MINISTRY OF BUSINESS,
INNOVATION & EMPLOYMENT
HĪKINA WHAKATUTUKI

Bibliometric Analysis of New Zealand Research Performance

Prepared by Motu Economic and Public Policy Research

May 2019

LSE 4784

New Zealand Government

MINISTRY OF BUSINESS,
INNOVATION & EMPLOYMENT
HĪKINA WHAKATUTUKI

Ministry of Business, Innovation and Employment (MBIE)

Hīkina Whakatutuki - Lifting to make successful

MBIE develops and delivers policy, services, advice and regulation to support economic growth and the prosperity and wellbeing of New Zealanders.

MBIE combines the former Ministries of Economic Development, Science + Innovation, and the Departments of Labour, and Building and Housing.

This MOTU research was funded by the New Zealand Ministry of Business, Innovation and Employment (MBIE).

www.mbie.govt.nz

0800 20 90 20

Information, examples and answers to your questions about the topics covered here can be found on our website www.mbie.govt.nz or by calling us free on 0800 20 90 20.

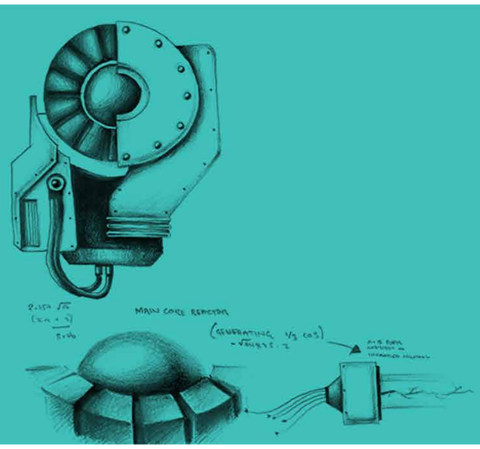
Disclaimer

This document is a guide only. It should not be used as a substitute for legislation or legal advice. The Ministry of Business, Innovation and Employment is not responsible for the results of any actions taken on the basis of information in this document, or for any errors or omissions.

ISSN 1176-2667 (Print), ISSN 1177-9047 (Online)

MAY 2019

BIBLIOMETRIC ANALYSIS OF NZ RESEARCH PERFORMANCE



An Executive Summary
Adam Jaffe and Kate Preston

Motu Economic and Public Policy Research
adam.jaffe@motu.org.nz and kate.preston@motu.org.nz

SUMMARY HAIKU

Bibliometrics.
Look at New Zealand science.
Good but not stellar.

INTRODUCTION

Bibliometrics is the use of data on research publications as indicators of the research output of researchers, research institutions or geographic/political areas. Whenever a metric or statistic is constructed, choices have to be made as to which data to use and what procedures to follow in constructing the metric. The use of bibliometrics to evaluate research endeavours is relatively new and has not been standardised. There is an academic literature on the issues in constructing bibliometric measures, but there is no widely accepted international agreement as to which issues of measurement are most important. Disagreement also prevails over whether it is appropriate to use confidence intervals on bibliometric data, which can be considered a population rather than a sample.

Our research involved stress-testing bibliometric measures of New Zealand (NZ) publication output across disciplines and over time to elucidate the consequences of different choices. Though the focus of this report is on NZ's performance as a nation, many of the insights will also be relevant to the bibliometric assessment of individual researchers, departments, and institutions. In addition, we discuss the arguments for and against using confidence intervals for the bibliometric measures and present bootstrapped confidence intervals around some of our results.

The measures we used to evaluate the quantity and impact of NZ's research are described in Table 1 below. They were calculated from Scopus Custom Data 2002-2015, a large database of abstracts and citations of peer-reviewed literature.

Table 1: Bibliometric Measures of NZ's Research Performance

Quantity of Research	
Total Publications	The number of publications with a NZ affiliation
Impact of Research	
Average MNCS	Normalised citations per NZ publication, interpreted as above the world average if greater than 1.
Total MNCS	Combines both the amount of research output (publications) and its average impact (citations per publication).
Fraction in top percentiles	The fraction of NZ publications in the top percentiles of the citation distribution of similar publications.

As there are multiple choices of calculation mode to make, it is not practical to explore all possible combinations of choices. Our approach was to explore each choice in isolation relative to a set of baseline results.

This work was funded by the New Zealand Ministry of Business, Innovation and Employment (MBIE).
The authors gratefully acknowledge this assistance.

The issues that we explored include:

- Types of publications to include
- Normalisation of citations by publication
- Comparing total output and impact to a benchmark
- Attributing internationally co-authored publications to countries
- Normalising citations of publications in multiple fields
- The effect of the integer nature of citations
- Measuring uncertainty
- An alternative assignment of publications to fields
- Consolidation of fields

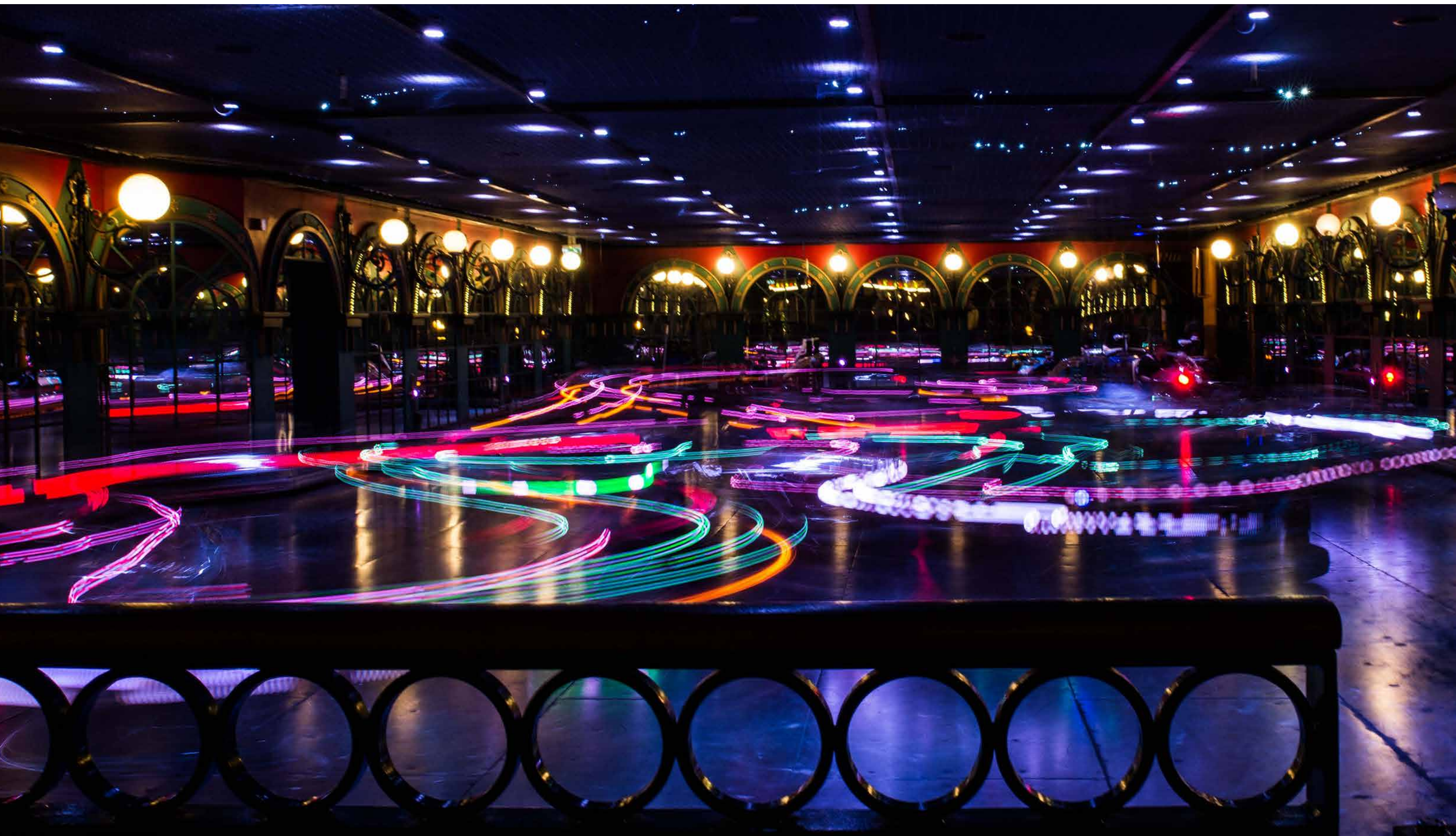
It is important to emphasise that in most cases there is not a ‘right’ or ‘wrong’ way to conduct the analysis. Rather, there are frequently equally valid alternative approaches, or approaches that inherently answer slightly different underlying questions. Therefore, except in a few cases where certain methods lead to inherently misleading results, our approach is to describe the consequences of different approaches, rather than identifying a correct approach.

KEY FINDINGS

The types of publications included in bibliometric analysis and how publication type and fields are accounted for in normalising citations are shown to have relatively small impacts on results for NZ’s bibliometric performance. Though small, one should be cautious that these impacts will be stronger for some fields than others.

Many NZ papers have both NZ and foreign authors. MBIE currently counts such papers as full NZ output. An alternative is to treat each of such papers as a fraction of a publication, with the fraction equal to the fraction of authors who are based in NZ. We find that NZ’s performance in terms of average MNCS across subjects tends to be worse when using fractional counting instead of full counting of authors, implying that the internationally co-authored papers are on average more highly cited than purely domestic papers.

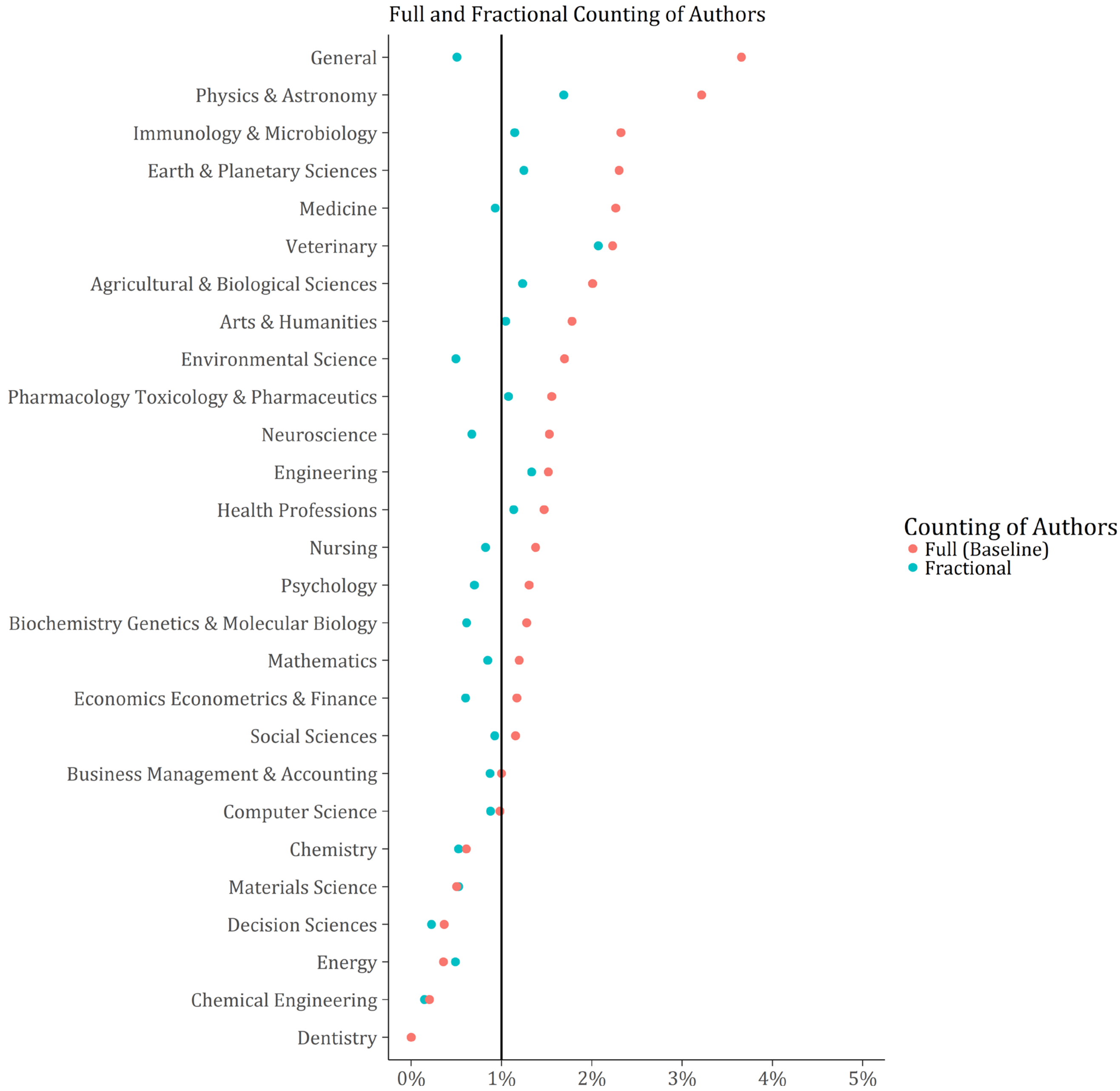
Whether internationally co-authored papers are viewed as a full or partial NZ output also has significant consequences for how the share of publications in top percentiles can be interpreted. In principle, the fraction of NZ publications

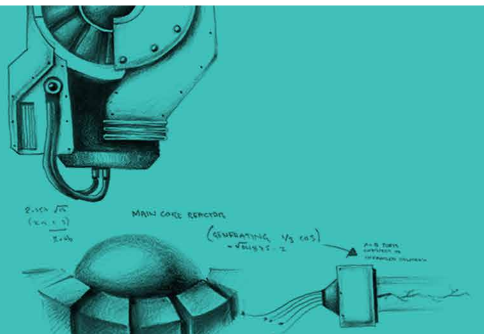


in top percentiles can be compared directly to the world at large so that, for example, having more than 1% of NZ publications in the top 1% of the world's citation distribution indicates that NZ has contributed a disproportionately high share of top publications. However, with full counting of authors, a paper with both NZ and U.S. authors that is in the top 1% is counted as a top-1% paper for both countries, meaning around the world there will be more than 1% of all papers in the 'top 1%'.

We show that when fractional counting is used, avoiding the double-counting issue, NZ's performance across disciplines on these percentile measures is much worse than it appeared with full counting.

Figure 1: Fraction of NZ Papers in Top 1% in 2012-2014

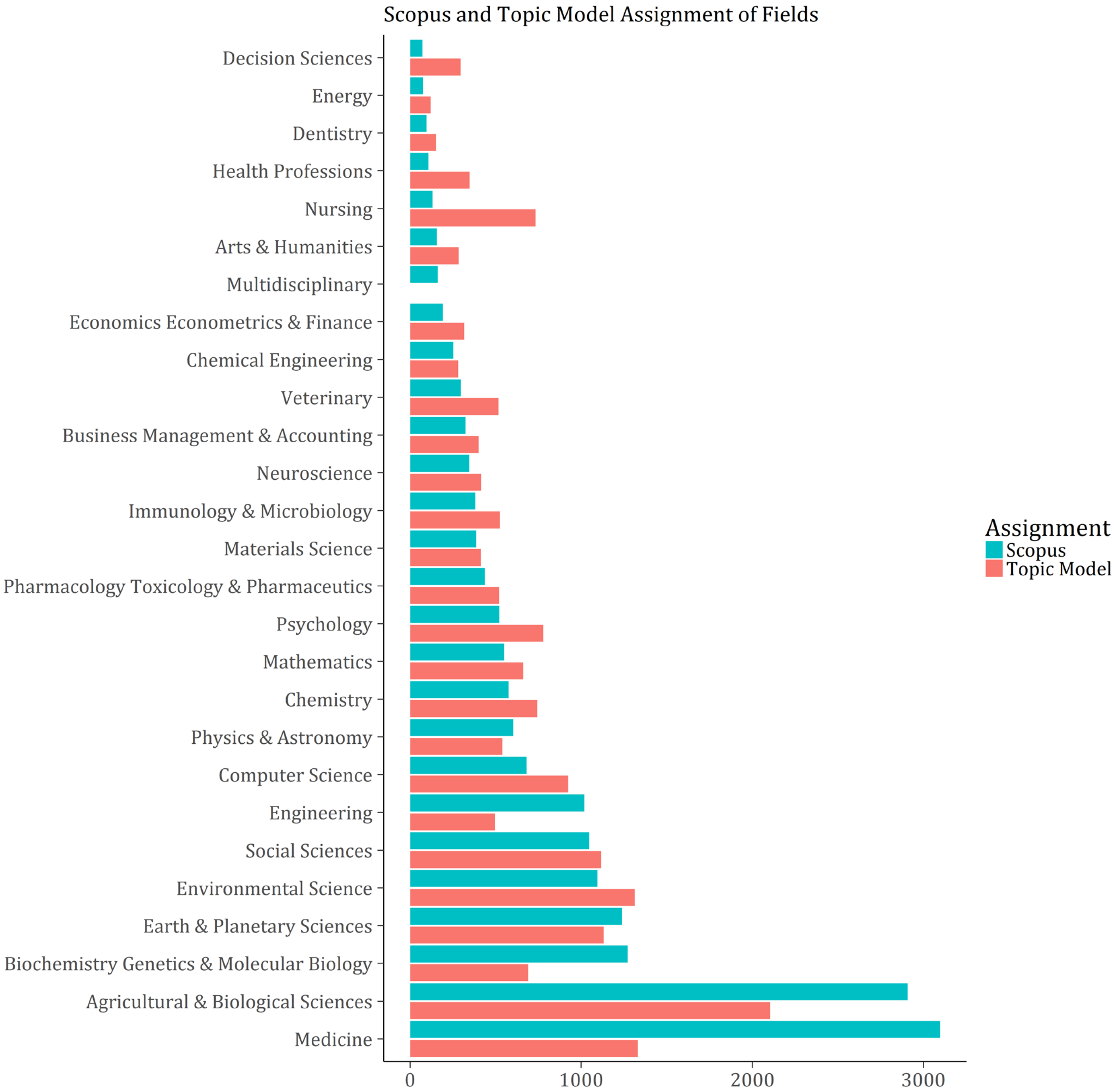


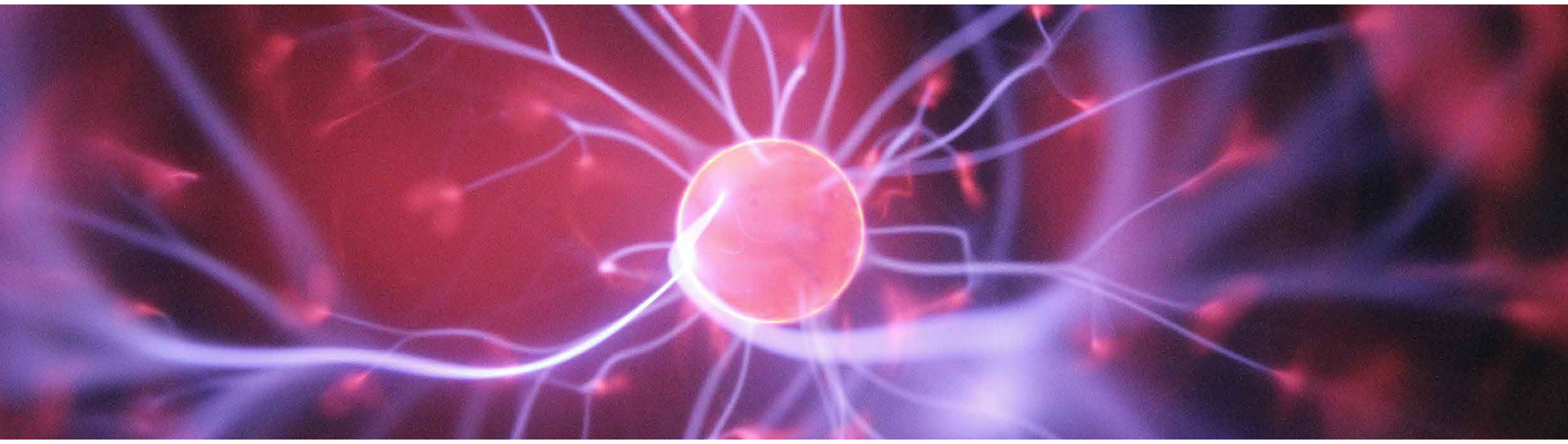


There is another problem in the calculation of fraction of publications in top percentiles that can distort the interpretation of results. Because citation data are discrete and many publications may have the same number of citations, it is possible, for example, that more than 1% of publications have at least as many citations as the threshold for being in the top 1% of the distribution. When calculating the fraction of publications above a certain threshold of the citation distribution, fractional weighting should be applied to publications with the same number of citations as the threshold so that the global fraction of publications in the top x% is exactly x% . We find that leaving out this adjustment leads to positively biased results for the share of NZ's publications that fall in the top percentiles, particularly in recent years when citation distributions tend to be more condensed.

Publications in the Scopus dataset are linked to disciplines according to the main fields covered by the journal in which they were published. Where publications are in journals that span multiple fields, we do not know to what extent the publication itself covers either of those fields. As an alternative, we explore the use of topic modelling to assign publications to fields using the words in their abstracts. The classifications produced differ somewhat from those based on Scopus journal assignments, but we do not have any external reference to tell us which is better.

Figure 2: **NZ's Total Publications 2002-2004**





While bibliometric measures are not based on samples in a statistical sense, they are nonetheless potentially sensitive to how the population corresponding to a particular question of interest is defined. We proposed using bootstrapped confidence intervals to characterize this sensitivity and demonstrated this for some of our results, but this is a topic that merits further work. We observe that, for fractions in the top percentiles, uncertainty of results increases the closer the threshold considered is to zero.

CONCLUSION

Because methodological choices underlying bibliometric analysis affect both statistics and their interpretation, we must stress the importance of making these choices transparent. The MBIE 2016 Science and Innovation System Performance Report leaves readers in the dark about how publications in multiple fields or with multiple authors are dealt with. Those readers who are concerned with such issues will find it unclear what conclusions can be drawn from the results, while those who are not could easily draw misinformed conclusions.

In terms of which metrics should be studied, it would not be wise to draw conclusions about NZ's research performance based on any of the metrics we use in isolation. In most disciplines, NZ research is above average, by which we mean that the average MNCS is greater than unity and the proportion of publications with above median citations is greater than 50%. But when we focus on the upper 10% or upper 1% of the citation distributions, NZ's share in most fields is below the world share when properly calculated. This suggests that NZ has a healthy proportion of good researchers but a disproportionately low concentration of international star researchers.

In future work, we will explore the career trajectories of individual NZ scientists, and their relationships with each other and international scientists, to try to understand the dynamics that underlie the distribution of outcomes.

It is important to emphasize that the measures of bibliometrics that we have employed look only at the outputs of the research process. What we would really like to understand is how these outputs relate to research inputs such as researchers, laboratory equipment and supplies. Indeed, in the absence of data on research inputs, none of the bibliometric measures in this report can be used to ascertain in which fields NZ researchers or NZ research investments are more effective or more efficient than others. The launching of the National Research Information System (NRIS) will create a basic data structure for research inputs. Linking those data to bibliometrics should create the potential to begin to answer these questions.

READ THE FULL VERSION OF THE WORKING PAPER BELOW

Bibliometric Analysis of New Zealand Research Performance: Measurement and Classification Issues

Adam Jaffe and Kate Preston

Motu Economic and Public Policy Research

May 2019

Document information

Author contact details
Adam Jaffe
Motu Economic and Public Policy Research, Queensland University of Technology, Te Pūnaha Matatini
adam.jaffe@motu.org.nz

Kate Preston
Motu Economic and Public Policy Research
kate.preston@motu.org.nz

Acknowledgements
This work was funded by the New Zealand Ministry of Business, Innovation and Employment (MBIE). We extend thanks to David Friggens and Șenay Yasar Saglam, both from MBIE, for their insights on this work and for their support with the data. We also thank our Motu colleagues Trinh Le, Dave Maré and Dean Hyslop for their helpful advice on several aspects of this report. For their helpful comments about publication data, we are grateful to Shaun Hendy from the University of Auckland and Jason Gush from Royal Society Te Apārangi.

Disclaimer
The findings and opinions expressed in this paper are those of the authors.

Motu Economic and Public Policy Research

PO Box 24390	info@motu.org.nz	+64 4 9394250
Wellington	www.motu.org.nz	
New Zealand		

Abstract

Bibliometric databases on publications and their citations offer the possibility to quantify the scale and impact of this key output of scientific research. But there are many decisions to be made about how to construct these bibliometric measures, and no global consensus on how best to do so. We consider several measures of research output across scientific disciplines in New Zealand and test their sensitivity to a number of methodological choices. Factors considered include which publications to include in the analysis, ways to count publications co-authored by both local and international authors, and procedures for normalising citations. We also explore the potential to assign publications to disciplines by semantic analysis and a way to construct confidence intervals for the measures. The findings provide insight into the extent to which these methodological considerations impact on bibliometric results, and their interpretation, and how uncertainty in the results can be quantified.

JEL codes

O39

Keywords

Bibliometrics, Citations, Normalisation, Counting method, Journal classification, Confidence intervals

Summary haiku

Bibliometrics.
Look at New Zealand science.
Good but not stellar.

Table of Contents

Tables and Figures

1 Introduction and Overview

Bibliometrics refers to the use of data on research publications, and citations made to research publications, as indicators of the research output of researchers, research institutions or geographic/political areas. Many countries and international organisations such as the OECD use bibliometric measures to shed light on the strength or success of research in various contexts. In particular, the Ministry of Business, Innovation & Employment (MBIE) has published a Science & Innovation System Performance Report (MBIE 2016) that provides various indicators of the success of the New Zealand (NZ) science enterprise.

Whenever a metric or statistic is constructed, choices have to be made as to which data to use and what procedures to follow in constructing the metric. Some statistics, such as those in the National Income Accounts that underlie calculation of GDP, have been used for many years, and much international effort has gone into standardising issues of definition and construction so that interpretation of these statistics is reliable and comparable across countries. The use of bibliometrics to evaluate research endeavours is newer and less standardised. There are no widely accepted international protocols. And while there is an academic literature on the issues that arise (Waltman 2016a), there is not even agreement as to which issues of measurement are most important. For these reasons, we were asked by MBIE to stress-test the metrics used in the System Performance Report, as well as related measures, in order to identify which issues of definition and construction are most important, and shed some light on the consequences of different choices. The focus of this report is on NZ's research performance as a nation by discipline and over time, but many of the insights will also be relevant to the bibliometric assessment of individual researchers, departments, and institutions.

Section 2 introduces the basic measures, and establishes a baseline against which variations in approach can be viewed. Sections 3 to 7 of the paper each explores a specific issue that arises in the construction of bibliometric measures, and characterizes to what extent and under what circumstances different choices seem to have a material effect on the implied conclusions. Section 8 focuses on the fact that the distribution of citations across papers is generally highly skewed. This skewness means that a single statistic (e.g. average number of citations received) is typically inadequate to characterize the overall distribution of results. This section illustrates the tendency for citation thresholds for the top 1%, 10%, and 50% of publications to be close to zero, and examines the importance of adjusting for the integer nature of citations in metrics of the fraction of publications in these upper percentiles.

Section 9 focuses on the question of the robustness or sensitivity of the results. This is an area that remains controversial. As discussed below, standard approaches to 'standard errors' or 'confidence intervals' may not be appropriate for bibliometric measures. Nonetheless, we believe it important to convey to uses that the metrics are indeed somewhat sensitive to issues of definition and construction, and may not be stable across measurement methods or over time. We suggest one approach to representing this sensitivity.

For many purposes motivating these analyses, we would like to construct output metrics for different research fields or disciplines. In fact, even when we do not tabulate metrics by discipline, this information is used in the normalisation process to make the measure of citations to one document comparable to the citations of another. Doing this requires associating published research papers with specific disciplines. Making this association raises both conceptual and practical problems. The conceptual issues arise from the inherently multidisciplinary nature of research and researchers. And the practical problems arise because there is no referee or central classification system that assigns a discipline or disciplines to each published paper. Most extant bibliometric analysis assigns papers to disciplines on the basis of the journal or proceedings in which the paper was published. The emergence in recent years of computer algorithms for natural language processing offers the potential to classify papers in disciplines based on systematic analysis of the text of the papers themselves. In Section 10, we explore such an alternative approach to disciplinary classification. We find that it can be done, and leads to results that differ modestly from results based on the traditional classification methods. Further work will be necessary to determine whether and under what conditions the text-based classification is more useful.

As we explore different choices in the construction of these metrics, it is important to emphasise that in most cases there is not a 'right' or 'wrong' way to conduct the analysis. Rather, there are frequently equally valid alternative approaches, or approaches that inherently answer slightly different underlying questions. Therefore, except in a few cases where certain methods lead to inherently misleading results, our approach is to describe the consequences of different approaches, rather than identifying a correct approach.

1.1 Bibliometric Measures

The measures used to evaluate the quantity and impact of NZ’s research are described in Table 1 below. They are calculated from Scopus Custom Data 2002-2015, extracted June 2016. Scopus is a large database of abstracts and citations of peer-reviewed literature which we treat as representative of the world’s total research output.[1] Quantity of research is simply reflected by the total number of publications, while the number of citations received by a publication is used as a proxy for its impact, captured in three different types of measure.

All citation-based measures are constructed relative to the citation performance of all publications from the same year, discipline, and publication type. This prevents the measures

[1] Though we recognise that Scopus does not cover all peer-reviewed literature, especially not in the arts, humanities, and social sciences, the focus of this study is on how best to use bibliometric databases like Scopus. Normalisation of measures by field (discussed below) corrects for the average deficiencies in Scopus across fields, but cannot of course rectify specific deficiencies.

from being distorted by differences in the opportunity each publication has had to be cited: earlier publications have had more time to be cited; different disciplines have different citation practices; and different publication types cover work of different nature which may attract more or less citations.

