



# Sources of convergence and divergence in university research quality: evidence from the performance-based research funding system in New Zealand

Robert A. Buckle<sup>1</sup> · John Creedy<sup>2</sup> · Norman Gemmell<sup>2</sup>

Received: 20 June 2021 / Accepted: 5 April 2022 / Published online: 16 May 2022  
© The Author(s) 2022

## Abstract

The introduction of performance-based research funding systems (PBRFS) in many countries has generated new information on their impacts. Recent research has considered whether such systems generate convergence or divergence of research quality across universities and academic disciplines. However, little attention has been given to the processes determining research quality changes. This paper utilises anonymised longitudinal researcher data over 15 years of the New Zealand PBRFS to evaluate whether research quality changes are characterised by convergence or divergence, and the processes determining those dynamics. A unique feature is the use of longitudinal data to decompose changes in researcher quality into contributions arising from the entry, exit and quality transformations of retained researchers, and their impacts on convergence or divergence of research quality across universities and disciplines. The paper also identifies how researcher dynamics vary systematically between universities and disciplines, providing new insights into the effects of these systems.

**Keywords** Education policy · Performance-based research funding systems · Research quality · Convergence · Universities

**JEL Classifications** I2 · I23 · I28 · L38

## Introduction

Performance-based research funding systems (PBRFS) are designed to improve research output, quality and impact, and to strengthen the accountability of universities for the use of public funds. They involve a range of incentives, created in particular by the metrics used to measure research quality, which are designed to encourage institutional and

---

✉ Robert A. Buckle  
bob.buckle@vuw.ac.nz

<sup>1</sup> School of Economics and Finance, Wellington School of Business and Government, Victoria University of Wellington, P. O. Box 600, Wellington 6140, New Zealand

<sup>2</sup> School of Accounting and Commercial Law, Wellington School of Business and Government, Victoria University of Wellington, P. O. Box 600, Wellington 6140, New Zealand

individual changes, although unintended consequences can also result.<sup>1</sup> In evaluating performance outcomes of a PBRFS, it is important to know if the design features led to the initially stronger universities capturing an increasing share of research funds, or to a catch-up process whereby initially weaker universities raised their standards at a faster rate. This relates to what is referred to as  $\beta$ -convergence, defined as a systematic tendency for lower-quality universities to experience relatively higher growth rates compared with higher-quality universities. A further type of convergence, called  $\sigma$ -convergence, refers to changes in the overall dispersion of university research quality.<sup>2</sup>

Despite the large literature on PBRFSs, very few systematic attempts have been made to assess their convergence properties. The first study to examine convergence formally in the context of a PBRFS is Buckle et al. (2020), who use information about the first two full rounds, in 2003 and 2012, of the New Zealand (NZ) PBRFS.<sup>3</sup> They found significant evidence for both  $\beta$ - and  $\sigma$ -convergence. Subsequently, Checchi et al. (2020) adopted the same technique to examine Italian universities, following the introduction of the Italian PBRFS (the Evaluation of Research Quality, VQR). They found convergence of research quality of Italian universities between the VQR rounds of 2004 and 2010, and between 2011 and 2014. Following the same basic method, Abramo and D'Angelo (2022) also provide evidence of convergence of the research quality of Italian universities and also find that the process of convergence is affected predominantly by a stronger rate of improvement of the initially lower research-quality universities.

The present paper extends this earlier work in several ways. The first, and most important extension, is that the precise sources of convergence are examined in detail. This is achieved by investigating the nature of the transitions of researchers involved in achieving quality improvements. Essentially, quality changes arise from staff turnover and quality transformations of those remaining within the same university. This exploration of the sources of convergence involves the development of a new decomposition method that identifies the separate contributions of exits, entrants and quality transformation.

Second, unlike the earlier analysis of Buckle et al. (2020), the paper is able to make use of the third full round carried out in 2018. This extension allows investigation of whether changes in research quality during the second period (2012–2018) were systematically related to the previous responses from 2003 to 2012. Third, more detailed information is available about researchers in each round, including gender and full-time equivalent (FTE) status, which allows for further control variables to be included in the analysis.

