Prestige and inequality in the ivory tower

It's no secret that we have an inequality problem within the hallowed walls of the academy. Much focus has been dedicated to problems of inequality—of status, of wage, of job security, of resulting social mobility, and beyond—mainly between tenured faculty and the growing precariat of contract teaching labour. The central importance of published research is often fingered as a central culprit in entrenching this inequality, and in this post I'll explore the inequality of citations via a citation distribution analysis. The analytical approach is borrowed from Thomas Piketty's *Capital in the Twenty-First Century*, his landmark work on economic inequality.

Inequality in academia

Why focus on inequality in academic publishing? Much like in economic inequality, early leaders are given the opportunities they need to continue advancing their publishing and networking, further extending their lead. Meanwhile, contract faculty live precarious lives and are expected to put in huge time commitments to keep up their research and publication output, without any direct remuneration for this work and with diminishing hopes of making the jump to secure employment. Performance as measured by publications and citations has a huge influence on professional and personal outcomes, and inequalities in this system tend to entrench themselves. So how intense is the inequality of citations? We often repeat in the bibliometrics community that we know citation distributions to be skewed. Accordingly, we warn about the pitfalls of using citation measures based on averages, which level off the peaks and valleys of underlying distributions. And yet average-based measures continue to be the indicators most often used. A major reason for this situation being perpetuated is likely that alternative approaches developed to date have yet to strike the critical—but extremely difficult—balance between methodological robustness on the one hand and the intuitive clarity needed to apply this metric in decision-making on the other.

Piketty's work on economic inequality

There have been many criticisms of Thomas Piketty's *Capital in the Twenty-First Century* (which should come as no surprise, given the attention it received in the popular press), but I have yet to read any criticism that calls out his work for being inaccessible to a non-academic audience. His analytical approach must tangle with a similar problem: averages of income and wealth pave over the intense inequalities in the underlying data. To overcome this problem, Piketty uses a quantile-based approach, dividing society into its top 10% (the Upper Class), the middle 40% (the Middle Class), and the bottom 50% (the Lower Class). The Upper Class gets further sub-divided to communicate how steep the inequality is at the top end of the income distribution, dissecting it into the top 0.1% (the Superlative Class), the next 0.9% (the Dominant Class) and the next 9% (the Well-to-do Class).

One of the most interesting revelations in the book is that until the 20th century, there was little difference in the economic situation of the bottom 40% and the middle 50%. That is to say, the proverbial middle class is truly a creation of the 20th century. A big question now is whether the 21st century will see this invention destroyed, whether inequality will once again rise to the point that any meaningful difference between the lower and middle classes is dissolved. But I digress. What, in *Capital*, did Piketty find using this approach? Looking at 2010 income data for the United States, the average salary was about $67,000 (see Table A-2). Piketty showed that the Upper Class earned half of all the income in the country, the Middle Class earned 30% of all US income, and the Lower Class earned 20% of all US income. Further dissecting the top end of that distribution, we see that the top 0.1% of earners accounted for 10% of all US income—nearly one hundred times the average income—while the next 0.9% raked in 10% as well, and the next 9% earned 30%.

Class	US economy 2010		
	Share of all income	Divergence factor (from overall avg)	Avg income (in thousands)
Upper Class Top 10%	50%	5.0	$335.0
Superlative Class Top 0.1%	10%	95.0	$6,365.0
Dominant Class Next 0.9%	10%	11.1	$744.4
Well-to-do Class Next 9%	30%	3.3	$223.3
Middle Class Next 40%	30%	0.8	$50.3
Lower Class Bottom 50%	20%	0.4	$26.8

Source: Prepared by Science-Metrix, data from Piketty 2014

In terms of disparities here, the top 1% of earners captured the same collective income as the bottom 50%. That is to say, the top 1% earned 20% of all US income, just as the bottom 50% earned 20%, even though there were fifty times as many earners in the Lower Class than there were in the top 1%. An average member of the superlative 0.1% earned about $6.4 million in 2010—250 times the $26,800 earned by the average member of the Lower Class. Occupy Wall Street, anyone?

Citation inequality, using Piketty's approach

How do things stand in the academic realm, with respect to citation distributions? The present analysis looks at citation scores for individual papers (not authors, institutions or journals) and considers papers published worldwide in chemistry between 2005 and 2013. Obviously, looking at citation inequality amongst papers isn't quite the right level to be looking at, as it's the inequality between *people* that drives most of our worries. Looking at papers is simply more convenient given available data and can therefore provide an early indication of whether citation inequality is worth investigating further, knowing that a full-fledged study would take more resources. This analysis is therefore better filed under "cheap and cheerful" than under "conclusive."

