The 'Disciplinary Effect' of the Performance-based Research Fund Process in New Zealand

Robert A. Buckle and John Creedy

WORKING PAPER 03/2019
February 2019

Working Papers in Public Finance

VICTORIA UNIVERSITY OF WELLINGTON
TE WHARE WĀNANGA O TE ŪPOKO O TE IKA A MĀUI

Chair in Public Finance
Victoria Business School

The Working Papers in Public Finance series is published by the Victoria Business School to disseminate initial research on public finance topics, from economists, accountants, finance, law and tax specialists, to a wider audience. Any opinions and views expressed in these papers are those of the author(s). They should not be attributed to Victoria University of Wellington or the sponsors of the Chair in Public Finance.

Further enquiries to:
The Administrator
Chair in Public Finance
Victoria University of Wellington
PO Box 600
Wellington 6041
New Zealand

Phone: +64-4-463-9656
Email: cpf-info@vuw.ac.nz

Papers in the series can be downloaded from the following website:
http://www.victoria.ac.nz/cpf/working-papers

Working Papers in Public Finance

The 'Disciplinary Effect' of the Performance-based Research Fund Process in New Zealand

Robert A. Buckle and John Creedy

Victoria University of Wellington

Abstract

This paper examines how the research quality of academic disciplines within New Zealand universities has evolved since the first Performance-based Research Fund (PBRF) assessment in 2003. The analysis uses a database consisting of an anonymous 'quality category' (QC) for each person assessed in the 2003 and 2012 PBRF assessment rounds. Individual researchers are assigned to academic discipline groups and the paper measures the distribution of researchers across disciplines and the discipline composition of universities. There has been little change in the distribution and their concentration within and across universities. However, exceptions are increases in the shares of medicine and agriculture, and a reduction in the share of education. Research Average Quality Scores are derived for discipline groups. All groups substantially increased their scores. Transition matrices show that there are significant differences in the dynamics of the various disciplines during the PBRF process. The paper shows that changes in the discipline composition of universities explains little of the proportional improvement of research quality among New Zealand universities.

Key words: academic disciplines, education policy, New Zealand universities, Performance-based Research Fund, productivity, research, transitions.

JEL classifications: I2; I23; I28.

Acknowledgements: We are grateful to Amber Flynn for assisting with engagement with New Zealand Tertiary Education Commission (TEC), and TEC for providing the data. We have benefited from discussions with Sharon Beattie and Shelly Biswell, and comments by Norman Gemmell on an earlier draft of this paper.

1. Introduction

The New Zealand Performance-Based Research Fund (PBRF) scheme was designed to unbundle the research component of Government funding of New Zealand tertiary education organisations (TEOs) and allocate funds based on research performance rather than the number of students. The Tertiary Education Commission explained the aims as follows: 'The primary purpose of the PBRF is to ensure that excellent research in the tertiary education sector is encouraged and rewarded. This entails assessing the research performance of tertiary education organisations (TEOs) and then funding them on the basis of their performance'.[1]

The scheme therefore introduced a new set of incentives facing individual researchers and university leaders: from this point of view it could be said to have a clear 'disciplinary effect' on the priority attached to research performance. Furthermore, to the extent that there are significant differences in the research environment among academic disciplines (relating for example to access to funding, ease of publication, and availability of outside labour market opportunities), different responses to the new incentives may be expected. Furthermore, combined with the varying costs associated with different disciplines, a possible strategy to increase the measured performance within a university involves a change in the allocation of funds among disciplines. Such responses could lead to changes both within and among universities in the concentration of academic disciplines: this is the second type of 'disciplinary effect' to which allusion is made in the title.

The purpose of this paper is to examine changes in the research quality of academic disciplines in New Zealand universities between the first Performance-based Research Fund (PBRF) assessment of research performance in 2003 and the 2012 round. It measures changes in the number of portfolios evaluated in each discipline, and considers whether there have been changes in the concentration of researchers by discipline among and within universities. The rates of improvement in average research quality of discipline groups such as core science, humanities, medicine, and so on are compared. The paper then uses transition matrices to measure the evolution of the quality of research within discipline groups and to assess whether there are differences in this process among the disciplines. The paper also examines the extent to which the improvements in the average research quality of staff in each university is affected by differences in the discipline composition of universities.

[1] See http://www.tec.govt.nz/funding/funding-and-performance/funding/fund-finder/performance-based-researchfund/. See also New Zealand Tertiary Education Commission (2002), Mahoney (2004), Ministry of Education (2012), and Smart and Engler (2013).

The analysis in this paper therefore extends the earlier work of Buckle and Creedy (2017, 2018a, 2018b and 2018c) which evaluated the evolution of research quality of staff for each New Zealand university between the 2003 and 2012 PBRF rounds, and assessed the metrics used in those processes. However, overall changes were considered without any disaggregation by discipline group.[2]

Section 2 briefly places the work in context by discussing previous findings relating to PBRF-type exercises and the New Zealand version. Section 3 describes the discipline composition of researchers at New Zealand universities, and examines how the number of researchers across disciplines changed between 2003 and 2012. Section 3 also examines how each university's share of discipline groups changed during the period and how the discipline composition of each New Zealand university changed. Section 4 measures the average quality of researchers in each discipline group in 2003 and 2012 and compares how these have changed in that period. Buckle and Creedy (2018a) found significant differences among universities in the rates of entry, exit and quality transition of staff. Section 5 provides a similar type of analysis of staff dynamics for each of nine discipline groups. This analysis reveals significant differences among disciplines in these types of quality transition. Section 6 then examines the extent to which changes in the composition and differences in proportional improvements of research quality by discipline groups have contributed to differences in the proportional changes in the average quality of researchers in each New Zealand university and the entire New Zealand university system. This is achieved by decomposing changes in Average Quality Scores (AQSs) for each university between the 2003 and 2012 PBRF rounds into contributions from pure quality changes and contributions from changes in the discipline composition of researchers. Conclusions are discussed in section 7.

2. Previous findings

Performance-based funding schemes have been introduced in many other countries. These schemes vary by coverage and assessment methods, which may be based on bibliometric data

[2] The dataset used here and in the earlier studies is not publicly available. It includes anonymous PBRF data provided by the Tertiary Education Commission (TEC) at the request of the authors following a confidentiality agreement. The data include, for each anonymous researcher, age, research subject area, university of employment, and PBRF quality category for each of the three PBRF assessment rounds undertaken in 2003, 2006 and 2012 in which a researcher's 'evidence portfolio' (EP) was submitted. The present paper concentrates on the changes between 2003 and 2012. The 2006 round was a partial round and universities could choose to submit a new evidence portfolio for researchers covering the previous six years of research or retain the quality category of those researchers who submitted portfolios in 2003 covering the six years prior to 2003. A fourth round was completed in 2018 but the results for that round were not available at the time the research for this paper was undertaken.

or peer review; see Ministry of Education (2013), de Boer *et al.* (2015), Wilsdon *et al.* (2015). The New Zealand scheme uses a peer-review assessment method and assesses individuals rather than groups.[3]

Despite the widespread introduction of university performance-based funding schemes, de Boer *et al*. (2015, p. 5) judge that, 'there still is not sufficient evidence on the effects of the systems and that our understanding of the proper design and implementation of performance agreements is still incomplete'. The present paper is a further contribution to understanding the effects of university performance-based funding schemes. It complements the earlier papers by Buckle and Creedy (2017, 2018a and 2018b) which evaluate responses to the PBRF rounds. The main conclusions from that research can be summarised as follows. There is qualified evidence that the introduction of the PBRF lead to a significant improvement in overall average research quality of staff at New Zealand's eight universities.[4] A dominant feature was a reduction in non-research-active staff in all universities. All universities had substantial proportional improvements in the average research quality of academic staff (that is, non-administration staff). There were considerable differences among universities in the extent of this improvement. The universities fell broadly into two groups consisting of the top 5 and the bottom 3. There was little movement between these two groups, and no systematic tendency for the average quality of researchers at each university to converge.[5] This persistence in rankings was due in part to a relatively high recruitment of higher quality researchers by the five higher-ranking universities, with relatively low exit rates from the group. The lower-ranked universities experienced relatively higher losses from the higher research-quality staff and they recruited a relatively higher proportion of lower research-quality staff.[6] Finally, there was substantial population ageing over the period because of a combination of an increase in the average age of entry and reduced exits from older age groups. This led to changes in both the age distribution within grades and the grade distribution within age groups.

---

[3] Buckle and Creedy (2018c) provide a detailed explanation and critique of the method used to assess the research quality of individual researcher's portfolios in the New Zealand PBRF system and the metrics used to rank and compare universities in the 2003, 2006 and 2012 rounds.