Hence, this paper presents techniques that identify the contributions of turnover and quality transformation of researchers to the growth in research quality following the introduction of a PBRFS. Turnover arises from exits and entrants of researchers, defined as follows. Entrants to a university consist of researchers who either move from another NZ university, or enter the NZ university system from outside it. Exits are researchers who leave a university, either transferring to another NZ university or moving outside the NZ university

<sup>1</sup> On the development of PBRFSs see, for example, de Boer et al. (2015), Hicks (2012), Kolarz et al. (2019), OECD (2010) and Wilsdon et al. (2015). Examples of assessments and critical evaluations of these schemes include Adams and Gurney (2010), Broadbent (2010), Buckle and Creedy (2019a, 2020), Buckle et al. (2021), Checchi et al. (2019), Hare (2003), Martin (2011), Payne and Roberts (2010), Woerlert and McKenzie (2018).

<sup>2</sup> On the use of these terms in the cross-country growth literature, see Quah (1993). These types of convergence are examined in detail in Buckle et al. (2020).

<sup>3</sup> There was an incomplete round in 2006.

system. When looking at NZ universities as a whole, the transfers among universities net out, and entrants and exits include only those moving to or from the NZ university system.

Each university is incentivised to manage the process of improvement in research quality, subject to constraints.<sup>4</sup> The nature of the incentives in the PBRFS design is likely to influence relative responses and therefore whether they generate convergence or divergence.<sup>5</sup> The benefits of the turnover of researchers of a particular quality vary by the type of university. In turn, the characteristics of a university's turnover, and quality transformation of incumbents, depend on its initial average research quality. These determine the process of convergence of research quality among all universities within the system.

The New Zealand scheme was designed to unbundle the research component of Government funding of New Zealand tertiary education organisations, and allocate the research component based on research performance rather than the number of students; see New Zealand Tertiary Education Commission (2019, p. 11). Three measures are used to allocate Government funding to support research at universities and other tertiary education organisations. The largest component, the focus of this paper, is Quality Evaluation.<sup>6</sup> The system substantially changed the incentives facing individuals, departments and universities. However, when the PBRFS was introduced, there was no explicit statement of whether an aim was to generate more concentration of higher-quality research.

'Growth rates of research quality and convergence' section begins by describing the research quality metric used by the NZ PBRFS. It then provides empirical results regarding  $\beta$ -convergence and  $\sigma$ -convergence, allowing for the influence of past changes. 'Contribution of exits, entrants and quality transformations to AQS changes' section reports the contributions of exits, entrants and quality transformation of researchers to research quality changes for universities and academic disciplines (subject areas), and ' $\beta$ -Convergence properties of contributions to AQS growth' section examines these component effects on convergence. 'Conclusions' section concludes.

## Growth rates of research quality and convergence

### The New Zealand measure of research quality

The assessment of research quality in the NZ PBRFS is based on the performance of all eligible individual researchers within each university.<sup>7</sup> The assessment is based on peer review of a range of research outputs over the previous 6 years. Each researcher is assigned to one of four quality categories (QCs), A, B, C and R, where the highest category is A, and R indicates an absence of significant research outputs. These QCs are used to compute

---

<sup>4</sup> On management responses to the introduction of a PBRFS, see Adams (2008) and Woelert and McKenzie (2018).

<sup>5</sup> In the UK context, Barker (2007, p. 6) suggests that the funding weights disproportionately rewarded universities which achieved high quality scores, and therefore led to divergence. However, convergence properties for the UK scheme cannot be tested formally, as in the present paper, because the metrics involve only a small number of qualitative categories, and not all academics needed to be included in evaluations.

<sup>6</sup> The other components are Research Degree Completions and External Research Income.