Table 1: Bibliometric Measures of NZ's Research Performance

Quantity of Research	
Total Publications	The number of publications with a NZ affiliation
Impact of Research	
Average MNCS	Normalised citations per NZ publication, interpreted as above the world average if greater than 1.
Total MNCS	Combines both the amount of research output (publications) and its average impact (citations per publication).
Fraction in top percentiles	The fraction of NZ publications in the top percentiles of the citation distribution of similar publications.[2]

The first two measures of impact rely on the Mean Normalised Citation Score (MNCS) as defined by Waltman et al. (2011), which is the number of citations normalised relative to the average for all publications from the same year, discipline, and publication type. Average MNCS gives an indication of the impact per publication, with a value of 1 corresponding to the impact of the global average publication because of the normalisation. Total MNCS reflects overall impact, but lacks an international benchmark. To gain an understanding of NZ's performance at different points of the citation distribution, we explore the fraction of publications in the 1st, 10th, and 50th (median) percentiles. In principle, the fraction of NZ publications in top percentiles can be compared directly to the world at large so that, for example, having more than 1% of NZ publications in the top 1% of the world's citation distribution indicates that NZ has contributed a disproportionately high share of top publications. As discussed further below, however, such direct comparison is only appropriate if papers with multiple authors and/or multiple disciplines are not double-counted so that the global share of publications found to be in the top x% is, indeed, x%.

1.2 Overview of issues investigated

As there are multiple choices of calculation mode to make, it is not practical to explore all possible combinations of choices. Our approach has been to explore each choice in isolation relative to a set of baseline results. The issues that we have explored are as follows:

[2] Similar publications are publications from the same year, discipline, and publication type, the same characteristics by which citations are normalised in MNCS.

1. **Types of publications to include -** Scopus tracks approximately 20 different types of publications, such as articles, books, reviews, conference proceedings, editorials, erratum and letters. In our baseline results, we include only articles, reviews, and conference proceedings, because this is what is currently preferred by MBIE. This choice could be criticised as unfair for certain disciplines which place more emphasis on alternative publication types. We find that expanding the analysis to cover all publication types makes a difference to NZ's results for some fields but not others.
2. **Normalisation of citations by publication types -** It is standard practice in bibliometrics to normalise citations by publication type. It is often noted that reviews tend to be relatively highly cited, because they conveniently summarise prior work in a field, and so reviews should only be compared with reviews. However, we argue that not all publication types need to be distinguished for the normalisation of citations. For example, if conference proceedings tend to be less cited than articles, then this could reflect true differences in impact. An alternative could be to compare publications of the same nature of work, categorised by whether they are new contributions to research or reviews. Our analysis shows this option makes a difference for some fields, but only on some metrics.
3. **Comparing total output and impact to a benchmark -** Total publications and total MNCS for NZ have no international benchmark. To gain an understanding of our comparative advantage on these measures, we present each subject's share in NZ's total publications (MNCS) relative to its share in global publications (MNCS).
4. **Attributing internationally co-authored publications to countries -** In our baseline analyses NZ is given full credit for papers with a mix of NZ and non-NZ authors because this is currently MBIE's preferred method. Our view is that the most intuitive way to 'count' papers with partial NZ-authorship is to treat each such paper as a fraction of a publication, with the fraction equal to the fraction of authors who are based in NZ. Below, we show that using fractional instead of full counting tends to considerably weaken NZ's performance on bibliometric measures.
5. **Normalising citations of publications in multiple fields -** Citations of publications spanning multiple fields are normalised relative to the harmonic mean of average citations in each field. As a result, the global average of each field's MNCS is not necessarily 1. Rather, the global benchmark of 1 is only valid for publications from all fields combined. We consider an alternative treatment of publications in multiple fields that ensures the average by field is 1, but it makes little difference to results for NZ.
6. **The effect of the integer nature of citations -** Citation data are discrete and many publications may have the same number of citations. Because of these characteristics of the data, it is possible, for example, that more than 1% of publications have at least as many citations as the threshold for being in the top 1% of the distribution. Hence, in our

baseline results, we make an adjustment suggested by Waltman and Schreiber (2013) to avoid distorting results for the fraction of publications in the top percentiles. Below, we show the importance of this adjustment.

7. **Measuring uncertainty -** There is a lack of consensus on whether it makes sense to use confidence intervals on bibliometric data, which can be considered a population rather than a random sample. We discuss the arguments for and against using confidence intervals and present bootstrapped confidence intervals around some of our results.
8. **An alternative assignment of publications to fields** - The only information we have regarding the discipline to which a publication contributes is the classification by Scopus of journals into All Science Journal Classifications (ASJC). Where publications are in journals spanning multiple fields, we do not know to what extent it covers either of those fields. Furthermore, some journals are categorised by Scopus as 'General' because they are multidisciplinary. As an alternative, we explore the use of topic modelling to assign publications to fields using the words in their abstracts.
9. **Consolidation of fields -** Aggregate fields into broader groups could help to attain more narrow confidence intervals. When it comes to the topic model, consolidating fields could also be sensible because it may struggle to differentiate some fields from each other. We explore the possibility of consolidating fields based on the extent to which fields overlap in the topic model results.

The issues 1 to 9 above are explored in detail in sections 3 to 11, respectively. Before turning to this analysis, our baseline results, against which each issue is considered against, are described in section 2 below.

2 Baseline results

The baseline results use the following methodology:

- **Citation window:** All citation information available is incorporated in our results. To limit citations up to a certain period would be throwing away information.
- **Publication type:** Only publications that are articles, reviews or conference papers are included.[3]
- **Field assignment and fractional counting:** Publications in the Scopus data are assigned to ASJC categories according to the main fields covered by the source in which they were published. ASJC has 27 top-level categories which are made up of 334 sub-level categories. For publications in multiple ASJC fields, we apply equal weighting across each

[3] Not included are publications of the following types: Abstract Report, Article in Press, Book, Business Article, Chapter, Conference Review, Editorial, Erratum, Letter, Note, Patent, Report, or Short Survey.

of the sub-level categories present. For presentation, results are aggregated across top-level ASJC categories with the sub-level weightings.[4]

- **Author share:** Any publication with a NZ affiliation is counted as a full publication for NZ.
- **Normalisation of citations:** Citations are normalised using the Mean Normalised Citations Score (MNCS) formula for publication *i*:

$$MNCS_i = \frac{c_i}{e_i}$$

where c_i is the total citations received by publication *i* and e_i is the average number of citations received by all publications of the same publication year, ASJC sub-field, and publication type as publication *i*. If publication *i* is assigned to more than one ASJC field, e_i is taken as the harmonic mean of each field's average number of citations for the relevant publication year and type.[5]

- **Fraction in top percentiles:** For each ASJC sub-field, we find the threshold number of citations needed to be in the top *x* percentile, accounting for the fractional weights of publications to each field.[6] For each publication, we count it as in the top x% of a sub-field if it has at least as many citations as the threshold for that field. In addition, publications with the same number of citations as the threshold for their field are weighted so that the global fraction of publications in the top *x*% is equal to *x*%. To understand the treatment of publications in multiple fields, consider a publication that is in two fields, A and B, whose number of citations is greater than the threshold for field A but not for field B. We would count this publication as one half of a publication in each field, and that half of a publication is counted as being in the top x% for field A, but not for field B.

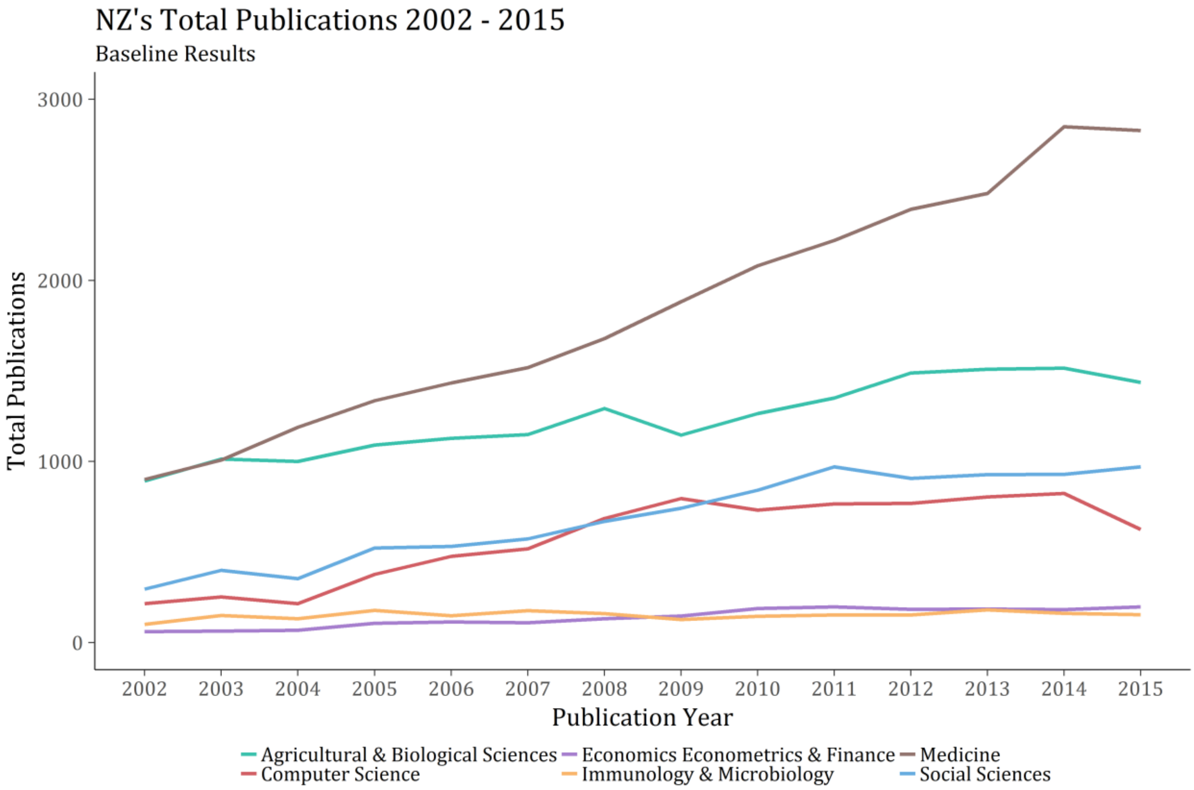
The baseline results following the above formula are presented below. Visualising results over 14 years of data and 27 fields is a challenge, so in each case we either show them for a selection of fields or of publication years.

[4] For example, a publication in a journal spanning two subjects in top-level field A and one subject in top-level field B is given a 66.67% weighting to field A and a 33.33% weighting to field B.
[5] Strictly speaking, we cannot calculate a normalized citation count for papers in a field where the average rate of citation is zero, since we normalize by dividing by the average. We get around this problem by defining the 'normalized' citation count for any paper with zero citations to be zero.
[6] For example, a publication in two fields with 10 citations counts as only half a publication in each field, but is still treated as having 10 citations in each.

2.1 Total publications

NZ's total publications over time are shown in Figure 1 below for six subjects. For most of these subjects there is an increasing trend over time, which is the most extreme for Medicine.

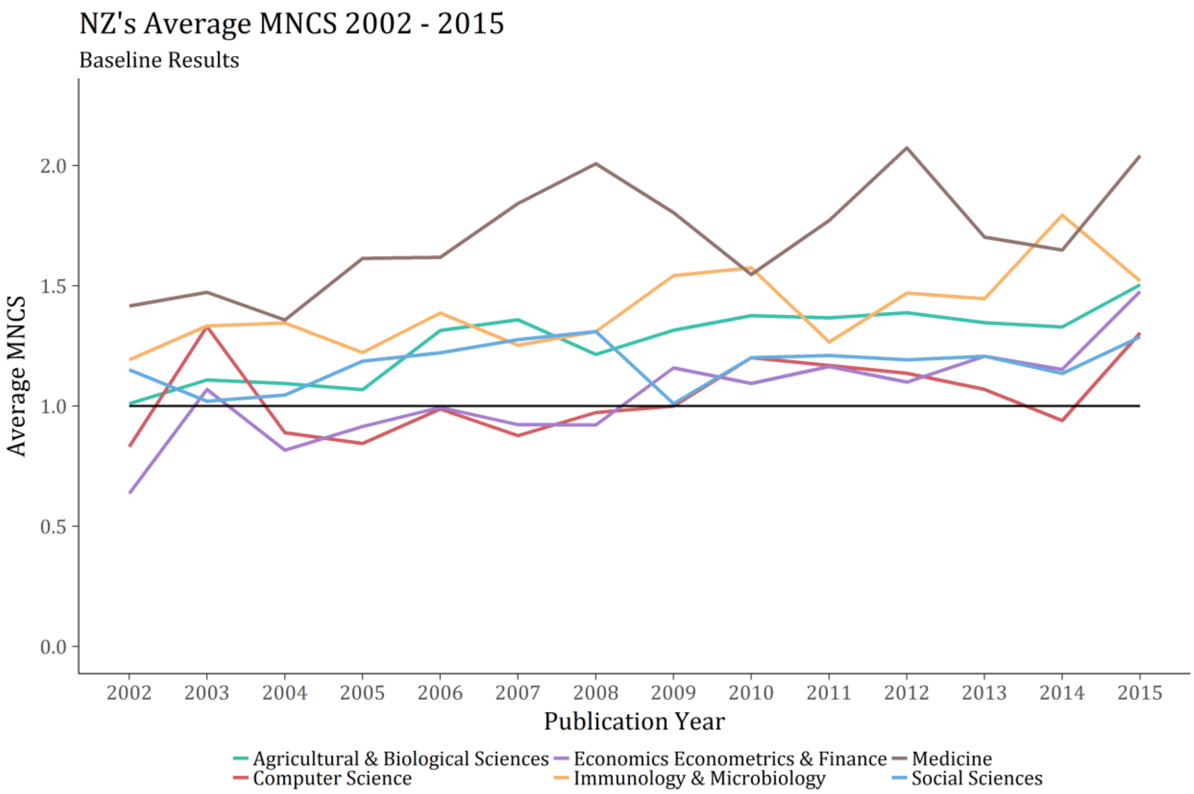
Figure 1: NZ's Total Publications 2002-2015 for Illustrative Fields – Baseline Results



2.2 Average MNCS

Figure 2 illustrates the average MNCS of NZ publications in each of the six example fields. We see that this tends to fluctuate over time, as opposed to the steady increase that we observed for total publications.

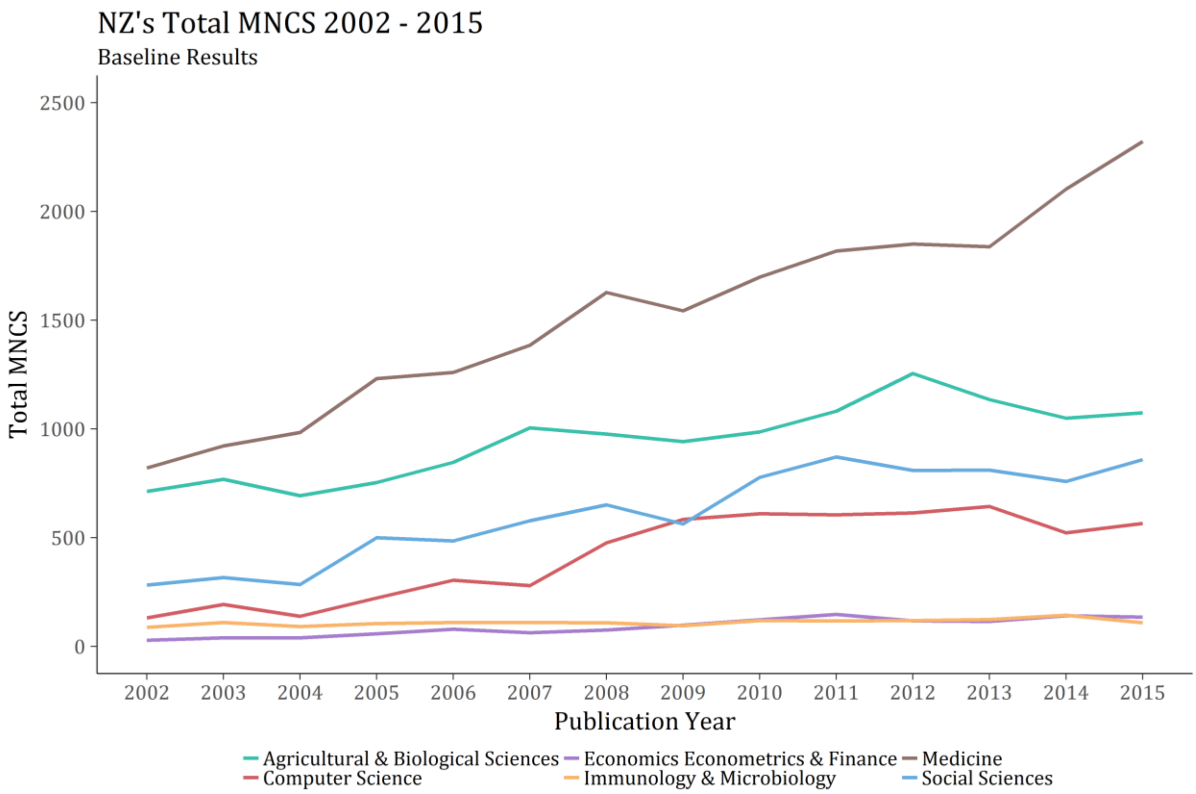
Figure 2: NZ's Average MNCS 2002-2015 for Illustrative Fields – Baseline Results



2.3 Total MNCS

Total normalised citations are set out in Figure 3 for the six illustrative fields. These trends are strongly aligned with the total number of publications in Figure 1.

Figure 3: NZ's Total MNCS 2002-2015 for Illustrative Fields – Baseline Results



2.4 Fraction of publications in top percentiles

The distribution of citations across papers is highly skewed, with the MNCS very sensitive to the presence of a small number of highly cited papers. Furthermore, there is evidence that it is indeed this small number of very-high-impact papers that are most important in the long run for advancement of science. For these reasons, it is useful to look at how NZ research performs in terms of production of papers at different points in the citation distribution.

Figure 4 to Figure 6 summarise the fraction of NZ publications in the top 50th, 10th and 1st percentiles of the world citation distribution from 2012 to 2014. In principle these figures can be interpreted as showing whether NZ scientists in a given field are producing more or fewer papers in the top of the citation distribution than their proportional share. That is, if more than 1% of NZ papers are in the top 1% of the world citation distribution, then our research is disproportionately high-impact. For this interpretation to be valid, however, it is crucial that the fraction of papers in the top 1% be calculated in a way that avoids double-counting across different groups of papers. If there is a paper with both NZ and U.S. authors that is in the top 1%, and this paper is counted as a top-1% paper for both countries, then around the world there will be more than 1% of all papers in the 'top 1%'. Similarly, there are papers that are categorized in multiple disciplines, and are in the top 1% of citations for one of their disciplines but not the others. If we count such papers as 'top 1%' then, again, more than 1% of all papers will be in the top 1%. If, overall, more than 1% of papers are assigned to this group, then NZ having more than 1% of its papers in this group does not necessarily indicate disproportionately high impact.

In Figures 4 to 6, we have avoided double-counting across disciplines by counting a paper that is in the top x% for one of its disciplines but not for others as only a fraction of a top x% paper. That is, a paper that is assigned to physics and chemistry and is top x% in physics but not in chemistry would contribute one-half of a paper to the top x% count. With respect to papers with both NZ and foreign authors, MBIE's standard practice is to count such papers as a full NZ paper, and we maintain this practice in these figures. This means that having more than x% of NZ's papers in the above-x% group is not necessarily indicative of a disproportionately high share. In Section 6 below, we explore the consequences of fractional counting of papers with both NZ and foreign authors.

While we cannot use these numbers to determine if NZ is outperforming the rest of the world in a given field, it is still informative to look at our relative performance across fields. It is interesting that our best performance is in General (the multidisciplinary category). On average, papers in this category are more highly cited than other fields, so the fact that this is the field in which NZ has the highest share of high-impact papers is a strong result showing that our very best papers are very good indeed. We will see, however, that these papers are also most likely to have foreign co-authors; when we correct for NZ's fractional contribution, our strength in this category mostly disappears (Figure Figure 11 to Figure Figure 13). We also observe over 2% of publications in the top 1% for Physics & Astronomy, Immunology & Microbiology, Earth & Planetary Sciences, Medicine, and Veterinary.

When we look at the fraction of publications in top 1%, 10%, and 50% over time, we see that the results are not particularly consistent. Comparing 2007 and 2012, we found that for many fields the results are quite different over the two years. The change is positive for some fields and negative for others, so does not appear to reflect a trend holding over NZ’s overall science system.

Figure 4: Fraction of NZ Papers in Top 50% 2012-2014 by Field – Baseline Results

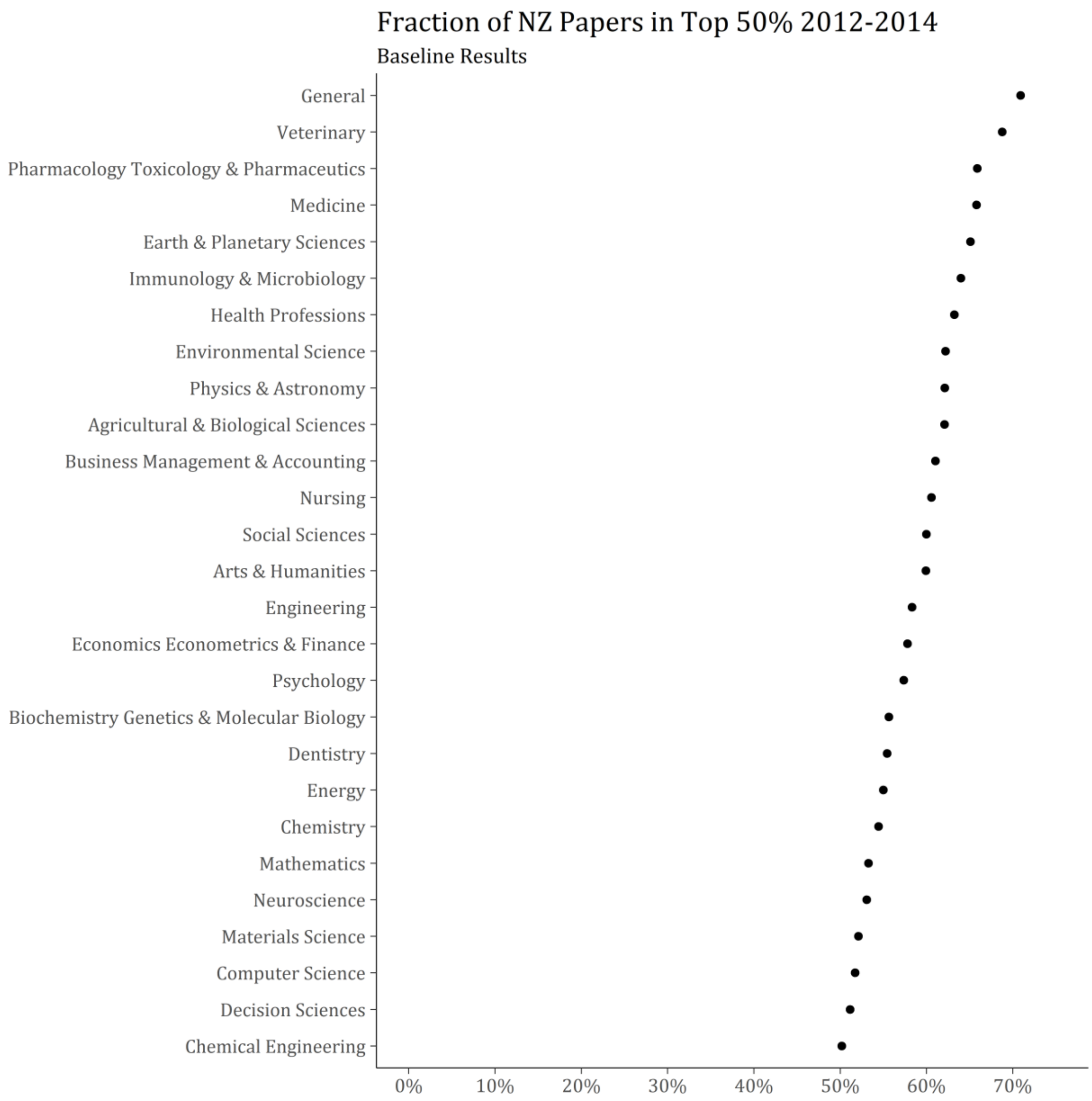


Figure 5: Fraction of NZ Papers in Top 10% 2012-2014 by Field – Baseline Results

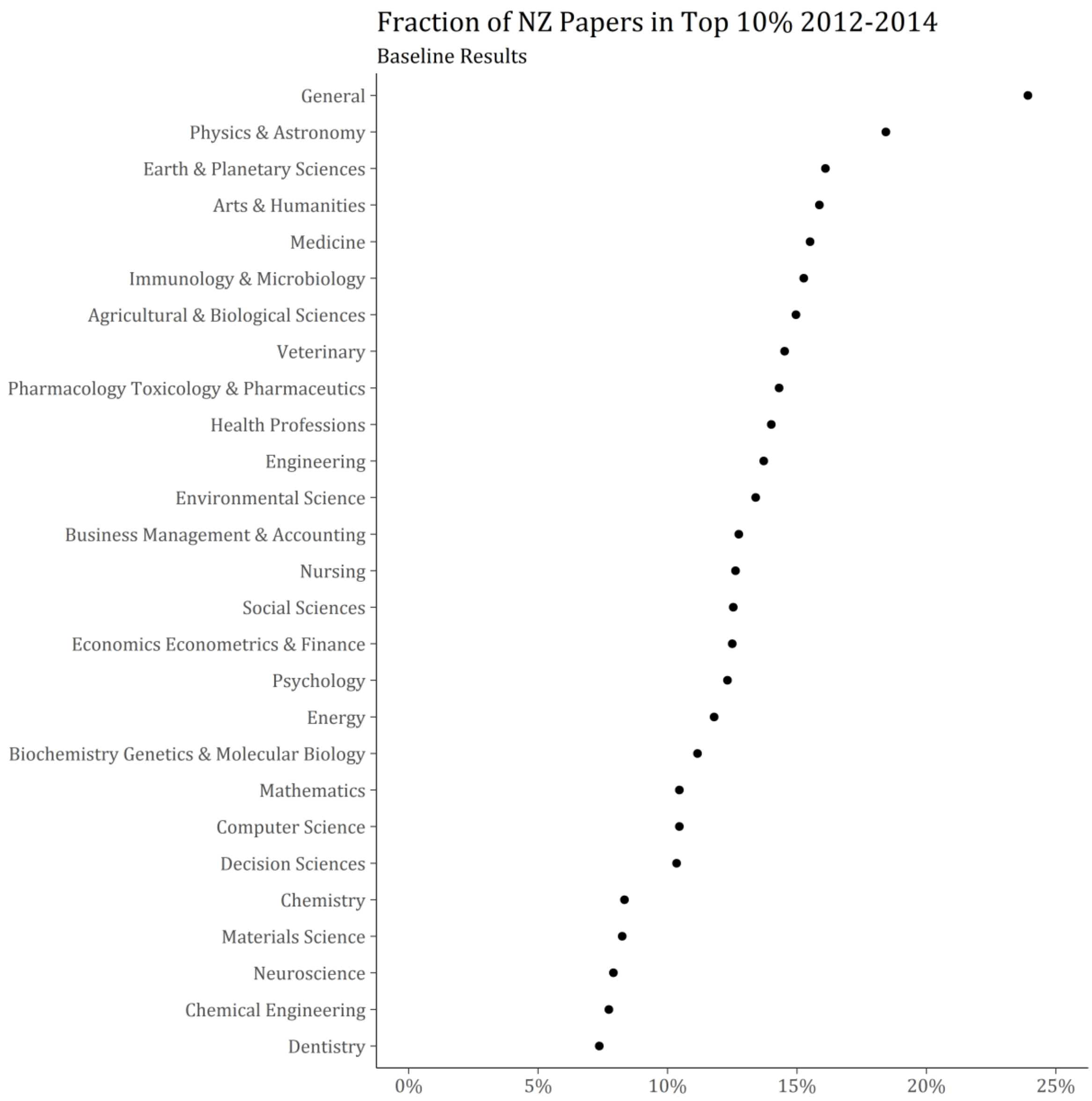
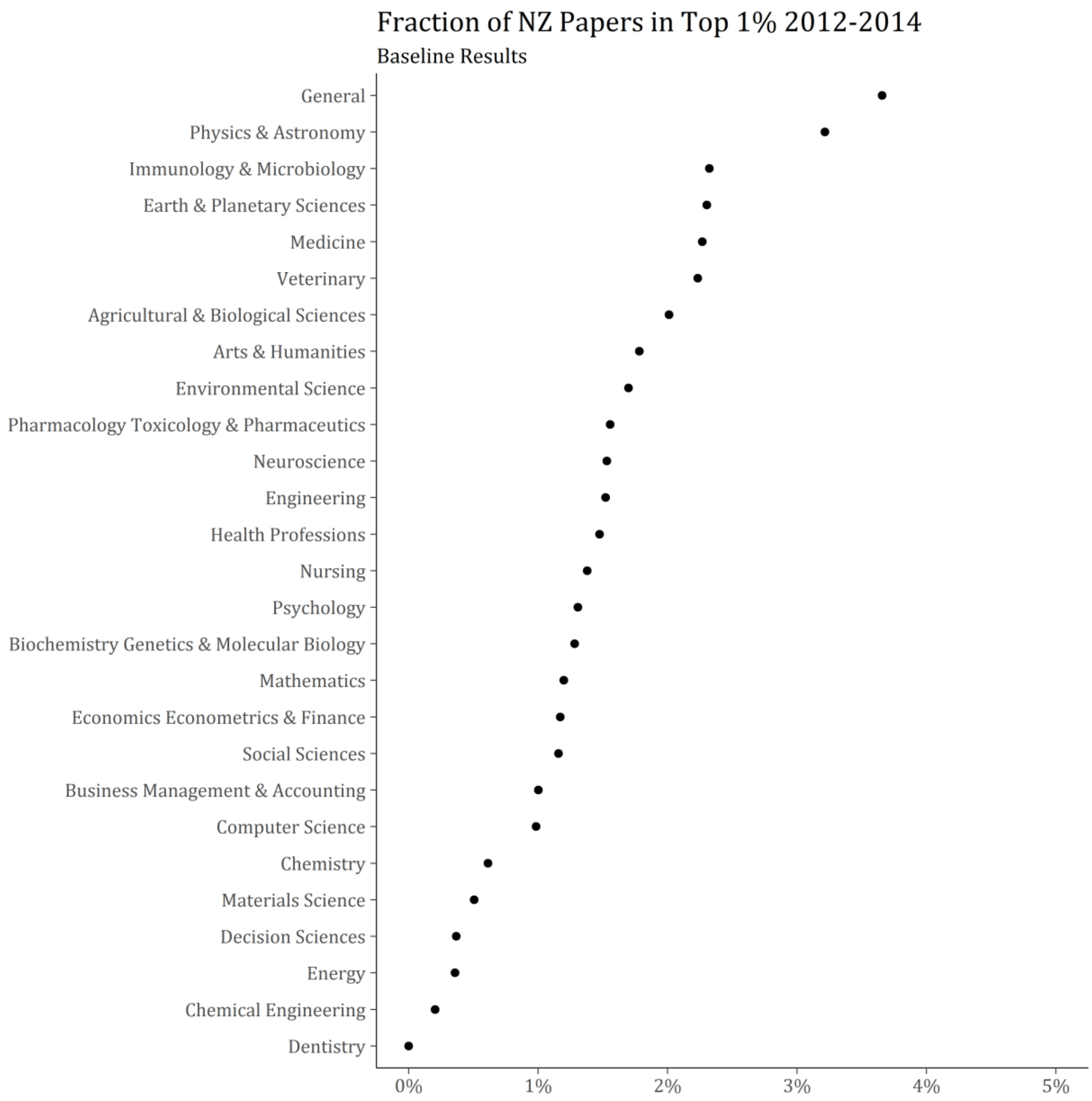


Figure 6: Fraction of NZ Papers in Top 1% 2012-2014 by Field – Baseline Results



3 Types of Publications to Include

Scopus covers a range of publication types, as listed in Table 2 below. The majority of the database is made up of articles. In our baseline results, we have included articles, conference papers, and reviews because this is the preferred choice at MBIE.