[4] This result is consistent with the finding of Gemmell *et al.* (2017) that research productivity in NZ universities increased markedly since the early 2000s.

[5] See also Buckle and Creedy (2018b, pp. 11-12 and Table 4). This conclusion is based on estimation of convergence of university AQSs where the denominator is all non-administration staff.

[6] This may be a consequence of the self-reinforcing nature of the funding formula used by TEC which reduces the effective price of A-researchers for those universities which are already better placed to attract top quality researchers, and which increases the relative marginal quality improvement of recruiting lower quality faculty by lower ranked universities (see Buckle and Creedy, 2018b, pp. 13-15 and Figure 3)

There are several reasons why there may also be significant differences among disciplines in the research level and average quality of improvement during the PBRF process. These differences could arise for example from differences in research methods, funding opportunities, and from alternative labour market opportunities and differences in the opportunity costs of careers in academia and research. A large number of studies have found that factors such as individual characteristics (age, experience, gender and personal preferences for research), institutional characteristics (performance management, commercial orientation of research, the significance of PhD programmes, and institutional mission and priorities), research methods (collaboration, team research, joint authorship), and the availability of contestable research funding, can all have significant influences on productivity. All these factors can vary among disciplines.

In an early study, Wanner *et al.* (1981) found that discipline was an important influence on the research environment and type of research output. By utilising data from the Changing Academic Profession (CAP) international survey database, Shin and Cummings (2010) concluded that staff collaboration, time spent on research, institutional goal-orientation and institutional mission influenced academic publishing, and these influences differed by academic discipline.[7] Similarly, in an assessment of the influences on the research productivity of Hong Kong academics, Jung (2012) found that the influence of individual and institutional factors on self-reported research output of academics differed significantly between the natural sciences, engineering, and medical science disciplines and the humanities, social sciences, and business disciplines.

Researcher recruitment and turnover are critical to research productivity in an institution or research group. Xu (2008) found substantial and systematic variations exist among different disciplines with respect to the main factors that determine staff turnover. Salaries and the presence of alternative career opportunities have a significant influence on the mobility of research staff and the ability of institutions to recruit research staff in the education field; see Ehrenberg *et al.* (1991). Differences in labour market opportunities can therefore influence differences in research performance among disciplines. Because these labour market opportunities vary among subjects and skills, those academic disciplines with the most valuable alternative opportunities are likely to have the greatest difficulty recruiting and

[7] Information about The Changing Academic Profession project is available from: https://melbourne-cshe.unimelb.edu.au/lh-martin-institute/research/projects/the-changing-academic-profession

maintaining high-calibre researchers. Boyle (2008) demonstrated that differences in labour market opportunities has had a significant effect on the relative research performance of academic disciplines as measured by the first New Zealand PBRF assessment round in 2003. He showed that disciplines in which researchers in New Zealand universities are most underpaid relative to other employment opportunities perform significantly worse than those in which the extent of underpayment is small.

3. The changing discipline composition of PBRF portfolios

3.1 Subject areas and discipline groups

The PBRF data for each anonymous researcher, provided by TEC, include a research subject area, of which there are 42 (see New Zealand Tertiary Education Commission, 2013). For the purpose of this paper, this information is used to allocate each researcher to one of nine discipline groups: Medicine; Engineering; Core Science; Management; Accounting, Finance and Economics (AFE); Humanities; Agriculture; Education; and Law. The choice of discipline groups and the allocation of subject areas to a discipline group was based on the typical make-up of New Zealand university faculties and is similar to the discipline groups formed by Jung (2012). The number of evidence portfolios in each of the nine discipline groups is shown in Table 1; details of the 42 subject areas and their allocation to discipline groups are given in Appendix Table A1.

The total number of portfolios submitted fell from 7041 in 2003 to 6652 in 2012, a fall of 5.5 per cent. However, there was substantial variation across discipline groups and subject areas within disciplines groups. The fifth column in Table 1 shows that increases in the number of evidence portfolios submitted occurred in only two discipline groups, Agriculture and Medicine. They fell in all other discipline groups. The largest declines occurred in Education, Humanities and Management.[8]

There were substantial changes in the composition of some of the discipline groups; see Appendix A. Although total portfolios in Medicine increased by only 4, Biomedical Science increased by 101 and Public Health increased by 51. However, some declined substantially,

[8] As explained in Buckle and Creedy (2018a, Section 3), the period since 2003 coincides with changes in Government tertiary education policy with respect to former Colleges of Education, and there were significant changes in Education staffing levels.

such as Other Health, Nursing, and Clinical Medicine. The predominant source of the increase for Agriculture occurred in Ecology, Evolution and Behaviour. Variation is evident, but generally not so substantial, within the disciplines which experienced declines in portfolio submissions, with some notable exceptions. In the Engineering group, declines occurred in all categories except Engineering and Technology. In the Management group, the category Management, Human Resource Management, Industrial Relations and Other Business declined, while a much smaller reduction occurred in Marketing and Tourism.

Table 1. Discipline groups and PBRF portfolios, 2003 and 2012

Discipline	2003: Number of portfolios	% share	2012: Number of portfolios	% share	Change in Number of portfolios	Change in % share
Medicine	1804	25.62	1808	27.18	4	1.56
Engineering	905	12.85	864	12.99	-41	0.14
Core Science	534	7.58	496	7.45	-38	-0.13
Management	473	6.72	398	5.98	-75	-0.74
Accounting, Finance and Economics	358	5.08	337	5.07	-21	-0.01
Humanities	1458	20.72	1354	20.35	-104	-0.37
Agriculture	548	7.78	652	9.81	104	2.03
Education	738	10.48	553	8.31	-185	-2.17
Law	223	3.17	190	2.86	-33	-0.31
TOTAL	7041	100	6652	100	-389	

3.2 Changes in discipline shares among and within universities

Changes in each university's share of portfolios submitted between 2003 and 2012, for each of the nine discipline groups, are shown in Table 2. For example, from the top left-hand entry, AUT's share of the total number of academics in Medicine who submitted evidence portfolios fell by 3.67 per cent. The total for each row is, by definition, equal to zero since they are share changes. The variations in the percentage changes in shares are indicated by standard deviations, shown in the final column (variations among universities, for each discipline) and row (variations among disciplines, within each university).

Table 2. Percentage changes in university shares of discipline portfolios, 2003 to 2012

Discipline	AUT	Lincoln	Massey	AU	CU	Otago	Waikato	VUW	Discipline standard deviation
Medicine	-3.67	0.22	-0.25	5.58	1.27	-2.42	-0.39	-0.34	2.74
Engineering	1.07	-0.27	-6.79	5.54	1.62	-3.36	1.19	1.01	3.67
Core Science	-0.38	-0.14	-5.13	3.59	0.14	0.57	-0.12	1.48	2.45
Management	-3.41	2.05	-2.96	-3.01	1.21	0.46	3.69	1.95	2.75
AFE	1.24	1.01	-1.33	-2.18	-0.32	-1.35	2.13	0.81	1.52
Humanities	-2.21	-0.06	-0.97	1.32	-0.42	3.14	-1.31	0.51	1.67
Agriculture	-0.17	-0.94	-4.49	5.79	-0.77	-0.38	0.45	0.51	2.83
Education	-2.03	-0.14	-5.91	1.79	4.31	1.63	-0.12	0.46	3.01
Law	1.91	-1.35	-4.91	4.91	-2.56	1.73	-0.67	0.93	3.04
Standard deviation	2.06	1.01	2.35	3.37	1.92	2.10	1.56	0.66	2.53

Note: AUT denotes Auckland University of Technology; Lincoln = Lincoln University; Massey = Massey University; AU = University of Auckland; CU = University of Canterbury; Otago = University of Otago; Waikato = University of Waikato; VUW = Victoria University of Wellington.

The major shifts in discipline shares are as follows. In Medicine, AUT and Otago experienced the largest share decline, offset predominantly by an increased share for Auckland (by about two standard deviations of the mean proportional change) and to a smaller extent Canterbury. In Engineering, the predominant changes were declines in shares for Massey and Otago offset primarily by a rising share for Auckland and to a lesser extent Canterbury. University shares of Agriculture, Education and Law portfolios also experienced relatively large changes. The largest proportionate changes for Agriculture and Law were the declines in Massey's shares of these disciplines, offset by Auckland's increased shares. For Education the largest shifts were a fall in Massey's share and a rise in Canterbury's share. The shift in university shares of other disciplines were as follows. In Core Science, the predominant changes were a decline for Massey offset predominantly by a rising share for Auckland; for Management, Lincoln, Waikato, VUW, Canterbury and Otago (to a smaller extent) increased shares while the shares for AUT, Massey and Auckland declined. For AFE and Humanities the variations were relatively small.