<sup>7</sup> Those eligible include all research and teaching staff who are employed on the PBRF census date under an employment agreement with a duration of at least 1 year, and are employed throughout the contract on at least a 0.20 FTE basis. Although precise information is not available, this appears to account for around 90% of total non-administration staff.

a quantitative performance score, referred to as an Average Quality Score (*AQS*) for each subject area and university, defined as follows.<sup>8</sup> Each individual,  $h$ , is given a cardinal score,  $G_h$ , depending on the QC: 10 for A; 6 for B; 2 for C; and 0 for R. The average quality score, *AQS*, is the employment-weighted arithmetic mean score.<sup>9</sup>

Define the full-time equivalent (FTE) employment weight of person,  $h$ , as  $e_h \leq 1$ , and let  $n$  denote the number of employees. The *AQS* for a university or discipline group is:

$$AQS = \frac{\sum_{h=1}^n e_h G_h}{\sum_{h=1}^n e_h} \quad (1)$$

The data used here include the QC assigned to every researcher who participated in any of the assessment rounds, in 2003, 2012 and 2018.<sup>10</sup> It also includes an anonymous identifier, age, gender, research subject area, university of employment, and FTE status, including whether the researcher exited or entered the entire NZ university system between rounds or transferred to or from another NZ university.

## $\beta$ -Convergence

A standard form of logarithmic regression specification between *AQS* changes and initial *AQS* to test for  $\beta$ -convergence from time  $t-1$  to  $t$  is:

$$\log AQS_{ijt} - \log AQS_{ijt-1} = c + \beta \log AQS_{ijt-1} + \varepsilon_{ijt} \quad (2)$$

The left-hand side measures the annualised proportional *AQS* growth rate for university  $j$ , and discipline  $i$ .<sup>11</sup> This enables the systematic component of change to be separated from other influences, including university-specific fixed effects and random shocks. Where three time periods are available, (2) can be extended to<sup>12</sup>:

$$\log AQS_{ijt} - \log AQS_{ijt-1} = \gamma + \beta_1 \log AQS_{ijt-1} + \beta_2 \log AQS_{ijt-2} + \varepsilon_{ijt} \quad (3)$$

In what follows,  $t$ ,  $t-1$  and  $t-2$  refer to 2018, 2012 and 2003. The full convergence effect,  $\beta$ , is measured by  $\beta = (\beta_1 + \beta_2)$ . Convergence tests can also be conditioned on a

<sup>8</sup> The assessment and scoring methods are described in more detail and critically evaluated in Buckle and Creedy (2019b).

<sup>9</sup> Two new categories C(NE) and R(NE) were introduced by the TEC for the 2012 and 2018 assessment rounds to signify if a researcher met the 'new and emerging' criteria. The score assigned by the TEC to R(NE) was 0 in 2012 and 2018, the same as in 2003. The score assigned by the TEC to C(NE) was 2 in 2012, the same as for C, and 4 in 2018 (New Zealand Tertiary Education Commission, 2019, p. 13). For the purposes of this paper the score assigned to C(NE) in 2018 is set equal to 2, to ensure consistency over time.

<sup>10</sup> The anonymised data used in this study are not publicly available and were provided by the NZ Tertiary Education Commission (TEC) following a confidentiality agreement.

<sup>11</sup> This form was used by Buckle et al. (2020) to examine changes from 2003 to 2012, but they did not need to annualise the growth rates because only one period was available. Academic subjects are collected into nine disciplines: medicine, engineering, core science, management, accounting finance & economics, humanities, agriculture, law and education; see Buckle and Creedy (2020) for details of the groups.

<sup>12</sup> This specification is consistent with one having the lagged (logarithm) of the 2012 *AQS*, and the growth from 2003 to 2012 (that is, the difference in the logarithms of the *AQS*s), on the right-hand side. The interpretation is that growth from 2012 to 2018 depends on the *AQS* level in 2012 and the previous annualised growth rate experienced from 2003 to 2012.

vector of additional variables, including the number of staff FTEs, the median age and the gender ratio of researchers by university and discipline.

