an open access journal

RESEARCH ARTICLE

Longitudinal variation in national research publication portfolios: Steps required to index balance and evenness

Jonathan Adams[1,2], Gordon Rogers[2], Warren Smart[3], and Martin Szomszor[2]

[1]The Policy Institute, King's College London, 22 Kingsway, London WC2B 6LE, UK
[2]Institute for Scientific Information, Clarivate Analytics, 160 Blackfriars Road, London SE1 8EZ, UK
[3]Tertiary Sector Performance Analysis Group, Ministry of Education, 33 Bowen St., Wellington 6140, New Zealand

Keywords: evenness, international collaboration, national comparisons, research balance, research diversity, research policy, Web of Science

ABSTRACT

National research diversity is explored through the balance of global and national papers in journal categories in the Web of Science (WoS) and Essential Science Indicators (ESI) and we examine the consequences of "normalizing" national publication counts against global baselines. Global balance across subject categories became more even as annual WoS indexing grew fourfold between 1981 and 2018, with a relative shift from biomedicine towards environment and technology. Change at the country level may have tracked this or been influenced by local policy and funding. We discuss choice of methods and indices for analysis: WoS categories provide better granularity than ESI; Lorenz curves are explored but found limiting; the Pratt index, Gini coefficient, and Shannon diversity are compared. At the national level, balance generally increases and is greatest in non-Anglophone countries, perhaps due to shifts in language and journal use. Two aspects of national change are revealed: the balance of actual WoS paper counts and the balance of counts normalized against world baseline. The broad patterns for these analyses are similar, but normalized data indicate relatively greater evenness. National patterns link to research capacity and regional networking opportunities, while international collaboration may blend national differences. A data set is provided for analytical use.

Citation: Adams, J., Rogers, G., Smart, W., & Szomszor, M. (2020). Longitudinal variation in national research publication portfolios: Steps required to index balance and evenness. *Quantitative Science Studies*, *1*(3), 1182–1202. https://doi.org/10.1162/qss_a_00073

DOI: https://doi.org/10.1162/qss_a_00073

Received: 16 October 2019
Accepted: 8 April 2020

Corresponding Author:
Jonathan Adams
jonathan.adams@kcl.ac.uk

Handling Editor:
Ludo Waltman

The MIT Press

1. INTRODUCTION

We explore "balance," a central aspect of disciplinary diversity, as represented by the categorical distribution of publications in journals indexed in the Web of Science (WoS). In this paper, we consider global context and national trends to establish baselines and we provide the data to enable others to use this indexing. Our longer-term purpose is to use these baselines to investigate whether national research policies, particularly national research assessment systems, are associated with change in research performance and diversity.

To avoid confusion about this objective, we need to distinguish the idea of balance (the spectrum between specialization and diversity) from the identification of specific strengths. For example, there have been prior studies of research strengths in terms of "comparative advantage" at regional level (Radosevic & Yoruk, 2014) and for specific city states (Horta, 2018). However, the limitations of such an approach in scientometrics (as opposed to econometrics), particularly regarding the use of the Balassa index (Balassa, 1979), have been outlined by Mansourzadeh, Shahmoradi, et al. (2019).

In this paper we are not identifying key specializations, although that could sensibly follow, but whether or not countries have tended to become more or less specialized over the last four decades. In our background work we plan to track shorter-term variances from secular trends, focusing particularly on three systems: the Australian Excellence in Research for Australia (ERA); the New Zealand Performance-Based Research Fund (PBRF: Adams, 2008); and the UK Research Assessment Exercise (RAE; in 2014 this became the Research Excellence Framework or REF). It is unquestionable that changes in the national research base are associated with these exercises (Adams, 2002, 2007; Adams, Cook, et al., 2000; Adams, Mount, & Smith, 2010; Wilsdon, Allen, et al., 2015). However, there has been no prior description of changing global and national research diversity, so in this report we develop a global baseline and examine the normalization of national profiles against that baseline to create a global context and background for subsequent work.

Diversity—a spread of resources or activity—is generally inferred to be a good thing, providing benefits through increased resilience to disruption, spreading and mitigating risks, and enhanced responsiveness to change (Page, 2011), but it can also result in underinvested priorities and a lack of focus. Its meaning, measurement, and role have been discussed in many disciplinary contexts, with notable developments among economists and ecologists. The measurement of the spread and diversity of wealth dates back over a century (Gini, 1909; Lorenz, 1905; Pareto, 1895), with emphasis on income inequalities. The analysis of the frequency and diversity of biological species has also had a long history (e.g., Fisher, Corbet, & Williams, 1943; Hill, 1973) that is now of significance to research on climate change (e.g., Seddon, Mace, et al., 2016). The approach we use is therefore not original but stands on this prior intellectual platform.

Stirling (2007) provides an excellent review of the background and arguments about diversity's value. As Stirling notes, diversity is a property of every system. Systems are not either diverse or not diverse: They are more or less diverse. A question that follows is how this might be measured, but a measure has little meaning without context. It is more informative to measure and compare cognate systems, where other relevant properties may also be apparent, and thus to inform policy and management discussions. To address this, Stirling draws a common framework with three basic—albeit interdependent—properties: **variety** (the number of categories into which the system is apportioned), **disparity** (the difference or distance between the categories), and **balance** (the pattern of apportionment across categories).

Diversity in the context of research could be considered in terms of input (using research grants), structure (using discipline-labeled units and departments), or output (using research publications). The most comprehensive and widely used information in research policy and management analyses is about publication outputs, particularly the outputs indexed by the Institute for Scientific Information (ISI) in what is now WoS. Leydesdorff, Wagner, and Bornmann (2019) have recently reviewed definitions of the components of diversity for such data (in the context of interdisciplinarity) and presented an innovative indicator (DIV) that operationalizes variety, disparity, and balance independently and then combines them. Van Dam (2019) has a related approach that accounts for the nature of disparities. However, we will focus solely on balance across the entire database because variety (the number and definition of categories) and disparity (the assignment of journals to one or more of those categories) are properties constrained in a managed publication database (see Section 2.1).

Abundance affects diversity and the research output of countries and institutions has risen over time. Matia, Amaral, et al. (2005) observed that the mean annual growth rate of publications is independent of the publication output of units at all levels of aggregation from countries to institutions to authors. Growth may occur without changing local balances of activity, or it could drive or be driven by diversification, or other system effects may be influenced by the emergence of new research institutions and exceptional national growth patterns, such as that in Asia-Pacific (Adams, 2013).

The size, and hence relative resource capacity, of an entity can influence diversity. Debates in the UK from the 1980s onwards about "selectivity and concentration" were underpinned by an assumption that large size (of research institutions and groups) conferred benefits not accessible to smaller units. Analysis showed "critical mass" ideas to be ill-founded (Adams et al., 2000) but some benefits may come not from size *per se* but from the colocation of multiple small groups that could interact synergistically (Johnston, 1994). Larger institutions can host more and larger disciplinary departments that could in turn support such diversity. Larger economies (or systems such as the European Research Area) can support a wider range of research and internationally competitive facilities.

The global research system does not exist in isolation. Historical, political, and cultural factors influence research choices at national and institutional level (Moed, 2006). For example, the "rise of China" is also a narrative about an evolving national portfolio, historically rooted in heavy industry and now expanding into a wider range of science and technology and onwards into social sciences (Bound, Saunders, et al., 2013). Everywhere, the balance of national activity is influenced by funding regimes that are in turn affected by a multiplicity of policy decisions. Those policy decisions are interdependent, influenced by competitiveness and emulation that is enabled more readily today by the internet than it was by mail and telephone in the 1980s.