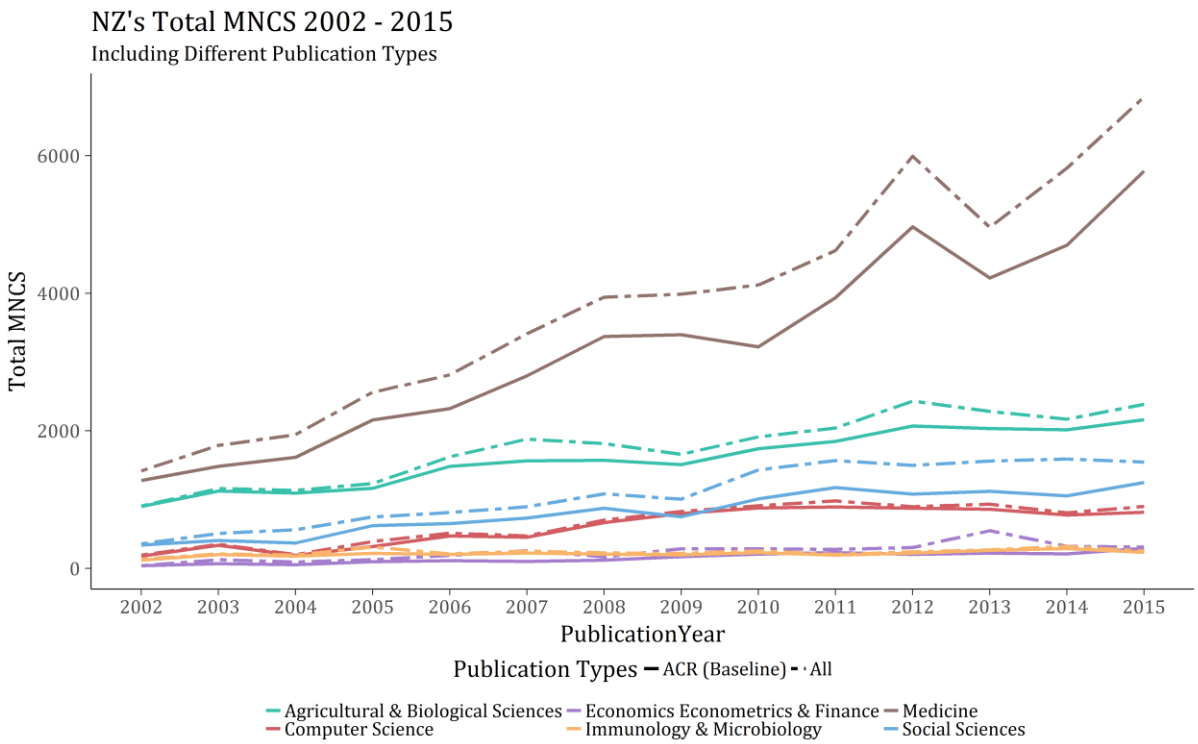
Table 2: Types of Publications in Scopus

Abstract Report	Editorial
Article	Erratum
Article in Press	Letter
Book	Note
Business Article	Patent
Chapter	Report
Conference Paper	Review
Conference Review	Short Survey

Excluding certain publication types could be unfair because different publication types are of greater importance to different disciplines. For example, the arts and humanities publish a great deal of work in books. Baseline results from including only articles, conference papers, and reviews (ACR) and results including all publication types (except errata[7]) are compared in Figure 7 for six subjects. Including all publication types does increase the total normalised citations to each field, and some fields are affected more than others (e.g. total MNCS increases a reasonable amount for Medicine and Social Sciences but hardly at all for either Computer Science or Immunology & Microbiology). Yet, the impact is small in terms of the overall trends observed, and how we view each field in relation to each other.

[7] Errata are acknowledgements and corrections of errors in previous work. We believe it is sensible to exclude this publication type from any analysis.

Figure 7: NZ’s Total MNCS 2002-2015 for Illustrative Fields – Including Different Publication Types



We also tested this comparison of publication types on NZ’s average MNCS (results not shown). Of the six example subjects, differences are minimal, except in the case of Economics, Econometrics & Finance. In many years, this field has a much higher average MNCS when including all publication types, relative to the baseline results. This confirms that the exclusion of some publication types will affect some fields more than others.

The question being asked of the data should guide the decision about which publication types to include. For example, to study new contributions to science, it could be sensible to eliminate publication types such as reviews.[8]

4 Normalisation of Citations by Publication Types

When normalising citations, standard practice is to normalise with respect to each publication type. This practice comes from the acknowledgement that different publication types are more or less likely to be cited and therefore a publication should not gain an advantage or disadvantage because of where it is published. However, we argue that not all publication types need to be distinguished for the normalisation of citations. If a publication has less outreach

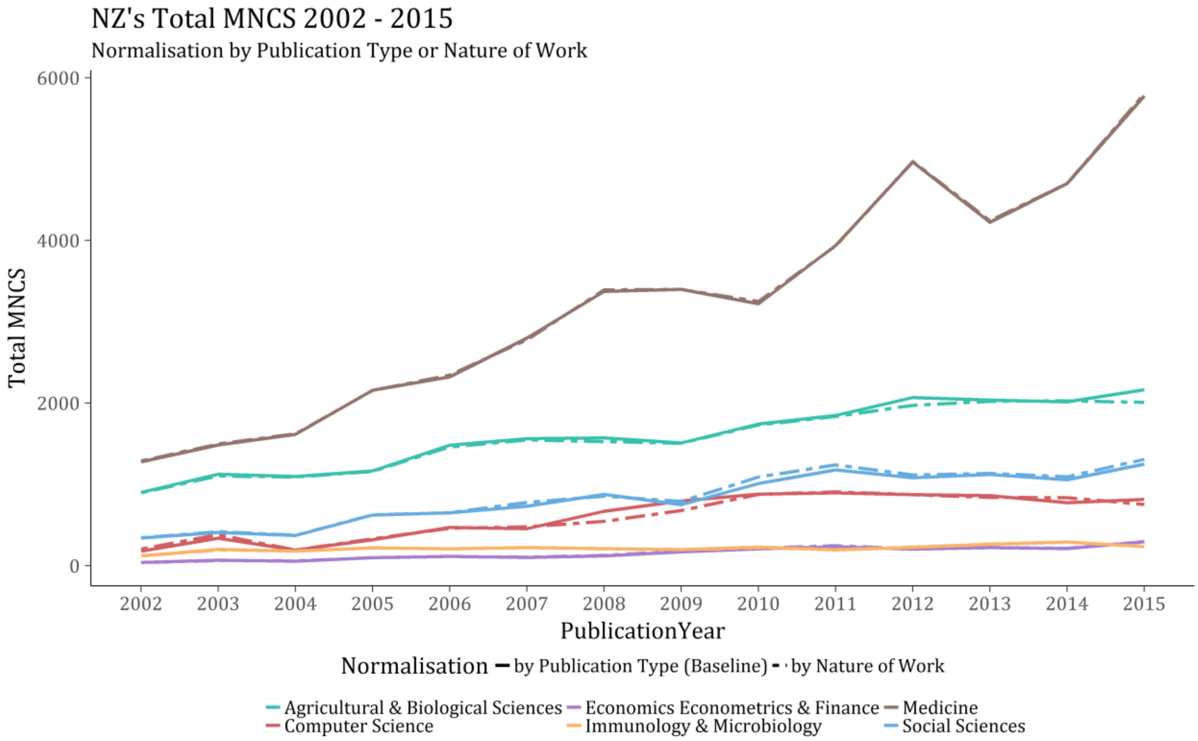
[8] Of the publication types listed in Table 2 above, we would define all categories except Conference Review, Editorial, Erratum, Letter, Review, and Short Survey as ‘new contributions’ to research. We compared our analysis using all publications to using this subset of ‘new contributions’. We found, similar to the above comparisons, that limiting the sample to new contributions has only a small effect on overall results.

because the source in which it is published is relatively inaccessible, then we see no reason for excluding that information when measuring impact. We instead suggest comparing publications of the same nature of work, categorised by whether or not they are new contributions to research. This would account for differences in the nature of work, but avoid eliminating meaningful differences in impact within those types of work.

We explored the effect of this normalisation strategy relative to our baseline results for total MNCS, average MNCS, and the fraction of publications in the top 1%. We counted articles and conference proceedings as new contributions, and reviews as a separate category. Overall, results are relatively insensitive to this choice. As shown in Figure 1Figure 8, the difference is trivial for total MNCS among the example subjects. The impact on average MNCS (not shown) is more significant, but still relatively small for the exemplar fields. The effect is substantial on the fraction of NZ publications in the top 1% for only a few fields (not shown).[9]

[9] The most notable differences, for 2002-2004, are that Arts & Humanities and Decision Sciences appear to have performed worse and Energy better under the alternative normalisation, compared to baseline results.

Figure 8: NZ's Total MNCS 2002-2015 for Illustrative Fields – Normalisation by Publication Type or Nature of Work



5 Comparing Total Output and Impact to a Benchmark

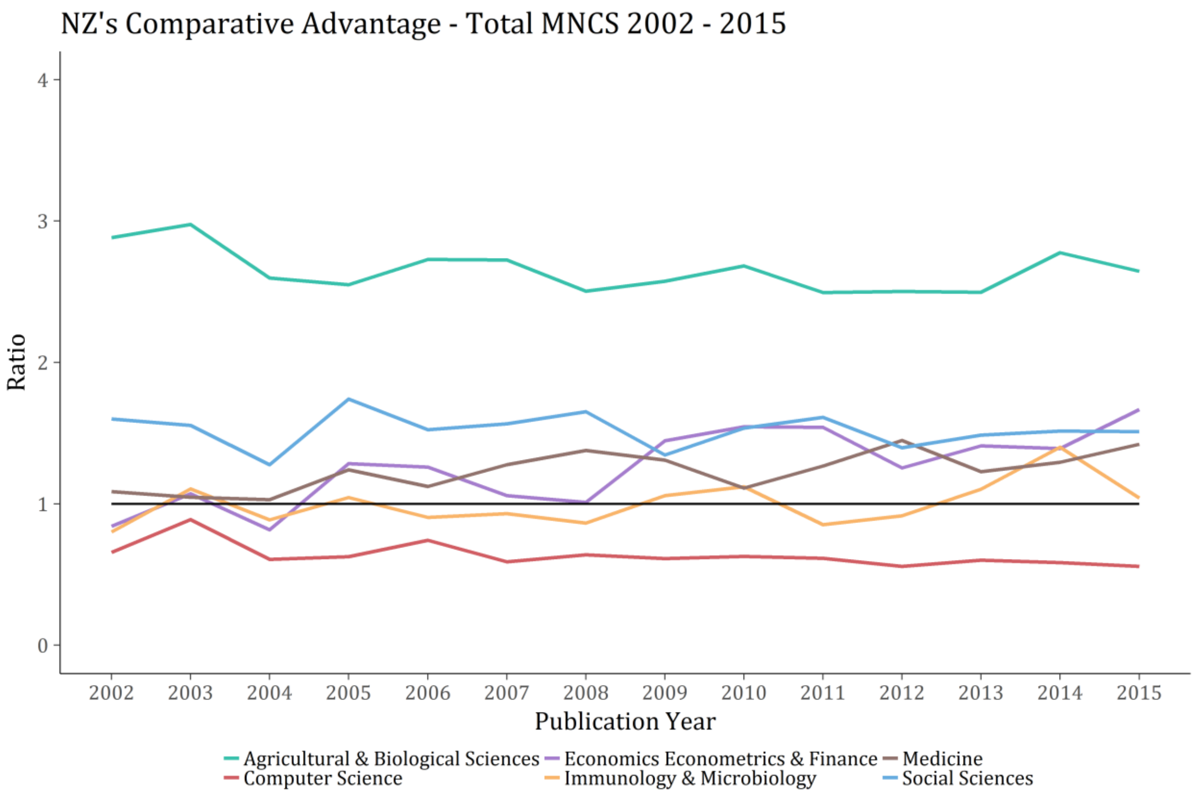
While metrics for NZ's total publications and MNCS allow us to make interesting comparisons across fields, these measures lack an international benchmark. A useful extension is to consider the ratio of each field's share in NZ's publications (MNCS) relative to its share in global publications (MNCS). MBIE (2016) refer to this ratio as NZ's comparative advantage in the Science & Innovation System Performance Report. It allows us to see, for example, whether NZ's large number of publications and citations in the field of Medicine is big in comparison to the share of global publications in this subject.

Figure 9 below shows NZ's comparative advantage in total MNCS for the six sample subjects. While in Figure 3 above we found that Medicine was the greatest contributor of these subjects to total citations, we see the proportion of our citations that come from Medicine is proportional to the share globally.[10] Meanwhile, our share in citations from Agriculture & Biological Sciences is around 2.5 to 3 times the global share each year.[11] On the other hand, NZ has relatively few citations from Computer Science compared to the world trend.

[10] We also found that the share of NZ's total publications in Medicine is similar to the share of global publications in the field.

[11] NZ's comparative advantage in terms of total publications is also high in Agricultural & Biological Sciences, but has been on a downward trend since 2002.

Figure 9: NZ’s Comparative Advantage – Total MNCS 2002-2015 for Illustrative Fields

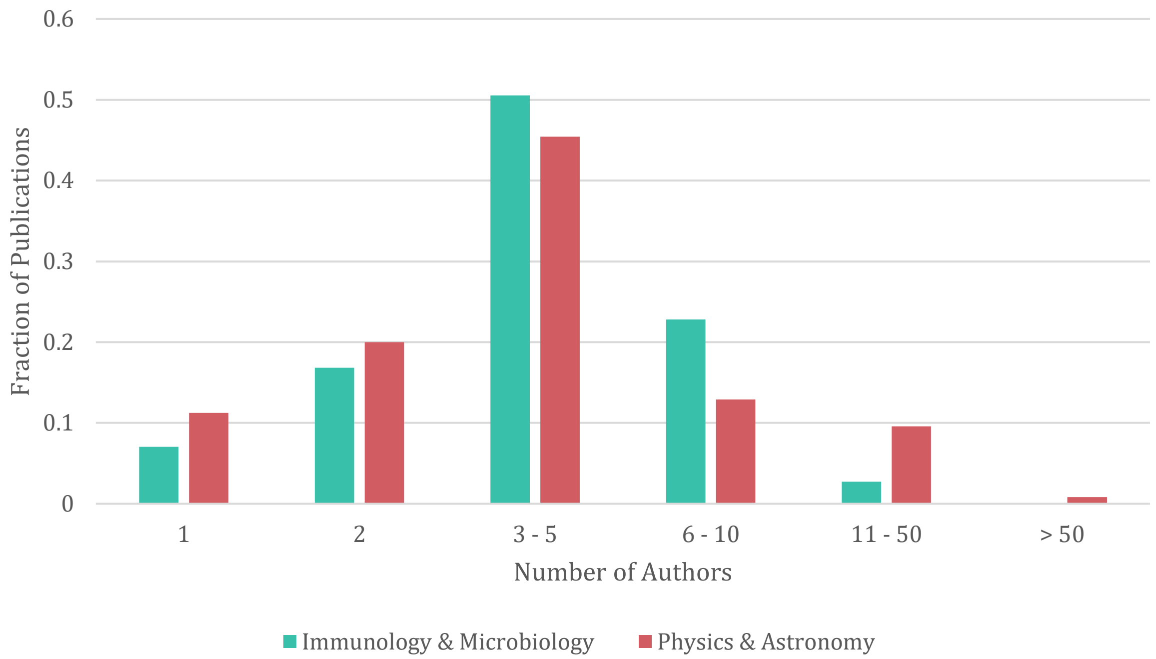


6 Attributing Internationally Co-Authored Publications to Countries.

Many scientific papers have multiple authors. As an illustration, Figure 10 below gives a sense of the distribution of number of authors for two fields with high rates of authorship.[12] Note that an issue that is often talked about - papers with dozens or scores of authors - is not a major issue. Less than 0.2% of papers have more than 50 authors. But what is a major issue is an author count in the high single digits: something like one-third of all papers in these fields have 3-5 authors and one-quarter to one-third have 6-10.

[12] In an early analysis, we found that the average number of authors per NZ publication was highest for Physics and Astronomy, with 5.23 authors per publication, followed by Immunology & Microbiology with 4.37 authors per publication 2002. This result was calculated on all publication types with no fractional counting for ASJC field.

Figure 10: Number of Authors per NZ Publication for Illustrative Fields, 2002



If these authors are divided between NZ and the rest of the world, how the NZ contribution to the paper is tabulated could make a significant difference. Fully counting publications with both NZ and non-NZ authors could be supported by the argument that even if NZ only contributes a fraction of the work toward a publication, it absorbs the full amount of knowledge that is generated. However, if the point of interest is how much NZ *contributed* to the scientific output, then full counting would overstate NZ's performance. Even if metrics using full counting are sensible for the research question being asked, as explained in Section 2.4 it may not be fair to compare them to a world benchmark that is based on equal weighting of each publication.

Our view is that the most sensible way to 'count' papers with partial NZ-authorship is to treat each such paper as a fraction of a publication, with the fraction equal to the fraction of authors who are based in NZ.[13] Aksnes, Schneider, and Gunnarsson (2012) compare citations scores with full and fractional counting for 23 countries, and find that all 23 countries have lower scores when publications are counted fractionally. This result suggests that internationally co-authored publications tend to be relatively highly cited.[14]

[13] Another possibility is to allocate credits to authors depending on their position in the by-line, but this is not always appropriate (e.g. where alphabetic ordering is used) and is still only an approximation of the contribution of each author.

[14] If this is the case, then if we were to weight publications N times for the number N of countries that they are associated with, then more than 1% of global publications would be in the top 1% of the citation distribution. By this reasoning, the world benchmark is not sensible when full counting is used.

By definition, fractional counting of authors reduces NZ's total publications and total MNCS for all fields. But we also find it has a significant negative effect on average MNCS for all six example fields (results not shown). This implies that NZ's publications with international co-authors tend to be more highly cited than our other publications in these fields.

Comparing the fraction of publications in the top 50% with full and fractional counting of authors further reinforces the above conclusion. These results are presented in Figure 13. As discussed in Section 2.4, the interpretation of the fraction of publications in the top x% to a global benchmark of x% is valid when fractional counting of both authors and disciplines is used. Hence, we display a black vertical line in representing the global benchmark for fractionally counted results in the figures below that display fractions in the top percentiles.

Figure 11 shows that the fraction above the median is reduced for all subjects when fractionalising the weights for the share of authors from NZ. The difference between full and fractional counting of authors is even more striking when it comes to the share of NZ papers in the top 10% and 1%, as illustrated in Figures 12 and 13. These figures highlight the extent to which the interpretation of results with full counting to the standard benchmark is misleading. In Figure 12 for example, the General category appears to have 24% of publications in the top 10% with full counting, which would could be mistakenly compared to a benchmark of 10%. With fractional counting, we show that in fact NZ can only claim to have 8% of General publications in the top 10%, implying that NZ underperforms on this measure compared to the world at large.

The effect is particularly large on fields that, with full counting, appear to be top performers on this metric. Indeed, once this correction is made, NZ stands out as having more than 1% of its papers in the top 1% in only nine of the 27 disciplines.. To reiterate, it may be appropriate for some purposes to count fractional-author NZ papers as a 'full' paper, but if one wants to count this way one must then compare to a correspondingly appropriate benchmark. Since counting this way puts well more than 1% of papers worldwide in the 'top 1%', one must use a benchmark greater than 1% to identify above-average performance.

Figure 11: Fraction of NZ Papers in Top 50% in 2012-2014 by Field – Full and Fractional Counting of Authors

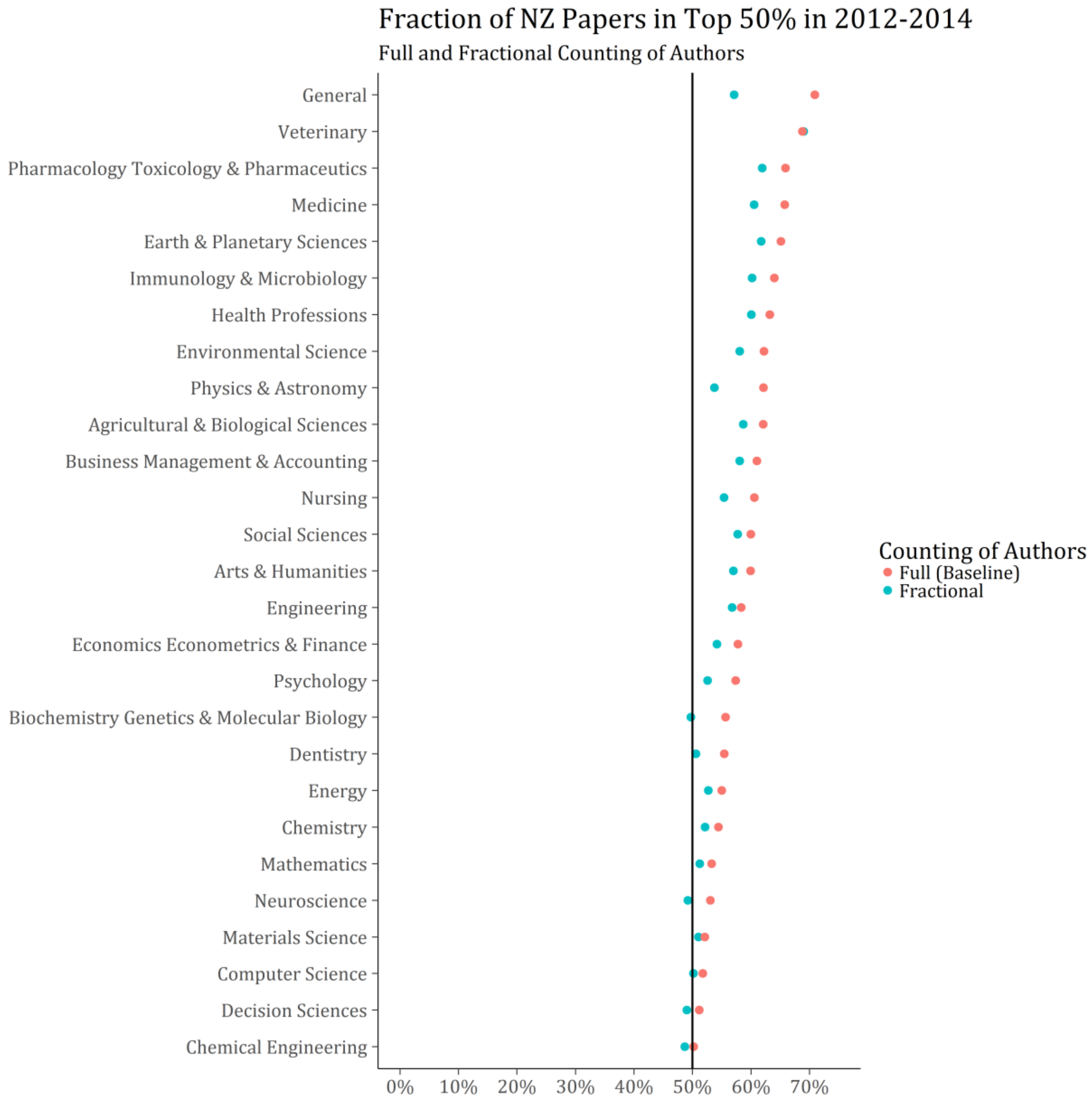


Figure 12: Fraction of NZ Papers in Top 10% in 2012-2014 by Field – Full and Fractional Counting of Authors

