with their strengths, and following their preferences can improve their future earnings. (MM 110)

I would word this slightly differently: student strengths are *inseparable* from their personal preferences, and their future wages, as actual individuals, will track their personal strengths, preferences, desires, passions etc. *more than* it tracks average salaries in a given field.

We already know that the best non-economic or non-pecuniary outcome comes from satisfying one's preferences or desires: we all want the pleasure in the everyday work that goes into one's countless hours on the job, so we know personal job satisfaction is important to having a good life. Love numbers and love accounting precision to be happy as an accountant. But there's a big policy implication in the fact that the chapter shows the same to be true for *pecuniary* outcomes. The best way to get the higher pay you want is for you to follow your preferences—"show me the money" means "pursue my passion," to use the dreaded word. To get a good wage, even a wage-fixated student has to figure out what their true desire is, and then pursue it systematically. And, to repeat, the pursuit of one's desire must be prior to responses to indicators like the average wage of accounting graduates because the extent to which a statistical indicator applies to you depends on how well *your desire* fits its vocation. To flip this around: The best way for a society to *lower the* wages of students is to bully them into supposedly high-value courses that they don't really like.

I offer a couple of interim conclusions from an agenda of metrics *reform*:

1. We can legitimately and indeed are compelled to ask the public and policymakers to understand the political economy, that is, the funding of higher education *rather than* missing this big picture by focusing on selectivity and other rankings that are funding's downstream effect.
2. We must help students, their families, the wider public identify their personal intellectual--existential interests and goals in attending university *before using metrics*. They must *first* have subjective autonomy in analyzing their own lives, and then use metrics in the spirit of a tool—in

the spirit of subordinating the metrical tool to what we might call the narrative of their existence.

2.

That's the *reform* agenda. It's in the mode of *clawback*. The idea is that we lost some qualitative insights to quantitative summary indicators, so we need to claw back the qualitative insights. This is a real practical gain: I do think students reading our book will make demystified and stronger decisions about which university to attend and why they are going. But reform doesn't challenge the underlying rules of metrics as a mode of performance management.

The rules of performance measurement have been set up over the past four decades. Most analysts who are willing to use the term *neoliberalism* agree that metrics are a core feature of the audit practices through which neoliberalism changes institutions. The rules of audit were explained clearly by Michael Power in *The Audit Explosion* (1994) and *The Audit Society* (1997), by Marilyn Strathern and colleagues in *Audit Culture* (2000), by scholars like Cris Shore and Susan Wright starting a quarter-century ago. Michael Power prophetically identified the secret power of metrics-based audit: the persisting gap between audit's metric and its object doesn't discredit audit but mandates *more*. To take our last example, showing the gap between the indicator of a median wage by course and a given student's future wage as rooted in their interests—showing this failure of the metric *produces more metrics*. This is exactly what our book has done: *these seven metrics don't work, so here are seven more!* Reform is in the mode of clawback. It's not in the mode of rebuilding.

What about something more like metrics revolution, or metrics abolition, or at least epistemic equality in which metrics are embedded entirely in social and personal discourses about educational goals? It sounds like a lot of work. But there's a simple reason why we have to do this work. *We can't even do the reforms I just mentioned if we don't abolish the ground rules that prevent linking metrics to their institutions, people, groups, systems, and decisions.* The function of the selectivity metric is to screen out the political economy of student instruction. The function of wages by major is to screen out the powers of desires

and preferences. It would amount to a metrics *revolution* to embed numerical information in its socio-cultural and psychological contexts.

To talk about this admittedly very quiet revolution, I’m going to switch to the more philosophical of the two books, *The Limits of the Numerical*. That is another collection of authors, all of whom favor having metrics but, once again, only in empowered interpretative contexts. The *Limits of the Numerical’s* final chapter, written by the economist Aashish Mehta and me, critiques the severing of the quantitative from the qualitative aims of education. The current situation—generally called neoliberalism—is rooted in a great conceptual mistake. So the metrics revolution that would reattach, resubordinate, indicators to their social contexts means the undoing of a somewhat arbitrary error.

Universities have a range of effects; universities also integrate disparate effects in ways that isolated training programs do not. The problem, we argue, has been that the field that dominates education policy, economics, has marginalized the effects of education--that it *agrees exist*--because it cannot *quantify* them. Economists have established four quadrants of effects of higher education, but more because of the rules of quantification than because of economists’ beliefs about education, the policy world sees only one. Higher education’s effects have come to look like this.

Higher Education Effects: Quantifiable

Table 10.1: An Economist’s Taxonomy of the Benefits of Higher Education, with examples

	Private Benefits	External Benefits
Pecuniary Benefits	Increased personal productivity Higher expected wages More stable employability Higher-earning partners	
Non-Pecuniary Benefits		

To repeat, economics acknowledges the three other categories of effects, the nonmonetary or non-pecuniary and the social effects of education. But the great

difficulty of quantifying them has meant that economics leaves nonpecuniary benefits to one side.

There's another problem too. The discipline of economics does not offer self-reflexivity about its methodological limits out loud. It doesn't say, "we know nonpecuniary effects are real and important, but we leave them out because our disciplinary methods cannot estimate them." Nor do they add, "We welcome help from complementary disciplines like philosophy who do have descriptions of these other effects." The public doesn't hear this message about economics' limits, and it doesn't hear it from economists themselves, the recognized experts on their own field. So the public gets no education at all on methods and their boundaries, and the need for qualitative-quantitative symbiosis.

Perhaps the most important result is that the public also hears nothing about nonpecuniary benefits. Our pilot study of the *Economics of Education Review* found that of the 62 articles we coded, nonpecuniary benefits were treated in exactly *one* of them.