Testing for possible differences across universities and disciplines in the growth (and convergence or divergence) of their *AQSs*, involves testing for the inclusion or exclusion in regressions of shift dummy variables, where  $D_i = 1$  for university  $i$ , and is zero otherwise. Slope dummies,  $S_i$ , equal to  $D_i \log AQS_{it-1}$  are also added. Similar dummies are defined for discipline groups. The parameters  $\gamma$  and  $\beta = (\beta_1 + \beta_2)$  in (3), along with their university-specific and discipline-specific equivalents, therefore respectively capture autonomous *AQS* growth and the rate of convergence ( $\beta < 0$ ) or divergence ( $\beta > 0$ ) of *AQSs* for each university and discipline. The parameter,  $\beta$ , may be interpreted as the ‘full’ rate of convergence over the whole period, including any university or discipline fixed effects.

The approach begins with all dummy variables included:  $2 \times 8$  university shift and slope dummies and  $2 \times 9$  discipline equivalents. These are progressively eliminated using the ‘general-to-specific’ (*Gets*) approach of Campos et al. (2005), Castle et al. (2011) and Hendry and Doornik (2014), whereby the variable with the lowest  $t$ -ratio is omitted first.<sup>13</sup> The regression is then re-run in a sequential process, until the most parsimonious specification of the data-generating process is obtained. This effectively treats the null hypothesis as  $\gamma_i \neq \gamma_j \neq \gamma$  and  $\beta_i \neq \beta_j \neq \beta$ , against the alternative of common values of  $\gamma$  and  $\beta$  across universities and disciplines.

Following the *Gets* approach, the convergence variables,  $\log AQS_{2012}$  and  $\log AQS_{2003}$ , always passed relevant  $t$ -tests, while dummy variables were progressively eliminated based on parameter  $t$ -tests. Table 1 shows  $t$ -ratios associated with the shift ( $D$ ) and slope ( $S$ ) dummy variables for each university and discipline added to regressions on (3), and the order in which they are omitted. Left-hand columns report results where the initial regression including all university and discipline dummies. As a check, right-hand columns report similar results but for initial regressions that include either all university dummies or all discipline dummies.

In both cases the table shows the  $t$ -ratio associated with a given dummy in the regression immediately prior to it being omitted. For example, in the left-hand columns of Table 1, the first variable eliminated was the shift dummy for Massey University (MU) with a  $t$ -value of 0.00. The regression was re-run omitting this variable. This led to the slope dummy for Medicine being identified as the lowest  $t$ -ratio and eliminated ( $t = 0.03$ ). This process was repeated until only variables with  $t$ -ratios  $> |3|$  were retained.<sup>14</sup> The only dummy variable meriting retention in the regression after this process is the slope dummy variable for the Accounting, Finance and Economics (AFE) discipline group, with a  $t$ -ratio of  $-3.48$ .<sup>15</sup>

<sup>13</sup> Campos et al. (2005, p. 2), for example, summarise the *Gets* approach as follows: ‘1. Ascertain that the general statistical model is congruent. 2. Eliminate a variable (or variables) that satisfies the selection (i.e., simplification) criteria. 3. Check that the simplified model remains congruent. 4. Continue steps 2 and 3 until none of the remaining variables can be eliminated’.

<sup>14</sup> A critical  $t$ -ratio = 3 was chosen following Castle et al. (2011) and Castle and Hendry (2014) who argue that substantial pre-testing and variable selection from many possible models increases the risk of retaining irrelevant variables. For example, with a critical significance level for  $t$ -tests of  $c_\alpha = 0.05$ , there is a 1 in 20 chance of an irrelevant variable being retained (with a threshold  $t > 2$ ) in the model on average. However, for  $c_\alpha = 0.01$  ( $t_\alpha \approx 2.6$ ) this becomes 1 in 100, and  $c_\alpha = 0.001$  ( $t_\alpha \approx 3.35$ ) implies 1 in 1000. Hence  $t_\alpha$  between 2.6 and 3.35 substantially reduce the risk of a false positive. Castle and Hendry (2014) recommend setting  $\alpha = \min(1/N, 1/T, 1\%)$ .