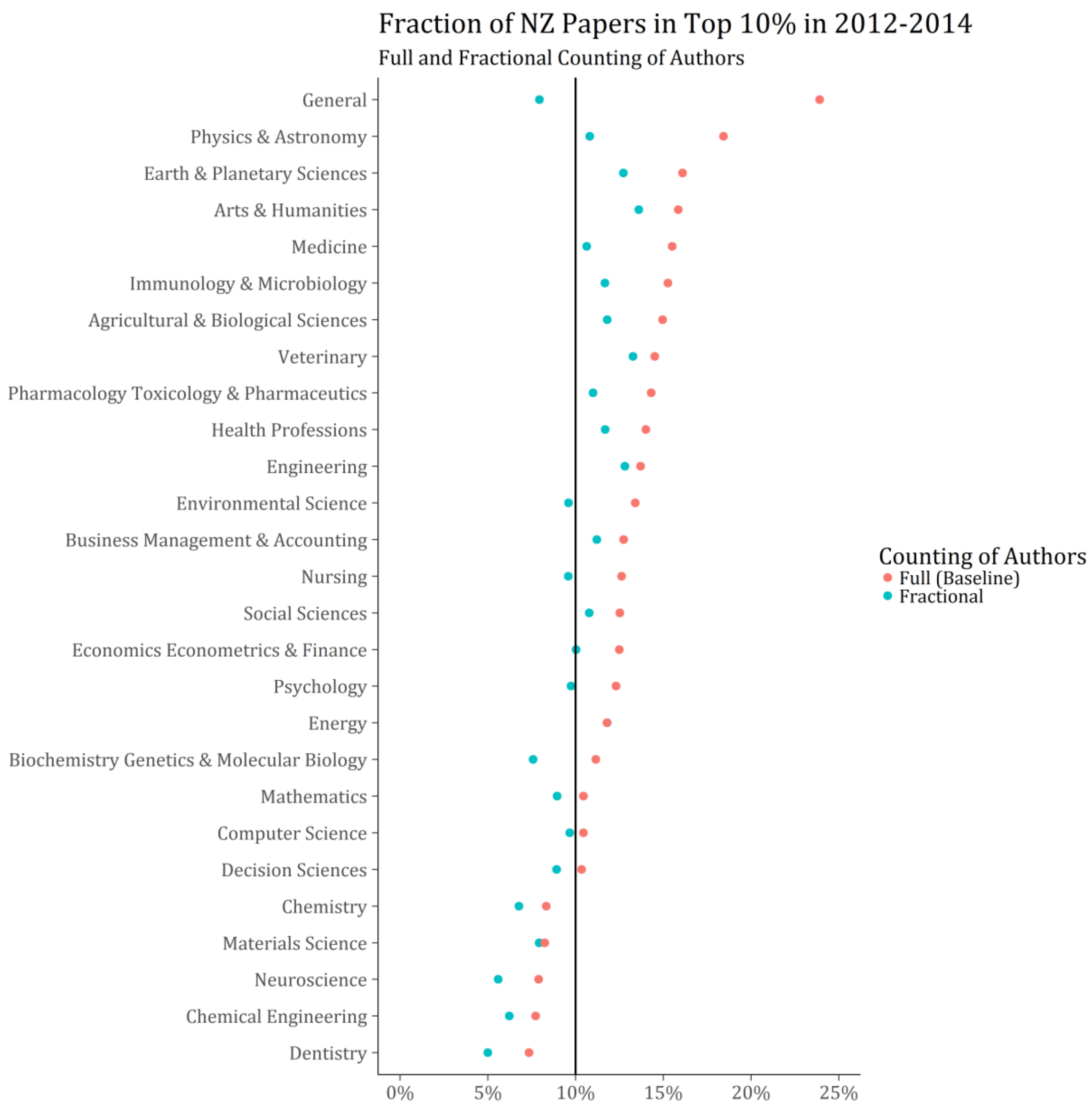
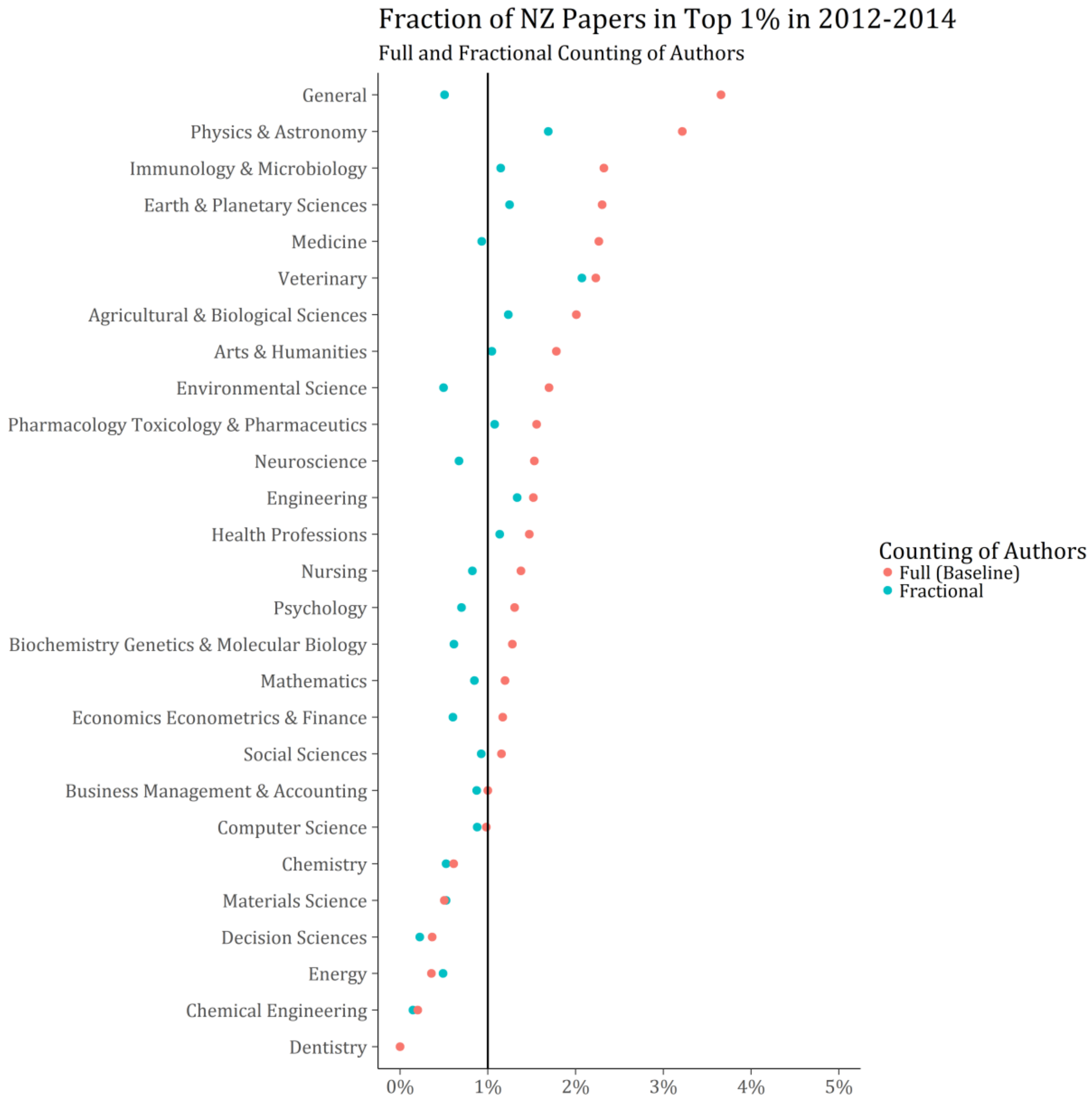
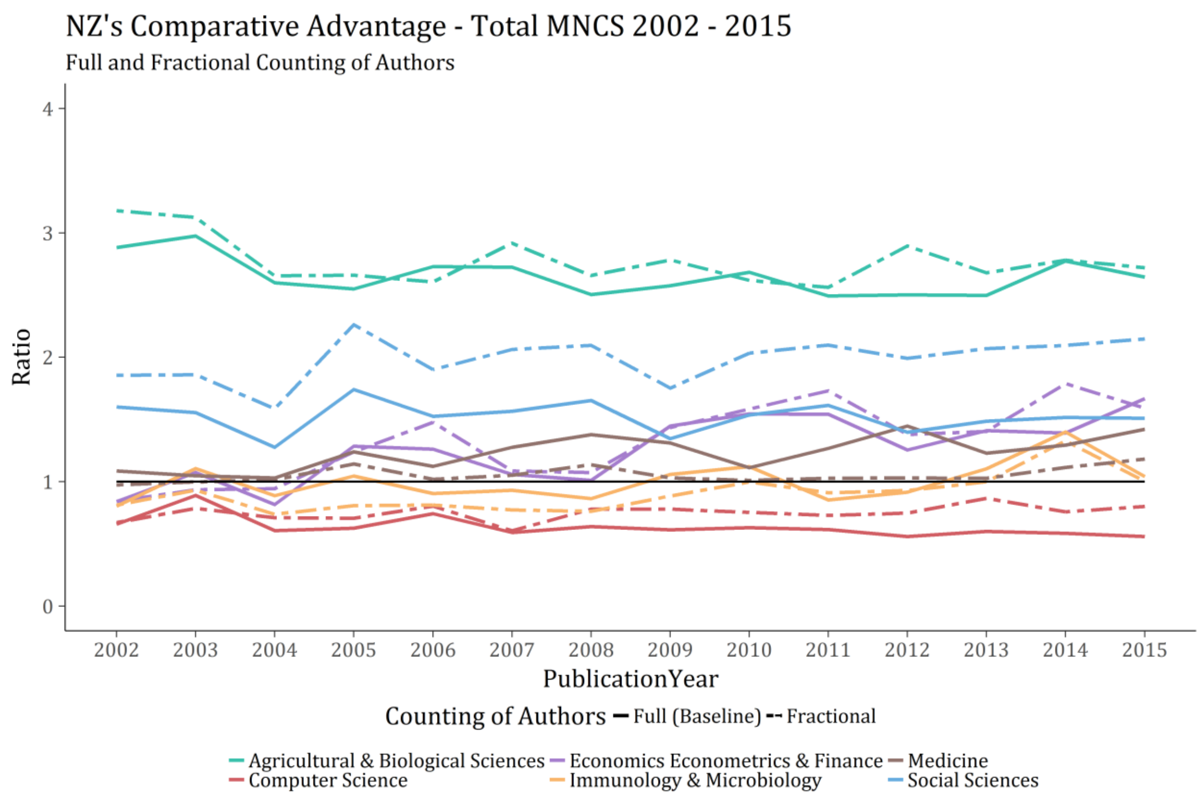


Figure 13: Fraction of NZ Papers in Top 1% in 2012-2014 by Field – Full and Fractional Counting of Authors



In Figure 14 we compare NZ's comparative advantage of total MNCS with and without fractional counting of authors. In some cases, NZ's relative advantage in MNCS is improved with fractional counting, while in others it is worse. Note that for Social Sciences, our relative share based on fractional counting consistently exceeds the relative share of MNCS based on full counting. This means that in this field, NZ researchers being a small fraction of an international team is less common than in other fields. Conversely, for Medicine, the dotted line is consistently above the solid line, meaning that NZ researchers are more often part of a larger international team in this field. Hence if we believe that NZ scientists who are a small part of a large international team are probably having less impact per publication, all else equal, than NZ scientists who comprise all or most of the authors on a paper, it is important to count papers fractionally.

Figure 14: NZ's Comparative Advantage – Total MNCS 2002-2015 for Illustrative Fields – Full and Fractional Counting of Authors



7 Normalising Citations of Publications in Multiple Fields

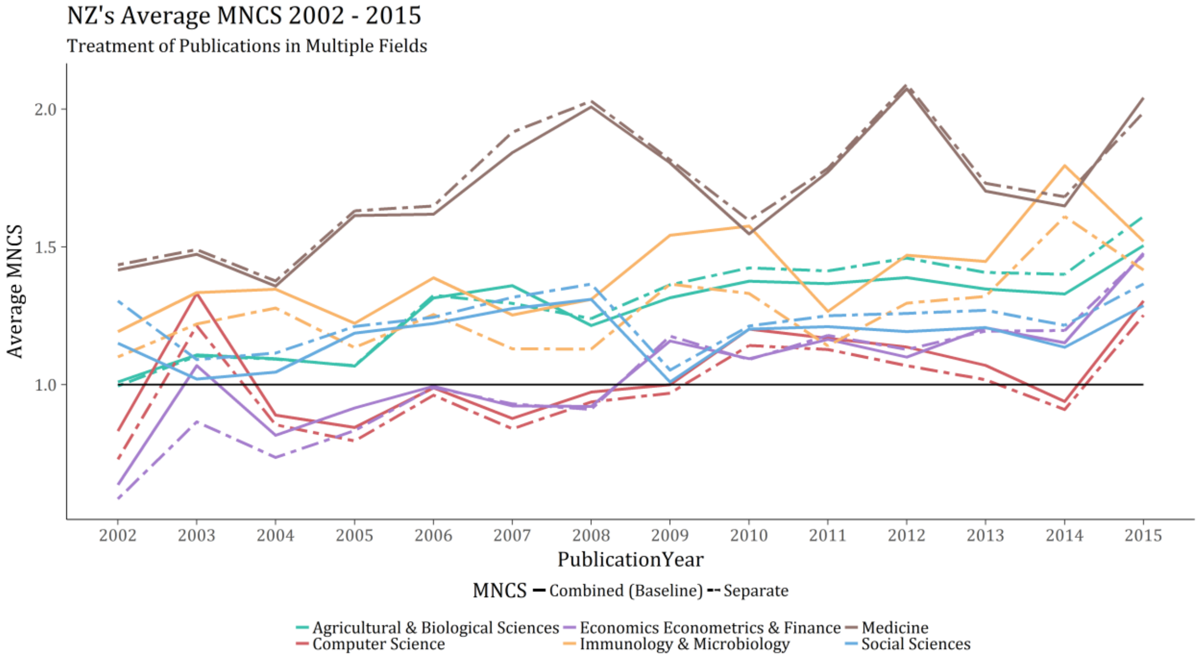
7.1 MNCS

In the baseline methodology, the MNCS of a paper in multiple fields is taken as the harmonic mean of its normalised citation score in each ASJC field. Therefore, the MNCS is the same in each field for the same paper. A problem with this method is that average MNCS over all publications in each field does not necessarily equal one, so the average MNCS for each field for NZ cannot be interpreted relative to this benchmark. We can only compare to a benchmark of one in the sense that it is the average MNCS for all publications in the world over all fields.

One way around this would be to renormalise each field's average MNCS for NZ against the average MNCS for global publications. However, we feel this would be overly complicated to explain to readers of the results. We propose an alternative way of thinking about publications in multiple fields, which is to think of the fraction in each field as a separate paper. Therefore, rather than taking the harmonic mean of the MNCS over each field, we can simply maintain the MNCS calculated for that specific field.

This proposed methodology makes little difference to NZ's total MNCS relative to baseline results (not shown). However, there are some clear differences to average MNCS for some fields, as shown in Figure 15 below. NZ's average MNCS in Immunology & Microbiology is consistently lower when publications are treated separately, suggesting that it tends to benefit by being paired with less highly cited fields in the baseline methodology.

Figure 15: NZ's Average MNCS 2002-2015 for Illustrative Fields – Treatment of Publications in Multiple Fields



If publications only ever belong to one of the fields covered by the source they are published in, then the proposed methodology provides the best guess for each of the potential fields, though we cannot identify which is the correct one. Conversely, if we think that a publication covers all of the fields, then the separate normalisation does not take into account the fact that the number of citations has likely been influenced by the citation practices of the other fields it belongs to. With the latter concern in mind, treating publications separately across fields is not the perfect solution to the issue at hand. In reality, both of these circumstances are probably true depending on the publication, so it is reassuring that we see little variation in the results using the two alternatives.

7.2 Fraction of publications in top percentiles

When it comes to our methodology for the fraction of papers in the top percentiles, we compare publications to the citation thresholds of each ASJC field separately. This is similar to the alternative we propose for MNCS in section 8.1, and similarly does not account for the fact that citations may be influenced by the citation practices of the other fields it belongs.

MBIE's preferred method combines the effect of each field, by fractionalising the contribution to the top x% over all fields. For example, if a paper is in two fields, and it is in the top 1% of the citation distribution for only one of those fields, then it would count as half of a paper in each field, of which a half of that half is counted as being in the top 1%. While this is not a variation that we have tested on our results, we warn that it will have similar implications to using the harmonic mean for MNCS. That is, the fraction of global publications in an ASJC field that are in the top 1% will not necessarily add to 1%, and so the benchmark is not field-specific.

8 Distribution of Citations

8.1 Characterization of the world distribution

The distribution of citations is highly skewed, with a large number of publications having zero or one citations, especially among those from the most recent years for which there has been less time for citations to accrue. Concurringly, the thresholds to be in the top percentiles lie relatively close to zero for many fields. Figure 16 to Figure 18 below show the threshold number of citations needed to be in the top 50th, 10th, and 1st percentiles respectively, for each group of similar publications in 2007 and 2012. Groups of similar publications are those published in the same year, of the same publication type and in the same sub-level ASJC field. We see that bulk of these citation thresholds are close to zero. The skew toward zero is stronger in 2012 than 2007 because publications have had less time to accrue citations.

Figure 16: Thresholds for 50th Percentile of Citation Distribution in 2007 and 2012

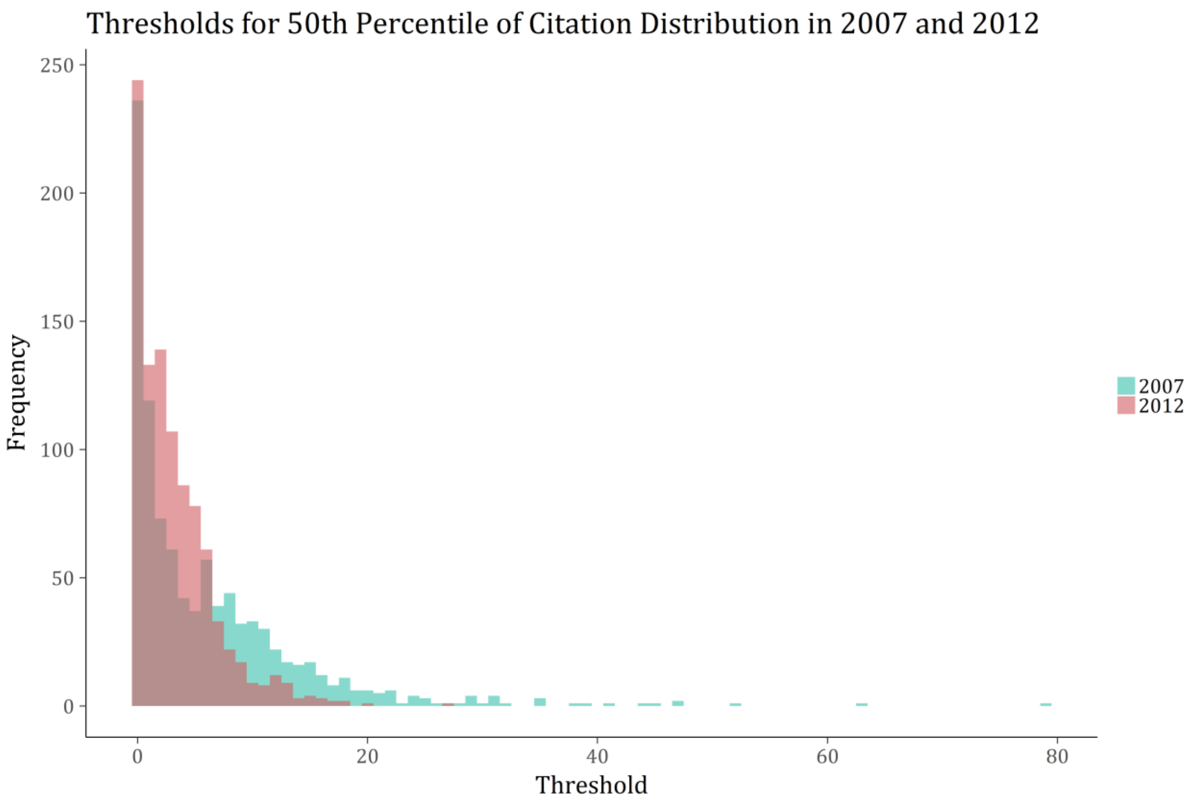


Figure 17: Thresholds for 10th Percentile of Citation Distribution in 2007 and 2012

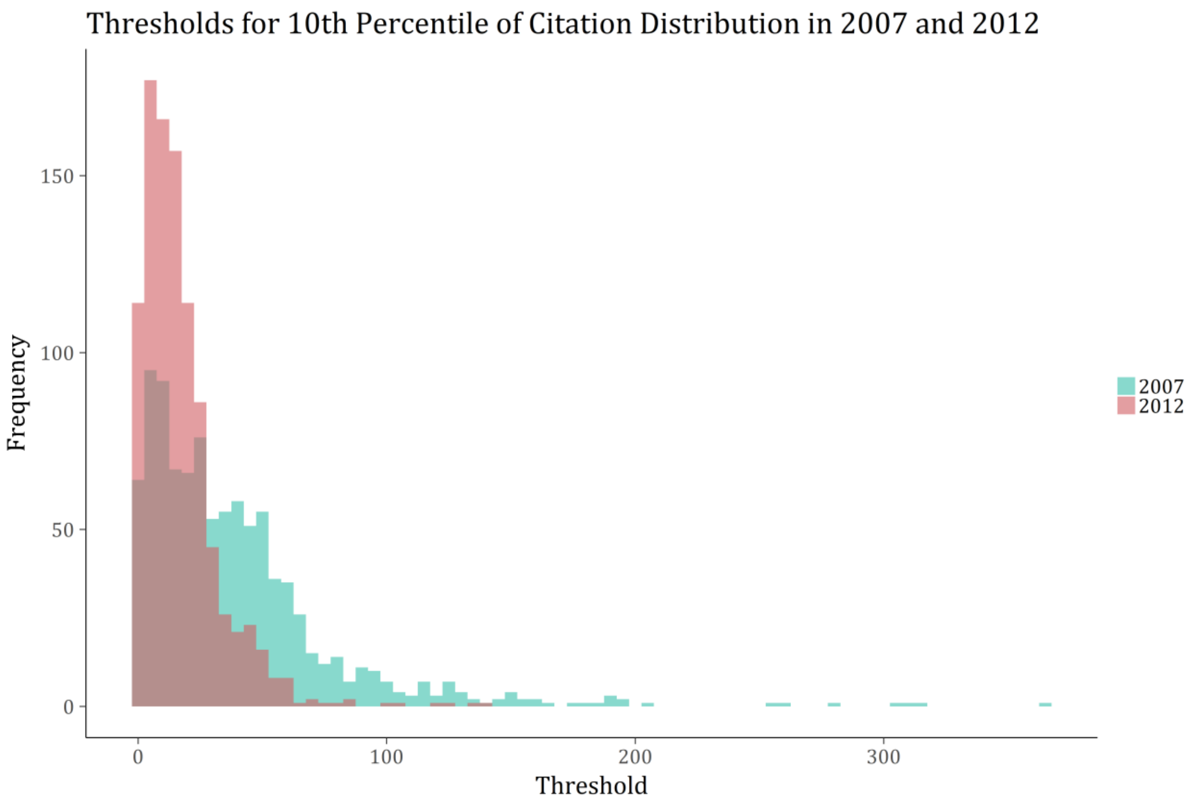
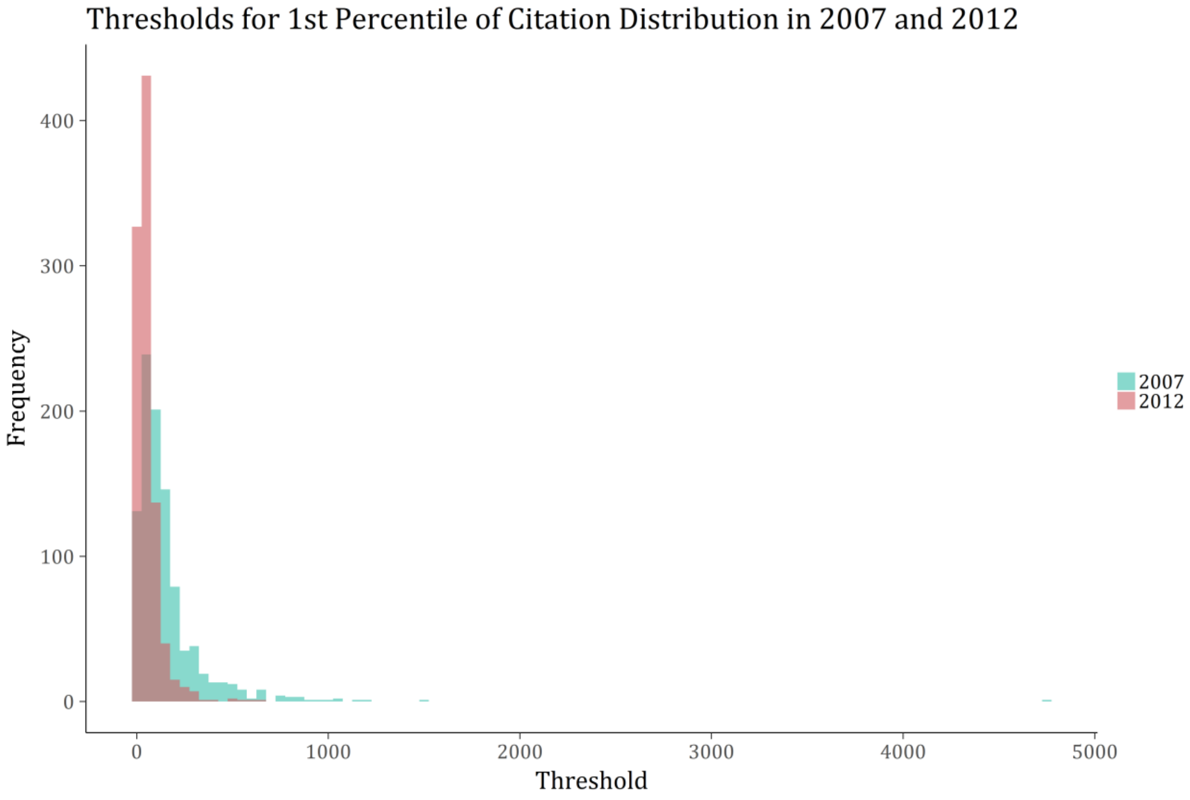


Figure 18: Thresholds for 1st Percentile of Citation Distribution in 2007 and 2012



8.2 The effect of the integer nature of citations

The discrete nature of the citation distribution, and the fact that many publications have the same number of citations, has implications that could bias the percentile measures upwards. To understand this problem, consider an example. Suppose a field has 100 publications, in which the top eight publications have 20 citations, and the top 12 publications have 19 citations. The threshold to be in the top 10% of the distribution is 19. However, 12% of publications have 19 citations.

In our baseline results, we deal with the above problem using an adjustment suggested by Waltman and Schreiber (2013). Fractional weights are applied to publications with citations equal to the threshold for their group. That is, for publications with the same number of citations as the threshold, we count them fractionally such that the fraction of global publications in the top *x*% indeed adds up to *x*%.

Though we have incorporated a solution in the baseline results, it is interesting to show the importance of this adjustment. In Figure 19 below, the effect of fractional counting at the threshold on the fraction of NZ papers in the top 1% is compared for two illustrative fields over time (baseline results use fractional counting). We see that the positive bias from full counting comes becomes substantial in more recent years - in around 2009 for Arts & Humanities and 2014 for Agricultural & Biological Sciences. This is because the citation distributions tend to be more condensed in more recent years, when citations have had only little time to accrue. In

Figure 20 the difference made by fractional counting of publications at the threshold is shown for all fields in calculating the fraction of publications in the top 1% for 2012-2014.

Figure 19: Fraction of NZ Papers in Top 1% 2002-2015 for Illustrative Fields - Full and Fractional Counting at Threshold

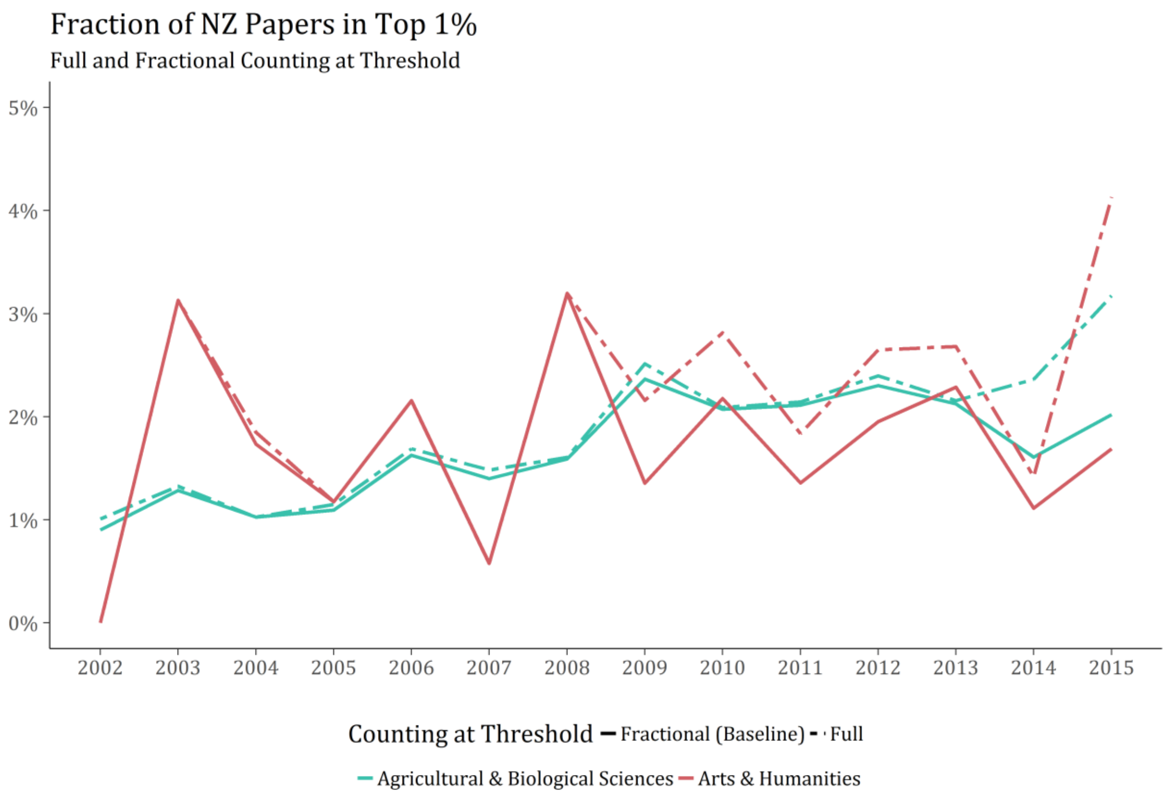
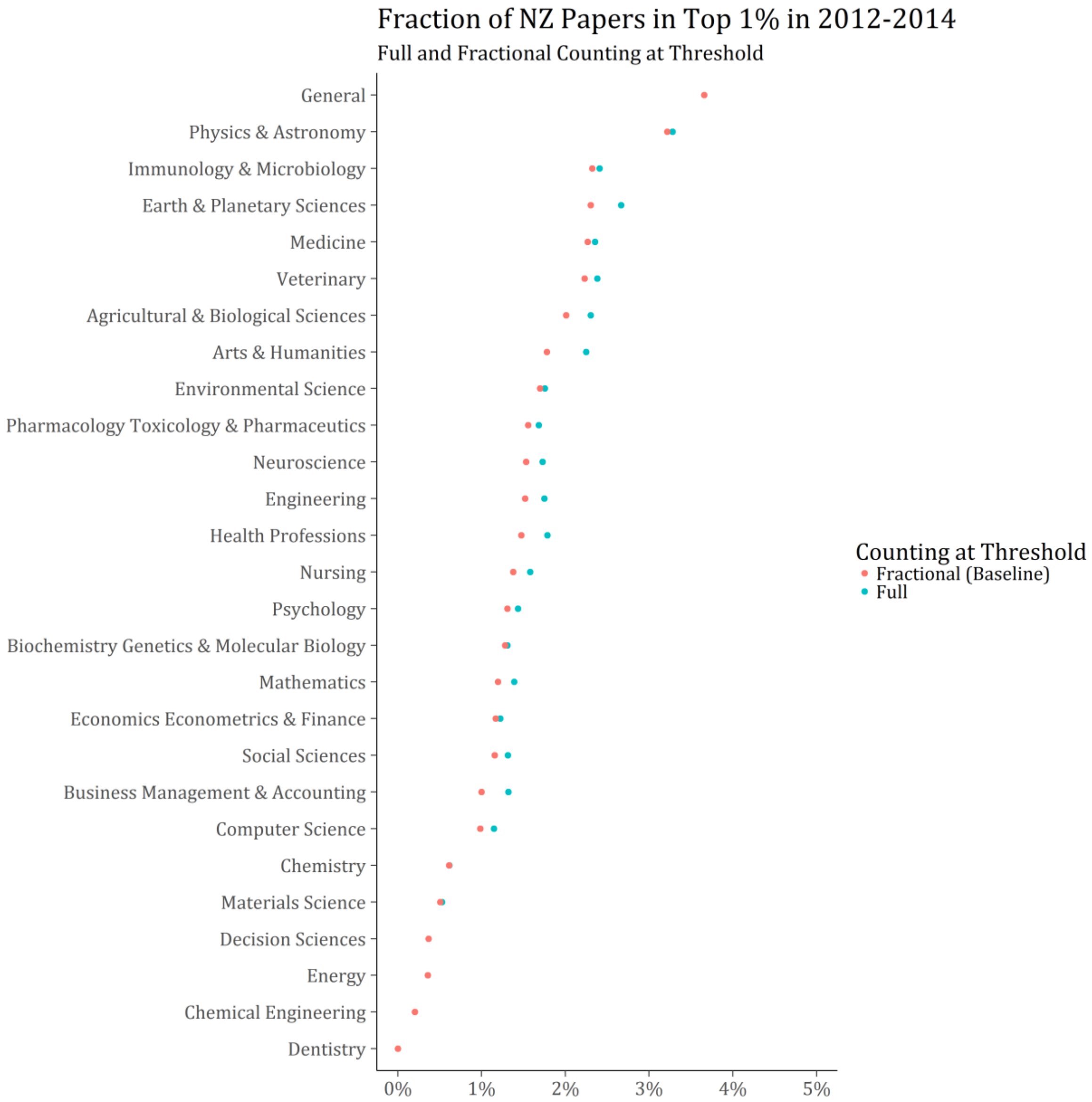


Figure 20: Fraction of NZ Papers in Top 1% in 2012-2014 by Field – Full and Fractional Counting at Threshold



9 Measuring Uncertainty

9.1 Interpretation of confidence intervals for bibliometric measures calculated on the population of papers

In classical statistics, we observe data that are selected, either randomly or non-randomly, from some larger population. This leads to the question of how confident we should be that parameters that are calculated using this sample are correct for the larger population from which the sample was drawn. 'Standard errors' and 'confidence intervals' are calculated and used to quantify the uncertainty about the true values of the parameters in the underlying population.