Research collaboration has grown enormously since 1981. More papers have multiple authors (King, 2012), the average number of authors continues to grow (Castelvecchi, 2015) and, for mature research economies, more than half of national output is coauthored with one or more international partners (Adams, 2012, 2013; Adams, Rogers, & Szomszor, 2019). The usefulness of the question "What is the research performance of the UK?" is debatable where only 40% of its papers are unique to that country. The outcome is not unidirectional: National and institutional diversity may increase because collaboration may facilitate activity constrained for an entity acting alone; but coauthorship means that portfolio differences are reduced. Through a parallel route, global diversity could be reduced by mutual convergence on priorities shared between national portfolios, which may not be wholly beneficial to human endeavor.

National research assessment policy interacts directly with management decisions (Sivertsen, 2017). For example, the large and well-funded US research base has no common institutional assessment program, though it does have a uniform national assessment process at project grant level through the National Science Foundation (NSF) and National Institutes for Health (NIH). By contrast, the UK has had cyclical institutional research assessment since 1986 that not only publishes graded outcomes but uses those grades to allocate research funding. Related systems have been introduced in Australia and New Zealand. Assessment has appeared elsewhere but tended to be more formative then summative (Sivertsen, 2017).

This study draws on a global database of research publications. To explore any policy-related questions an analyst needs to know how the global system as a whole has evolved, what the balance has been between research areas, and how the global pattern is reflected in difference and similarity at national levels. In this paper, we first quantify secular trends in the overall global background in research balance; second, we describe and compare broad trends at national level and comment on national variance.

2. METHODS

2.1. Data Source

This analysis uses the WoS databases. The data extract was limited to the core collection of the three principal WoS citation indices: Science, Social Science, and Arts & Humanities. Emerging Sources were not included.

As noted earlier, diversity can be seen in terms of variety (the number of categories), disparity (the distance between categories), and balance (the spread of activity across categories) (Stirling, 2007). In WoS, the *variety* of categories is driven by decisions about the assignment of journals, which were originally grouped by the degree to which they cross-reference one another and thereby represent cognate communities. *Disparity* is unclear and neither uniform nor constant: Some categories are rather similar, even overlapping; others are quite separate. Furthermore, the constitution of categories changes over time: Journals evolve content in response to their field; titles are added or deleted as their impact waxes and wanes; new topics evolve. Variety and disparity may therefore change over time, but they do so in a global way and these changes may be considered as common, background variance throughout the database.

An important caveat here is that the impact of a global change may fall unevenly at national level: We have not yet identified a process to track and analyze such variance.

Balance may be affected by the number of items annually indexed. The global balance of papers assigned (via journal assignment) to each specific WoS subject category provides a benchmark or reference point to which more specific data sets can be compared. Note that an index based on raw national counts could differ from one based on national- or institutional-level data contextualized annually against the global baseline. In other words, if there is managed, structural disparity between categories and their growth across years, is the national share of papers in category A more or less than its share of papers in the database as a whole? Both index options must therefore be considered.

Two categorical systems are available for the data indexed in WoS. In both cases, each category contains all documents from a stated set of journals. Essential Science Indicators (ESI) has one multidisciplinary and 21 subject-based categories. It does not cover the humanities and the visual and performing arts, where journal papers are at best a patchy guide to significant research. ESI is an application with a focus on the most highly cited part of the global literature and has high analytical value when considering national scientific portfolios. However, the coarse granularity (of, for example, clinical medicine, engineering, and physics) may obscure the evolution of and shifts in research activity at a more detailed disciplinary level. ESI data were explored but not subsequently used.

WoS journal categories cover 254 research disciplines (at October 2019) across all subject areas. A criterion for inclusion is that the journal should publish English titles and abstracts. Because the language constraint means that the necessarily more parochial research areas of the arts, humanities, and much of social science are not effectively covered for an analysis that requires global comparability, the WoS category set was reduced to 194 categories where there is more general international journal coverage and usage, largely around science and technology (Appendix 1).

The document types used for the analysis are substantive, original academic research contributions in the form of articles and reviews. These are referred to as *papers*.

No fractional attribution is applied. A paper is assigned to a country's tally if there is at least one author address for that country on the paper.

A paper is assigned as a whole count to a category tally if the journal was assigned to that category. No fractional attribution for journals with multiple assignments was applied.

Global and national counts of papers were converted to proportions of the respective annual total. Overall growth in volume means that it is essential to use relative numbers (proportions) rather than absolute numbers when considering the evenness of paper counts across categories.

2.2. Country Coverage

Our background work has focused on three countries: Australia, New Zealand, and the United Kingdom. However, to provide a reference set for global background analysis and for the use of others, nine other countries of varying size and research capacity were added for this analysis: Canada, China, Denmark, France, Germany, Japan, Norway, Switzerland, and the United States.

2.3. Time Period

Detailed publication records in WoS are available in a common format and at a standard suitable for analysis from 1981. Earlier data are available, but not all serials are complete and records benefit less comprehensively from good curation. The study period therefore covers the 38 years from 1981–2018.

2.4. Indexing Publication Balance

The specific numerical value of an index of publication balance (or diversity) for a country or institution carries no practical information. Useful management policy information comes from comparisons, across time or between countries, about the trajectory of balance, diversity, or its relative value.

The disparity between categories may vary, possibly on an annual basis, as journals are added to and deleted from the database but new additions are also backfilled. Variation in disparity would need to be evaluated at an article level, because the contents of the journals also vary over time. For the purposes of this exercise the categories will be taken to have constant disparity, which is likely to be broadly correct across a very large system of 15,000+ journals.

Rousseau, van Hecke, et al. (1999; see also Nijssen, Rousseau, & van Hecke, 1998) have argued that a Lorenz curve is a strong representation of evenness. We examine this in our analysi but will show that, while academically valuable and informative, this approach has limited practical comparative value for multiple, complex sets (here covering 38 years and 12 countries). To make sense of real-world data, we have to reduce our description to an index value because a series of visual profiles for each year for each entity would have constrained a publishable analysis to a small set of comparisons.

Our choice of a preferred index for evenness was guided by Moed (2006). Moed, referring to the methods set out in Egghe and Rousseau (1990), made use of Pratt's Index (Egghe, 1987; Pratt, 1977) in comparing the concentration of published articles among universities (Moed, 2006, Table 8). The Pratt Index is in fact a later variant of the widely used Gini coefficient and has a similar arithmetic value (Carpenter, 1979). Both are indices of disciplinary specialization, which indicate the extent to which a publication portfolio (of a country or an institution) is highly specialized (high index values indicating specialism and a lack of balance and evenness) or evenly spread across categories (low index values indicating an even distribution).

We compared the outcomes of initial analyses using the Pratt Index, the Gini coefficient, (Gini, 1909), the Shannon entropy index (Spellerberg & Fedor, 2003) and Simpson's diversity index (Simpson, 1949). Hill (1973) points out that Shannon is intermediate between Simpson's index and a count of varieties and the theoretical background to the relationship between these is addressed elsewhere (Nijssen et al., 1998; Keylock, 2005). Our interest was purely empirical and illustrative.

2.5. Normalization of Balance Against a Global Baseline

WoS categorical publication counts are innately uneven (discipline categories naturally vary in size), so it is informative to compare both absolute counts for any entity and the balance of those counts when "normalized" against the global baseline. We suggest that the research diversity of any country thus has two components carrying useful information:

- research choices in line with the consensus of the global research base, computed from the proportion of *original publication counts* across categories; and
- research choices specific to that country, dictated by resource and policy factors, computed *after normalization* of original publication proportions by comparison with global proportions.

We consider balance, as a key aspect of diversity, through both analytical paths.