Table 3. Percentage changes in the discipline composition of universities, 2003 to 2012

Discipline	NZ total	AUT	Lincoln	Massey	AU	CU	Otago	WU	VUW	Within university discipline standard deviation
Medicine	1.56	-4.91	1.55	3.36	3.84	3.34	-0.83	-0.61	-0.66	2.98
Engineering	0.14	5.96	-2.71	-2.69	1.14	1.15	-2.24	2.43	1.07	2.99
Core Science	-0.13	0.14	-0.56	-1.36	-0.32	-1.07	0.03	0.01	0.79	0.69
Management	-0.74	-1.62	1.88	-0.75	-1.65	0.11	-0.42	2.13	0.19	1.43
AFE	-0.02	2.35	0.39	0.56	-0.86	-0.51	-0.40	1.62	0.18	1.10
Humanities	-0.35	-1.21	-1.31	1.64	-1.54	-3.18	2.80	-2.90	-0.42	2.07
Agriculture	2.02	0.79	2.73	3.99	2.53	0.79	1.06	2.31	2.01	1.12
Education	-2.17	-2.51	-0.49	-3.94	-2.89	0.77	-0.02	-4.40	-2.71	1.89
Law	-0.31	1.00	-1.48	-0.81	-0.25	-1.39	0.01	-0.61	-0.46	0.80
Standard deviation	1.15	3.11	1.78	2.65	2.15	1.85	1.37	2.42	1.33	

Associated with these shifts in university shares of discipline portfolios are changes in the discipline composition of each university, shown in Table 3. The total for each column is zero, since the changes are for shares of disciplines for each university. The impact on the aggregate discipline shares of total portfolios is shown in the first column of Table 3. The standard deviation of change in proportionate shares of discipline groups for the entire New Zealand university system was only 1.15 per cent (see the last row and first column of the table) and, as expected from the discussion of Table 1, the variation is dominated by increases in the shares for Medicine and Agriculture, and the decline of Education's share.

The within-university variation is more substantial. The largest variations in discipline composition occurred in AUT, Massey, Waikato and Auckland, followed by smaller variations for Canterbury, Lincoln. Otago and VUW display much less variation in discipline group composition. The sources of the changes in discipline composition vary across universities. However, variations in shares of Medicine, Engineering and Education, and to a lesser extent Agriculture and Humanities, tend to be the major sources of variation in discipline composition for those universities displaying the largest change in composition between 2003 and 2012. Core Science and Law tend to display the smallest variation in discipline shares of universities.

3.3 Concentration of disciplines

The question arises of how changes in discipline shares affected the concentration of disciplines within and among universities. This is considered using the normalised Herfindahl-Hirschman Index (*HHI*), defined as follows. Let p_i denote the proportional share of group i, with $i = 1,...,N$, and define *H* as:

$$H = \sum_{i=1}^{N} p_i^2 \tag{1}$$

Then $HHI = \left(H - 1/N\right)/\left(1 - 1/N\right)$. Values of the index are calculated for the nine discipline groups, for university shares of evidence portfolios in each of the discipline group, and for the discipline composition of each of the eight universities. The results are shown in Table 4, together with the change for the 42 research subject areas.

Table 4. Herfindahl-Hirschman concentration indices for disciplines and universities

		2003 *HHI*	2012 *HHI*	Change
Research subject areas (*N* = 42)	Total	0.014	0.012	-0.002
Discipline groups (*N* = 9)	Total	0.051	0.056	0.005
University shares of discipline groups (*N* = 8)	Medicine	0.172	0.180	0.008
	Engineering	0.052	0.063	0.011
	Core Science	0.078	0.094	0.016
	Management	0.033	0.019	-0.014
	AFE	0.018	0.012	-0.006
	Humanities	0.030	0.036	0.006
	Agriculture	0.074	0.055	-0.019
	Education	0.058	0.057	-0.001
	Law	0.055	0.086	0.031
University composition by discipline groups (*N* = 9)	AUT	0.085	0.071	-0.014
	Lincoln	0.147	0.170	0.023
	Massey	0.030	0.045	0.015
	AU	0.069	0.086	0.017
	CU	0.055	0.047	-0.008
	Otago	0.261	0.259	-0.002
	WU	0.057	0.035	-0.022
	VUW	0.056	0.052	-0.004

A dominant feature of these results is the generally low degree of concentration, both of disciplines within universities and of university shares of discipline groups. However, of discipline areas across universities, Medicine is the most concentrated. Otago and Lincoln are the most concentrated in terms of staff members of discipline groups. The higher concentration at Otago and Lincoln reflects the relatively high proportion of staff members in Medicine at Otago and Agriculture at Lincoln. Although Auckland also has a relatively high proportion of portfolios in Medicine and Victoria has a relatively high proportion of portfolios in Humanities, in general there is little specialisation either within or among universities.

Considering the changes in concentration between 2003 and 2012, the PBRF exercise has clearly had a negligible effect on the initial degree of specialisation. At the level of the 42 subject areas there was a small decline in concentration, but at the level of the nine discipline groups there was a small increase in concentration. The *HHI* for university shares of each discipline group indicate that the increase in overall concentration was the result predominantly of increases in concentration amongst universities of portfolios in Medicine, Engineering, Core Science, Humanities and Law, offset by decreasing concentration of portfolios in Agriculture and Management and to a smaller extent AFE and Education. Changes in discipline composition within individual universities is also evident but varies. Lincoln, Massey, and Auckland universities became more concentrated, Waikato and AUT less concentrated, while much smaller reductions in concentration occurred at Canterbury, Otago and VUW.

4. Changes in Average Quality Scores (AQSs) by discipline

This section examines the change in research quality by discipline groups over the period, 2003 to 2012. Research quality of each discipline group is measured by using the quality category assigned to each researcher by the PBRF assessment process to derive Average Quality Scores (AQS) for each discipline groups. The AQS is the metric used by the Tertiary Education Commission to describe the research quality of universities and research areas. Subsection 4.1 briefly describes the complex process used in arriving at an AQS for any group. Subsection 4.2 reports changes in AQSs for the separate discipline groups and provides a preliminary examination of the contributions to changes in AQSs for each discipline group.

4.1 Definition of the AQS

For each portfolio, a quality category, QC, is determined by a panel assigned to a subject area or group of subject areas.[9] The relevant subject panel assesses the quality of each portfolio and assigns a score from 0 to 7 for each of three categories: these are 'research output'; 'peer esteem'; and 'contribution to research environment'. These three scores, s_i, are given weights, w_i, of 0.70, 0.15 and 0.15. The total score, S, for an individual is obtained by multiplying the weighted sum of the s_i values by 100. Hence:

$$S = 100\sum_{j=1}^{3} w_j s_j \tag{2}$$

Thus the maximum individual score is 700. A letter grade is then assigned depending on the assessed total as follows: R for scores 0 to 199; C for scores between 200 and 399; B for scores from 400 to 599; and A for scores from 600 to 700.[10] A numerical score, G, is then assigned to each letter grade: 10 for an A; 6 for a B; 2 for a C; and 0 for R. A university's average quality score, AQS, is the employment-weighted mean score, which can range from zero to 10. Define the employment weight of person i as $e_i \leq 1$, and let n denote the relevant number of employees in a university. The average quality score is:

$$AQS = \frac{\sum_{i=1}^{n} e_i G_i}{\sum_{i=1}^{n} e_i} \tag{3}$$

Since the grade for R-type staff is equal to zero, their number affects only the denominator in equation (3).

AQSs can also be derived for each of the discipline groups described in Table 1. This is done by using the same weights or numerical scores, G, applied by the TEC, to the QCs achieved by each individual researcher and by using the information indicating their subject area to

---

9 The assessment and scoring method used in the New Zealand PBRF system for the 2003, 2006 and 2012 rounds are described in more detail and critically evaluated in Buckle and Creedy, 2018a, 2018b and 2018c.

10 The recognition that new researchers may take time to establish their research, publications, and academic reputations led to the introduction in 2006 of the new categories, C(NE) and R(NE). These categories applied to new and emerging researchers who did not have the benefit of a full six-year period. The following analysis does not distinguish the NE categories, since numerical scores are not affected.

assign their score to one of the nine discipline groups described in Table 1.[11] However, there is an important qualification that should be recognised when comparing estimated AQSs between 2003 and 2012.