For the bibliometric analyses described herein, it is not clear whether the dataset is a sample from some larger population, or the data we have constitute the relevant population. Thus, when we calculate the MNCS for a discipline or for NZ, or the fraction of papers that exceed the 1% threshold, it can be argued that there is no inherent uncertainty associated with these estimates, because we calculate them based on the complete population of papers. They are arithmetically accurate and perfectly precise measures (Schneider 2016).

Other analysts have proposed that even though the data represent, by some definition, the complete population, it is more appropriate to think of them as samples from a hypothetical population. We can, for example, think of the neuroscience papers published in NZ in a given year as 'standing in' for different contexts, such as all of the papers that might have been written by NZ scientists working in neuroscience, or the set of papers that would be classified as neuroscience under different classification schemes. Or, more profoundly, we could argue that we really do not care at all about papers per se, what we care about is new knowledge, and the papers imperfectly and somewhat randomly correspond to the 'population' of bits of new knowledge. Under this sample-from-a-hypothetical-population view of our data, it is appropriate to explore the degree of confidence that is appropriate in using estimates from those data (Williams and Bornmann 2016a).

The problem with this approach is that strictly speaking analyses predicated on this model of the world must specify the precise mechanism that connects our data to the hypothetical population it is representing. And we rarely if ever are able to do that. In the absence of a precise model of the relationship between the data and the hypothetical population, there is no truly rigorous way to determine the correct formula or method to estimate standard errors or confidence intervals (Waltman 2016b).

In reality, however, it is rarely the case in any statistical analysis that the model of the relationship between the data and the population is completely specified. In other words, the charge that the calculation of confidence intervals for bibliometrics lack adequate conceptual foundation could be levied at the vast majority of statistical analyses carried out (Williams and Bornmann 2016b).

In the case at hand, we wish to draw inferences about the magnitude and quality of science in NZ, how these vary across disciplines and time, and how they compare to other countries. We do not, in fact, know exactly how our measures derived from bibliometrics correspond to those underlying concepts of interest. We cannot, therefore, derive rigorously the appropriate measures of our uncertainty around the accuracy of our measures. This leaves us with a choice: we can calculate confidence intervals based on an intuitively sensible and widely used method (bootstrapping) even though we cannot prove that this is really the right method. Or we can eschew the use of confidence intervals and simply report the calculated statistics.

The one thing we know for sure is that there is some degree of uncertainty about the accuracy of our measures in terms of their relationship to the underlying concepts. It seems perverse to use the fact that the precise relationship between the measures and the concepts cannot be specified as a reason to avoid presenting any indication of the uncertainty. Therefore, we present our main results for MNCS and fractions of papers above various citation thresholds with confidence intervals. These confidence intervals should not be interpreted literally as a precise probability statement about whether the 'true' value of a given concept is above or below some given number. Rather, they should be viewed as approximate indicators of our degree of confidence in the results. As long as they are looked at this way, we believe that including confidence intervals is both more honest and more useful for decision makers than presenting point estimates without confidence intervals.

9.2 Use of bootstrapping to calculate confidence intervals

Bootstrapping is a method for calculating standard errors and confidence intervals for calculated parameters whose underlying statistical properties are not known. Bootstrapping methods work by repeatedly and randomly varying the data used in the calculations, and simply measuring how the calculated parameters change as the data are varied. One draws multiple random samples from the data (sampling with replacement), and calculates the parameter of interest on each sample dataset. This repeated calculation yields a range of estimates. There are various ways to construct confidence intervals from this range; the one we use is the percentile method – taking the 10th and 90th percentiles of the calculated statistics as the upper and lower bounds of the 90% confidence interval.

As discussed in the previous section, the statistical meaning of such confidence intervals in our context is not clear. We are not suggesting that the confidence intervals reported below can or should be interpreted in a statistical or hypothesis-testing sense. They are intended only as qualitative indicators of the robustness of the calculated figures. We suggest that where these ranges are large, the point estimates should be viewed as not very robust, in the sense that they are likely quite sensitive to exactly how groups of outputs are defined over time and disciplines, and to movement of particular researchers in and out of the underlying populations.[15]

After our work was mostly completed, we attended a presentation by the MBIE Science and Innovation Trends Team that proposed a related concept of 'stability' intervals for bibliometric measures. This presentation showed that their method for constructing these

[15] In our bootstrapping calculations, we have taken the cutoff values (derived from the world population) of papers) as given, rather than also sampling on the world population. As discussed further below, the dominant factor in determining the breadth of confidence intervals is the size of the sampled populations. Since the world populations are large, it is unlikely that treating them also as potentially varying would change the results.

stability measures is superior to bootstrapping, based on Monte Carlo simulations that they have carried out. We have not undertaken to compare any of our bootstrapped confidence intervals to stability intervals based on the MBIE concepts. As noted, we intend these intervals only to be used as qualitative indicators of the sensitivity of results to the underlying data. Refinement of these presentations based on alternative methods for characterizing this sensitivity could be the focus of further work.

9.3 Results with bootstrapped confidence intervals

The three figures below (Figure 21Figure 23) show the fraction of NZ publications in the top 50%, 10% and 1% with 90% percentile bootstrap confidence intervals for 2002-2004 and 2012-2014. The bootstrapping was done with 5,000 replications and stratified by ASJC field.

Figure 21: Fraction of NZ Papers in Top 50% 2002-2004 and 2012-2014 by Field – Baseline Results with 90% Confidence Intervals

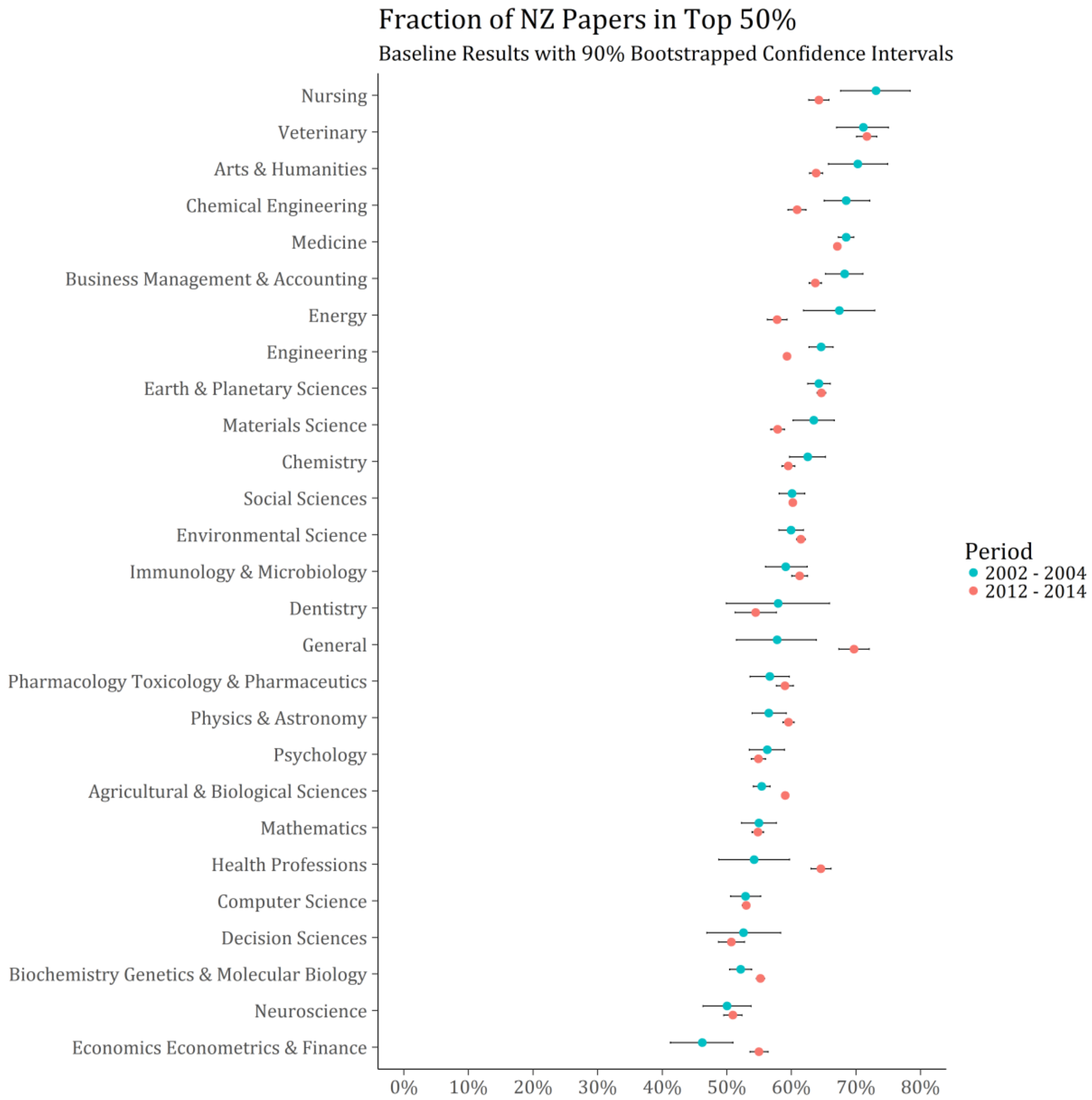


Figure 22: Fraction of NZ Papers in Top 10% 2002-2004 and 2012-2014 by Field – Baseline Results with 90% Bootstrapped Confidence Intervals

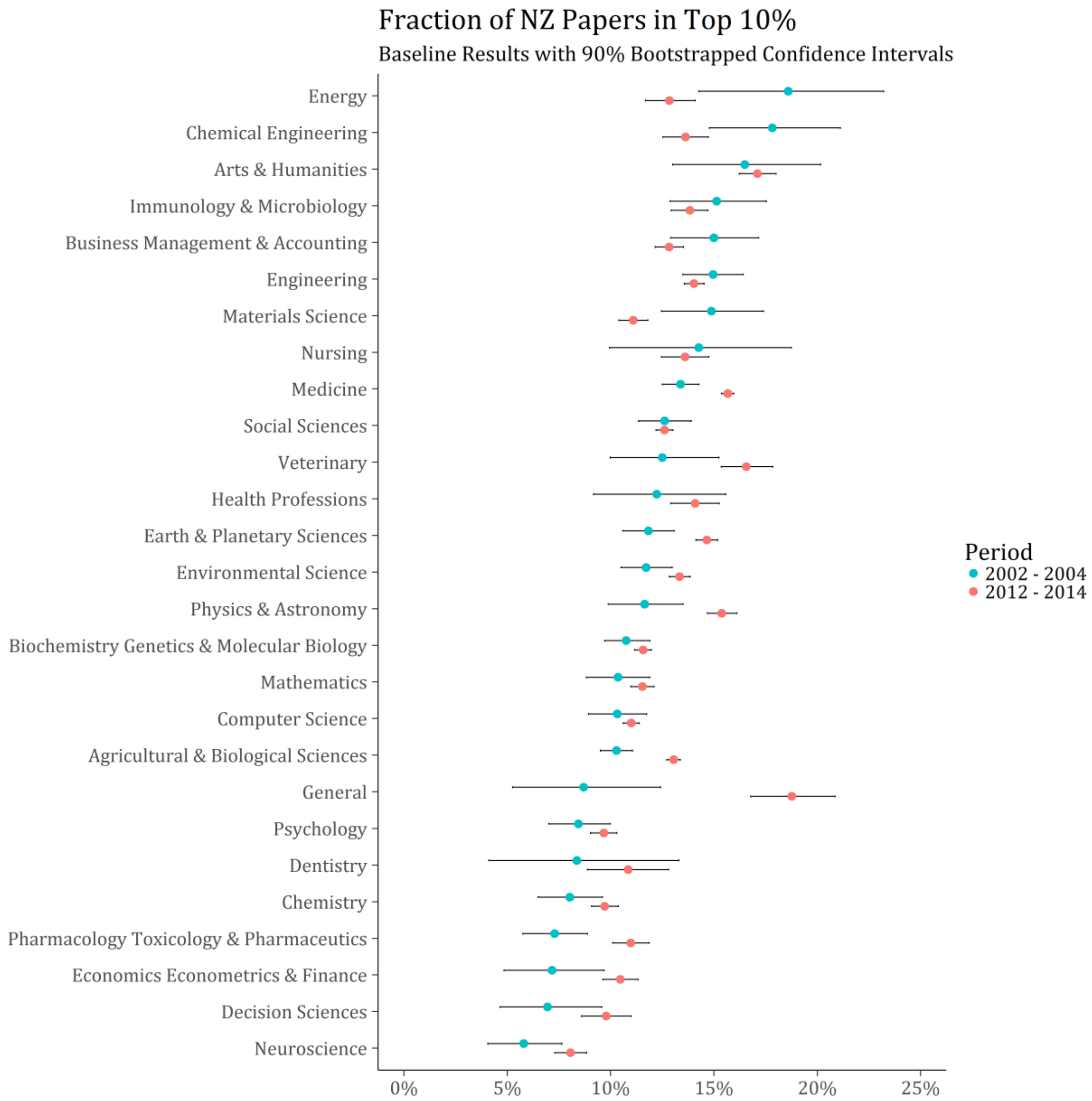
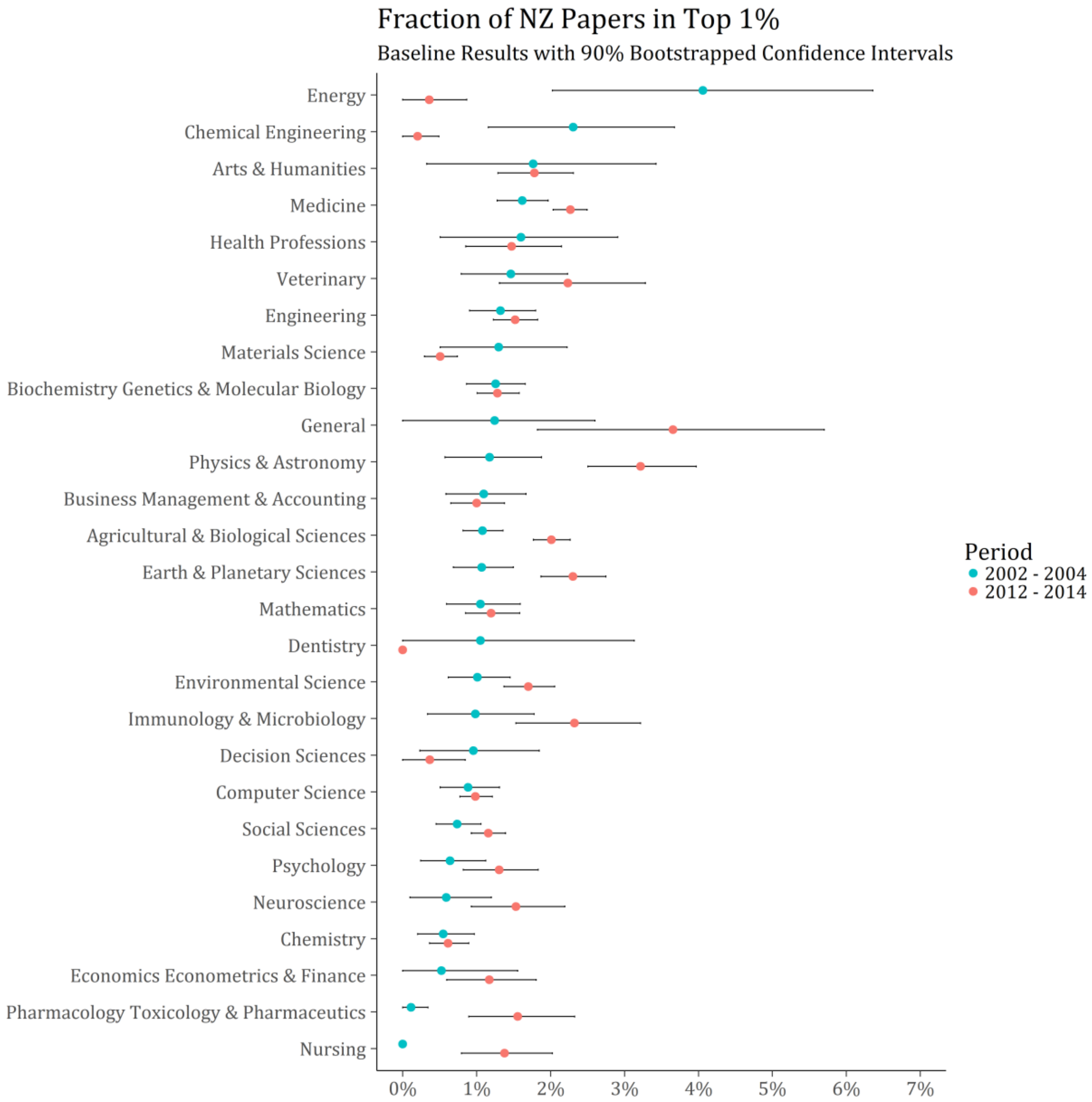


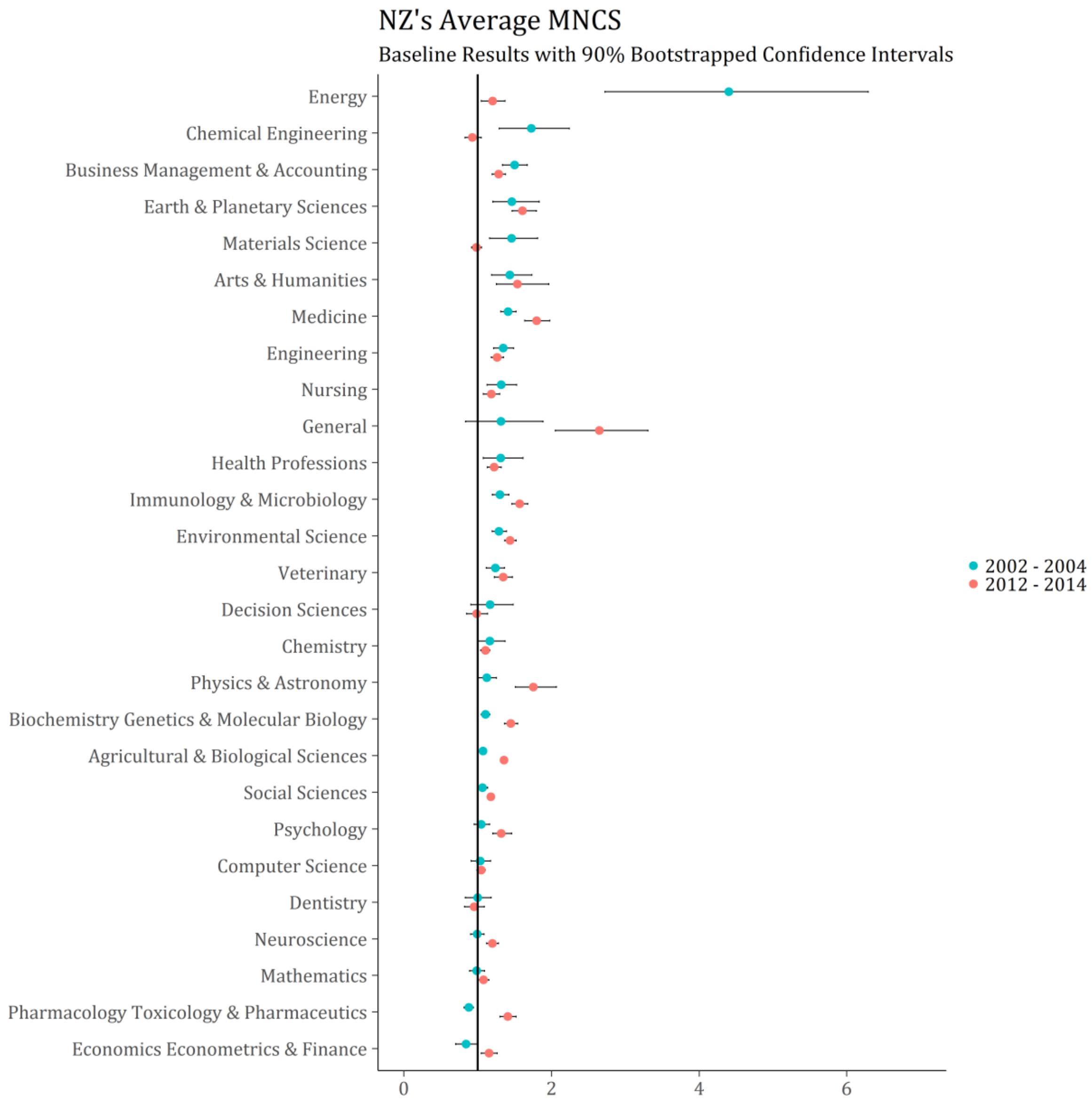
Figure 23: Fraction of NZ Papers in Top 1% 2002-2004 and 2012-2014 by Field – Baseline Results with 90% Bootstrapped Confidence Intervals



Confidence intervals for the more recent period of 2012-2014 are tighter than those for 2002-2004. This will be a result of the larger sample sizes in more recent years, as the number of publications in the dataset has increased greatly over time. It is important to emphasise that these error bands reflect precision, and we know that the accuracy of our measures is in fact better for the earlier results. We also see that confidence intervals tend to be wider the closer the threshold that we consider is to zero.

Finally, we have produced confidence intervals for NZ's average MNCS for the same two different periods used in above results, shown in Figure 24 below. For most fields, the results are fairly precise. We see that the majority of fields have an average MNCS that is statistically significantly higher than 1.

Figure 24: NZ's Average MNCS 2002-2004 and 2012-2014 by Field – Baseline Results with 90% Bootstrapped Confidence Intervals



9.4 Determinants of size of confidence intervals

In section 9.3 we saw that estimates based on more publications tend to have narrower confidence intervals. It may be useful to know the approximate number of publications needed in a year-field group in order to have a small enough confidence interval. The figures below (Figure 25 to Figure 27) show the relationship between the width of the confidence interval and the number of publications for the fraction of NZ publications in the top 50th, 10th, and 1st

percentile of the citation distribution for data from 2002-2004. These relationships could be used as a guide for the sample size needed for a certain degree of precision.

Figure 25: Width of Confidence Interval of Fraction of NZ's Publications in Top 50% in 2002-2004

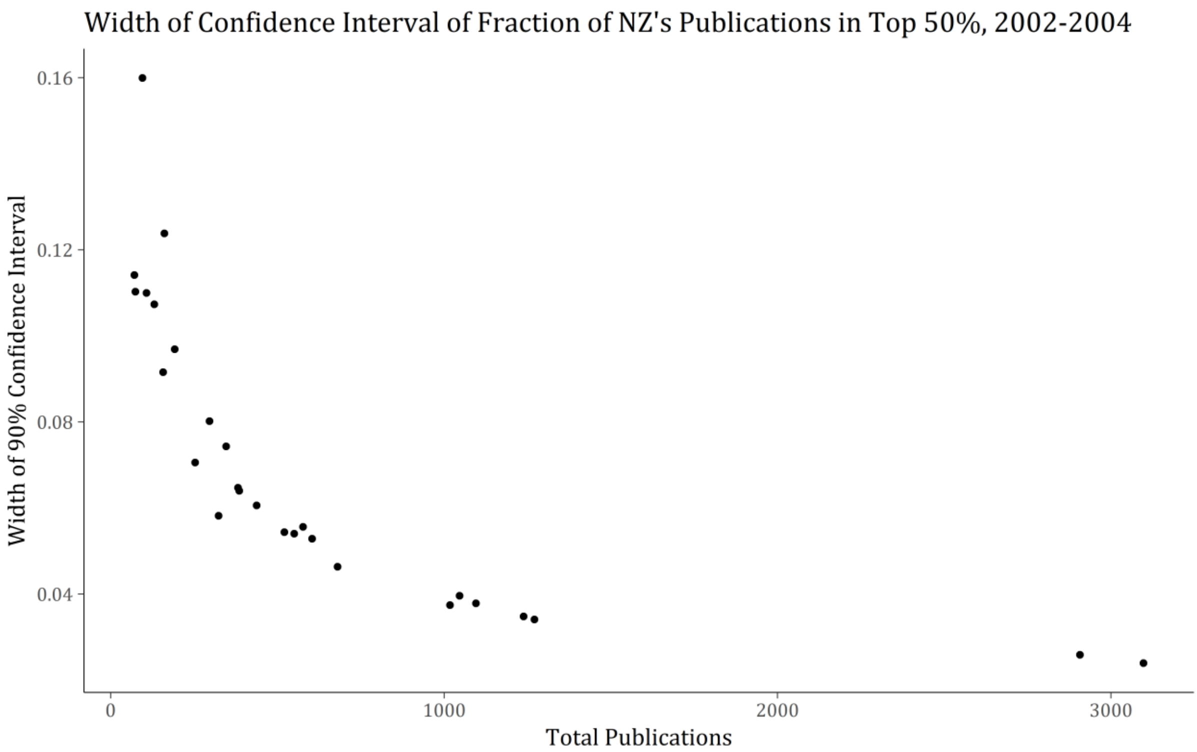


Figure 26: Width of Confidence Interval of Fraction of NZ's Publications in Top 10% in 2002-2004

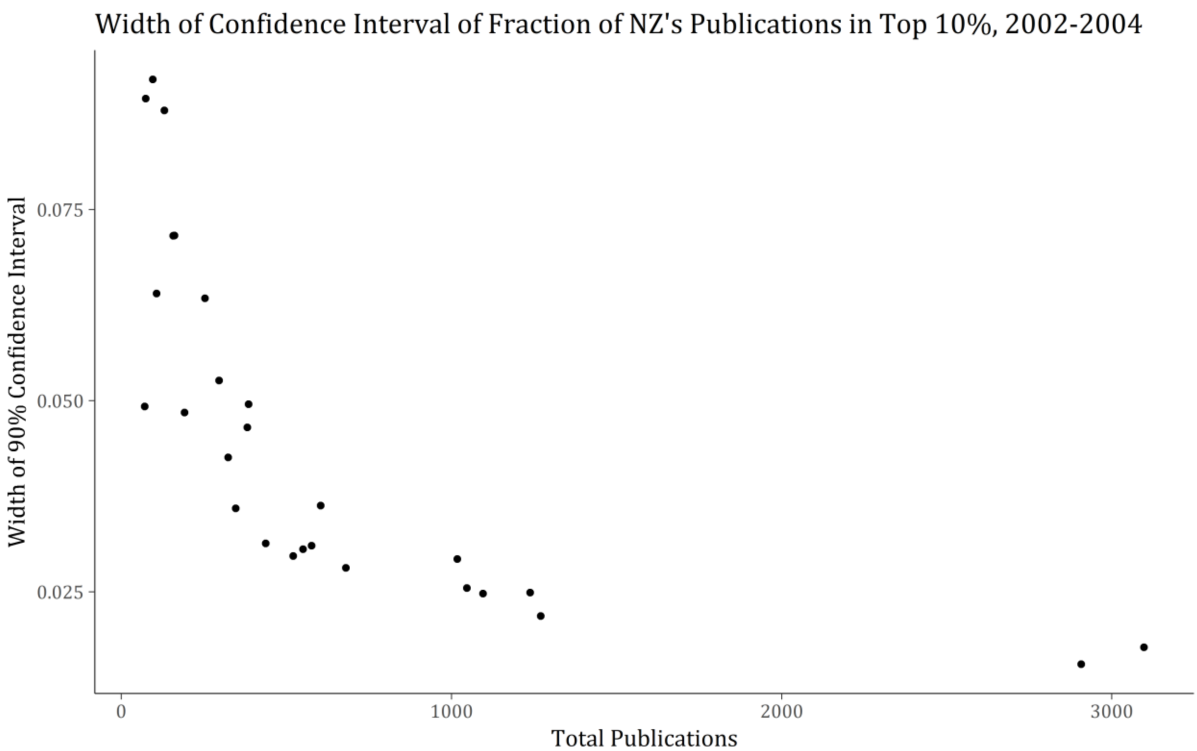
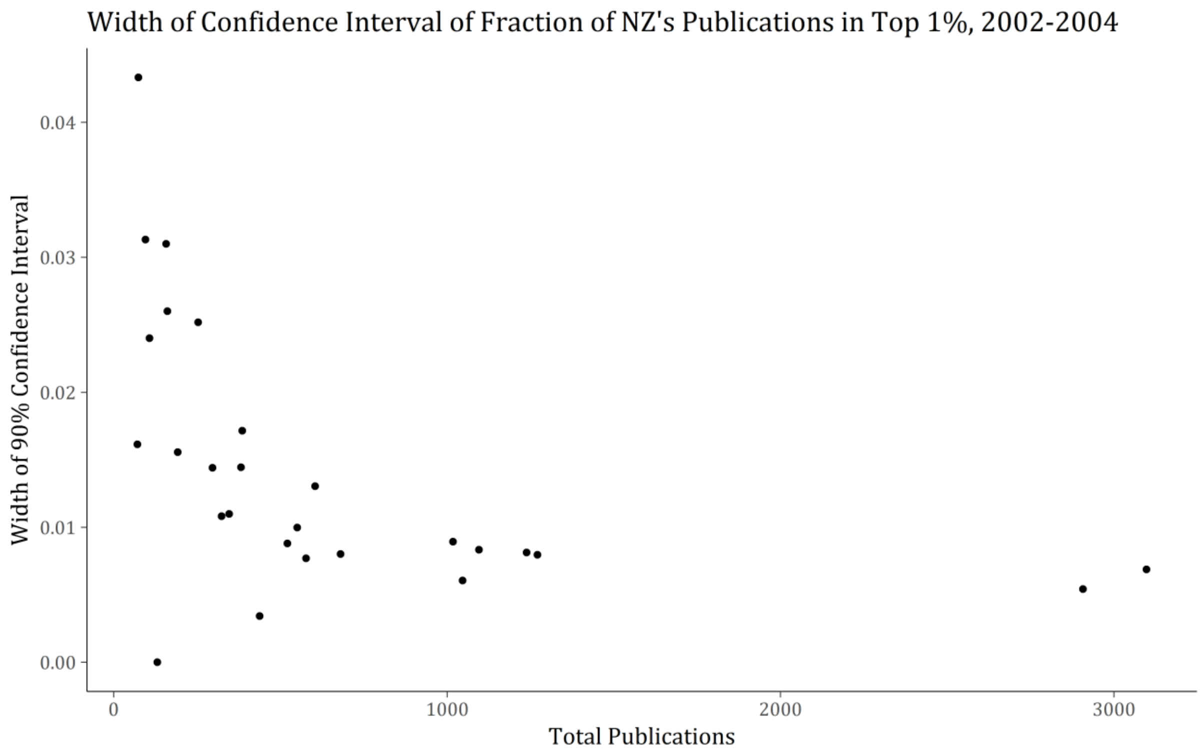


Figure 27: Width of Confidence Interval of Fraction of NZ's Publications in Top 1% in 2002-2004



10 An Alternative Assignment of Publications to Fields

Linking publications to the fields within which they lie is important not only for distinguishing NZ's strengths and weaknesses across fields, but also to account for different citation practices. Citation distributions vary by field, and therefore a fair assessment of the quality of an individual publication requires comparing its level of citations to publications in the same field. Publications in the Scopus dataset are linked to ASJC fields according to the main fields covered by the journal in which they were published. There are 27 parent-level ASJC fields and 334 sub-fields in the Scopus dataset. A publication can be attributed to multiple sub-fields which can cross over multiple parent-level fields. Furthermore, if a publication spans disciplines outside of those the journal is tagged in, we are unable to detect this.