To compute the global balance of papers, the total annual numbers of papers (articles and reviews) are counted by category and by year. We then examined the annual spread of papers at the category level and aggregated these into broad macrodisciplinary groups to create an overview of changing subject proportions. We also calculated a set of Lorenz curves (Rousseau et al., 1999) to track the annual variation in database balance across categories at a finer-grained level.

For index-based analysis, the category-specific proportion of the total annual count of papers is calculated. The rank (within year) of the relative volume in each category is also calculated. A global publication activity index (PAI) is derived by Moed from the sum of the product of rank and proportion using a formula based on Bookstein and Yitzhaki (1999, Eq. 3). The relationship between the publication balance for an individual country (or institution or other entity) and the global baseline is computed by calculating the ratio between the national and global papers in each category and dividing this by the ratio of total national and global papers for the same year. The values are then reduced to proportions of the sum of ratios (i.e., so the annual total sums to 1.0).

For country i among $1 - m$ countries and publications in category c of $1 - n$ categories, denote the number of items in the cth category by x. The relative number of items C_i in the cth category is calculated as

$$C_i = x_c / \left(\sum_{k=1}^{n} x_k \right).$$

To define the Pratt Index we assume that the x_c (and therefore also the C_i) are ordered decreasingly. Then

$$q = \sum_{i=1}^{n} i\, C_i.$$

And the Pratt Index, P, is defined as

$$P = 2 \left(\frac{N+1}{2} - q \right) / (N-1).$$

From this we can easily calculate the Gini coefficient (G):

$$G = C \times \frac{(N-1)}{N}.$$

For further background, the reader is referred to Egghe and Rousseau (1990, especially Section 2.5).

The total number of papers are dispersed across categories according to the assignment of journals and the numbers of papers published in each journal. We do not, consequently, expect that there will be an "even" spread of papers. This question is whether a particular sample has a distribution that is more or less concentrated than the global "null model." Moed (2006) describes this as the disciplinary PAI. For both the ESI and WoS calculations, the country-specific data are therefore compared with the global data as a reference baseline.

The relationship between the publication balance for an individual country (or institution or other entity) and the global baseline is computed by calculating the ratio between the national and global papers in each category and dividing this by the ratio of total national and global papers for the same year. The values are then reduced to proportions of the sum of ratios (i.e., so the annual total sums to 1.0). The relevant index is then calculated as before.

2.6. Evenness and Granularity

Having determined a preferred indicator and established global reference sets, we examined the relative merits of more or less granular data categories. The choice is between the 194 WoS categories and the 21 nonmultidisciplinary ESI categories.

For Australia, New Zealand, and the United Kingdom, we calculated the category-specific values by normalizing the country data against the global data by category and year. This was done for both the ESI data set and the WoS data set for the period 1981–2018. We found that the coarser-grained ESI data suggest a greater level of evenness than the WoS data (see Section 3). This was unsurprising but required confirmation. The broader ESI categories absorb differences in absolute volume and relative change across subsumed specialist categories and this obscures real change in emergent fields. All subsequent analysis therefore made use of the WoS data to expose more information.

3. RESULTS

3.1. Evolution of the Web of Science

In 1981 there were 614,608 articles and reviews (academic research papers) in WoS. This annual count of papers rises to exceed 1 million for 1994 and 2 million for 2011, and reaches 2,731,368 for 2018: a more than fourfold expansion.

The global system is an unpredictable aggregation of national changes, but Figure 2 shows that a broadly similar balance between the proportions of papers by category in WoS has been maintained across the 38-year span from 1981 to 2018. The figure uses, for illustrative purposes, macrodisciplinary subject categories originally developed for UK comparative national research performance indicators. This schema includes Biological sciences (covering both Organismal and Molecular) and Physical sciences (both Chemistry and Physics).

While WoS is intended broadly to represent the balance of leading global research publications, the indexed database may or may not mirror specific growth and change in the real world. In fact, the database exhibits relative expansion in Engineering & Technology, which increased in share from 8.4% in 1981 to 17.1% of the database in 2018 (the biggest relative growth is in Nanosciences). Environmental (5.1 to 8.6%) and Mathematical sciences (3.2 to 4.1%) both also expanded. There

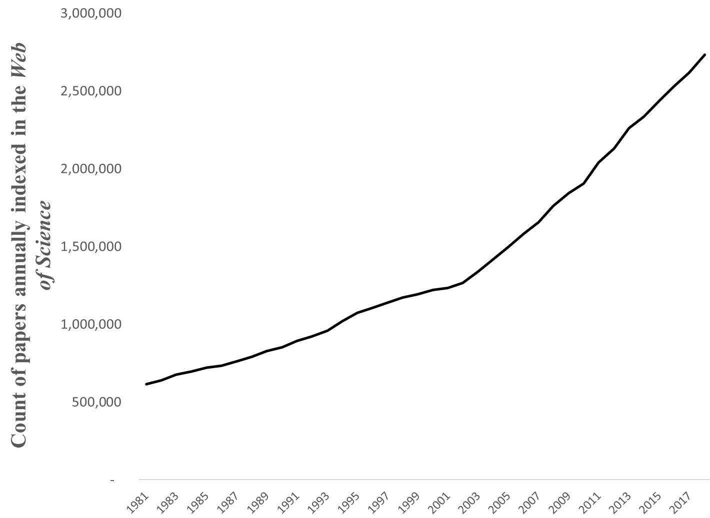

Figure 1. The number of articles and reviews (academic research papers) annually indexed in the Web of Science across the period 1981 to 2018.

was a relative diminution in Clinical sciences (down from 22.3% to 18.3%), with the biggest relative fall in General & Internal Medicine. Biological sciences fell from 20.7% to 14.8% over the period.

The effect of these changes in the balance of publications across WoS categories can be visualized in a Lorenz curve, which reveals the inequality or unevenness of a distribution. The distribution approaches a continuous function where we have 194 categories ranked in ascending order of proportions within year. A series of annual Lorenz curves could show the progressive shift in unevenness for a data set, but in practice, for these WoS data, this becomes uninformative because the 38 annual curves closely overlie one another and become indistinguishable. The data series shown here are therefore reduced to the first and last annual sets and decadal milestones in between (Figure 3).

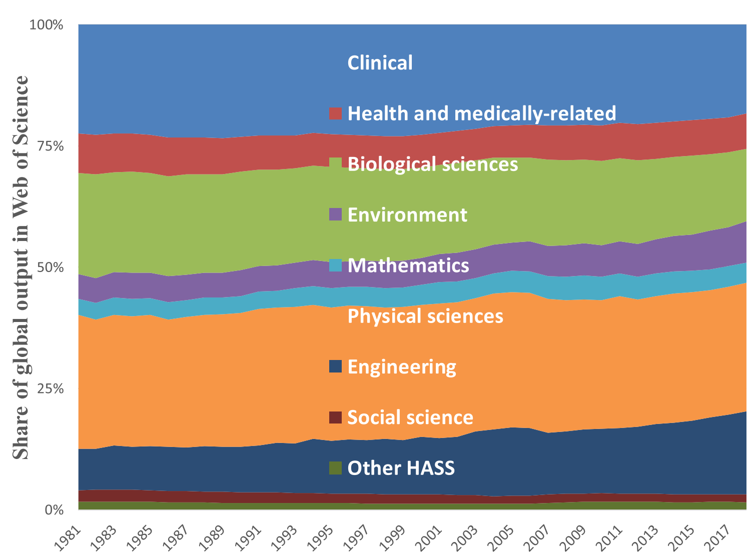

Figure 2. The global proportion of papers in the Web of Science across the period 1981 to 2018. Journal categories are grouped by macrodisciplinary areas. Humanities and most social sciences (HASS) are not included in this analysis as many journals are regional and only represent a part of the typical research output of those fields (see Appendix 1).