The average paper in chemistry over this period received 18 citations. However, recycling Piketty's approach to dissecting these findings, we can explore the skewed distribution of the underlying data. The Upper Class once again received over 45% of the capital—in this case, the social and professional capital of citations rather than the financial capital of incomes. The Middle Class received about 45% as well, with just under 10% remaining for the Lower Class.

Dissecting the Upper Class, we see that the top 0.1% of papers captured 3% of all citations, while the next 0.9% captured 10% and the Well-to-do Class of papers captured about 34%.

Class	Chemistry papers 2005–2013		
	Share of all citations	Divergence factor (from overall avg)	Avg citations
Upper Class Top 10%	47%	4.7	84.5
Superlative Class Top 0.1%	3%	31.4	568.2
Dominant Class Next 0.9%	10%	10.6	192.4
Well-to-do Class Next 9%	34%	3.8	68.3
Middle Class Next 40%	44%	1.1	20.0
Lower Class Bottom 50%	9%	0.2	3.3

Source: Prepared by Science-Metrix, data from Scopus (Elsevier)

What does this mean in terms of number of citations? An Upper Class paper received on average about 85 citations, while a Middle Class paper received 20, and a Lower Class paper received about 3. The top end of the distribution is very skewed, though, as we are already aware. A Superlative paper received nearly 570 citations, on average, while a Dominant paper received nearly 200, and a Well-to-do paper received on average just under 70.

Comparing academic and economic inequality

How does the level of citation inequality for these chemistry papers fare relative to the level of income inequality for the US? In broad terms, the Upper Class as a whole fared marginally less well in collecting citations as they did in collecting income. The academic Middle Class was far better off than the economic Middle Class, doing nearly 50% better. The Lower Class, however, struggled much more in the academic sphere than in the general economy; the bottom 50% of papers accounted for only 10% of all citations, whereas the bottom 50% of earners accounted for 20% of all income. At the very top end, the Superlative Class was far less imposing in the academic sphere than it was in the economic sphere, while the Dominant Class fared approximately just as well, and the Well-to-do Class did slightly better in academia than in the open waters of the total economy.

The methodological approach here might have an important influence on assessments of the Lower Class. More than 10% of chemistry papers were [never](#) [cited](#), and almost another 10% received only one citation. These papers all fall within our Lower Class of the citation distribution analysis. On the assessment of incomes, however, do we count those who are so marginalized from the economy that they actually have no income (or practically none) to declare? My suspicion is that the data source used for economic analyses actually leaves these people out entirely, just as our measures of unemployment leave out those people who have become so disheartened that they give up even looking for work. These blind spots have an important effect on our ability to measure the bottom end of economic distributions. Are there analogous challenges at the very bottom end at the academic distribution?

These analyses suggest—to me—that there is an inequality problem in academia that warrants further investigation. Certainly, if the distribution of incomes is skewed enough to warrant shifting to an approach like Piketty's, then the distribution of citation outcomes warrants similar attention. Of course, calling a set of papers "Lower Class" politicizes things enormously, but given that these citation outcomes are important determinants of professional advancement, job security and economic well-being, it could be argued that the issue is deeply political already.

Addendum

This article was first published by the author at [ScienceMetrics](#) on the 18th December 2017.

References

Alvaredo, F., Chancel, L., Piketty, T., Saez, E., Zucman, G. (2017). World Inequality Report. World inequality lab. [Link](#).

Alvaredo, F., Chancel, L., Piketty, T., Saez, E., Zucman, G. (2017). Inequality is not inevitable – but the US 'experiment' is a recipe for divergence. The Guardian. [Link](#).

Alvarez, M. (2017). Academe's Prestige Problem–We're all complicit in perpetuating a rigged system. The Chronicle of Higher Education. [Link](#).

DeNavas-Walt, C., Proctor, B. D., Smith, J. C. (2011). Current Population Reports, P60-239, Income, Poverty, and Health Insurance Coverage in the United States: 2010, U.S. Government Printing Office, Washington, DC. [Link](#).

Campbell, D., Tippett, Ch., Côté, G., Roberge, G., Archambault, E. (2016). An approach for the condensed presentation of intuitive citation impact metrics which remain reliable with very few publications. 21st International Conference on Science and Technology Indicators - STI 2016. Book of Proceedings. [Link](#).

Lok, C. (2016). Science's 1%: How income inequality is getting worse in research. Nature News. [Link](#).

Marginson, S. (2016). Higher education and growing inequality. Academic Matters. [Link](#).

Neate, R. (2017). World's richest 0.1% have boosted their wealth by as much as poorest half. The Guardian. [Link](#).

Van Noorden, R. (2017). The science that's never been cited. Nature Feature. [Link](#).

Piketty, T., & Goldhammer, A. (2014). Capital in the twenty-first century. Cambridge Massachusetts: The Belknap Press of Harvard University Press.

Stephenson, G. C. (2016). How to improve conditions for the academic precariat. University World News. [Link](#).