Though 334 ASJC categories would seem to offer a fine-grained level of classification, many journals are tagged with broad disciplines such as 'Agriculture and Biological Sciences (all)' and 'Agriculture and Biological Sciences (miscellaneous)'. In the most extreme case, there is a 'General' category (also referred to as multidisciplinary).[16] This field encompasses some of the world's most highly cited journals, including Nature and Science.

Because of these limitations of the data, MBIE asked us to explore the possibility of allocating publications in the Scopus dataset to fields using semantic analysis of the abstracts. To

[16] To add to the confusion, journals which are assigned to multidisciplinary categories are still sometimes assigned to other categories as well.

do this we used topic modelling, a machine learning tool that searches for thematic content from text. With topic modelling we were able to predict the share of each publication's abstract that belongs to each ASJC field. We then used this information, in place of the original Scopus classification, in assessing NZ's research performance across fields. The topic modelling process and results are summarised in the following sections.

10.1 Topic modelling approach

The standard topic model is Latent Dirichlet Allocation (LDA), which aims to find the topics exhibited within a set of documents, and the distribution over these topics that individual documents span. The resulting set of topics can be described by the most probable terms that they are made up of. The model can then predict topic distributions from a new set of topics based on its learning of which terms are associated with which topics in the 'training dataset'.

We employed Labelled LDA (Ramage et al. 2009), a variant of LDA which is supervised by existing labels on the training set of documents. The labelled LDA is designed to predict the share of attribution to each topic when documents have multiple labels. This model could allow us to use existing information about the potential topic coverage of the publications. Labelled LDA also enables us to preserve the ASJC structure, rather than discovering sets of words from which we would have to infer new topics. This consistency makes it simple for us to compare new results to the previous findings.

Using open-source software called Stanford Topic Modeling Toolbox (TMT)[17], we applied a labelled LDA to a training dataset of abstracts labelled with the parent-level ASJC fields.[18] The training dataset included abstracts of all publications except any from a journal tagged as multidisciplinary. The model predicts the share of words of each abstract associated with each of its ASJC labels. In this process, the model 'learns' the words most associated with each label. We then let the trained model predict the share of words across fields for every publication in our dataset, without constraining predictions to fall within the original labels. Relaxing this constraint allows the model to detect topics outside of the main foci of the journal a publication is published and meant we could apply this inference to publications in multidisciplinary journals.

In practice, we ran this topic model process separately for each year of our data. For 2002-2005, we use all non-multidisciplinary publications in the training of the model. As the sample size increases greatly year-by-year, we ran into memory problems using the full sample for training from 2006. We therefore chose to randomly select 50% of the publications in the

[17] We used version 0.4.0, available at https://nlp.stanford.edu/software/tmt/tmt-0.4. TMT was produced in 2009-2010 and is no longer updated. The major drawback of this was that we had to revert to an older version of Java (version 7). This was because TMT is written in Scala which itself runs on Java, did not seem to be compatible with the newer version. The only other open-source software we could find which offers labelled LDA was Mallet (see http://mallet.cs.umass.edu) but we could find no documentation on how to use it.

[18] We selected the collapsed variational Bayes approximation to the LDA objective to train the model. The alternative in Stanford TMT is a collapsed Gibbs sampler.

dataset from 2006 onwards for training. We were still able to apply these trained models to the full dataset. We believe that with the huge corpora of text, a 50% sample will be more than sufficient to learn the words associated with different topics.

For each publication, the topic model predicts the fraction of words in each of the 26 parent ASJC fields (after excluding 'General'). There are various ways that we could use this information to assign publications to fields. The simplest option would be to assign each publication to the field in which the topic model predicts it has the largest share of words. At the other extreme, we could use the entire distribution predicted by the topic model.

We found that a large portion of the publications were predicted to have a relatively high proportion of their words in a single topic. Over 30% of publications have a highest score of at least 90%. Spot checking of several publications which received high scores in one or two particular fields had sensible allocations (according to our non-expert judgement).

However, there are a reasonable number of publications with much lower maximum scores, which must be spread across a number of fields. Some of this could be because some publications are truly multidisciplinary, but it could also be because the topic model struggles to detect the fields represented by the words in some abstracts. We looked at the abstracts of a handful of publications that received low scores across several fields, and felt that the latter is of some concern.

To reduce the weight given to a publication in a wrongly predicted field, we opted to use full distribution of predicted word shares, other than very low scores. Any score lower than 5% was replaced with a 0, then the weights in the remaining fields were normalized to sum to 100%.[19] We then linked these new weights across fields back to the Scopus dataset and used them to redo our analysis of NZ's research performance.[20]

10.1.1 Exclusion of Common Words

Common words are usually removed from text before a topic model is applied because it can be difficult for the model to distinguish the topics of such words. In initial topic model attempts, we excluded the 30 most common words in the dataset.[21] However, we found a number of very common words turned out in the top keywords within each field, and these were very often common to more than one field. To address any concern that this would affect the ability of the topic model to differentiate topics, in addition to the exclusion of the most common 30 words, we excluded a list of common English words from further topic models. The "stop-words" we applied were from a list of 524 common English words (including individual letters) used by the

[19] This simplifies our dataset as we move to analyzing the number and impact of publications across fields.
[20] Because our topic model makes predictions at the parent-level of ASJC categories, we lose the ability to use benchmarks based on groups at a finer level when we use the topic model inference. Options to incorporate the second-level ASJC categories are discussed later.
[21] Other adjustable settings we used in the topic model were to ignore characters that are not letters or numbers, drop words with less than four characters, drop words in less than four documents, and drop documents with fewer than 6 words.

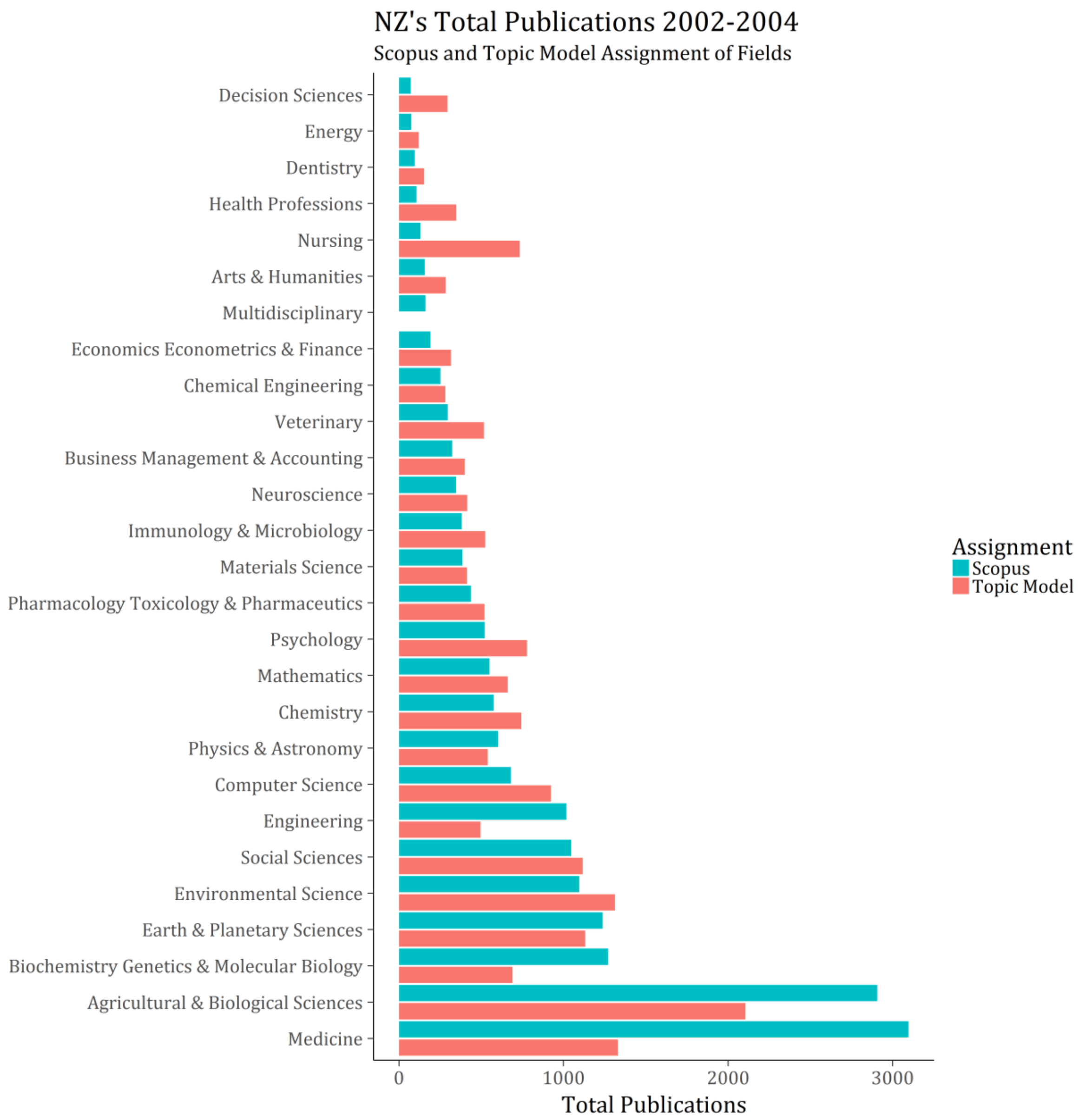
Mallet "MAchine Learning for LanguagE Toolkit".[22] Schofield, Magnusson, and Mimno (2017) find that stop word removal has little impact on topic model inference, and consistent with this we found the change made very little difference to our end results.

10.2 Results of topic modelling

The first thing we consider is the number of publications observed in each field. Figure 28 compares field assignment by Scopus and the topic model for NZ publications from 2002-2004.[23] We observe moderate differences in the number of publications for many disciplines. Among some fields, however, the difference is particularly large, e.g. for Agricultural & Biological Sciences and Medicine. We find similar patterns by field for global publications. Though Medicine remains the world's largest field, under topic modelling it consists of almost half of the number of publications as it did by Scopus allocation of fields. While we would expect some change in the allocations by the topic model, changes by proportions as large as this are surprising.

[22] Information about Mallet can be found at http://mallet.cs.umass.edu/. An online version of the "stop-word" list is provided by https://github.com/mengjunxie/ae-lda/blob/master/misc/mallet-stopwords-en.txt
[23] Total publications in each discipline are weighted sums of the fractional count of publications in each field. Only articles, conference proceedings and reviews are included in this analysis.

Figure 28: NZ's Total Publications 2002-2004 by Field – Scopus and Topic Model Assignment of Fields



A transition matrix split over Appendix Table 1A and 1Bshows how publications in each original Scopus field are allocated to fields by the topic model for the 2002-2004 data.[24] This is shown as a fraction of the papers originally classified in that field. Substantial proportions of publications from Medicine journals are predicted by the topic model to belong to the Health Professions and Nursing.[25] While this overlap could be because of true interdisciplinary work, it is also possible that the topic model struggles to differentiate between closely related topics. From a policy perspective, it may be feasible to combine groups that would have similar

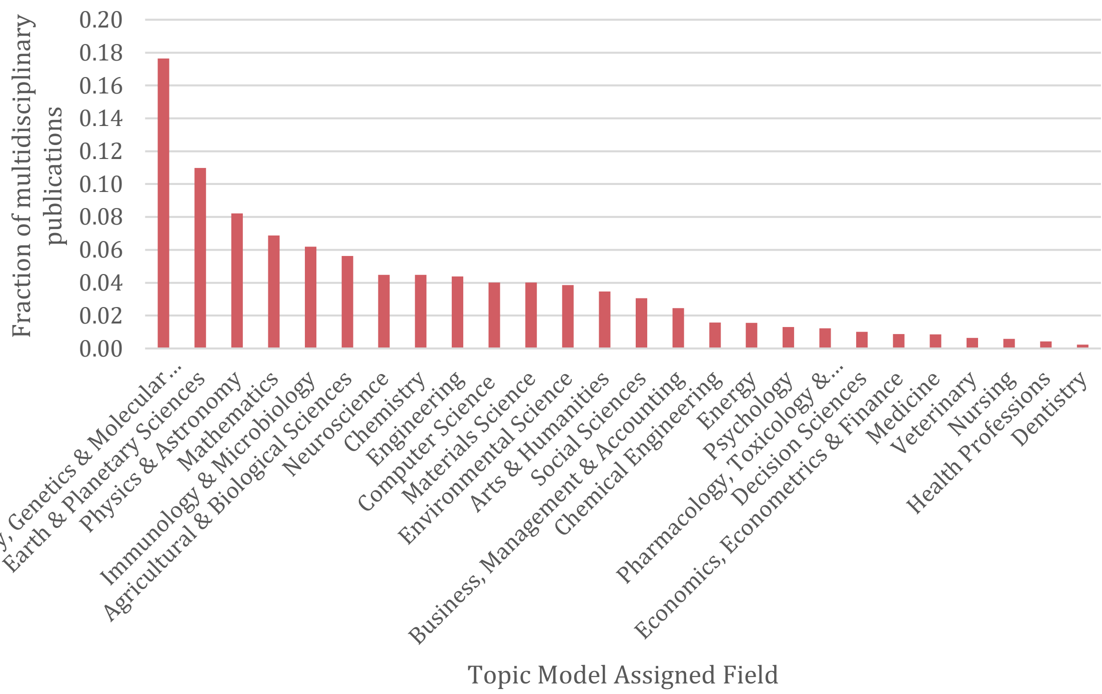
[24] Note that this transition matrix covers all publication types, because all publication types were included in the topic model training and inference.
[25] Considerable shares of publications in Medicine journals were also allocated to Immunology & Microbiology, Psychology, and Pharmacology, Toxicology & Pharmaceutics.

terminology anyway. However, there may be differences in the citation practices at a finer level that ought to be accounted for.

The transition matrix split over Appendix Table 2A and 2B is the same as that in Appendix Table 1A and 1B but uses the full distribution of topic model predictions, instead of imposing the 5% minimum to be in a field. Comparing these two matrices, the 5% minimum has little effect on the results.

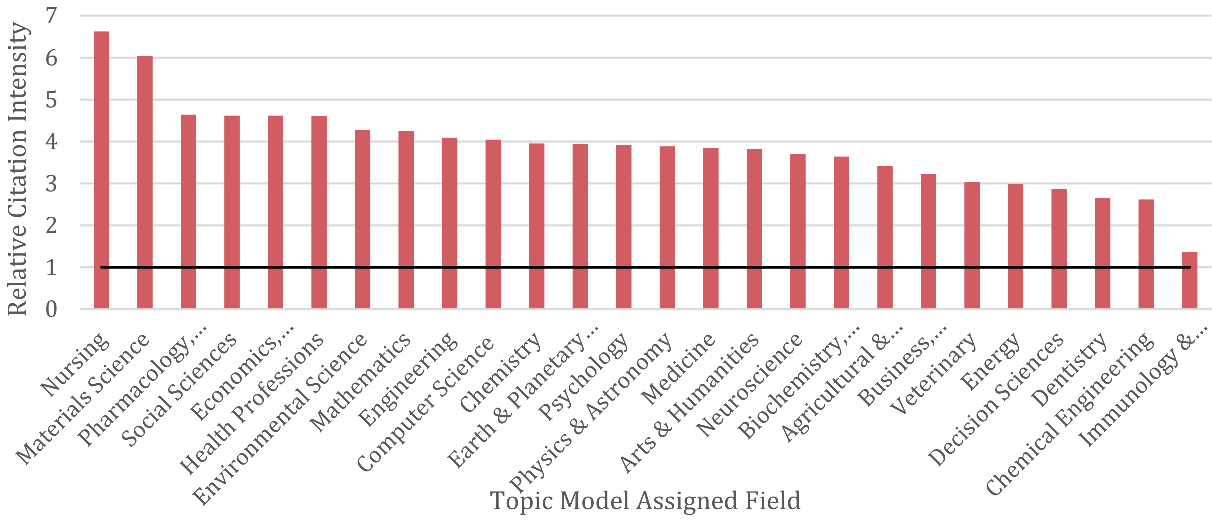
The field assignment of publications originally classified as General (or multidisciplinary) by Scopus are set out in Figure 29 for global publications from 2002-2004, as a share of the total group of General publications.[26] The shares do not appear to correlate with total publications in each field. For instance, Medicine is the largest field (according to both topic model and Scopus assignment of fields), yet makes up relatively few of the General publications. Figure 30 then shows the relative citation intensity of these multidisciplinary publications to non-multidisciplinary publications allocated by the topic model to the same field. Multidisciplinary publications are relatively highly cited in all fields.

Figure 29: Topic Model Field Assignment of Multidisciplinary Publications, 2002-2004



[26] All publication types are included in this analysis.

Figure 30: Relative Citation Intensity of Multidisciplinary Publications to Non-Multidisciplinary Publications by Topic Model Field Assignment, 2002-2004



10.3 Distribution of NZ research output based on topic-modelled field assignments

Results for the fraction in top percentiles with Scopus and topic model field-assignment are compared in Figure 31 to Figure 36 for two illustrative periods. Under Scopus field allocation, the calculations are carried out at the ASJC sub-fields and then aggregated into the broader structure. In the topic model, everything is done at the broad level. Assigning fields by topic modelling produces results that differ somewhat from classification based on Scopus journal assignments, but we cannot be sure which version of results is closer to the truth.

Figure 31: Fraction of NZ Papers in Top 50% 2002-2004 by Field – Scopus and Topic Model Assignment of Fields

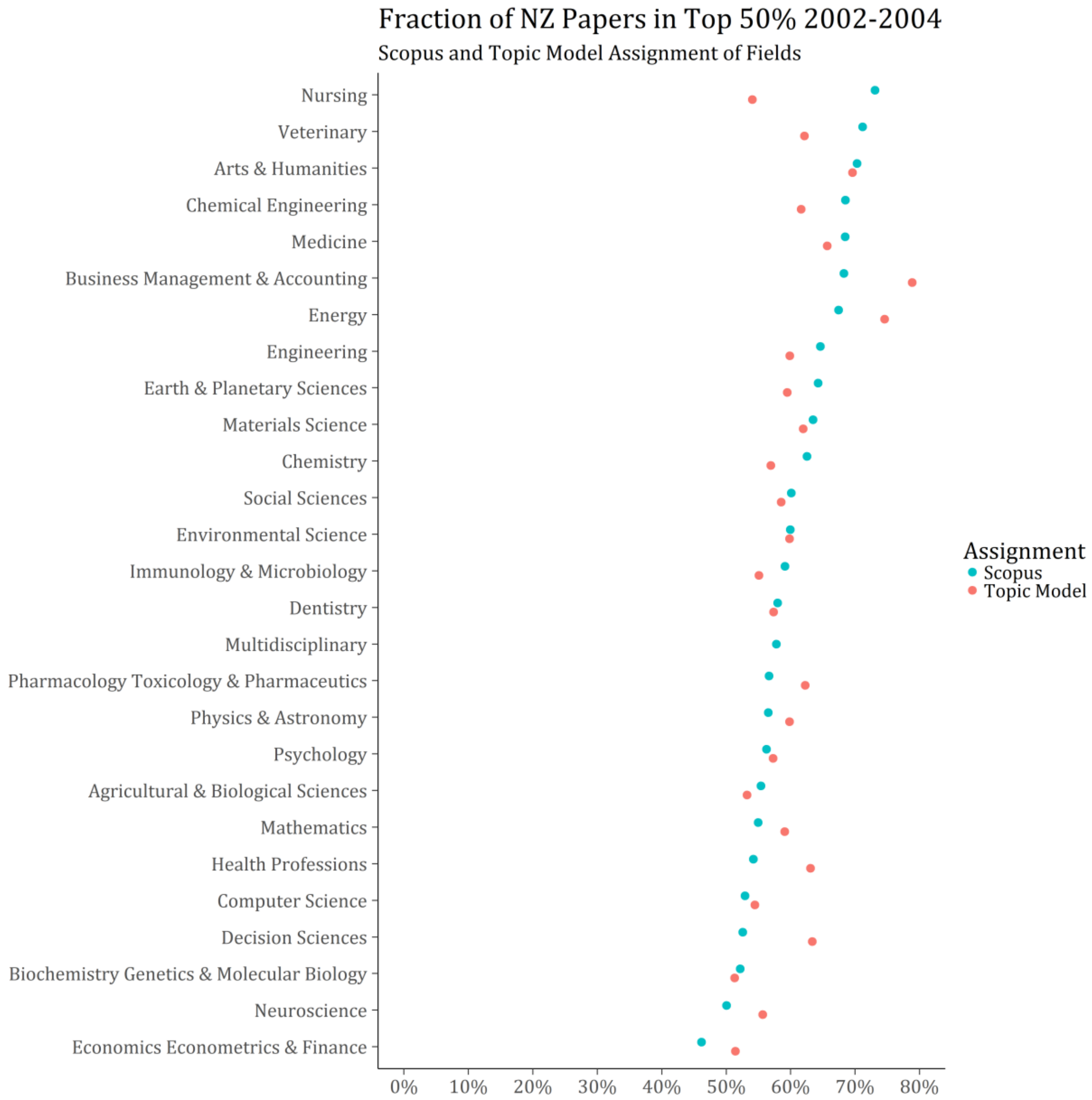


Figure 32: Fraction of NZ Papers in Top 10% 2002-2004 by Field – Scopus and Topic Model Assignment of Fields

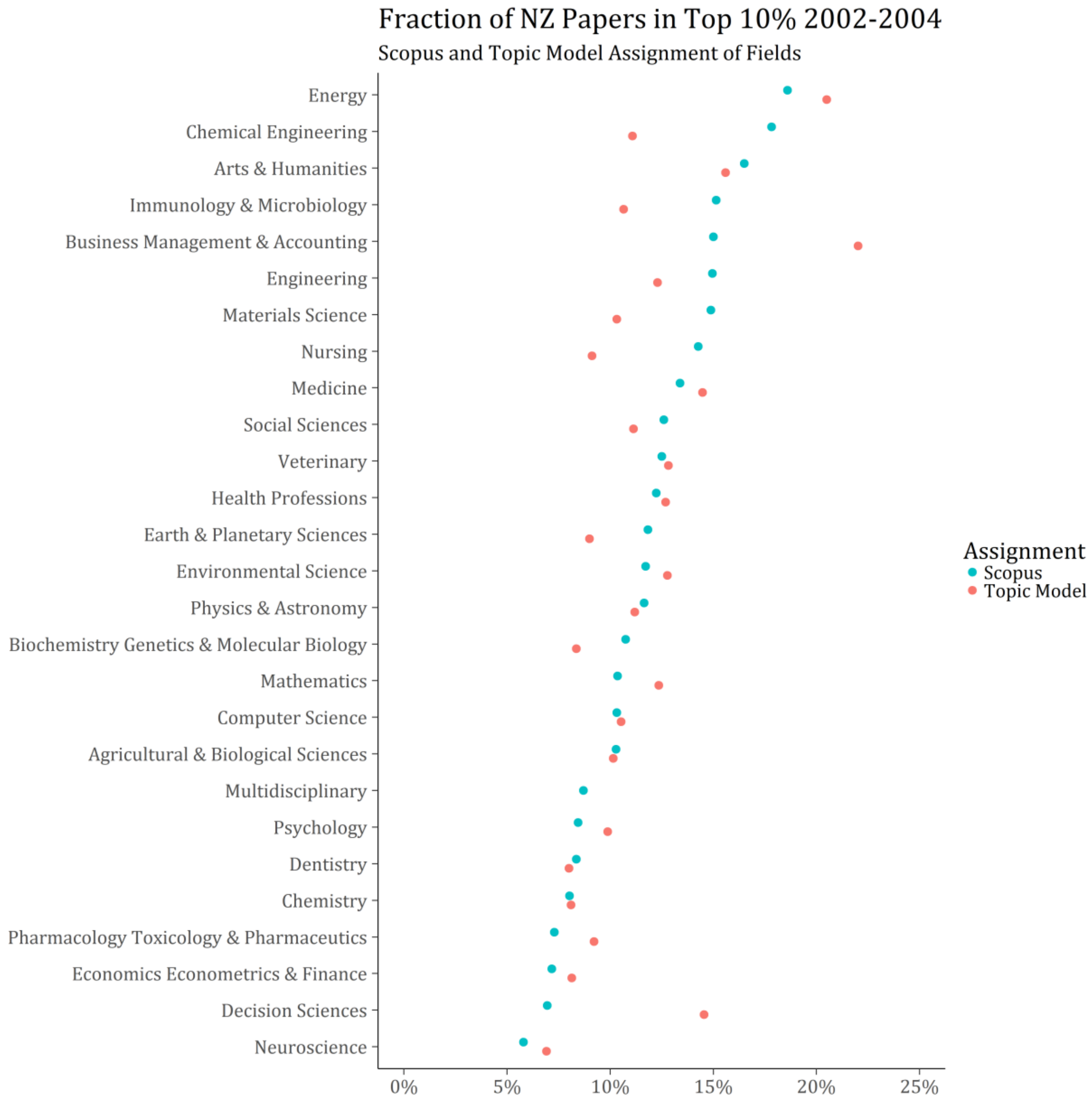


Figure 33: Fraction of NZ Papers in Top 1% 2002-2004 by Field – Scopus and Topic Model Assignment of Fields

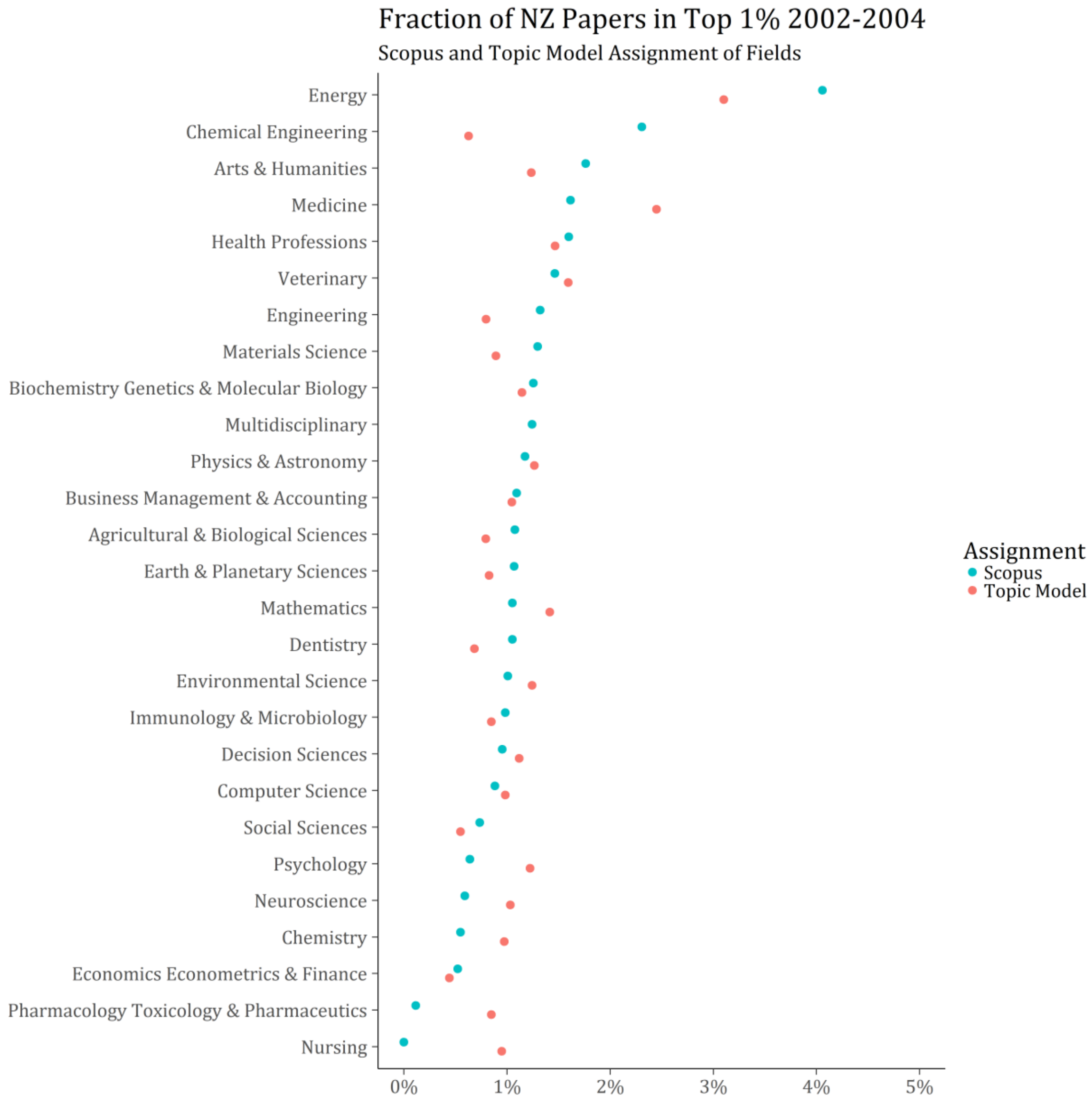


Figure 34: Fraction of NZ Papers in Top 50% 2012-2014 by Field – Scopus and Topic Model Assignment of Fields