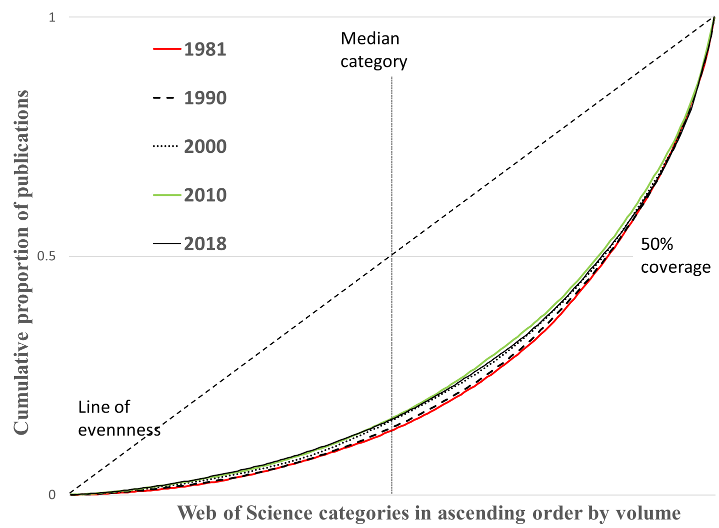

Figure 3. Lorenz curves of cumulative proportion of publications indexed annually in the Web of Science. Data are shown for selected years at approximately decadal intervals. Lines to locate the median category and 50% coverage are indicated; the principal diagonal would track maximum evenness across all categories. The extent to which the data curve departs from this line indicates the degree of unevenness in the spread of data across categories.

The changes in evenness appears to be relatively small but complex. There was a shift towards greater evenness between 1981 and 2010 but a slight shift back from this through 2018, though retaining greater evenness than at the outset. The Lorenz curve provides a valuable "first impression" but does not bring out the detail of change nor readily reveal the dynamics over the period.

The distance of the curve from perfect evenness can be indexed by two simple indicators that derive directly from the data: the proportion of publications covered by half of the categories (i.e., the cumulative proportion at the ranked median) and the minimum number of categories required to include 50% of annual publications. We found that these vary in synchrony. The proportion covered at the ranked median is compared with a formal index below.

3.2. Comparisons Between Indices

The Lorenz curve is illustrative but provides no analytical power, and its illustrative value breaks down with large numbers of years or countries. A summary index is therefore necessary for any large-scale or complex comparative analysis.

A preliminary comparison of alternative indices (but based on ESI categories, which are fewer than those in WoS) was applied to the smaller New Zealand national data set to determine which index was more informative for this project. The Pratt and Gini indices produce nearly identical results for this sample, differing only by a factor related to the number of categories (= $(N - 1)/N$), and are therefore substitute choices. The difference between them is trivial for 194 WoS categories. The general profile of the Shannon index was similar, with synchronous peaks, troughs, and plateaus but a smaller range of index values. For a test set of UK data, index variation only became clear on an expanded axis. Simpson's Index had similar constraints. Figure A2.1 illustrates the indicator profiles over time.

Leydesdorff et al. (2019) chose the Gini coefficient and it has been pointed out that Gini has precedence over Pratt. We therefore focused on the Gini coefficient as indicated in Section 2.

3.3. Global Diversity

WoS has grown fourfold over the study period (Figure 1) and balance has changed at a coarse level (Figure 2). The Lorenz curves for 1981–2018 show that the overall distribution of papers across categories became slightly more even through 2000 and then shifted back after 2010 (Figure 3). Some of the earlier change is likely to be due to the relative growth noted in technology and environmental sciences and the relative reduction in medicine and biology: implicitly a shift in relative volume from larger to smaller categories. Later change is less readily explained but may be due to a global realignment.

We can now analyze the detailed change in evenness across the WoS database using a formal index. Here, comparison of the Gini coefficient for these global data with a parameter derived from the annual Lorenz curves (Figure 3) demonstrates that Gini and the proportion of papers included at the category median varied in synchrony. As a greater cumulative proportion of papers is absorbed by the least abundant half of data categories, rising from around 13% to over 16%, so evenness increases. As evenness increases, so the value of the Gini coefficient falls, from 0.54 to around 0.5 (Figure 4).

A question of interest at this point is, of course, whether the Gini coefficient actually provides any additional information than the far simpler computation of Lorenz median coverage. We draw attention to the exceptional regularity of the Lorenz curves (Figure 3) based on a large number of relatively scarce categories. We are therefore cautious, anticipating that while the Lorenz "shortcut" appears to be a good descriptor of evenness in these data, it may not do so effectively for small categorical counts or for highly skewed abundance data.

Figure 4 shows that global evenness increases between 1981 and 2008. After this there is some indication of a plateau and then a possible shift back towards specialization, which will need to be re-evaluated at a later date. The overall global pattern is, as noted, the summed outcome of many national changes, themselves a synthesis of local and institutional decisions, driven by change in disciplinary opportunities and priorities. When we examine individual entities, we do so in this generic context. *Ceteris paribus*, we should expect that all countries and institutions

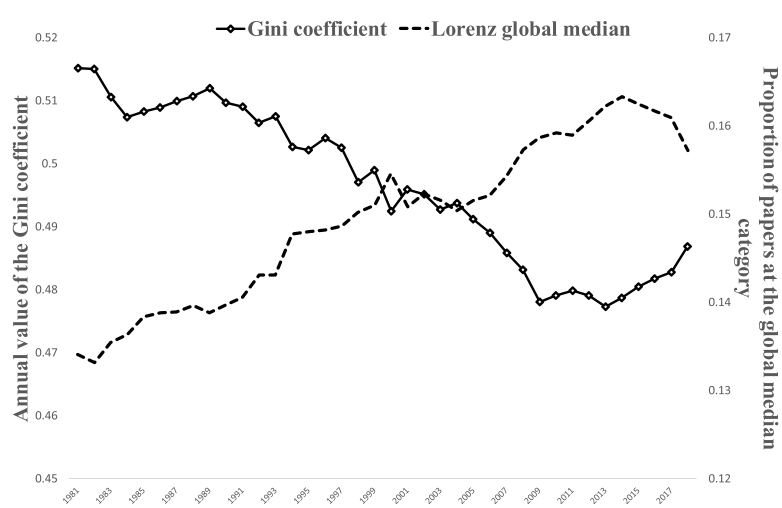

Figure 4. Count and variation in the diversity of articles and reviews annually indexed in the Web of Science, grouped by journal category and measured by the Gini coefficient and by the proportion of papers covered by the least abundant half of WoS categories (value at the global median on the Lorenz curve) in each year.

share the common component towards less specialized portfolios: A test of that and of variance from the global pattern is therefore of interest.

3.4. Granularity

Having established a global baseline, the relative paper counts as normalized against that baseline (proportion by country divided by proportion for world, for each category) were computed for three countries (Australia, New Zealand, and the United Kingdom) for both ESI and WoS databases. All three countries have relatively more even publication data distribution than the world baseline. In all three instances, the finer-grained data of the WoS categories results in greater evenness than for ESI. The shape of the alternative profiles for each country are, however, relatively similar. In very broad terms, New Zealand's diversity increases and the United Kingdom's diversity falls over time, while the two profiles for Australia converge but remain within more narrow bounds than for the other countries. (Figure A2.2)

Because it reveals greater variance, the granularity of the WoS data was preferred as the basis for subsequent analysis. The ESI data are informative, but the breadth of the categories suppresses the detail of change. Note that the overall change in global WoS diversity lies between indicator values of 0.54 and 0.5, whereas the change in values for normalized data (i.e., diversity after the global component has been discounted) for New Zealand and for the United Kingdom are greater than this.