A late change was made in 2012 regarding the assessment process when the TEC determined that universities could choose to not submit portfolios for researchers deemed to be of R-quality. Therefore, the estimated AQSs for groups of researchers (whether the groups are universities or subject areas or disciplines) are not strictly comparable with those calculated for 2003. To account for differences among universities in their treatment of R-quality researchers in 2012, Buckle and Creedy (2017, 2018a and 2018b) used total full-time equivalent non-administration staff as the denominator. The university-level ratios of evidence portfolios to total non-administration staff varied substantially in 2003 and although the ratio increased for all universities (from about 60 per cent in 2003 to 70 per cent in 2012), there was still a substantial variation among universities in 2012 (see Buckle and Creedy, 2018b, pp. 29-30). However, data for total non-administration staff (either full-time equivalent or head-count) are not available at the level of subject areas or discipline groups. Since the discipline composition also varies substantially among universities (which may imply that the treatment of R-quality researchers may also vary by discipline), the estimation and analysis of changes in discipline group AQSs that follows should be viewed with this caveat in mind.

4.2 Changes in AQSs of discipline groups.

As discussed in section 2, there are many reasons why there may be significant differences in the level and quality of research activity among disciplines. These differences can arise from different research cultures and incentives. For example, prior to the introduction of the PBRF scheme New Zealand Colleges of Education for a long period placed a different priority on research than disciplines in universities. Amongst disciplines that have had a long tradition of research, differential access to research funding and the costs of undertaking research can vary. There are also differences in other labour market opportunities which have been shown to influence research quality. Differences in publication practices (such as rejection rates,

[11] The employment weight per researcher, e_i is not available from the data set used in this study and hence in deriving each discipline AQS, $e_i = 1$. for each portfolio (i.e., for each i). In the calculation of university AQSs, Buckle and Creedy (2018a, section 4.1, Table 3) show that employment weighting makes only a small difference to the calculation of the university-level and New Zealand total AQSs and the ranking of university AQSs.

number of journals available, the importance of journal versus book publishing) might influence the assessment of research quality when evaluation is carried out in mixed peer-review panels.

It is therefore not surprising that the average quality of research varied substantially amongst disciplines when the first PBRF round was undertaken. This is shown by the AQSs for 2003 displayed in the top row of Table 5. The score for Education is between one third and a half of the scores for most other disciplines. Some disciplines which tend to have considerable scope for other labour market opportunities, such as Management, also had relatively low AQSs in the initial 2003 PBRF round. Boyle (2008) has demonstrated that subject areas in which researchers in New Zealand universities are most underpaid, relative to salaries in USA universities, performed significantly worse in the 2003 PBRF than those in which the extent of underpayment is small.

Similarly, the scope to improve the quality of research varies amongst disciplines and institutions. By tying funding to research performance, the introduction of PBRF schemes has changed the incentives for university managers and research staff. Woelert and McKenzie (2018) find that Australian universities have tended to replicate the national performance indicators used in the national PBRF scheme in their individual level performance management frameworks for academic staff.

The potential for improvement in research capability depends on the type of performance management schemes that existed previously, and the scope to change the structure of departments and units within universities. Disciplines with initially relatively high proportions of R-quality staff could be expected to have greater scope to improve the average research quality, but this is also influenced by other factors such as revenue growth and the growth of students and the pressures on teaching requirements. It also depends on the relative costs of attracting A and B-quality staff, and opportunities and costs of improving the research capability of the existing staff.

There are reasons to expect that, following the introduction of a PBRF-type funding scheme, higher growth rates of research quality would be evident amongst disciplines with relatively lower initial AQSs. For disciplines and universities with a low initial AQS, an increase in C-quality staff can be beneficial, but for those disciplines with higher AQSs, recruitment of Cs may lower the AQS (see the illustration in Buckle and Creedy, 2018b, pp. 35-37). Moreover, recruitment of A and B-quality staff is relatively costly and hence disciplines and universities with initial high research capability may face diminishing returns from continuing to invest in higher-quality research capability.

Table 5. Growth of discipline group average quality scores (AQS), 2003 to 2012

	All	Med	Eng	CS	Mment	AFE	Hum	Agric	Educ	Law
AQS 2003	2.81	2.79	2.94	3.77	2.17	2.42	3.13	3.61	1.33	3.00
AQS 2012	4.52	4.43	4.51	5.34	3.93	4.01	4.69	4.95	3.79	4.94
% growth	60.8	58.6	53.4	41.7	81.2	65.6	49.6	37.2	185.1	64.8
Rank	n.a.	5	6	8	2	3	7	9	1	4
Growth in AQS if a common score were assigned to A, B & C categories										
% growth	45.9	12.6	45.3	25.1	68.3	59.1	33.9	17.8	149.9	38.3
Rank	n.a.	9	4	7	2	3	6	8	1	5
Growth in AQS above that which would result from a common score for A, B & C categories										
% growth	14.9	46.0	8.1	16.6	12.9	6.5	15.7	19.4	35.2	26.5
Rank	n.a.	1	8	5	7	9	6	4	2	3

Note: The AQSs shown in this table are based on a calculation that uses total submitted portfolios as the denominator. This explains why the calculations for "All" differs and is higher than the calculations for "NZ" in Table 4 in Buckle and Creedy (2018b) which uses total non-administration staff as the denominator.

Table 5 shows the AQSs (based on submitted portfolios) for each discipline group for 2003 and 2012, along with their respective proportional growth rates. There is substantial variation in the growth rates of discipline AQSs between 2003 and 2012. The growth rates range from 185 per cent for Education to 37 per cent for Agriculture. It is shown in Section 5 that Education experienced a significantly different pattern of changes between 2003 and 2012 compared with those for all disciplines combined, and this difference is explained primarily by the movements of R-quality researchers. On the other hand, transition patterns in Agriculture were not markedly different from the average for all disciplines (as shown in Table 9) despite being the discipline that experienced the highest aggregate growth rate of researchers, as shown in Table 1.

Improvements in research quality can therefore arise from removal of lower-quality researchers and from changing the distribution of researchers across the quality categories. The relative importance of these processes is examined in more detail in Section 5, but a preliminary assessment of the contribution from the distribution of research quality changes for each discipline group is provided in Table 5. The lower two panels of Table 5 show the hypothetical growth rates of discipline AQSs if each quality category is assigned a common numerical score, G, for all non-R researchers; that is, G is 6 for A, B and C researchers. The middle panel shows the growth in discipline group AQSs, and the bottom panel shows the growth in AQSs above that which would result from a common numerical score for *G*. That is, the bottom panel shows the difference between the growth rates of discipline AQSs shown in the top panel and the middle panel; it therefore shows the contribution of the differences between the distribution across quality categories of the stocks of researcher portfolios in 2003 and 2012.

For all disciplines combined, the growth in AQS of submitted portfolios was 60.8 per cent.[12] The contribution from transforming the distribution of the stock of researchers was only 14.9 per cent. Clearly, the reduction in the proportion of R-quality researchers (either by a net reduction of R-quality researchers via non-submission of a portfolio or exits exceeding entrants) has been the major factor contributing to the rise in the overall AQS. However, there is considerable variation among disciplines. Medicine and Agriculture sourced over fifty per cent of the improvement in the AQS for those disciplines from a change in the distribution of the stock of researchers across the quality categories: Medicine sourced 78 per cent (i.e., 46/58.6) and Agriculture 52 per cent (i.e., 19.4/37.2). For all other disciplines the proportion of the improvement was less than fifty per cent.[13] This variation in the sources of improvement in AQSs suggests that the transitions between 2003 and 2012 varied among disciplines.

5. Quality category transitions by research discipline

This section examines the differences among disciplines in the transitions amongst the various quality categories between 2003 and 2012. As discussed above, there are several factors that could cause differences in the ability of universities to improve the research quality of disciplines and therefore it may be expected that the transitions vary by discipline. If there are differences in the quality improvement of disciplines, the changes in the discipline composition of universities could be a factor contributing to the large differences among universities in the average quality improvement that occurred between 2003 and 2012.

[12] This growth refers to the growth in AQS where the denominator for expression (2) is the sum of portfolios submitted in each of the 2003 and 2012 rounds. It does not include non-administration staff for whom no portfolio was submitted in 2012 because these cannot be identified by discipline. The growth in AQS for "All" when total non-administration staff are included in the denominator (as in Table 4 of Buckle and Creedy, 2018b) is 91 per cent.

[13] The changes are: Law (41 per cent), Core Science (40 per cent), Humanities (32 per cent), Education (19 per cent), Management (16 per cent), Engineering (15 per cent), AFE (10 per cent).

Table 5 is suggestive of significant differences and the purpose of this section is to identify more precisely the differences among disciplines in quality category transitions.