Here's the table from the book reflecting what both fields *do* accept as real effects of higher ed.

Higher Education Effects: Combined

Table 10.1: An Economist's Taxonomy of the Benefits of Higher Education, with examples

	Private Benefits	External Benefits
Pecuniary Benefits	Increased personal productivity Higher expected wages More stable employability Higher-earning partners	Increased productivity of/ with /for co-workers Increased earnings of employers Higher national income? ~~~~~
Non-Pecuniary Benefits	Increased enjoyment of culture Better emotional relationships	More vibrant democracy Greater environmental awareness

This is a big improvement over the previous figure. But it isn't in circulation, as it's promoted neither by economists of education nor by humanities scholars nor by university officials.

Meanwhile, inside their disciplines, humanities scholars, particularly philosophers, have been evolving the nonpecuniary and social effects of education for well over two thousand years. Our chapter text offers a compressed inventory of the last two centuries. A few highlights: The university does two things at once, which is to devote itself entirely to the progress of knowledge while also focusing on a related thing, the development of the self as a bearer of knowledge who must also learn how to create knowledge. This applies to students but also to teachers and researchers. The university must provide the environment which makes this production of knowledge and subjects of knowledge possible. That involves at least two features: complete intellectual freedom to pursue the truth wherever it leads; and freedom from coercion in the form of financial need. The U.S. discourse around “free college” reflects an understanding of the university environment that was developed in 1790s Prussia, not 1960s Berkeley. The result of the proper rigor in aims, structures, and practices is the creation of knowing subjects of a certain kind, who grasp that learning is (1) a *process* that is (2) unending (mastery will never take the form of a final understanding), and (3) active, reflexive, personal engagement that is inseparable from self-creation.

And then this is evolved over the course of the 19th and 20th and 21st centuries in various stages. The land-grant movement, marred as a settler-colonial “land grab” movement, expanded in the name of equal educational access across disciplines and all social groups. (This was implicit philosophically in Fichte). Settler control of this model had to be removed, through long argument and practical movements. In the United States this took one hundred years, measured from the end of the Civil War in 1865 to the early federal educational civil rights protections of 1965. It was accelerated by Black activism in particular, which advocated both full access as a matter of civil rights and social justice and also for diversification of subjects and epistemologies-- Black Studies and feminist standpoint theory were crucially important.

A further stage we discuss is the use of critical theory to link the creation of knowledge to the formation of identity. This puts what is wrongly dismissed as identity politics at the core of *all* of the disciplines, which need finally to take account of each other’s methods and self-understandings. For example, Judith Butler uses Foucault to define critique as *both* refusing subordination to an established authority and the obligation to produce a self (232). Critique entails

the continuous self-reflection about the categories and processes of knowledge production (233). Knowledge creation means “we have to keep deciding”—about relations between quant and qual, about the social effects of our research, about how we are going to talk about the limits of our methodologies.

All this explodes the table.

Higher Education Effects: Full Humanities

Table 1: An Economist's Taxonomy of the Benefits of Higher Education, with examples

	Private Benefits	External Benefits
Pecuniary Benefits	Increased personal productivity Higher expected wages More stable employability Higher-earning partners	Increased productivity of/ with /for co-workers Increased earnings of employers Higher national income?
Non-Pecuniary Benefits	Increased enjoyment of culture Better emotional relationships Intellectual freedom Freedom from economic coercion Understanding learning as (1) a process that is (2) unending and (3) an active personal engagement inseparable from self-creation (Bildung) Critique seen as obligation to continuous methods self-reflection AND to produce a self capable of this (refusing subordination to authority)	More vibrant democracy Greater environmental awareness Public comfort with other peoples' intellectual freedom Public benefits of general expectation of freedom from economic coercion Public benefits of learning as described at left Equitable uses of learning across diverse social groups Democratization of the intellect leading to pluralized epistemologies and equitable relations

Now we're really getting somewhere with the creation of knowledge and the effects of the university. Our chapter shows that in fact we—any of us working together-- *can* overcome this division between the cultures. It expresses urgency about actually doing this.

Let me conclude. It's no secret that universities in the Anglophone countries are stuck in the mud. Neither the wider public nor policymakers think of them warmly as nurturers of personal dreams, or as essential to a transformed future, or even as reliable engines of economic growth. (The more society talked about higher ed commercialization, the more stagnation it got in productivity growth.) But I ask you to consider whether a metrics revolution offers a road out. (I set aside for the moment the question of metrics abolition as something I need to conceptualize better than I have to this point.)

A metrics revolution would ask the public to understand the basis of university rankings in the political economy of higher education as a public system. It would ask students of all ages to formulate their intellectual and existential goals prior to using metrics as an evaluative tool.

It would define higher education as *primarily* a set of *non-pecuniary* individual and social effects. Students and instructors have the obligation to produce an autonomous knowing subject and the cognitive capabilities that allow that subject to operate without subordination to an established authority. This autonomous subject is profoundly social, so requires two historic features of universities: face-to-face interactions and the integration of knowledge across disciplines.

My view is that this university, which openly aims at forming qualitative individual and social capabilities, would be much more *popular* than the wage-focused university defined by floating metrics. But we can get this full university only if we're willing to re-subjugate metrics within a regime of diverse knowledges interacting on the basis of epistemic equality. And *that* will happen, in turn, only if professionals are willing to recapture the ground rules by which metrics are applied. This recapture—I've called it a metrics revolution--will be viewed by managers as a kind of class struggle against them, which it indeed would be. All I can say in advance, on the basis of what I've said today, is that it will be worth it.