<sup>15</sup> Table 3 also indicates that if a more conventional  $t > 2$  was used to retain dummy variables, a shift dummy for Medicine ( $t = 2.41$ ), a slope dummy for core Science ( $t = 2.49$ ) and a slope dummy for Education ( $t = 2.57$ ) would be retained.

The same final outcome is obtained using the approach in the right-hand columns which, for example, tests the null hypothesis,  $\beta_i \neq \beta$ , while maintaining  $\beta_j = \beta$ , and vice versa.

Details of the final regression are reported in Table 2.<sup>16</sup> These reveal a strong common initial rate of  $\beta$ -convergence across all universities and disciplines (except AFE) of  $\beta_1 = -0.1167$ .<sup>17</sup> The initial AFE convergence rate, at  $\beta_1 = -0.1321$  ( $-0.1167-0.0154$ ), is quantitatively similar to, if statistically different from, the average rate of  $-0.1167$  across all disciplines. There are several possible reasons why the AFE discipline group could have a different estimated rate of convergence. This may reflect differences in international market conditions and rates of staff turnover for AFE researchers (see, for example, Boyle, 2008; Ehrenberg et al., 1991; Xu, 2008), a difference in attitudes and responses to the PBRFS incentives (Shin & Cummings, 2010), differences in the availability of contestable research funding for AFE researchers, or a difference in assessment standards adopted by the PBRFS assessment panel for the AFE group. In the absence of suitable data, these issues cannot be explored here.

Table 2 suggests that the inclusion of  $\log AQS_{2003}$  in the specification is statistically justified and slightly reduces the longer-run annual convergence rate; with  $\beta_2 = 0.022$  and hence  $\beta = -0.0947$  ( $t = -10.994$ ). The positive sign on  $\log AQS_{2003}$  reflects the fact that the degree of convergence in the first period is negatively associated with convergence in the second period, thereby modifying the full-period degree of convergence. This slightly slower rate of convergence when estimated over the whole 2003–2018 period is unsurprising given the rapid estimated convergence rates for 2003–2012 (reported in Buckle et al., 2020), such that by 2018, average research quality scores were much closer across universities and disciplines than they had been in 2003 when the PBRFS was introduced.

The value of  $\beta$ , while small in absolute terms, implies a high rate of convergence, in terms of differential average growth rates, conditional on the initial  $AQS$ . These are clearly illustrated in the cross-plot of the 87 individual observations in Fig. 1, which shows the relationship between  $\log AQS_{2012}$  and actual  $AQS$  growth. The  $\log AQS$  observations in 2012 can be seen to range from as low as 0.55 ( $AQS = 1.74$ ) to highs up to 1.88 ( $AQS = 6.54$ ). Similarly, differences across research units in the rate of growth of  $AQS$  over the 2012–2018 period are substantial, ranging from negative annual growth rates of  $-0.06$  to large positive growth of 0.12. The degree of  $\beta$ -convergence therefore reflects relatively rapid convergence, or ‘catch-up’.

Figure 1 reveals the negative relationship across individual research units, implying a strong  $\beta$ -convergence tendency. In addition, the highlighted observations (white dots) for average values (across disciplines) within the each of the eight universities confirm a convergence tendency at the aggregated university level. Thus, for example, AUT and Lincoln University (LU) had the lowest initial  $AQS$  values and the fastest  $AQS$  growth, whilst the initially-leading university, VUW, had the slowest (near zero)  $AQS$  growth.