5.1. Transitions for all disciplines combined

First, consider the transitions among the various categories, along with entrants and exits, for all disciplines combined. These are shown for movements from 2003 to 2012 in Table 6, which taken from Buckle and Creedy (2018a, Table 8). The flows are from rows to columns, and the transition proportions are given in parentheses immediately beneath the frequencies. The flows for those remaining in the same category (the diagonal entries in the matrix) are highlighted in bold.[14]

These overall flows show that a low recruitment rate (just under 6 per cent) and a high exit rate (just over 70 per cent) of Rs is a strong feature of transitions over the period from 2003 to 2012.[15] The majority of entrants between 2003 and 2012 are classed as Bs and Cs in 2012 (at 35 and 52 per cent respectively), and a low proportion of entrants are classified as A-researchers in 2012 (7.5 per cent). For any university and discipline with an AQS in 2003 below 2, the recruitment of a C-quality researcher would increase their score. Although Table 5 shows that only education has an AQS below 2 in 2003, these values are based on the number of portfolios submitted, rather than all non-administration staff. As shown in Buckle and Creedy (2018b, p. 33), when all non-administration staff are assigned an R grade, there were four universities with an overall AQS below 2 in 2003. It is not possible here to distinguish non-administration staff by discipline, so an equivalent adjustment is not possible. However, it is clear that there were widespread benefits from recruiting C-quality researchers, particularly those judged to have substantial growth potential. Among all universities and disciplines, the recruitment of B-quality researchers always improved the 2003 AQS.

Of those classed as C-researchers in 2003, a high proportion (46 per cent) had exited by 2012. It is perhaps likely that many of these moves were to another university and may have involved promotion, but this information is not available from the data. The Bs also experienced high outward mobility (34 per cent), and a substantial proportion (29 per cent) of

[14] Details about movements from one NZ university to another NZ university were not available. Hence, they appear in the table simply as being among the exits and entrants.

[15] However, as explained in Buckle and Creedy (2018a), the proportions for exits are somewhat misleading because some of the individuals could simply have been given new contracts which meant that they avoided the need to submit a portfolio.

those who were As in 2003 had exited. There can be a large proportion of As who left between 2003 and 2012, since the denominator (initial number of As) is much smaller.

Table 6. Matrix of flows: All disciplines combined, 2003 to 2012

	Category in 2012					
Category in 2003	A	B	C	R	Exits	Total
A	**242**	71	2	0	127	442
	(0.548)	(0.161)	(0.005)		(0.287)	
B	320	**681**	166	0	605	1772
	(0.181)	(0.384)	(0.084)		(0.341)	
C	68	624	**552**	29	1101	2374
	(0.029)	(0.263)	(0.233)	(0.012)	(0.464)	
R	0	147	472	**113**	1721	2453
		(0.060)	(0.192)	(0.046)	(0.702)	
Entrants	238	1099	1641	187		3165
	(0.075)	(0.347)	(0.518)	(0.059)		
Total	868	2622	2833	329	3554	10206
	(0.085)	(0.257)	(0.278)	(0.032)	(0.348)	

Transitions from one research quality category to another vary by initial QC. Just under 20 per cent of the 2003 R's moved upwards to become Cs in the same institution by 2012, and 6 per cent moved upward to B, though again this may overstate the actual rate of improvement (since the denominator excludes those who did not submit a portfolio, although this was less important in 2003). Upward movements within the same institution came mainly from those who were initially C (about 26 per cent to B and 3 per cent to A quality), and to a lesser extent B (18 per cent to A quality) researchers in 2003.

The aggregate discipline (and university) transitions discussed in section 5.1 and shown in Table 6 can be further aggregated into exits, entrants, and transformations to show how the total stock of PBRF portfolios changed between 2003 and 2012. This is done in Table 7, which shows the distribution of the total portfolios submitted in 2003 by quality category, and how exits and entrants, and transitions from one QC to another by research staff who submitted portfolios in both periods, influenced the change between 2003 and 2012.[16]

Table 7. Exits, entrants and quality transformations

Quality Category	2003 portfolios	Exits	Entrants	Transformations	2012 portfolios
A	442	127	238	315	868
	(0.06)	(0.04)	(0.07)		(0.13)
B	1772	605	1099	356	2622
	(0.25)	(0.17)	(0.35)		(0.39)
C	2374	1101	1641	-81	2833
	(0.34)	(0.31)	(0.52)		(0.43)
R	2453	1721	187	-590	329
	(0.35)	(0.48)	(0.06)		(0.05)
Totals	7041	3554	3165	0	6652

There are several striking features evident from Table 7. There was clearly considerable turnover of research staff.[17] Nearly half the exits were R-quality. The proportion of researchers exiting is consistently lower for higher-quality researchers. On the other hand, most of the Entrants were C and B-quality researchers (52 and 35 per cent respectively). Both characteristics can increase a university and discipline 2003 AQS, as explained above. Transformations in the quality of researchers who submitted portfolios in both periods show a clear pattern of an increase of 671 B and A-quality researchers offset predominantly by a large decline in R-quality researchers. Consequently, the distribution of QCs changed markedly. Comparison of the proportions of total portfolios in each QCs shows a large fall in the proportion of R-quality portfolios from 35 per cent of the total in 2003 to only 5 per cent in 2012, offset by increases in the proportions of all other QCs in 2012. As is evident from Tables 6 and 7, changes in AQSs (whether for research areas, discipline groups, or universities) can arise both from an improvement in the research quality of individuals who remain in the university system, and the replacement of lower with higher-performing researchers.

[16] Details about movements from one NZ university to another NZ university were not available. Hence, the exits and entrants for discipline groups are likely to be overstated somewhat.

[17] Summary measures of turnover are presented in Appendix B.

5.3. Transitions for individual disciplines

This subsection examines the extent of heterogeneity among discipline groups in recruitment and transitions. Table 7 provides information about the flows, from 2003 to 2012, between the different research quality categories, along with the entrants and exits, for each discipline: the actual flows are shown in the left-hand side of the table. It is possible to test whether the pattern of transitions and exits and entrants differ (statistically) significantly from those obtained by taking all universities combined, as in Table 6. The right-hand side of Table 9 shows, for each discipline, the difference between the actual flows over the period and the flows which would result from starting with their actual stocks in 2003 and applying the transition and exit rates for all disciplines, as shown in Table 6. Using these hypothetical flows as 'expected' values, a standard chi-square test (based on 16 degrees of freedom) can be carried out.

The resulting chi-squared values are reported in each case on the same line as the discipline name. The appropriate chi-square values, for type I errors of 0.05 and 0.10, and for 16 degrees of freedom, are 26.30 and 23.54 respectively. Two discipline group transitions do not differ significantly at the 5 per cent level of significance from those for all disciplines combined. These are Agriculture (chi-squared value of 23.63) and Management (24.61), although they are just significantly different at the 10 per cent level. The disciplines with the largest chi-square values are Education (75.40) and Medicine (72.54). The chi-square values for Engineering (40.65), Core Science (49.43), AFE (33.93), Humanities (36.42) and Law (32.50) are smaller but significant.

Table 9. Actual and hypothetical quality category transitions by discipline

	Actual transitions					Differences between actual and hypothetical transitions				
	A	B	C	R	Exits	A	B	C	R	Exits
Medicine	*Chi-squared = 72.54*									
A	**77**	15	0	0	14	19	-2	-1	0	-16
B	102	**178**	40	0	136	20	3	-3	0	-20
C	19	167	**134**	4	296	1	4	-10	-4	9
R	0	33	100	**18**	471	0	-4	-20	-11	35
Entrants	54	248	545	74		-15	-72	67	20	0

Engineering		*Chi-squared = 40.65*								
A	**35**	13	0	0	19	-2	2	0	0	0
B	29	**89**	28	0	83	-12	1	7	0	5
C	10	63	**79**	3	154	1	-18	7	-1	11
R	0	18	33	**9**	240	0	0	-25	-5	30
Entrants	25	174	243	13		-9	16	7	-14	0
Core Science		*Chi-squared = 49.43*								
A	**33**	10	0	0	16	1	1	-1	0	-1
B	39	**66**	19	0	51	7	-1	3	0	-9
C	2	39	**33**	3	109	-3	-10	-10	1	22
R	0	2	12	**2**	98	0	-5	-10	-3	18
Entrants	27	99	107	3		9	17	-15	-11	0
Management		*Chi-squared = 24.61*								
A	**6**	5	0	0	7	-4	2	0	0	2
B	6	**38**	10	0	33	-10	5	2	0	3
C	2	43	**48**	2	67	-3	1	10	0	-8
R	0	17	51	**9**	129	0	5	11	0	-16
Entrants	10	51	91	9		-2	-5	8	-1	0
AFE		*Chi-squared = 33.93*								
A	**7**	0	0	0	7	-1	-2	0	0	3
B	7	**34**	13	0	32	-9	1	5	0	3
C	2	18	**28**	4	53	-1	-10	4	3	4
R	0	7	34	**11**	101	0	-2	5	4	-7
Entrants	20	53	84	15		7	-7	-5	5	0
Humanities		*Chi-squared = 36.42*								
A	**39**	15	1	0	38	-12	0	1	0	11
B	70	**156**	40	0	165	-8	-10	0	0	18
C	15	150	**133**	8	221	0	11	10	2	-23
R	0	26	95	**16**	270	0	2	17	-3	-16
Entrants	47	245	275	23		3	40	-31	-12	0
Agriculture		*Chi-squared = 23.63*								
A	**22**	9	1	0	13	-3	2	1	0	0
B	45	**63**	12	0	59	13	-6	-5	0	-2
C	9	57	**49**	1	111	3	-3	-4	-2	6
R	0	5	11	**4**	77	0	-1	-8	0	9
Entrants	31	143	175	15		4	17	-14	-7	0
Education		*Chi-squared = 75.40*								
A	**15**	3	0	0	7	1	-1	0	0	0
B	11	**26**	3	0	26	-1	1	-3	0	3