3.5. Gross Categorical Evenness of National Publications

The diversity of national research activity, as reflected in evenness calculated by the Gini coefficient for counts of papers across 194 WoS categories and computed on the actual numbers of papers untransformed by reference to the global baseline, increases across time for all countries except the United States (Figure 5).

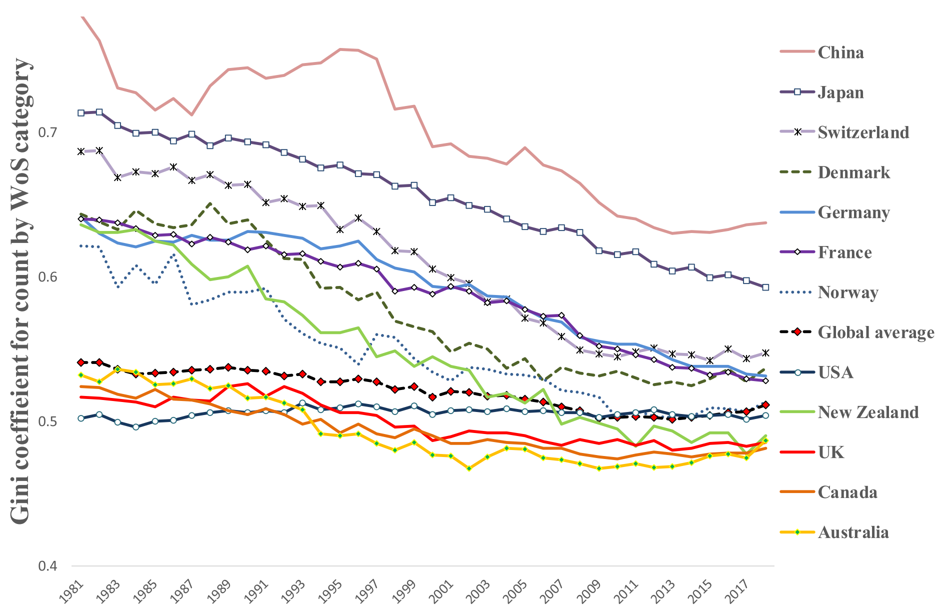

Figure 5. Gini coefficient values based on untransformed publication frequency data for the disciplinary evenness of papers calculated using the Web of Science (194 categories) data set. The global baseline (Figure 4) is shown for reference. Papers had at least one national author for the assigned country; no fractional attribution was used.

In 1981, the global research base was dominated by the G8 group of advanced research economies, which accounted for around 70% of all outputs then indexed by the WoS. By 2018, this was no longer the case and they accounted for less than half of indexed outputs. The United States was by far the largest economy and had enjoyed a relatively high level of sustained R&D investment since 1945, but its share of global output fell from a dominant 36% in 1981 to a little over 25% in 2018. Much of that change in balance (not volume) can be attributed to the rise of Asian research economies and particularly China (Adams & Wilsdon, 2006; Bound et al., 2013).

Three groups of countries emerge from this analysis and change is greatest for the group with the least even portfolios in 1981, hence with the greatest potential for change in the direction of the global system.

- Asia-Pacific. The countries with the least diverse research portfolios indexed in the WoS database in 1981 are Japan and China, and their subsequent shift to greater evenness is similar (China shifts by 14 decimal points from 0.78 to 0.64; Japan by 12 points from 0.71 to 0.59).
- European mainland. The non-Anglophone group initially had more even research portfolios than the Asia-Pacific group, but shifted by a similar degree to them (Denmark, France, Germany, Norway: 11 points; Switzerland: 14 points) and may be converging on a similar degree of evenness.
- Anglophone. The Anglophone countries, with the clear exception of New Zealand, had a more even distribution of papers indexed in the WoS database in 1981 than the global benchmark. Excepting the United States, their research portfolios became slightly more even but most changed by only 3–4 points over the period to 2018. The more marked change has been New Zealand's convergence on others in this group.
- The United States stands out because its research portfolio has become less even, both compared to itself historically and compared to the other Anglophone countries.

The trajectory of Japan is unexpected. It was a major G8 research economy in the 1980s and initiated successful internationally collaborative research initiatives such as the Human Frontier Science Program (https://www.hfsp.org/). Thus, its lower evenness, between that of the EU countries and of China, and its only later shift towards typical European research balance, would not have been predicted.

3.6. Normalized Categorical Evenness of National Publications

The analysis in Figure 5 explicitly does not take account of the underlying global pattern and its secular variation (Figure 4). A different and complementary perspective on the changes in evenness for each country will be revealed by normalizing the national categorical paper counts against that global distribution.

The Gini coefficient calculated for national papers across 194 WoS categories, computed on the numbers of papers for each country in each category and then transformed by reference to the global baseline, confirms that evenness increases across time for most countries except the United States and, more notably, the United Kingdom (Figure 6).

The arithmetic values of the Gini coefficient should be noted. With the exception of China, the index value is markedly reduced for all countries when the data are normalized for the curves in Figure 6 compared to the same country in Figure 5. In other words, there is greater unevenness in the data before normalization against the world baseline. This is unsurprising, because it implies a common component of unevenness associated with the overall global data (e.g., more medical than technology papers) that obscures the relative diversity of each national profile.

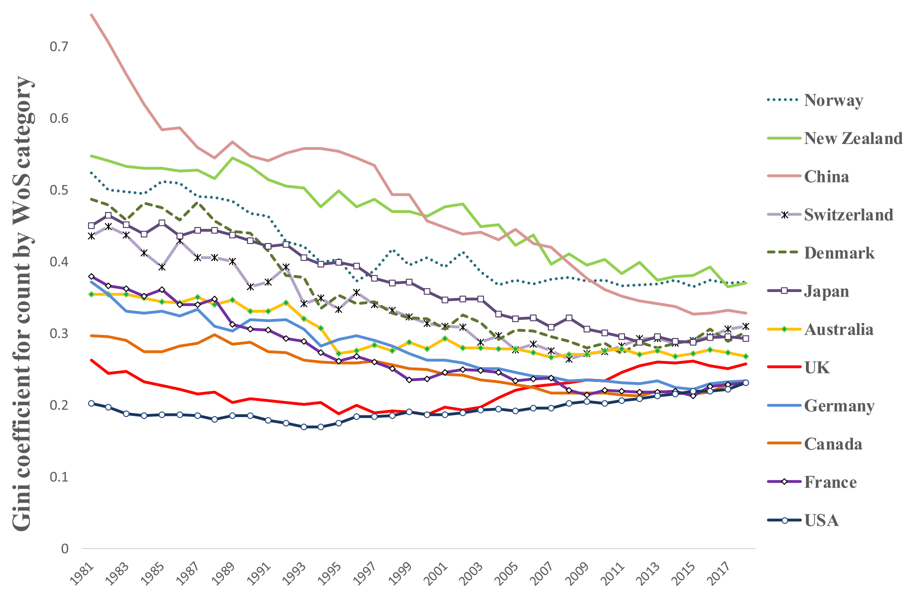

Figure 6. Gini coefficient values based on publication frequency data normalized against global baseline for the disciplinary diversity of papers with at least one national author for a country, calculated using the Web of Science (194 categories) data set.

For most countries, there is also a greater overall tendency for the curves in Figure 5 to converge and to do so at a lower level of unevenness (Gini coefficient between 0.2 and 0.3). China's trajectory is notable in evolving from the most specialized to an evenness close to that of Switzerland, Denmark, and Japan. Japan itself has more similar research evenness to EU countries throughout the period.