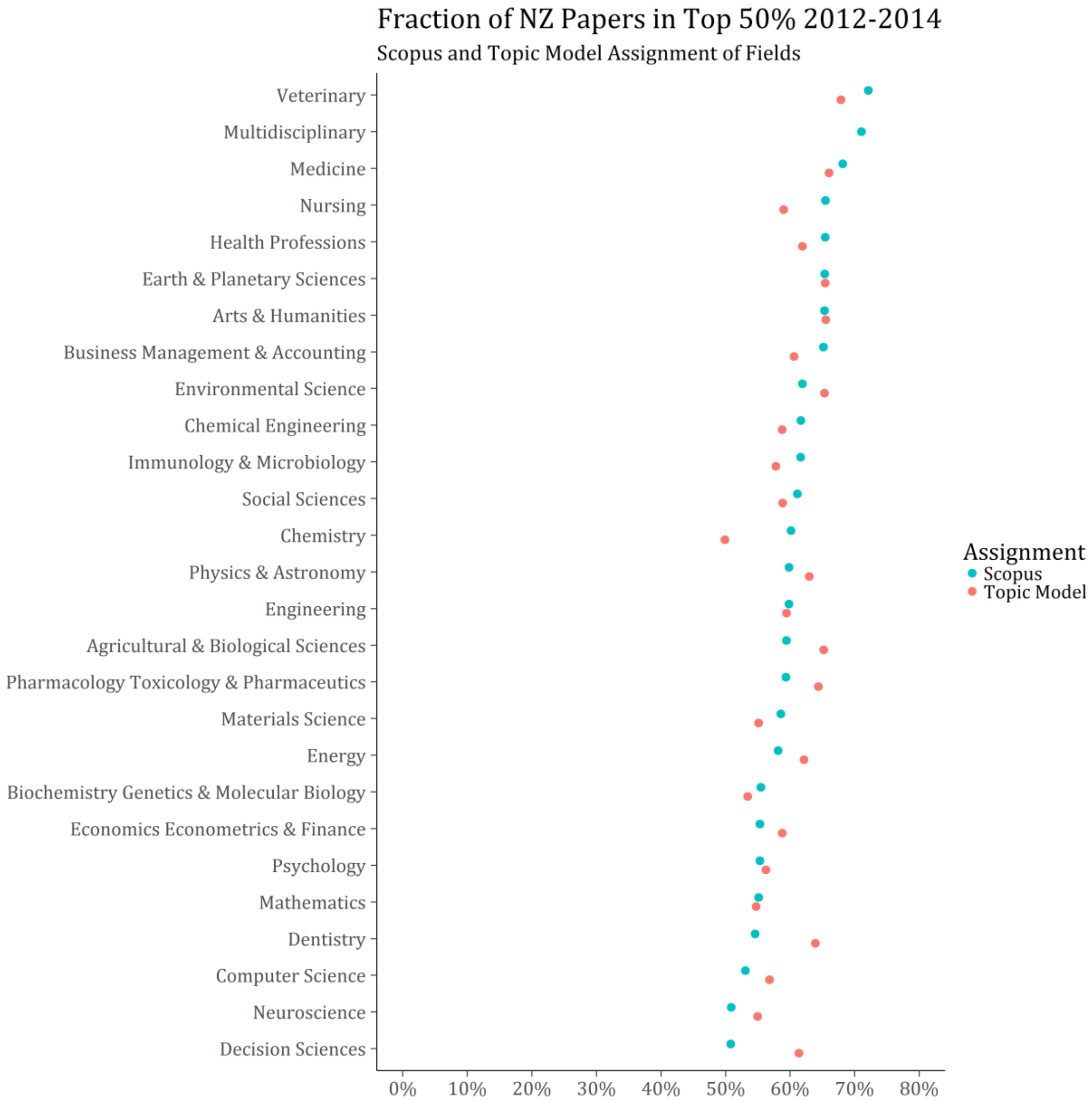


Figure 35: Fraction of NZ Papers in Top 10% 2012-2014 by Field – Scopus and Topic Model Assignment of Fields

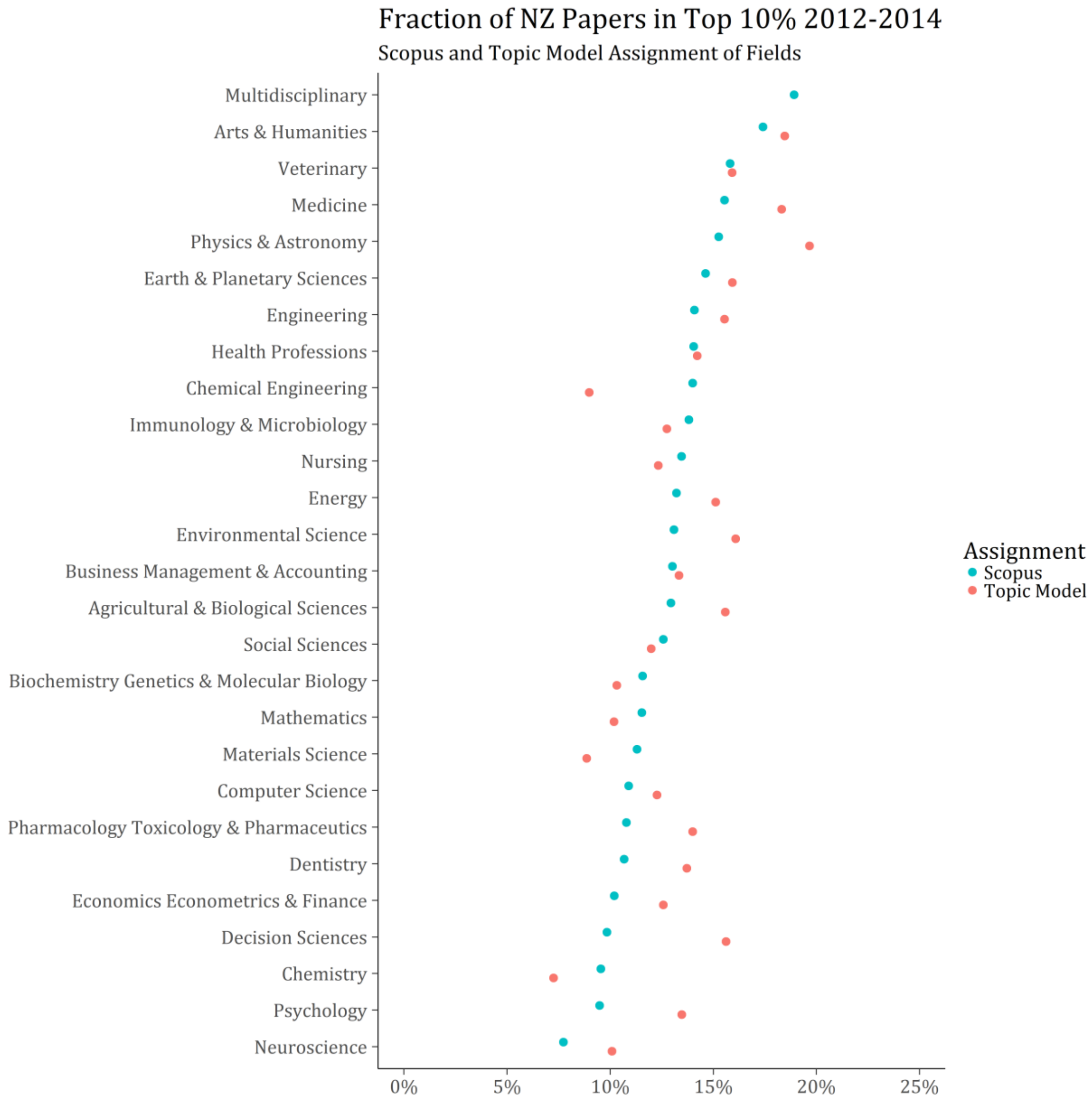
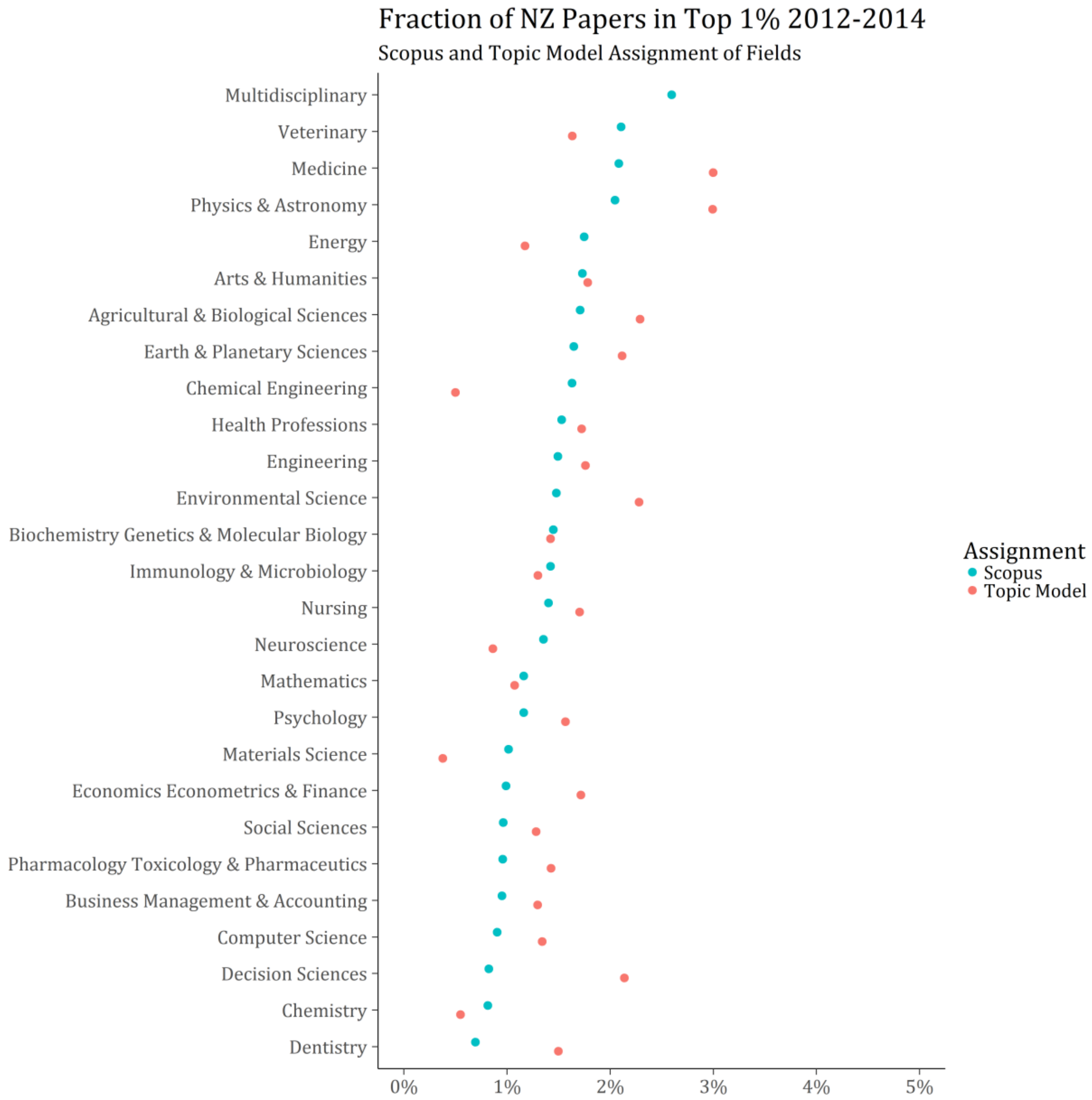


Figure 36: Fraction of NZ Papers in Top 1% 2012-2014 by Field – Scopus and Topic Model Assignment of Fields



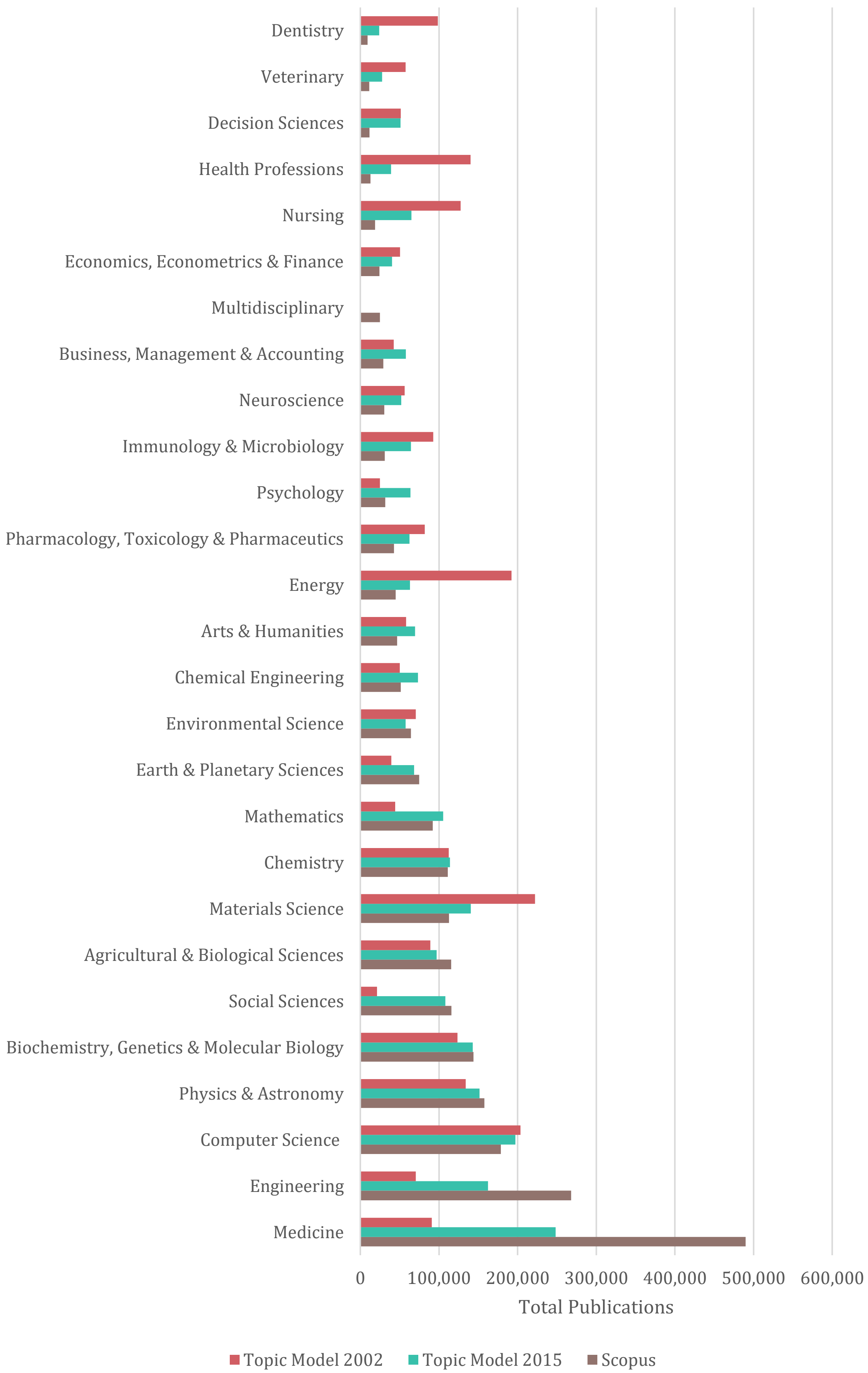
While we cannot be sure which version of results is closer to the truth, their similarity among most fields provides a level of confidence that allocating publications to topics according to the journals they are published in does not greatly distort our interpretation of the impact across fields. This result is surprising given the topic model did have a relatively significant effect on the quantity of publications across fields. Topic model results for each year are provided in the supplementary materials to this report.

10.4 Sensitivity to Training Dataset

The approach we have taken has been to carry out inference on each year of data using a model trained on the same data. Where the dataset was too large for training purposes, we used a 50% sample of that year's publications. Our reason for training on new data in each year is due to the concern that the words pertaining to disciplines could evolve over time.

If the key words within topics do not change quickly over time, then it could be more practical to train the model only once using a single year, which could then be applied for inference on all years of data. To see the extent to which this makes a difference to results, we compare field assignment in 2015 using a model trained on 2002 data to one trained on a 50% sample of 2015 data. Total publications (including all publication types) by field in 2015 are shown in Figure 37 under these two alternatives of field assignment and with the original Scopus field classification. We see that the allocation of publications is highly sensitive to the year in which the data is trained.

Figure 37: Total Publications in 2015 by Different Methods off Field Assignment



11 Consolidation of Fields

In the previous section we raised the concern that topic modelling might be limited in its ability to distinguish between topics which use similar terminology. We are also interested in aggregating fields into broader groups because of the wide confidence intervals around some of our estimates, particularly for fields with a small number of publications.

In this section we consider a systematic way in which we could combine fields based on the intensity with which we observe fields overlapping within publications, as allocated by the topic model. This intensity would increase not only with the true overlap of fields within publications, but also with word similarity that is not due to field overlap.[27] While conceptually we would not want to combine fields that are not truly appearing in the same publications, it would be necessary to solve the problem of the topic model being unable to separate topics using similar terminology.

We define a measure of overlap between pairs of fields using the topic model predictions of field allocation for each publication (after the imposition of the 5% minimum). The idea is, for each publication, to multiply the fractions allocated to each of the two fields of interest. The average of this overlap, weighted by the combined significance of the two fields for each paper, is taken as the measure overlap between the two fields overall.

The overlap $O_{kk'}$ calculation for fields k and k' is defined as follows. Call w_{ik} the fractional weight of assignment of paper i to field k, where $0 \leq w_{ik} \leq 1$ and $\sum_k w_{ik} = 1\ \forall i$. We calculate for all k and k':

$$O_{kk'} = \frac{\sum_i (w_{ik} + w_{ik'})(w_{ik} \cdot w_{ik'})}{\sum_i (w_{ik} + w_{ik'})}$$

It follows that $0 \leq O_{kk'} \leq 0.25$ and $O_{kk'} = 0.25$ if and only if every paper for which $w_k > 0$ or $w_{k'} > 0$ has $w_k = w_{k'} = 0.5$.

A matrix of weighted overlap of field assignments for 2002-2004 is presented in Appendix 3A and 3B. The overlap between fields is generally small given that the maximum possible overlap between fields is 0.25. We observe that in most cases, $O_{kk'}$ is less than 0.01, while in some cases it is as high as 0.02-0.03. The range is very small (from close to 0 to about 0.032). Pairs with relatively high overlap include, for instance, Arts & Humanities with Social Sciences and Physics & Astronomy with Materials Sciences, which are not the most intuitive matches. Therefore, we do not think this is a very useful measure to guide the consolidation of fields.

[27] The reason for this is that if an abstract contains words common to two fields, but truly belongs only to one of those fields, the topic model would still spread the publication across both of them.

12 Limitations of bibliometric data

In this report we focus on determining which choices of methodology make a difference in analysing bibliometric data. It should be acknowledged that bibliometric databases, including Scopus, have limitations themselves. Not all publications are included in these databases, and coverage can vary by field. Furthermore, not all research outputs are in the form of publications, not all impact is reflected in citations, and not all citations reflect true impact.

It is important to note that citations take time to accrue and so are less reliable indicators of impact when there is only a short time period between the time of publications and the time that citations are counted. Therefore, one should be cautious of interpreting bibliometric analyses on recent time periods, including those within this report.

13 Conclusion

Bibliometrics provides a valuable tool for assessing the magnitude and potential scientific impact of NZ research. As with any tool, however, one must choose judiciously how to use it. In most cases, this is not a matter of a 'correct' versus an 'incorrect' metric. It is rather a matter of thinking carefully how to match the appropriate metric to the underlying question being asked. We have explored the sensitivity of various measures to different ways of treating the underlying data. The major themes that emerge from this analysis include:

- Many NZ papers have both NZ and foreign authors, and these foreign-co-authored papers are on average more highly cited than purely domestic papers and, as a result, measures of average impact tend to turn out lower when fractional counting is used instead of full counting. Whether such papers are viewed as a full or partial NZ output also has significant consequences for inference regarding the fraction of NZ's publications in top percentiles. Fractions of publications in the top x% are often compared to a global benchmark of x%, but this is not a valid interpretation when full counting of authors is used.
- When calculating the fraction of publications above a certain threshold of the citation distribution, fractional weighting should be applied to publications with the same number of citations as the threshold so that the global fraction of publications in the top x% is exactly x%.
- We are inferring the performance of researchers in different fields based on their publications, but researchers in a given field may publish in other fields, and the assignment of papers to fields is inherently imperfect. Advances in computational semantics offer the potential for richer and possibly more precise classifications of papers by discipline, but it is not yet clear how well these methods work. They produce

classifications that differ somewhat from those based on Scopus journal assignments, but we do not have any external reference to tell us which is better.

- The distribution of citations across papers is highly skewed, so that measures of central tendency such as MNCS carry only limited information. Looking at the upper tail of the citation distribution may be more informative, but such measures have to be constructed carefully, and the relatively small number of papers in this tail make these measures less stable or robust.
- While bibliometric measures are not based on samples in a statistical sense, they are nonetheless potentially sensitive to how the population corresponding to a particular question of interest is defined. We have proposed a way of characterizing this sensitivity, but this is a topic that merits further work.
- The types of publications included in bibliometric analysis and how publication type and fields are accounted for in normalising citations are shown to have relatively small impacts on results for NZ's bibliometric performance. Though small, one should be cautious that these impacts will be stronger for some fields than others.

Given the methodological choices underlying bibliometric analysis affect both the figures and their interpretation, we must stress the importance of making these choices transparent. The 2016 Science and Innovation System Performance Report (MBIE 2016) leaves readers in the dark about how publications in multiple fields or with multiple authors are dealt with. Those readers who are concerned with such issues will find it unclear what conclusions can be drawn from the results, while those who are not could easily draw misinformed conclusions. Furthermore, any discussion of how to interpret results needs to be executed with great care. It is, for example, alarmingly common for the fraction of publications in the top x% to be compared to a benchmark of x% in contexts where, as described in this report, this interpretation is in fact not appropriate (see for example MBIE (2016, 11)).

In terms of which metrics should be studied, it would not be wise to draw conclusions about NZ's research performance based on any of the metrics used in this report in isolation. Measures of the quantity and impact of research within a field are best understood together, as components of total impact. Furthermore, it is useful to consider measures of impact at various points of the citation distribution and measures of comparative advantage provide an additional perspective on NZ's position in the world. Indeed, in most disciplines, NZ research is above average, by which we mean that the average MNCS is greater than unity and the proportion of publications with above median citations is greater than 50%. But when we focus on the upper 10% or upper 1% of the citation distributions, NZ's share in most fields is below the world share when properly calculated. This suggests that NZ has a healthy proportion of good researchers but a disproportionately low concentration of international star researchers. As emphasized above, whether this is where we want to be or a situation we want to change is a question that cannot be answered with bibliometrics. In future work, we will explore the career trajectories of individual NZ scientists, and their relationships with each other and international scientists, to try to understand the dynamics that underlie the distribution of outcomes.

It is important to emphasize that the measures of bibliometrics that we have employed look only at the *outputs* of the research process. What we would really like to understand is how these outputs relate to research inputs such as researchers, laboratory equipment and supplies. We produce a lot of papers and citations in medicine, but we also spend a lot of money on medical research. On the other side, we may have a disproportionately high share of top papers in a field because we produce very few papers overall in that field and have a handful of very good ones. It's not clear that a situation like that has any real significance for NZ. As noted in the work of (Abramo and D'Angelo 2014, 2016), we cannot really judge the effectiveness of a research establishment without looking at the outputs relative to the inputs.

Indeed, in the absence of data on research inputs, none of the bibliometric measures in this report can be used to ascertain in which fields NZ researchers or NZ research investments are more effective or more efficient than others. The launching of the National Research Information System (NRIS) will create a basic data structure for research inputs. Linking those data to bibliometrics will ultimately create for the first time the potential to begin to answer these questions.

References

Abramo, Giovanni, and Ciriaco Andrea D'Angelo. 2014. "How Do You Define and Measure Research Productivity?" *Scientometrics* 101 (2): 1129–1144.

———. 2016. "A Farewell to the MNCS and like Size-Independent Indicators." *Journal of Informetrics* 10 (2): 646–651.

Aksnes, Dag W., Jesper W. Schneider, and Magnus Gunnarsson. 2012. "Ranking National Research Systems by Citation Indicators. A Comparative Analysis Using Whole and Fractionalised Counting Methods." *Journal of Informetrics* 6 (1): 36–43.

Ministry of Business, Innovation & Employment. 2016. "Science & Innovation System Performance Report."

Ramage, Daniel, David Hall, Ramesh Nallapati, and Christopher D. Manning. 2009. "Labeled LDA: A Supervised Topic Model for Credit Attribution in Multi-Labeled Corpora." In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 1-Volume 1*, 248–256. Association for Computational Linguistics. http://dl.acm.org/citation.cfm?id=1699543.

Schneider, Jesper W. 2016. "The Imaginarium of Statistical Inference When Data Are the Population: Comments to Williams and Bornmann." *Journal of Informetrics* 10 (4): 1243–48. https://doi.org/10.1016/j.joi.2016.09.011.

Schofield, Alexandra, Mans Magnusson, and David Mimno. 2017. "Pulling Out the Stops: Rethinking Stopword Removal for Topic Models." In *EACL 2017*, 432. http://www.aclweb.org/anthology/E/E17/E17-2.pdf#page=464.

Waltman, Ludo. 2016a. "A Review of the Literature on Citation Impact Indicators." *Journal of Informetrics* 10 (2): 365–391.

———. 2016b. "Conceptual Difficulties in the Use of Statistical Inference in Citation Analysis." *Journal of Informetrics* 10 (4): 1249–52. https://doi.org/10.1016/j.joi.2016.09.012.

Waltman, Ludo, and Michael Schreiber. 2013. "On the Calculation of Percentile-Based Bibliometric Indicators." *Journal of the Association for Information Science and Technology* 64 (2): 372–379.

Williams, Richard, and Lutz Bornmann. 2016a. "Sampling Issues in Bibliometric Analysis." *Journal of Informetrics* 10 (4): 1225–32. https://doi.org/10.1016/j.joi.2015.11.004.

———. 2016b. "Sampling Issues in Bibliometric Analysis: Response to Discussants." *Journal of Informetrics* 10 (4): 1253–57. https://doi.org/10.1016/j.joi.2016.09.013.