C	5	62	**34**	0	67	0	18	-5	-2	-11
R	0	32	123	**38**	286	0	3	31	16	-50
Entrants	19	59	93	30		4	-11	-11	18	0
Law		*Chi-squared = 32.50*								
A	**8**	1	0	0	6	0	-2	0	0	2
B	11	**31**	1	0	20	0	7	-5	0	-2
C	4	25	**14**	4	23	2	7	-2	3	-10
R	0	7	13	**6**	49	0	3	-2	3	-4
Entrants	5	27	28	5		0	5	-6	1	0

Notes: This table shows the actual and hypothetical flows of quality categories 2003 to 2012 (from rows to columns) for each discipline for all New Zealand universities. The appropriate chi-square values, for type I errors of 0.05 and 0.10, and for 16 degrees of freedom, are 26.296 and 23.542 respectively.

Further inspection of Table 9 reveals that the discipline group transitions differ from each other in different ways.[18] Medicine and Education are the disciplines with the largest chi-squared values. Education experienced the largest total and proportionate decline in portfolios submitted and has the largest chi-squared value. This appears to be due primarily to the transitions involving R-quality researchers: a relatively small number of R's (and Cs) exiting, a large number entering, and a large number remaining as R-quality. These are all relative to the numbers expected if the transition proportions were the same as for all disciplines combined. The effect of these transitions was offset somewhat by a relatively large number of Rs and Cs transitioning to a higher quality.

Medicine experienced a small increase in portfolios submitted but the transition pattern also differs significantly from the pattern for all disciplines combined shown in Table 6. Medicine is the most concentrated discipline and the change in concentration primarily involves four universities (an increase in the shares for Auckland and Canterbury and declining shares for AUT and Otago) which may partly explain this different transition pattern. Medicine has a relatively low exit and low entry of A and B-quality researchers but relatively high exit and high entry of C and R-quality researchers. Transitions to higher quality research was relatively high amongst higher grade researchers but relatively low amongst lower grade researchers.

[18] In comparing performance relating to the exits of R-researchers, it must be kept in mind that this figure is distorted for reasons discussed above.

The next largest chi-squared value is for Core Science. This appears to be attributable to a relatively large exit and low entry of C and R-quality researchers and a relatively low exit and high entry of A and B-quality researchers, compared to the transition proportions for all disciplines combined. The transitions of C and R-quality researchers to a higher quality was relatively low. The fourth highest chi-squared value is for Engineering and the differences from the transitions for the combined disciplines pattern are similar to the differences displayed by Core Science, particularly the large exit of R and C-quality researchers, relatively low entry of R and high entry of B-quality researchers. Engineering was not as successful as Core Science in attracting A-quality entrants. Humanities also experienced relatively high exit and low entry of C and R-quality researchers and a very high entry of B-quality researchers.

Although Agriculture had the largest increase in portfolios submitted between 2003 and 2012, the transition pattern is similar to that for all disciplines combined and hence it has a low chi-squared value. However, it does display a relatively high exit and low entry of C and R-quality researchers and high entry of A and B-quality researchers. Management, AFE and Law also have relatively low chi-squared values and therefore have transitions relatively close to the pattern for the disciplines combined.

6. Contributions of changes in quality and discipline composition to changes in university research quality

Section 3 has shown that there were marked changes in the changes in the shares of portfolios submitted by discipline group and changes in the distribution of discipline portfolios across the eight New Zealand universities between the 2003 and the 2012 PBRF assessments. Section 4 has shown that there are large differences in discipline level AQSs and these differences arise from significant differences in the transition patterns across the nine disciplines, as shown in section 5. The question arises as to the extent to which the marked differences in the improvement in AQSs between 2003 and 2012 are attributable to changes in quality of researchers, for a given distribution of discipline portfolios, and from a change in the distribution of discipline portfolios, for a given quality. This section examines the relative contribution of these two factors to the AQS for each New Zealand university.

6.1 A decomposition of the contributions to a university AQS

Suppose there are systematic differences between (as well as within) disciplines in their average research quality, as measured by the PBRF process. This means that a university can improve its AQS, and hence funding, by changing the composition, by discipline, of its research staff, or by changing the quality of staff in each discipline (either by turnover or introducing methods of enhancing research performance of existing staff). In practice, changes are likely to arise from a combination of both methods. The question arises: to what extent can the change in AQS for a university be attributed to the changing quality of its researchers or the changing composition (by discipline) of its staff?

The decomposition method follows the general approach suggested by Shorrocks (2013).[19] Consider a single university. There are *D* disciplines. The proportion of staff in discipline, *j*, at time, *t*, is denoted by $p_{j,t}$. Let the AQS in discipline, *j*, at time, *t*, be $q_{j,t}$. This is measured by the PBRF process as an average of individual numerical values assigned to QC as shown by expression (2). Changes in $q_{j,t}$ reflect turnover and changes experienced by staff who remain in the same organisation. However, the following is concerned only with changes in the 'research quality' of the discipline, $q_{j,t}$: the way in which this is achieved is examined using transition matrices.

The AQS of the university, Q_t in period, *t*, is equal to:

$$Q_t = \sum_{j=1}^{D} p_{j,t} q_{j,t} \tag{4}$$

In the following, it is convenient to write $Q_t = Q(q_t, p_t)$, where q_t and p_t are the vectors, $[q_{1,t}, ..., q_{D,t}]$ and $[q_{1,t}, ..., q_{D,t}]$. The change in quality from period *0* to period *1*, denoted $\Delta Q = Q_1 - Q_0$, is given by:

$$\Delta Q = Q(q_1, p_1) - Q(q_0, p_0) \tag{5}$$

[19] Further applications of the method can be found in, for example, Bargain (2012), Creedy and Hérault (2015) and Ball and Creedy (2016).

This can be expressed as:

$$\Delta Q_A = \left[Q(q_1, p_1) - Q(q_0, p_1) \right] + \left[Q(q_0, p_1) - Q(q_0, p_0) \right] \quad (6)$$

The first term reflects the change in AQS attributed to the changing quality of its staff, given the discipline composition in period *1*. The second term reflects the change in AQS attributed to the changing discipline composition of the university, given the quality of staff in period 0.

However, the change in AQS may also be decomposed using an alternative perspective, as follows:

$$\Delta Q_B = \left[Q(q_1, p_0) - Q(q_0, p_0) \right] + \left[Q(q_1, p_1) - Q(q_1, p_0) \right] \quad (7)$$

The first term reflects the change in AQS attributed to the changing quality of its staff, given the discipline composition in period 0. The second term reflects the change in AQS attributed to the changing discipline composition of the university, given the quality of staff in period *1*. There is no special presumption in favour of using (6) or (7). Hence, one approach is to take the arithmetic mean of the components.

Estimates of the decompositions are shown in Table 10. The estimated contributions are the mean of the relative contributions derived using expressions (7) and (8). The top panel of Table 8 shows the contributions of changes in quality and changes in discipline composition to each university's AQS derived for all portfolios submitted, including R-quality portfolios. The top panel shows that for all universities, most of the improvement in AQS between 2003 and 2012 arose from improvements in the quality of researchers for which portfolios were submitted.

The mean of the proportional contributions from quality improvement to the improvement in university AQSs is 98 per cent. The contributions range from 99 per cent for AUT and Canterbury to 94 per cent for Waikato. The contributions arising from composition change is very small. If the focus were simply to increase the AQS of a university, it may be expected that more emphasis would have been placed on changing the discipline composition by increasing recruitment into higher-quality disciplines such as Core Science and Agriculture, and contracting staff levels in lower-quality disciplines. Yet Education is the only discipline which has lost relatively large numbers, and there were, in the early years of the PBRF, substantial institutional changes involving the merger of colleges of education with universities.