New Zealand and Norway (both relatively small but mature research economies) appear to have somewhat more concentrated (less even, Gini coefficient around 0.37) portfolios than others. Denmark provides an interesting contrast (Gini coefficient around 0.3). While Denmark is also relatively small, it is embedded in the European Framework Programme as a full EU member, which Norway (as an Associated Country) is not. New Zealand has no formal regional research network to which it might refer. Thus, the benefits for a small entity of being in a larger structured organization, hypothesized for units at institutional level, seem supported. Note that Switzerland is also an EU Associated Country, but has the advantage of hosting large communal scientific facilities.

The U.S. research portfolio shows a rise in Gini coefficient, with decreasing relative evenness of indexed papers, throughout the period. The rise is both gradual and steady and could be due either to domestic factors or to rising competition for publication opportunities in the indexed journal set, particularly from countries in Asia but also in mainland Europe as they shift towards Anglophone publishing (see Section 4).

The trajectory of the United Kingdom is exceptional. The data in Figure 6 indicate that the United Kingdom's research portfolio was already less specialized (more evenness than most) in 1981. It is the only country for which the research portfolio becomes more even from 1981 through to the 1990s and then becomes less even in the period after 2000. This will be analyzed elsewhere.

4. DISCUSSION

This paper sets out a background to the components of and trends in evenness (across journal-based categories) as a key aspect of the diversity of research papers (articles and reviews) in

WoS. Diversity is informative and important because it both provides flexibility and capacity for response to research challenges and reflects the impact of policy. Appropriate indexing is therefore essential to support effective management and the background in and data set provided with this paper will now enable others to perform similar analyses.

The analyses reported here suggest that "indexing research diversity" is neither necessarily simple nor direct and that interpretation of changes in evenness and specialization may be challenging. Nonetheless, valuable information about the global research system does emerge. We note the significance of variety (the number of categories) and disparity (the distance between categories): Variety changed very little over the period; disparity requires a detailed analysis at publication level (because journals were not only added but changed in content). We therefore focused in this paper solely on the evenness component.

The historical policy for journal coverage (used by ISI and Thomson Reuters) has been that journal papers must include a title and abstract in English. In 1981 this would have restricted coverage for many important non-Anglophone journals at a time when translation services were limited. The consequence, at a time of slower and more limited global communications, may have been that some output of significant non-Anglophone research economies was missed, so only partial content was covered (e.g., in Earth Sciences: Rey-Rocha & Martin-Sempere, 2004).

During 1981 to 2018, a period when the annual intake to the database increased more than fourfold (Figure 1), the categorical evenness of global (baseline) WoS research publication data appeared to increase (Figure 2). Superficially, the change appears small at a global level when tracked in a series of Lorenz curves (Figure 3) and by a plot of annual Gini coefficient values (Figure 4), which could be seen as a comment on the management of the database. However, there was in fact greater growth of physical sciences than life sciences (Figure 2), and some specific categories within the data set (194 of 254 "global" journal categories were analyzed here, excluding more parochial areas in the humanities and parts of the social sciences) grew substantially. Overall, there was a relative shift from biomedical sciences to technology-based areas, environment, and mathematics. This may explain the difference in trajectory for the United States, with exceptional growth in the research budget of the NIH (now equal to the combined budgets of the NSF, Department of Energy, and Environmental Protection Agency). Thus, U.S. investment decisions ran counter to the global trend.

We employed two related methodologies for data management: indexing the original proportions of papers by category (Figure 5) and then indexing the proportions after normalization against the global baseline (Figure 6). Across countries and across these two methods, the overall impression is that the diversity/evenness of research publication output increased consistently at all levels. It is obvious to note the difference between the Anglophone and non-Anglophone groups.

- Publication diversity (on an evenness measure) increased and countries converged from 1981 through to 2018.
- The Anglophone countries had research portfolios that were, in 1981, relatively similar, but more diverse than other countries and overall global data. The United States was more diverse than the global baseline in 1981 but actually lost diversity (possibly for budget reasons as noted) and by 2018 was similar to that global average. (Figure 5).
- Change is most clear in the marked and consistent shift to more diverse portfolios evidenced by the non-Anglophone European and the two large Asia-Pacific research economies.
- Specific national characteristics are revealed when research portfolios are normalized against the global baseline. Normalized baselines are more diverse (more even) because

of the innate structural unevenness in the distribution of journals and papers across categories in the global baseline. (Figure 6)

- Size-related factors. The underlying diversity (Figure 6) of the French and German portfolios is very similar to that of Canada and initially that of Australia, and later that of the United States. By contrast, New Zealand has greater unevenness and similar diversity to Norway.
- Networks. Denmark has similar diversity to Japan, although its economy is much smaller. Two-thirds of its output in the last 5 years has had an international coauthor, 80% of which were within the European Research Area, and this had the effect of increasing its capacity.

It is evident that analysis of the original data and of the normalized data each reveal features of the national and global systems. A complete explanation for the overall global patterns will be complex because it must absorb many individual national policy and funding factors. Global diversity and the spread of publications by category is made up of many national contributions, which could be quite distinct (different policy environments) or share a common agenda (shared research community). A common trend towards much more even research portfolios (less selective, more balanced) appears when the data are disaggregated by country. The greater the level of relative selectivity in 1981, the more change there has been (and can be) so that nations converge on a similar index value by 2018 (Figure 5). This similar index value does not mean that they have similar portfolios, only that they have a similar balance of more and less specialized disciplines.

Two factors may be particularly influential.

- Change in global communications (both physical in terms of air travel to conferences and technical in terms of networking) reinforces the benefits of rapid exchange of information supported by the common use of English among researchers. This will have enabled the non-Anglophone Europeans and the Asian economies to submit a higher proportion of their research to "international" journals to expose their discoveries more widely. This would drive a shift in indexed publication diversity which may, in reality, be a signal of a shift in publishing practice.
- International research collaboration has undoubtedly increased and it is increasingly common for half or more of a country's output to have at least one nondomestic coauthor (Adams, 2013). Thus, a nation's publication portfolio will potentially diversify through partnership and at the same time tend to look more like an international average on which all nations converge.

This shows that both global factors and national policy factors shape research diversity. Analysis relevant to both is therefore required for informed interpretation. The WoS categories are managed, change in content and grow in volume over time. They do so in response to global changes, of which any analytical target (a country or an institution) would be a part. This analysis confirms that it is essential to be aware of that part of research diversity (whether variety, disparity or balance) which is also part of the global research community and that part which is linked to national (or institutional) research policy and culture.

ACKNOWLEDGMENTS

We thank the anonymous referees: of an earlier version of this paper for drawing our attention to the historical relationship between the Gini coefficient and the later Pratt index and for other constructive and informative comments and of the recent version for helping to clarify presentation of our purpose in this analysis. We also thank our ISI colleagues, especially David Pendlebury and

Josh Schnell, for discussing the meanings of balance, evenness, and diversity in the context of research activity and portfolios as captured in WoS.

AUTHOR CONTRIBUTIONS

Jonathan Adams: Methodology, Project administration, Supervision, Visualization, Writing—original draft, Writing—review & editing. Gordon Rogers: Data curation, Formal analysis, Investigation. Warren Smart: Conceptualization, Data curation, Investigation, Validation, Writing—review & editing. Martin Szomszor: Investigation, Methodology, Resources, Software, Validation, Writing—review & editing.

COMPETING INTERESTS

Jonathan Adams, Gordon Rogers, and Martin Szomszor are employees of the Institute for Scientific Information, which is a part of Clarivate Analytics. Clarivate Analytics is the owner of WoS, to which this paper refers.