14 Appendices

Appendix Table 1A: Transition of 2002-2004 publications from original classification in Scopus to topic model classification, as a fraction of the papers originally classified in that field

	Topic model allocated field													
Scopus allocated field	**Ag & Bio**	**Arts**	**Bio chem**	**Busi ness**	**Chem Eng**	**Chem istry**	**Comp Sci**	**Decis Sci**	**Earth**	**Econ**	**Energ y**	**Engin**	**Envir**	**Immu n**
General	0.06	0.03	0.18	0.02	0.02	0.04	0.04	0.01	0.11	0.01	0.02	0.04	0.04	0.06
Agricultural & Biological Sciences	0.48	0.01	0.06	0.01	0.02	0.02	0.01	0.01	0.03	0.01	0.01	0.01	0.12	0.04
Arts & Humanities	0.00	0.72	0.00	0.02	0.00	0.00	0.01	0.00	0.01	0.01	0.00	0.01	0.01	0.00
Biochemistry, Genetics & Molecular Biology	0.04	0.00	0.40	0.00	0.01	0.10	0.04	0.01	0.00	0.00	0.00	0.01	0.01	0.08
Business, Management & Accounting	0.00	0.03	0.00	0.41	0.02	0.00	0.05	0.05	0.00	0.13	0.02	0.04	0.01	0.00
Chemical Engineering	0.01	0.01	0.01	0.11	0.27	0.11	0.02	0.02	0.01	0.01	0.10	0.06	0.02	0.01
Chemistry	0.02	0.00	0.03	0.01	0.05	0.60	0.01	0.00	0.01	0.00	0.01	0.01	0.01	0.01
Computer Science	0.00	0.01	0.00	0.04	0.00	0.01	0.56	0.06	0.01	0.01	0.01	0.10	0.00	0.00
Decision Sciences	0.00	0.01	0.00	0.08	0.00	0.00	0.14	0.39	0.00	0.05	0.01	0.03	0.00	0.00
Earth & Planetary Sciences	0.02	0.01	0.00	0.01	0.02	0.02	0.03	0.01	0.56	0.01	0.06	0.06	0.08	0.00
Economics, Econometrics & Finance	0.00	0.03	0.00	0.06	0.00	0.00	0.01	0.03	0.00	0.63	0.03	0.00	0.01	0.00
Energy	0.00	0.00	0.00	0.03	0.08	0.03	0.02	0.02	0.03	0.02	0.49	0.09	0.02	0.00
Engineering	0.00	0.01	0.00	0.04	0.02	0.01	0.19	0.04	0.02	0.01	0.05	0.32	0.01	0.00
Environmental Science	0.11	0.01	0.01	0.03	0.05	0.03	0.02	0.02	0.06	0.02	0.05	0.04	0.38	0.02
Immunology & Microbiology	0.05	0.00	0.11	0.00	0.01	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.55
Materials Science	0.00	0.00	0.00	0.05	0.04	0.09	0.02	0.01	0.01	0.00	0.02	0.07	0.01	0.00
Mathematics	0.00	0.01	0.00	0.00	0.00	0.00	0.16	0.04	0.01	0.01	0.00	0.03	0.00	0.00
Medicine	0.01	0.01	0.08	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.01	0.06
Neuroscience	0.01	0.01	0.08	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.02
Nursing	0.01	0.02	0.01	0.02	0.00	0.00	0.01	0.01	0.00	0.01	0.00	0.00	0.00	0.01
Pharmacology, Toxicology & Pharmaceutics	0.02	0.00	0.10	0.02	0.01	0.12	0.01	0.01	0.00	0.00	0.00	0.00	0.02	0.05

	Topic model allocated field													
Scopus allocated field	**Ag & Bio**	**Arts**	**Biochem**	**Business**	**Chem Eng**	**Chemistry**	**Comp Sci**	**Decis Sci**	**Earth**	**Econ**	**Energy**	**Engin**	**Envir**	**Immun**
Physics & Astronomy	0.00	0.00	0.00	0.00	0.01	0.06	0.03	0.01	0.05	0.00	0.01	0.08	0.00	0.00
Psychology	0.00	0.05	0.00	0.02	0.00	0.00	0.02	0.01	0.00	0.01	0.00	0.00	0.00	0.00
Social Sciences	0.01	0.13	0.00	0.04	0.00	0.00	0.04	0.02	0.01	0.07	0.01	0.02	0.02	0.00
Veterinary	0.06	0.01	0.02	0.01	0.00	0.00	0.00	0.01	0.00	0.01	0.00	0.00	0.01	0.09
Dentistry	0.00	0.01	0.03	0.01	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.01	0.00	0.03
Health Professions	0.00	0.01	0.02	0.01	0.00	0.00	0.03	0.01	0.00	0.01	0.00	0.01	0.00	0.01

Appendix Table 1B: Transition of 2002-2004 publications from original classification in Scopus to topic model classification, as a fraction of the papers originally classified in that field

	Topic model allocated field												
Scopus allocated field	**Materials**	**Math**	**Med**	**Neuro**	**Nurs**	**Pharmacy**	**Physic**	**Psych**	**Social**	**Vet**	**Dent**	**Health Prof**	
General	0.04	0.07	0.01	0.04	0.01	0.01	0.08	0.01	0.03	0.01	0.00	0.00	1.00
Agricultural & Biological Sciences	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.02	0.05	0.00	0.01	1.00
Arts & Humanities	0.01	0.01	0.00	0.00	0.01	0.00	0.01	0.03	0.12	0.00	0.00	0.00	1.00
Biochemistry, Genetics & Molecular Biology	0.01	0.01	0.08	0.05	0.01	0.06	0.01	0.01	0.01	0.01	0.01	0.02	1.00
Business, Management & Accounting	0.02	0.01	0.00	0.00	0.02	0.00	0.00	0.04	0.12	0.01	0.00	0.00	1.00
Chemical Engineering	0.11	0.03	0.01	0.00	0.01	0.01	0.03	0.00	0.01	0.00	0.00	0.01	1.00
Chemistry	0.09	0.01	0.00	0.00	0.00	0.03	0.08	0.00	0.00	0.00	0.00	0.00	1.00
Computer Science	0.01	0.10	0.00	0.01	0.01	0.00	0.03	0.01	0.03	0.00	0.00	0.01	1.00
Decision Sciences	0.00	0.20	0.00	0.00	0.01	0.00	0.01	0.01	0.06	0.00	0.00	0.00	1.00
Earth & Planetary Sciences	0.03	0.03	0.00	0.00	0.00	0.00	0.05	0.00	0.02	0.00	0.00	0.00	1.00
Economics, Econometrics & Finance	0.00	0.04	0.00	0.00	0.01	0.00	0.00	0.02	0.11	0.00	0.00	0.00	1.00
Energy	0.06	0.01	0.00	0.00	0.00	0.00	0.04	0.00	0.02	0.00	0.00	0.00	1.00
Engineering	0.07	0.06	0.00	0.00	0.01	0.00	0.10	0.00	0.02	0.00	0.00	0.01	1.00
Environmental Science	0.02	0.01	0.01	0.00	0.01	0.02	0.01	0.01	0.06	0.01	0.00	0.00	1.00
Immunology & Microbiology	0.00	0.00	0.09	0.01	0.02	0.03	0.00	0.00	0.01	0.03	0.00	0.01	1.00
Materials Science	0.50	0.02	0.00	0.00	0.00	0.00	0.15	0.00	0.00	0.00	0.00	0.00	1.00
Mathematics	0.01	0.62	0.00	0.00	0.00	0.00	0.06	0.01	0.01	0.00	0.00	0.00	1.00
Medicine	0.00	0.00	0.43	0.03	0.09	0.05	0.00	0.05	0.02	0.02	0.02	0.07	1.00
Neuroscience	0.00	0.01	0.09	0.51	0.02	0.05	0.01	0.10	0.01	0.01	0.00	0.04	1.00
Nursing	0.00	0.00	0.13	0.01	0.56	0.02	0.00	0.07	0.08	0.00	0.00	0.03	1.00
Pharmacology, Toxicology & Pharmaceutics	0.01	0.00	0.07	0.06	0.05	0.38	0.00	0.02	0.01	0.01	0.00	0.01	1.00
Physics & Astronomy	0.10	0.08	0.00	0.00	0.00	0.00	0.54	0.00	0.00	0.00	0.00	0.01	1.00
Psychology	0.00	0.01	0.02	0.06	0.06	0.01	0.00	0.60	0.08	0.00	0.00	0.01	1.00

	Topic model allocated field												
Scopus allocated field	**Materials**	**Math**	**Med**	**Neuro**	**Nurs**	**Pharmacy**	**Physic**	**Psych**	**Social**	**Vet**	**Dent**	**Health Prof**	
Social Sciences	0.00	0.01	0.01	0.00	0.05	0.00	0.00	0.08	0.46	0.00	0.00	0.00	1.00
Veterinary	0.00	0.00	0.06	0.01	0.03	0.03	0.00	0.01	0.02	0.55	0.01	0.03	1.00
Dentistry	0.02	0.00	0.07	0.01	0.06	0.01	0.00	0.02	0.01	0.01	0.66	0.02	1.00
Health Professions	0.00	0.01	0.14	0.03	0.13	0.02	0.01	0.06	0.04	0.01	0.01	0.42	1.00

Appendix Table 2A: Transition of 2002-2004 publications from original classification in Scopus to topic model classification using full distribution, as a fraction of the papers originally classified in that field

	Topic model allocated field													
Scopus allocated field	**Ag & Bio**	**Arts**	**Biochem**	**Business**	**Chem Eng**	**Chemistry**	**Comp Sci**	**Decis Sci**	**Earth**	**Econ**	**Energy**	**Engin**	**Envir**	**Immun**
General	0.06	0.04	0.18	0.02	0.02	0.04	0.04	0.01	0.11	0.01	0.01	0.04	0.04	0.07
Agricultural & Biological Sciences	0.47	0.01	0.06	0.01	0.02	0.02	0.01	0.01	0.03	0.01	0.01	0.01	0.12	0.04
Arts & Humanities	0.00	0.70	0.00	0.02	0.00	0.00	0.02	0.01	0.01	0.01	0.00	0.01	0.01	0.00
Biochemistry, Genetics & Molecular Biology	0.04	0.00	0.40	0.00	0.01	0.10	0.03	0.01	0.00	0.00	0.00	0.01	0.01	0.08
Business, Management & Accounting	0.00	0.03	0.00	0.40	0.02	0.00	0.05	0.05	0.00	0.13	0.02	0.04	0.01	0.00
Chemical Engineering	0.01	0.01	0.01	0.10	0.27	0.11	0.02	0.02	0.01	0.01	0.10	0.06	0.02	0.01
Chemistry	0.02	0.00	0.03	0.01	0.05	0.59	0.01	0.01	0.01	0.00	0.01	0.01	0.01	0.01
Computer Science	0.00	0.01	0.00	0.03	0.00	0.01	0.55	0.06	0.01	0.01	0.01	0.10	0.00	0.00
Decision Sciences	0.00	0.01	0.00	0.08	0.00	0.00	0.13	0.39	0.00	0.05	0.01	0.03	0.00	0.00
Earth & Planetary Sciences	0.02	0.01	0.00	0.01	0.02	0.02	0.03	0.01	0.55	0.01	0.06	0.06	0.07	0.00
Economics, Econometrics & Finance	0.00	0.03	0.00	0.06	0.00	0.00	0.01	0.03	0.00	0.61	0.04	0.00	0.01	0.00
Energy	0.00	0.01	0.00	0.03	0.08	0.03	0.02	0.02	0.03	0.02	0.47	0.10	0.02	0.00
Engineering	0.00	0.01	0.00	0.04	0.02	0.01	0.18	0.04	0.02	0.01	0.05	0.32	0.01	0.00
Environmental Science	0.11	0.01	0.01	0.03	0.05	0.04	0.02	0.02	0.06	0.02	0.04	0.04	0.37	0.02
Immunology & Microbiology	0.05	0.00	0.11	0.00	0.01	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.02	0.54
Materials Science	0.00	0.01	0.00	0.04	0.04	0.09	0.02	0.01	0.01	0.00	0.02	0.07	0.01	0.00
Mathematics	0.00	0.01	0.00	0.01	0.01	0.00	0.16	0.05	0.01	0.01	0.00	0.03	0.00	0.00
Medicine	0.01	0.01	0.08	0.01	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.01	0.01	0.06
Neuroscience	0.01	0.01	0.08	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.01	0.00	0.02
Nursing	0.01	0.02	0.01	0.02	0.00	0.00	0.01	0.01	0.00	0.01	0.00	0.00	0.00	0.01
Pharmacology, Toxicology & Pharmaceutics	0.02	0.00	0.10	0.02	0.01	0.12	0.01	0.01	0.00	0.00	0.00	0.00	0.02	0.05
Physics & Astronomy	0.00	0.00	0.00	0.00	0.01	0.05	0.04	0.01	0.05	0.00	0.01	0.08	0.00	0.00
Psychology	0.00	0.05	0.00	0.02	0.00	0.00	0.03	0.01	0.00	0.01	0.00	0.01	0.00	0.00

	Topic model allocated field													
Scopus allocated field	**Ag & Bio**	**Arts**	**Biochem**	**Business**	**Chem Eng**	**Chemistry**	**Comp Sci**	**Decis Sci**	**Earth**	**Econ**	**Energy**	**Engin**	**Envir**	**Immun**
Social Sciences	0.01	0.13	0.00	0.04	0.00	0.00	0.04	0.02	0.01	0.07	0.01	0.01	0.02	0.00
Veterinary	0.06	0.01	0.02	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.00	0.01	0.09
Dentistry	0.00	0.01	0.03	0.01	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.01	0.00	0.03
Health Professions	0.00	0.01	0.02	0.01	0.00	0.00	0.03	0.01	0.00	0.01	0.01	0.01	0.00	0.01

Appendix Table 2B: Transition of 2002-2004 publications from original classification in Scopus to topic model classification using full distribution, as a fraction of the papers originally classified in that field

	Topic model allocated field												
Scopus allocated field	**Materials**	**Math**	**Med**	**Neuro**	**Nurs**	**Pharmac**	**Physic**	**Psych**	**Social**	**Vet**	**Dent**	**Health Prof**	
General	0.04	0.07	0.01	0.05	0.01	0.01	0.08	0.01	0.03	0.01	0.00	0.00	1.00
Agricultural & Biological Sciences	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.02	0.05	0.00	0.01	1.00
Arts & Humanities	0.01	0.01	0.00	0.00	0.01	0.00	0.01	0.03	0.12	0.00	0.00	0.00	1.00
Biochemistry, Genetics & Molecular Biology	0.01	0.01	0.08	0.05	0.01	0.06	0.01	0.01	0.01	0.01	0.01	0.02	1.00
Business, Management & Accounting	0.02	0.01	0.01	0.00	0.02	0.00	0.00	0.04	0.12	0.01	0.00	0.00	1.00
Chemical Engineering	0.11	0.03	0.01	0.01	0.01	0.01	0.03	0.00	0.01	0.00	0.01	0.01	1.00
Chemistry	0.09	0.01	0.00	0.00	0.00	0.03	0.08	0.00	0.00	0.00	0.00	0.00	1.00
Computer Science	0.01	0.10	0.00	0.01	0.01	0.00	0.03	0.01	0.03	0.00	0.00	0.01	1.00
Decision Sciences	0.00	0.20	0.00	0.00	0.01	0.00	0.01	0.01	0.05	0.00	0.00	0.00	1.00
Earth & Planetary Sciences	0.03	0.03	0.00	0.00	0.00	0.00	0.05	0.00	0.02	0.00	0.00	0.00	1.00
Economics, Econometrics & Finance	0.00	0.04	0.00	0.00	0.01	0.00	0.00	0.02	0.11	0.00	0.00	0.00	1.00
Energy	0.06	0.01	0.00	0.00	0.00	0.00	0.04	0.00	0.02	0.00	0.00	0.01	1.00
Engineering	0.07	0.06	0.00	0.00	0.01	0.00	0.09	0.00	0.02	0.00	0.00	0.01	1.00
Environmental Science	0.02	0.01	0.01	0.00	0.01	0.03	0.01	0.01	0.06	0.01	0.00	0.00	1.00
Immunology & Microbiology	0.00	0.00	0.09	0.01	0.02	0.03	0.00	0.01	0.01	0.03	0.01	0.01	1.00
Materials Science	0.49	0.02	0.00	0.00	0.00	0.00	0.15	0.00	0.00	0.00	0.00	0.00	1.00
Mathematics	0.01	0.61	0.00	0.00	0.00	0.00	0.07	0.01	0.01	0.00	0.00	0.00	1.00
Medicine	0.00	0.00	0.42	0.03	0.09	0.05	0.01	0.05	0.02	0.02	0.02	0.07	1.00
Neuroscience	0.00	0.01	0.09	0.50	0.02	0.06	0.01	0.10	0.01	0.01	0.00	0.04	1.00
Nursing	0.00	0.00	0.12	0.01	0.55	0.02	0.00	0.07	0.08	0.01	0.01	0.03	1.00
Pharmacology, Toxicology & Pharmaceutics	0.01	0.00	0.07	0.06	0.05	0.37	0.00	0.02	0.01	0.01	0.00	0.01	1.00
Physics & Astronomy	0.10	0.08	0.00	0.00	0.00	0.00	0.53	0.00	0.00	0.00	0.00	0.01	1.00
Psychology	0.00	0.01	0.02	0.06	0.06	0.01	0.00	0.59	0.08	0.00	0.00	0.01	1.00

	Topic model allocated field												
Scopus allocated field	Materials	Math	Med	Neuro	Nurs	Pharmac	Physic	Psych	Social	Vet	Dent	Health Prof	
Social Sciences	0.00	0.01	0.01	0.00	0.05	0.00	0.00	0.08	0.46	0.00	0.00	0.01	1.00
Veterinary	0.00	0.00	0.06	0.01	0.03	0.03	0.00	0.01	0.02	0.54	0.02	0.03	1.00
Dentistry	0.02	0.00	0.06	0.01	0.06	0.01	0.00	0.02	0.01	0.01	0.64	0.02	1.00
Health Professions	0.00	0.01	0.14	0.03	0.12	0.02	0.02	0.06	0.04	0.01	0.01	0.41	1.00

Appendix Table 3A: Weighted overlap of field assignments for 2002-2004

	Ag & Bio	Arts	Bio chem	Busi ness	Chem Eng	Chem istry	Comp Sci	Decis Sci	Earth	Econ	Energy	Engin	Envir	Immun
Agricultural & Biological Sciences	-	0.0042	0.0112	0.0025	0.0051	0.0063	0.0013	0.0016	0.0072	0.0022	0.0017	0.0010	0.0285	0.0137
Arts & Humanities	0.0042	-	0.0011	0.0088	0.0008	0.0007	0.0050	0.0024	0.0044	0.0045	0.0022	0.0019	0.0031	0.0012
Biochemistry, Genetics & Molecular Biology	0.0112	0.0011	-	0.0006	0.0014	0.0126	0.0035	0.0012	0.0005	0.0003	0.0003	0.0007	0.0010	0.0267
Business, Management & Accounting	0.0025	0.0088	0.0006	-	0.0126	0.0008	0.0093	0.0151	0.0013	0.0171	0.0142	0.0086	0.0036	0.0009
Chemical Engineering	0.0051	0.0008	0.0014	0.0126	-	0.0157	0.0011	0.0029	0.0036	0.0024	0.0193	0.0075	0.0094	0.0023
Chemistry	0.0063	0.0007	0.0126	0.0008	0.0157	-	0.0018	0.0014	0.0023	0.0002	0.0026	0.0015	0.0054	0.0053
Computer Science	0.0013	0.0050	0.0035	0.0093	0.0011	0.0018	-	0.0210	0.0043	0.0026	0.0030	0.0226	0.0021	0.0016
Decision Sciences	0.0016	0.0024	0.0012	0.0151	0.0029	0.0014	0.0210	-	0.0022	0.0104	0.0078	0.0111	0.0049	0.0012
Earth & Planetary Sciences	0.0072	0.0044	0.0005	0.0013	0.0036	0.0023	0.0043	0.0022	-	0.0010	0.0073	0.0099	0.0199	0.0006
Economics, Econometrics & Finance	0.0022	0.0045	0.0003	0.0171	0.0024	0.0002	0.0026	0.0104	0.0010	-	0.0097	0.0010	0.0043	0.0006
Energy	0.0017	0.0022	0.0003	0.0142	0.0193	0.0026	0.0030	0.0078	0.0073	0.0097	-	0.0161	0.0098	0.0005
Engineering	0.0010	0.0019	0.0007	0.0086	0.0075	0.0015	0.0226	0.0111	0.0099	0.0010	0.0161	-	0.0026	0.0005
Environmental Science	0.0285	0.0031	0.0010	0.0036	0.0094	0.0054	0.0021	0.0049	0.0199	0.0043	0.0098	0.0026	-	0.0045
Immunology & Microbiology	0.0137	0.0012	0.0267	0.0009	0.0023	0.0053	0.0016	0.0012	0.0006	0.0006	0.0005	0.0005	0.0045	-
Materials Science	0.0028	0.0009	0.0016	0.0048	0.0154	0.0217	0.0011	0.0019	0.0045	0.0003	0.0076	0.0153	0.0023	0.0009
Mathematics	0.0013	0.0029	0.0010	0.0006	0.0048	0.0027	0.0211	0.0149	0.0060	0.0048	0.0013	0.0169	0.0020	0.0006
Medicine	0.0017	0.0018	0.0147	0.0009	0.0009	0.0005	0.0011	0.0013	0.0007	0.0010	0.0008	0.0011	0.0014	0.0176
Neuroscience	0.0029	0.0016	0.0179	0.0003	0.0004	0.0009	0.0033	0.0009	0.0007	0.0003	0.0003	0.0014	0.0006	0.0050
Nursing	0.0014	0.0058	0.0009	0.0084	0.0003	0.0002	0.0038	0.0049	0.0004	0.0066	0.0019	0.0014	0.0020	0.0038

	Ag & Bio	Arts	Bio chem	Busi ness	Chem Eng	Chem istry	Comp Sci	Decis Sci	Earth	Econ	Energy	Engin	Envir	Immun
Pharmacology, Toxicology & Pharmaceutics	0.0068	0.0006	0.0223	0.0015	0.0026	0.0126	0.0005	0.0011	0.0002	0.0004	0.0007	0.0004	0.0031	0.0127
Physics & Astronomy	0.0009	0.0012	0.0019	0.0009	0.0060	0.0163	0.0052	0.0017	0.0119	0.0005	0.0046	0.0245	0.0009	0.0005
Psychology	0.0030	0.0130	0.0011	0.0044	0.0003	0.0003	0.0058	0.0039	0.0007	0.0044	0.0005	0.0013	0.0013	0.0014
Social Sciences	0.0022	0.0307	0.0011	0.0199	0.0005	0.0005	0.0086	0.0074	0.0023	0.0199	0.0060	0.0034	0.0087	0.0021
Veterinary	0.0146	0.0012	0.0053	0.0013	0.0012	0.0011	0.0004	0.0009	0.0006	0.0013	0.0009	0.0004	0.0045	0.0141
Dentistry	0.0020	0.0013	0.0031	0.0012	0.0015	0.0006	0.0008	0.0009	0.0009	0.0005	0.0007	0.0018	0.0010	0.0030
Health Professions	0.0018	0.0012	0.0030	0.0010	0.0014	0.0010	0.0051	0.0016	0.0013	0.0005	0.0011	0.0051	0.0009	0.0013

Appendix Table 3B: Weighted overlap of field assignments for 2002-2004

	Materials	Math	Med	Neuro	Nurs	Pharmac	Physic	Psych	Social	Vet	Dent	Health Prof
Agricultural & Biological Sciences	0.0028	0.0013	0.0017	0.0029	0.0014	0.0068	0.0009	0.0030	0.0022	0.0146	0.0020	0.0018
Arts & Humanities	0.0009	0.0029	0.0018	0.0016	0.0058	0.0006	0.0012	0.0130	0.0307	0.0012	0.0013	0.0012
Biochemistry, Genetics & Molecular Biology	0.0016	0.0010	0.0147	0.0179	0.0009	0.0223	0.0019	0.0011	0.0011	0.0053	0.0031	0.0030
Business, Management & Accounting	0.0048	0.0006	0.0009	0.0003	0.0084	0.0015	0.0009	0.0044	0.0199	0.0013	0.0012	0.0010
Chemical Engineering	0.0154	0.0048	0.0009	0.0004	0.0003	0.0026	0.0060	0.0003	0.0005	0.0012	0.0015	0.0014
Chemistry	0.0217	0.0027	0.0005	0.0009	0.0002	0.0126	0.0163	0.0003	0.0005	0.0011	0.0006	0.0010
Computer Science	0.0011	0.0211	0.0011	0.0033	0.0038	0.0005	0.0052	0.0058	0.0086	0.0004	0.0008	0.0051
Decision Sciences	0.0019	0.0149	0.0013	0.0009	0.0049	0.0011	0.0017	0.0039	0.0074	0.0009	0.0009	0.0016
Earth & Planetary Sciences	0.0045	0.0060	0.0007	0.0007	0.0004	0.0002	0.0119	0.0007	0.0023	0.0006	0.0009	0.0013
Economics, Econometrics & Finance	0.0003	0.0048	0.0010	0.0003	0.0066	0.0004	0.0005	0.0044	0.0199	0.0013	0.0005	0.0005
Energy	0.0076	0.0013	0.0008	0.0003	0.0019	0.0007	0.0046	0.0005	0.0060	0.0009	0.0007	0.0011
Engineering	0.0153	0.0169	0.0011	0.0014	0.0014	0.0004	0.0245	0.0013	0.0034	0.0004	0.0018	0.0051
Environmental Science	0.0023	0.0020	0.0014	0.0006	0.0020	0.0031	0.0009	0.0013	0.0087	0.0045	0.0010	0.0009
Immunology & Microbiology	0.0009	0.0006	0.0176	0.0050	0.0038	0.0127	0.0005	0.0014	0.0021	0.0141	0.0030	0.0013
Materials Science	-	0.0041	0.0007	0.0005	0.0002	0.0024	0.0299	0.0003	0.0004	0.0006	0.0042	0.0014
Mathematics	0.0041	-	0.0006	0.0012	0.0005	0.0004	0.0234	0.0019	0.0010	0.0004	0.0006	0.0017
Medicine	0.0007	0.0006	-	0.0094	0.0223	0.0183	0.0009	0.0078	0.0024	0.0098	0.0085	0.0326
Neuroscience	0.0005	0.0012	0.0094	-	0.0023	0.0192	0.0017	0.0161	0.0007	0.0045	0.0027	0.0135
Nursing	0.0002	0.0005	0.0223	0.0023	-	0.0059	0.0002	0.0284	0.0250	0.0030	0.0037	0.0079
Pharmacology, Toxicology & Pharmaceutics	0.0024	0.0004	0.0183	0.0192	0.0059	-	0.0006	0.0040	0.0009	0.0103	0.0028	0.0055
Physics & Astronomy	0.0299	0.0234	0.0009	0.0017	0.0002	0.0006	-	0.0006	0.0006	0.0003	0.0008	0.0034
Psychology	0.0003	0.0019	0.0078	0.0161	0.0284	0.0040	0.0006	-	0.0174	0.0019	0.0022	0.0061
Social Sciences	0.0004	0.0010	0.0024	0.0007	0.0250	0.0009	0.0006	0.0174	-	0.0016	0.0008	0.0010
Veterinary	0.0006	0.0004	0.0098	0.0045	0.0030	0.0103	0.0003	0.0019	0.0016	-	0.0047	0.0061
Dentistry	0.0042	0.0006	0.0085	0.0027	0.0037	0.0028	0.0008	0.0022	0.0008	0.0047	-	0.0099
Health Professions	0.0014	0.0017	0.0326	0.0135	0.0079	0.0055	0.0034	0.0061	0.0010	0.0061	0.0099	-

Recent Motu Working Papers

All papers in the Motu Working Paper Series are available on our website https://motu.nz, or by contacting us on info@motu.org.nz or +64 4 939 4250.

19-09 Kerr, Suzi, and Catherine Leining. 2019. 'Paying for Mitigation: How New Zealand Can Contribute to Others' Efforts."

19-08 Kerr, Suzi, and Catherine Leining. 2019. "Uncertainty, Risk and Investment and the NZ ETS."

19-07 Leining, Catherine and Suzi Kerr. 2019. 'Managing Scarcity and Ambition in the NZ ETS."

19-06 Kerr, Suzi, Juan-Pablo Montero, Ruben Lubowski, Angela Cadena, Mario Londoño, Adriana Cavallo, Lisa Lafferty, Soffia Alarcon, Oscar Rodriguez, and Angela Solanilla. 2019. "Designing a prototype emissions trading system for Colombia." (forthcoming)

19-05 Maré, David C and Jacques Poot. 2019. "Valuing Cultural Diversity." (forthcoming)

19-04 Kerr, Suzi, Steffen Lippert and Edmund Lou. 2019. "Financial Transfers and Climate Cooperation."

19-03 Fabling, Richard and David C Maré. 2019. "Improved productivity measurement in New Zealand's Longitudinal Business Database." (forthcoming)

19-02 Sin, Isabelle and Judd Ormsby. 2019. "The settlement experience of Pacific migrants in New Zealand: Insights from LISNZ and the IDI"

19-01 Benjamin Davies and David C Maré. 2019. "Relatedness, Complexity and Local Growth."

18-16 Hendy, Jo, Anne-Gaelle Ausseil, Isaac Bain, Élodie Blanc, David Fleming, Joel Gibbs, Alistair Hall, Alexander Herzig, Patrick Kavanagh, Suzi Kerr, Catherine Leining, Laëtitia Leroy, Edmund Lou, Juan Monge, Andy Reisinger, Jim Risk, Tarek Soliman, Adolf Stroombergen, Levente Timar, Tony van der Weerdan, Dominic White and Christian Zammit. 2018. "Land-use modelling in New Zealand: current practice and future needs."

18-15 White, Dominic, Niven Winchester, Martin Atkins, John Ballingall, Simon Coates, Ferran de Miguel Mercader, Suzie Greenhalgh, Andrew Kerr, Suzi Kerr, Jonathan Leaver, Catherine Leining, Juan Monge, James Neale, Andrew Philpott, Vincent Smart, Adolf Stroombergen, and Kiti Suomalainen. 2018. "Energy- and multi-sector modelling of climate change mitigation in New Zealand: current practice and future needs."

18-14 Preston, Kate, David C Maré, Arthur Grimes and Stuart Donovan. 2018. "Amenities and the attractiveness of New Zealand cities."

18-13 Alimi, Omoniyi, David C Maré and Jacques Poot. 2018. "Who partners up? Educational assortative matching and the distribution of income in New Zealand."

18-12 Fabling, Richard. 2018. "Entrepreneurial beginnings: Transitions to self-employment and the creation of jobs."

18-11 Fleming, David A and Kate Preston. 2018. "International agricultural mitigation research and the impacts and value of two SLMACC research projects." (also a Ministry for Primary Industries publication)

18-10 Hyslop, Dean and David Rea. 2018. "Do housing allowances increase rents? Evidence from a discrete policy change."

18-09 Fleming, David A., Ilan Noy, Jacob Pástor-Paz and Sally Owen. 2018. "Public insurance and climate change (part one): Past trends in weather-related insurance in New Zealand."

18-08 Sin, Isabelle, Kabir Dasgupta and Gail Pacheco. 2018. "Parenthood and labour market outcomes." (also a Ministry for Women Report)

18-07 Grimes, Arthur and Dennis Wesselbaum. 2018. "Moving towards happiness."

18-06 Qasim, Mubashir and Arthur Grimes. 2018. "Sustainable economic policy and well-being: The relationship between adjusted net savings and subjective well-being."

18-05 Clay, K Chad, Ryan Bakker, Anne-Marie Brook, Daniel W Hill Jr and Amanda Murdie. 2018. "HRMI Civil and Political Rights Metrics: 2018 Technical Note."

LSE 4784