Table 10. Quality and composition change contributions to improvements to each university's improvement in research quality, 2003 to 2012

University	AQS 2003	Increase in AQS	Proportion attributed to: Quality	Composition
AUT	0.73	2.28	0.99	0.014
Lincoln	2.49	1.07	0.96	0.04
Massey	2.06	2.05	0.98	0.02
Auckland	3.56	1.29	0.96	0.04
Canterbury	3.54	1.03	0.99	0.01
Otago	3.08	1.69	0.98	0.02
Waikato	2.93	1.36	0.94	0.06
VUW	3.06	2.25	0.97	0.03
All	2.81	1.71	0.98	0.02

As explained in section 4.1, on the eve of the 2012 PBRF round, the TEC determined that universities could choose to not submit portfolios for researchers deemed to be of R-quality. As shown in Buckle and Creedy (2018a and 2018b) when the denominator of expression (2) for the derivation of AQS includes all non-administration staff (and therefore research staff for whom no portfolio was submitted), the AQS is substantially reduced. Some indication of the possible effect on the decomposition of this change can be obtained by setting all non-portfolios equal to R. However, this cannot be done precisely, because the number of full-time equivalent persons is available rather than the total: the AQSs obtained above are all calculated using the number of portfolios rather than the full-time equivalents, used by the TEC.[20] Using the full-time equivalent numbers as an approximation for the additional number of R staff, the resulting decompositions are shown in Table 11.

As expected, this adjustment reduces the contribution of quality change to the improvement in the AQSs calculated for all non-administration staff. The differences from values in Table 10 are shown in the final column of Table 11. The overall contribution of quality improvement falls from 98 per cent to 84 per cent and the contribution of the change in

[20] Buckle and Creedy (2018a, p. 4, Table 1) report information about the number of full-time equivalent portfolios, NP, and non-administrative staff, NT. In the calculations here, the number of staff for whom a portfolio was not submitted in 2012 is approximated by NT-NP.

discipline composition increases from 2 per cent to 14 per cent. Moreover, there is much more variation among universities in the relative contributions of quality and composition change when the AQS is derived for all non-administration staff.

Table 11. Quality and composition change contributions to improvements to each university's improvement in research quality: 2003 to 2012 (Setting all non-portfolios equal to R)

University	AQS 2003	Increase in AQS	Proportion attributed to: Quality	Composition	Reduction from including all non-portfolios as Rs
AUT	0.32	1.09	0.95	0.05	0.04
Lincoln	1.78	0.84	0.88	0.12	0.08
Massey	1.38	1.60	0.87	0.13	0.10
Auckland	2.64	1.10	0.85	0.15	0.10
Canterbury	2.72	1.56	0.56	0.44	0.43
Otago	2.31	1.34	0.94	0.06	0.04
Waikato	1.88	1.39	0.65	0.35	0.30
VUW	2.42	1.83	0.96	0.04	0.01
All	1.95	1.44	0.84	0.16	0.14

7. Conclusions

This paper has evaluated the change in the research quality of research discipline groups in New Zealand universities between 2003 and 2012 following the introduction of the New Zealand PBRF scheme. Earlier research found that the introduction of the PBRF scheme promoted a significant improvement in average research quality of staff at New Zealand's eight universities. There were significant differences among universities in the extent of this improvement. There are reasons to expect there could also be significant differences in the extent of improvement of research capability at the discipline group level. Moreover, if universities differ in their discipline composition of research staff then differences among disciplines in the extent of their improvement in research quality of staff may be a factor contributing to differences among universities.

There have been substantial changes in the discipline composition of New Zealand universities during the PBRF period. For example, Agriculture portfolios increased by 19 per cent and its share of all discipline portfolios increased by over 2 per cent; Medicine portfolios increased slightly and its share increased by about 1.6 per cent; Education portfolios fell by 25 per cent and its share fell by over 2 per cent; all other disciplines experienced a fall in portfolios submitted and experienced small changes in their share of total portfolios. These shifts in discipline portfolio shares resulted in relatively small changes in the discipline composition of each university and small changes in the concentration across universities in the distribution of disciplines level research portfolios.

The AQSs for all disciplines increased substantially over the period, and the difference between the highest and lowest discipline AQS has fallen substantially. The quality category transformations differed significantly across most of the disciplines. Some disciplines were able to improve the average quality score by achieving a higher proportion of recruitment and retention of higher quality researchers than others. This was a feature of Core Science in particular. Some disciplines were more successful at transforming existing staff into higher quality, including previous B and C-quality staff to higher quality researchers. This was a feature of Medicine. Others achieved relatively high rates of upward transformation for previously lower quality researchers. This was a feature of Management, Humanities, and Education.

The total changes in university AQSs were decomposed into those resulting from improvements in average quality of staff and changes in discipline composition. For the total university system, the contribution from the change in discipline composition to the growth in AQS (based on submitted portfolios) was only 2 per cent. Hence, quality improvement of staff rather than discipline composition changes is the dominant source of improvements to university level research quality. When adjustments are made to the contributions from staff research quality improvement by including all non-administration staff and assigning an R-quality to those for whom no portfolio was submitted, the extent of quality improvement is reduced. The average reduction for all university AQSs arising from this adjustment is about 14 per cent. However, the effect varies substantially among universities. For VUW the effect is only 1 per cent because of the high proportion of all non-administration staff for whom research portfolios were submitted. In contrast, the impact for Canterbury is 43 per cent and for Waikato it is 30 per cent.

The objective of the PBRF was to increase the research quality of staff in the tertiary education sector in New Zealand. However, it is not clear whether the architects of the scheme intended this to result in changes in the allocation of resources among disciplines, and among universities. This paper has shown that there have been differences in quality transformation processes and in the rates of improvement of research quality among disciplines, but these have not generally led to greater discipline concentrations either within or among universities. Furthermore, changes in the discipline compositions of universities contributed only a small proportion of the improvement in the average research quality of staff. The predominant source of improvement in research quality has come from improving the research quality of staff in all disciplines.

Appendix A: Composition of Discipline Groups

Appendix Table A1. Discipline groups and PBRF portfolios, 2003 and 2012

Discipline	2003: Number of portfolios	% share	2012: Number of Portfolios	% share	Change in Number of portfolios	Change in % share
Medicine	**1804**	**25.62**	**1808**	**27.18**	**4**	**1.56**
Biomedical	175	2.49	276	4.15	101	1.66
Clinical Medicine	307	4.36	260	3.91	-47	-0.45
Dentistry	55	0.78	39	0.59	-16	-0.19
Molecular, Cellular & Whole Organism Biology	388	5.51	391	5.88	3	0.37
Nursing	130	1.85	70	1.05	-60	-0.8
Other Health Studies (incl.Rehab. Therapies)	245	3.48	171	2.57	-74	-0.91
Pharmacy	0	0	27	0.41	27	0.41
Psychology	229	3.25	242	3.64	13	0.39
Public Health	211	2.99	262	3.93	51	0.94
Sport & Exercise Science	64	0.91	70	1.05	6	0.14
Engineering	**905**	**12.85**	**864**	**12.99**	**-41**	**0.14**
Architecture, Design, Planning, Surveying	128	1.82	107	1.61	-21	-0.21
Computer Science, Info Tech, Info Sciences	352	4.99	274	4.12	-78	-0.87
Design	78	1.11	62	0.93	-16	-0.18
Engineering & Technology	347	4.93	421	6.33	74	1.4
Core Science	**534**	**7.58**	**496**	**7.45**	**-38**	**-0.13**
Chemistry	194	2.76	181	2.72	-13	-0.04
Physics	110	1.55	114	1.71	4	0.16
Pure & Applied Mathematics	144	2.05	117	1.76	-27	-0.29
Statistics	86	1.22	84	1.26	-2	0.04
Management	**473**	**6.72**	**398**	**5.98**	**-75**	**-0.74**
Management, HR, IR & Other Businesses	307	4.36	247	3.71	-60	-0.65
Marketing & Tourism	166	2.36	151	2.27	-15	-0.09
Accounting, Finance and Economics	**358**	**5.08**	**337**	**5.07**	**-21**	**-0.01**
Accounting & Finance	200	2.84	183	2.75	-17	-0.09
Economics	158	2.24	154	2.32	-4	0.08
Humanities	**1458**	**20.72**	**1354**	**20.35**	**-104**	**-0.37**
Anthropology & Archaeology	59	0.84	76	1.14	17	0.3
Communications, Journalism & Media Studies	87	1.24	86	1.29	-1	0.05