FUNDING INFORMATION

There was no external funding for this project.

DATA AVAILABILITY

An application for the use of the Web of Science data in indexing global diversity may be made to the authors.

REFERENCES

Adams, J. (2002). Research assessment. *Science*, *296*, 805.

Adams, J. (2007). *The use of bibliometrics to measure research quality in UK higher education institutions*. London: Universities UK. http://jfmela.free.fr/jfmblog/wp-content/bibliometrics.pdf

Adams, J. (2008). *Strategic review of the performance-based research fund: The assessment process*. A report to the New Zealand Tertiary Education Council. https://www.tec.govt.nz/assets/Publications-and-others/PBRF-Publications/c240a83d4b/PBRF-strategic-review-of-asessment-process-2008-review.pdf

Adams, J. (2012). Collaborations: The rise of research networks. *Nature*, *490*, 335–336.

Adams, J. (2013). The fourth age of research. *Nature*, *497*, 557–560.

Adams, J., Cook, N., Law, G., Marshall, S., Mount, D., … Stephenson, J. (2000). *The role of selectivity and the characteristics of excellence*. Report of a consultancy study for the Higher Education Funding Council for England. HEPU, University of Leeds.

Adams, J., Mount, D. R., & Smith, D. N. (2010). *The future of the UK university research base*. London: Universities UK. http://www.universitiesuk.ac.uk/Publications/Pages/Thefutureofresearch.aspx

Adams, J., Rogers, G., & Szomszor, M. (2019). *The Annual G20 Scorecard—Research Performance 2019*. London: Institute for Scientific Information. https://clarivate.com/g/the-annual-g20-scorecard-research-performance-2019/

Adams, J., & Wilsdon, J. (2006). *The new geography of science: UK research and international collaboration*. A report by Evidence and Demos, September 2006. Leeds: Evidence.

Balassa, B. (1979). The changing pattern of comparative advantage in manufactured goods. *The Review of Economics and Statistics*, *61*(2), 259–266.

Bookstein, A., & Yitzhaki, I. M. (1999). Own-language preference: A new measure of "relative language self-citation." *Scientometrics*, *46*, 337–348.

Bound, K., Saunders, T., Wilsdon, J., & Adams, J. (2013). *China's absorptive state: Research innovation and the prospects for China-UK collaboration*. London: NESTA.

Carpenter, M. P. (1979). Similarity of Pratt's measure of class concentration to the Gini index. *Journal of the American Society for Information Science*, *30*, 108–110.

Castelvecchi, D. (2015). Physics paper sets record with more than 5,000 authors. *Nature News*. https://doi.org/10.1038/nature.2015.17567

Egghe, L. (1987). Pratt's measure for some bibliometric distributions and its relation with the 80/20 rule. *Journal of the American Society for Information Science*, *38*, 288–297.

Egghe, L., & Rousseau, R. (1990). Elements of concentration theory. *Informetrics*, *89/90*, 97–137.

Fisher, R. A., Corbet, A. S., & Williams, C. B. (1943). The relation between the number of species and the number of individuals in a random sample of an animal population. *Journal of Animal Ecology*, *12*, 42–58.

Gini, C. (1909). Il diverso accrescimento delle classi sociali e la concentrazione della ricchezza. *Giornale degli Economisti*, *11*, 37.

Hill, M. O. (1973). Diversity and evenness: A unifying notation and its consequences. *Ecology*, *54*, 427–432.

Horta, H. (2018). The declining scientific wealth of Hong Kong and Singapore. *Scientometrics*, *117*(1), 427–447. https://doi.org/10.1007/s11192-018-2845-0

Johnston, R. (1994). Effects of resource concentration on research performance. *Higher Education*, *28*, 25–37.

Keylock, C. J. (2005). Simpson diversity and the Shannon–Wiener index as special cases of a generalized entropy. *Oikos*, *109*, 203–207.

King, C. (2012). Multiauthor papers: Onward and upward. *Science Watch* (Thomson Reuters). Available at: http://archive.sciencewatch.com/newsletter/2012/201207/multiauthor_papers/

Leydesdorff, L., Wagner, C. S., & Bornmann, L. (2019). Interdisciplinarity as diversity in citation patterns among journals: Rao-Stirling diversity, relative variety, and the Gini coefficient. *Journal of Informetrics*, *13*, 255–269.

Lorenz, M. O. (1905). Methods of measuring the concentration of wealth. *Quarterly Publications of the American Statistical Association*, *9*, 209–219.

Mansourzadeh, M. J., Shahmoradi, B., Dehdarirad, H., & Janavi, E. (2019). A note on using revealed comparative advantages in scientometrics studies. *Scientometrics*, *121*(1), 595–599. https://doi.org/10.1007/s11192-019-03207-8

Matia, K., Amaral, L. A. N., Luwel, M., Moed, H. F., & Stanley, H. E. (2005). Scaling phenomena in the growth dynamics of scientific output. *Journal of the American Society for Information Science and Technology*, *56*, 893–902.

Moed, H. F. (2006). *Bibliometric ranking of world universities*. CWTS Report 2006-01, pages 1–36.

Nijssen, D., Rousseau, R., & van Hecke, P. (1998). The Lorenz curve: A graphical representation of evenness. *Coenoses*, *13*, 33–38.

Page, S. E. (2011). *Diversity and complexity*. Princeton, NJ: Princeton University Press.

Pareto, V. (1895). La legge della domanda. *Giornale degli Economisti*, *10*, 59–68.

Pratt, A. D. (1977). A measure of class concentration in bibliometrics. *Journal of the American Society for Information Science*, *28*, 285–292.

Radosevic, S., & Yoruk, E. (2014). Are there global shifts in the world science base? Analysing the catching up and falling behind of world regions. *Scientometrics*, *101*(3), 1897–1924. https://doi.org/10.1007/s11192-014-1344-1

Rousseau, R., van Hecke, P., Nijssen, D., & Bogaert, J. (1999). The relationship between diversity profiles, evenness and species richness based on partial ordering. *Environmental and Ecological Statistics*, *6*, 211–223.

Seddon, N., Mace, G. M., Naeem, S., Tobias, J. A., Pigot, A. L., … Walpole, M. (2016). Biodiversity in the Anthropocene: Prospects and policy. *Proceedings of the Royal Society B: Biological Sciences*, *283*, 1–9. http://dx.doi.org/10.1098/rspb.2016.2094

Simpson, E. H. (1949). Measurement of diversity. *Nature*, *163*, 688.

Sivertsen, G. (2017). Unique, but still best practice? The Research Excellence Framework (REF) from an international perspective. *Palgrave Communications*, *78*, 1–6. https://doi.org/10.1057/palcomms.2017.78

Spellerberg, I. F., & Fedor, P. J. (2003). A tribute to Claude Shannon (1916–2001) and a plea for more rigorous use of species richness, species diversity and the "Shannon–Wiener" Index. *Global Ecology and Biogeography*, *12*, 177–179.

Stirling, A. (2007). A general framework for analysing diversity in science, technology and society. *Interface, a Journal of the Royal Society*, *4*, 707–719. https://doi.org/10.1098/rsif.2007.0213

van Dam, A. (2019). Diversity and its decomposition into variety, balance and disparity. *Royal Society Open Science*, *6*(7), 190452. https://doi.org/10.1098/rsos.190452

Wilsdon, J., Allen, L., Belfiore, E., Campbell, P., Curry, S., … Johnson, B. (2015). *The Metric Tide: report of the independent review of the role of metrics in research assessment and management*. Bristol: Higher Education Funding Council for England. https://doi.org/10.13140/RG.2.1.4929.1363

APPENDIX 1

Web of Science journal categories (194 of 254) included in the analyses reported in this paper. Some humanities, arts, and social science categories were excluded because they have strong regional rather than global affiliations.