English Language & Literature	111	1.58	86	1.29	-25	-0.29
Foreign Languages & Linguistics	190	2.69	161	2.42	-29	-0.27
History, H of Art, Classics & Curatorial Ss	189	2.68	161	2.42	-28	-0.26
Human Geography	61	0.87	68	1.02	7	0.15
Māori Knowledge & Development	104	1.48	98	1.47	-6	-0.01
Music, Literary Arts & Other Arts	114	1.62	116	1.75	2	0.13
Philosophy	68	0.97	68	1.02	0	0.05
Political Science, Int. Relations & Public Policy	94	1.34	107	1.61	13	0.27
Religious Studies & Theology	26	0.37	27	0.41	1	0.04
Sociology, Social P, Social W, Crim & Gender Ss	222	3.15	180	2.71	-42	-0.44
Theatre & Dance, Film, Television & Multimedia	47	0.67	52	0.78	5	0.11
Visual Arts & Crafts	86	1.22	68	1.02	-18	-0.2
Agriculture	**548**	**7.78**	**652**	**9.81**	**104**	**2.03**
Agriculture & Other Applied Biological Sciences	158	2.24	159	2.4	1	0.16
Earth Sciences	146	2.08	167	2.51	21	0.43
Ecology, Evolution & Behaviour	177	2.51	269	4.04	92	1.53
Veterinary Studies & Large Animal Science	67	0.95	57	0.86	-10	-0.09
Education	**738**	**10.48**	**553**	**8.31**	**-185**	**-2.17**
Law	**223**	**3.17**	**190**	**2.86**	**-33**	**-0.31**
TOTAL	**7041**	**100**	**6652**	**100**	**-389**	

Appendix B: Turnover rates in discipline groups

This appendix examines the rate of turnover among disciplines. The information about changes in the numbers of staff in different discipline groups and universities cannot alone indicated the extent to which turnover exists.

The rate of turnover, *T*, from period 1 to 2 can be expressed as the number of exits, *X*, as a proportion of the average number of people over the period. If N_1 and N_2 are the numbers in each period, then:

$$T = \frac{2X}{\left(N_1 + N_2\right)} \qquad \text{(B1)}$$

Turnover rates for all universities combined, by discipline group, are shown in Table B1. This shows considerable turnover, of about 50 per cent, of university staff over the period. There is little variation among disciplines, although the exception is the group with the lowest turnover rate, Core Science, with a rate of 19 per cent. Turnover clearly contributed the improvement in AQSs. The following section explores this in more depth by looking in detail at the entrants, exits and quality changes of those who remained in the same university over the period.

Table B1. The rate of research staff turnover by discipline, 2003 to 2012

	N (2003)	*N* (2012)	Turnover (%)	Change in *N* (%)
Medicine	1804	1808	51	0
Engineering	905	864	56	-5
Core Science	534	496	19	-7
Management	473	398	54	-16
AFE	358	337	56	-6
Humanities	1458	1354	49	-7
Agriculture	548	652	43	19
Education	738	553	60	-25
Law	223	190	47	-15
ALL	7041	6652	52	-6

References

Ball, C. & Creedy, J. (2016). Inequality in New Zealand 1983/84 to 2012/13. *New Zealand Economic Papers*, 50:3, 323-342.

Bargain, O. (2012). Decomposition analysis of distributive policies using behavioural simulations. *International Tax and Public Finance*, 19, 708-731.

Boyle, G. (2008). Pay Peanuts and Get Monkeys? Evidence from Academia. *The B.E. Journal of Economic Analysis & Policy*, 8:1, (Contributions), Article 21.

Buckle, R.A. & Creedy, J. (2017). The evolution of research quality in New Zealand Universities as measured by the Performance-Based Research Fund process, Victoria University of Wellington Working Papers in Public Finance (No. WP11/2017). Wellington.

Buckle, R. A. & John Creedy, J. (2018a). The evolution of research quality in New Zealand universities as measured by the performance-based research fund process. *New Zealand Economic Papers*, DOI: 10.1080/00779954.2018.1429486

Buckle, R. A. & Creedy, J. (2018b). The Impact on Research Quality of Performance-Based Funding: The Case of New Zealand's PBRF Scheme. *Agenda*, 25:1, 25-48.

Buckle, R. A. & John Creedy, J. (2018c). An evaluation of metrics used by the Performance-based Research Fund process in New Zealand. *New Zealand Economic Papers*, DOI:10.1080/00779954.2018.1480054

Creedy, J. (2017). Alternative Distributions for Inequality Comparisons. *The Australian Economic Review*, 50:4, 484–497.

Creedy, J., & Hérault, N. (2015). Decomposing inequality and social welfare changes: The use of alternative welfare metrics. *Fiscal Studies*, 36, 157-180.

de Boer H., Jongbloed B., Benneworth P., Cremonini L., Kolster R., Kottmann A., Lemmens-Krug K., & Vossenstey H. (2015). Performance-based funding and performance agreements in fourteen higher education systems. *Report for the Ministry of Education, Culture and Science, Centre for Higher Education Policy Studies*, C15HdB014, *Universiteit Twente, Enschede*. Retrieved from www.utwente.nl/cheps.

Ehrenberg, R., Kasper, H., & Rees, D. (1991). Faculty turnover at American colleges and universities: Analyses of AAUP data. *Economics of Education Review*, 10(2), 99–110.

Gemmell, N., Nolan, P., & Scobie, G. (2017). Estimating quality-adjusted productivity in tertiary education: Methods and evidence for New Zealand. In Victoria University of Wellington Working Papers in Public Finance (No. WP17/2017). Wellington.

Jung, J. (2012). Faculty Research Productivity in Hong Kong across Academic Discipline. *Studies in Higher Education,* 2:4, 1-13.

Mahoney, P. (2004) Performance-based research funding. Wellington: Parliamentary Library Background Note, 2004/07.

Ministry of Education (2012). *A history and overview of the PBRF*. Wellington: Ministry of Education. Retrieved from https://education.govt.nz/assets/Documents/Further-education/Policies-and-strategies/Performance-based-research-fund/PBRFHistoryAndOverview.pdf

Ministry of Education (2013). *An international comparison of performance-based research funding schemes*. Retrieved from http://www.minedu.govt.nz/NZEducation/EducationPolicies/TertiaryEducation/PolicyandStrategy/~/media/MinEdu/Files/EducationSectors/TertairyEducation/PBRF/InternationalComparison.pdf

New Zealand Tertiary Education Commission (2002). *Investing in excellence: The report of the performance-based research fund working group*. Wellington: Ministry of Education and Transition Tertiary Education Commission.

New Zealand tertiary Education Commission (2013). Performance-Based Research Fund: Evaluating Research Excellence – the 2012 Assessment. Wellington: tertiary Education Commission (Available from Evaluating Research Excellence - the 2012 Assessment).

OECD (2010). Performance-based funding for public research in tertiary education institutions: Workshop proceedings. OECD. http://www.oecd-ilibrary.org/

Sabharwal, M. (2013). Comparing Research Productivity Across Disciplines and Career Stages. *Journal of Comparative Policy Analysis: Research and Practice*, 15:2, 141-163.

Shin, J. C., & Cummings, W. K. (2010). Multilevel analysis of academic publishing across disciplines: research preference, collaboration, and time on research. *Scientometrics*, 85:2, 581-594.

Shorrocks, A. F. (2013), Decomposition procedures for distributional analysis: A unified framework based on the Shapley value. *Journal of Economic Inequality*, 11, 100–126.

Smart, W., & Engler, R. (2013). In pursuit of excellence: Analysing the results of New Zealand's PBRF quality evaluations. Wellington: Ministry of Education. Retrieved from www.educationcounts.govt.nz.

Wanner, R. A., Lewis, L. S., & Gregorio, D. I. (1981). Research Productivity in Academia: A Comparative Study of the Sciences, Social Sciences and Humanities. *Sociology of Education*, 54:4, 238-253. http://dx.doi.org/10.2307/2112566.

Wilsdon J., Allen L., Belfiore E., Campbell P., Curry S., Hill S., . . . Johnson B. (2015). *The metric tide: Report of the independent review of the role of metrics in research assessment and management*. doi:10.13140/RG.2.1.4929.1363.

Woerlert, P., & McKenzie, L. (2018). Follow the money? How Australian universities replicate national performance-based funding mechanisms, *Research Evaluation*, 27:1, 184–195. https://doi.org/10.1093/reseval/rvy018

Xu, Y. J. (2008). Faculty Turnover: Discipline-Specific Attention is Warranted, *Research in Higher Education*, 49:1, 40–61

About the Authors

Robert Buckle is Emeritus Professor at Victoria Business School, Victoria University of Wellington, New Zealand.
Email: bob.buckle@vuw.ac.nz

John Creedy is Professor of Public Finance at Victoria Business School, Victoria University of Wellington, New Zealand.
Email: john.creedy@vuw.ac.nz