Acoustics	Engineering, multidisciplinary	Obstetrics & gynaecology
Agricultural economics & policy	Engineering, ocean	Oceanography
Agricultural engineering	Engineering, petroleum	Oncology
Agriculture, dairy & animal science	Entomology	Operations research & management science
Agriculture, multidisciplinary	Environmental sciences	Ophthalmology
Agronomy	Environmental studies	Optics
Allergy	Ergonomics	Ornithology
Anatomy & morphology	Evolutionary biology	Orthopaedics
Andrology	Fisheries	Otorhinolaryngology
Anaesthesiology	Food science & technology	Palaeontology
Anthropology	Forestry	Parasitology
Area studies	Gastroenterology & hepatology	Pathology
Astronomy & astrophysics	Genetics & heredity	Pediatrics
Audiology & speech-language pathology	Geochemistry & geophysics	Peripheral vascular disease
Automation & control systems	Geography	Pharmacology & pharmacy
Behavioral sciences	Geography, physical	Physics, applied
Biochemical research methods	Geology	Physics, atomic, molecular & chemical
Biochemistry & molecular biology	Geosciences, multidisciplinary	Physics, condensed matter
Biodiversity conservation	Geriatrics & gerontology	Physics, fluids & plasmas
Biology	Gerontology	Physics, mathematical
Biophysics	Green & sustainable science & technology	Physics, multidisciplinary
Biotechnology & applied microbiology	Health care sciences & services	Physics, nuclear
Cardiac & cardiovascular systems	Haematology	Physics, particles & fields
Cell & tissue engineering	Horticulture	Physiology
Cell biology	Imaging science & photographic technology	Plant sciences
Chemistry, analytical	Immunology	Polymer science

(*continued*)

Chemistry, applied	Infectious diseases	Psychiatry
Chemistry, inorganic & nuclear	Information science & library science	Psychology
Chemistry, medicinal	Instruments & instrumentation	Psychology, applied
Chemistry, multidisciplinary	Integrative & complementary medicine	Psychology, biological
Chemistry, organic	International relations	Psychology, clinical
Chemistry, physical	Limnology	Psychology, developmental
Clinical neurology	Logic	Psychology, experimental
Communication	Management	Psychology, mathematical
Computer science, cybernetics	Marine & freshwater biology	Psychology, multidisciplinary
Computer science, hardware & architecture	Materials science, biomaterials	Psychology, psychoanalysis
Computer science, information systems	Materials science, ceramics	Public, environmental & occupational health
Computer science, interdisciplinary applications	Materials science, characterization & testing	Quantum science & technology
Computer science, software engineering	Materials science, coatings & films	Radiology, nuclear medicine & medical imaging
Computer science, theory & methods	Materials science, composites	Rehabilitation
Construction & building technology	Materials science, multidisciplinary	Remote sensing
Critical care medicine	Materials science, paper & wood	Reproductive biology
Crystallography	Materials science, textiles	Respiratory system
Demography	Mathematical & computational biology	Rheumatology
Dentistry, oral surgery & medicine	Mathematics	Robotics
Dermatology	Mathematics, applied	Social sciences, mathematical methods
Development studies	Mathematics, interdisciplinary applications	Soil science
Developmental biology	Mechanics	Spectroscopy
Ecology	Medical informatics	Statistics & probability
Economics	Medical laboratory technology	Substance abuse
Electrochemistry	Medicine, general & internal	Surgery
Emergency medicine	Medicine, research & experimental	Telecommunications
Endocrinology & metabolism	Metallurgy & metallurgical engineering	Thermodynamics
Energy & fuels	Meteorology & atmospheric sciences	Toxicology
Engineering, aerospace	Microbiology	Transplantation
Engineering, biomedical	Microscopy	Transportation

(*continued*)

Engineering, chemical	Mineralogy	Transportation science & technology
Engineering, civil	Mining & mineral processing	Tropical medicine
Engineering, electrical & electronic	Multidisciplinary sciences	Urban studies
Engineering, environmental	Mycology	Urology & nephrology
Engineering, geological	Nanoscience & nanotechnology	Veterinary sciences
Engineering, industrial	Neuroimaging	Virology
Engineering, manufacturing	Neurosciences	Water resources
Engineering, marine	Nuclear science & technology	Zoology
Engineering, mechanical	Nutrition & dietetics	

APPENDIX 2

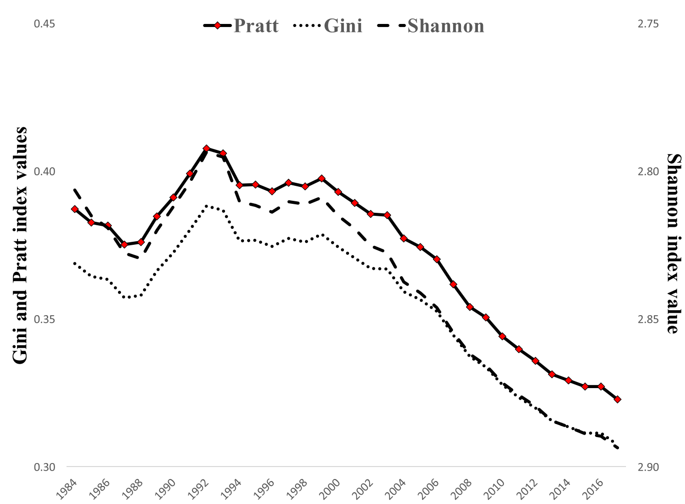

Figure A2.1. The publication diversity of the New Zealand research base using Essential Science Indicator categories and comparing the Pratt, Gini, and Shannon diversity indices (index values displayed as average for five-year rolling windows by middle year of window).

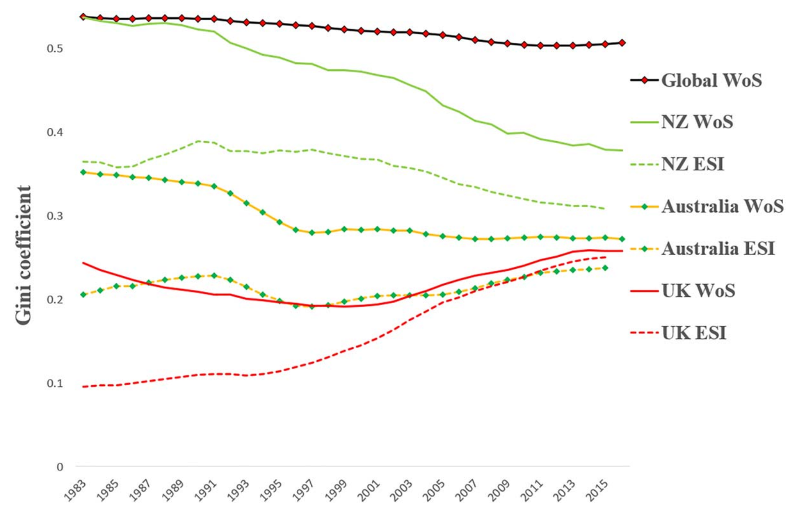

Figure A2.2. Comparison between Gini coefficient based on publication frequency data normalized against global baseline for the disciplinary diversity of papers with at least one national author for each of three countries, calculated using the Essential Science Indicator (ESI, 21 categories) and Web of Science (WoS, 194 categories) data sets. The global baseline calculated from the WoS data (Figure 5) is shown for reference. Data are presented as average index values for five-year rolling windows by middle year of window.