<sup>16</sup> Regressions in Table 2 use 87 observations: 70 for the 8 universities and 9 disciplines within universities, less two missing observations for Law and Education at Lincoln, plus 17  $AQS_{ij}$  values, averaged across each university and each discipline. This enables parameters for each university’s growth and convergence to be compared directly with the average across all universities or disciplines rather than adopting one university and discipline as the omitted variable. However, adjusted- $R^2$ s must be interpreted cautiously since they are somewhat inflated by the inclusion of individual observations and their cross-university or cross-discipline averages.

<sup>17</sup> The estimated common convergence rate above is an annual rate, implying a convergence rate over the 6 years, from 2012 to 2018, of  $-0.7002$ , which compares with  $-0.722$  reported by Buckle et al. (2020) for 2003–2012 (and where lagged  $AQS$  values could not be included).

**Table 1** General-to-specific modelling of *AQS* growth

All universities and disciplines				(1) All universities; then		
Unit	Dummy	Order omitted	<i>t</i> -ratio	(2) All disciplines		
				Dummy	Order omitted	<i>t</i> -ratio
LU	<i>D</i>	17	-1.27	<i>D</i>	16	1.73
	<i>S</i>	19	1.09	<i>S</i>	1	0
AU	<i>D</i>	15	1.25	<i>D</i>	4	0.20
	<i>S</i>	7	-0.29	<i>S</i>	3	-0.47
CU	<i>D</i>	26	0.07	<i>D</i>	5	0.46
	<i>S</i>	25	-1.50	<i>S</i>	6	-0.33
OU	<i>D</i>	14	0.90	<i>D</i>	7	1.03
	<i>S</i>	13	-0.98	<i>S</i>	8	-0.22
WU	<i>D</i>	23	1.66	<i>D</i>	10	0.20
	<i>S</i>	24	-1.12	<i>S</i>	9	-0.97
VUW	<i>D</i>	9	-0.63	<i>D</i>	2	0.32
	<i>S</i>	10	0.17	<i>S</i>	12	-1.14
MU	<i>D</i>	1	0	<i>D</i>	15	0.36
	<i>S</i>	6	-0.23	<i>S</i>	14	-1.43
AUT	<i>D</i>	28	1.77	<i>D</i>	13	1.33
	<i>S</i>	29	-1.75	<i>S</i>	11	-0.86
Medicine	<i>D</i>	30	<b>-2.41</b>	<i>D</i>	2	-0.05
	<i>S</i>	2	0.03	<i>S</i>	9	-1.50
Engineering	<i>D</i>	12	0.71	<i>D</i>	3	0.36
	<i>S</i>	3	-0.05	<i>S</i>	1	-0.02
Core Science	<i>D</i>	33	0.30	<i>D</i>	17	0.30
	<i>S</i>	32	<b>-2.49</b>	<i>S</i>	16	-2.49
Management	<i>D</i>	5	0.19	<i>D</i>	8	-0.23
	<i>S</i>	11	0.64	<i>S</i>	7	-1.22
AFE	<i>D</i>	16	1.37	<i>D</i>	5	0.57
	<i>S</i>	-	<b>-3.48</b>	<i>S</i>	-	<b>-3.48</b>
Humanities	<i>D</i>	4	0.04	<i>D</i>	15	2.21
	<i>S</i>	21	1.47	<i>S</i>	4	-0.23
Agriculture	<i>D</i>	18	-0.93	<i>D</i>	12	-1.74
	<i>S</i>	22	1.44	<i>S</i>	13	1.44
Law	<i>D</i>	20	1.19	<i>D</i>	14	2.27
	<i>S</i>	8	-0.49	<i>S</i>	6	-0.81
Education	<i>D</i>	31	-2.3	<i>D</i>	11	1.75
	<i>S</i>	27	<b>2.57</b>	<i>S</i>	10	1.66

*D* shift; *S* slope; *LU* Lincoln University, *AU* Auckland University, *CU* Canterbury University, *OU* Otago University, *WU* Waikato University, *VUW* Victoria University of Wellington, *MU* Massey University, *AUT* Auckland University of Technology

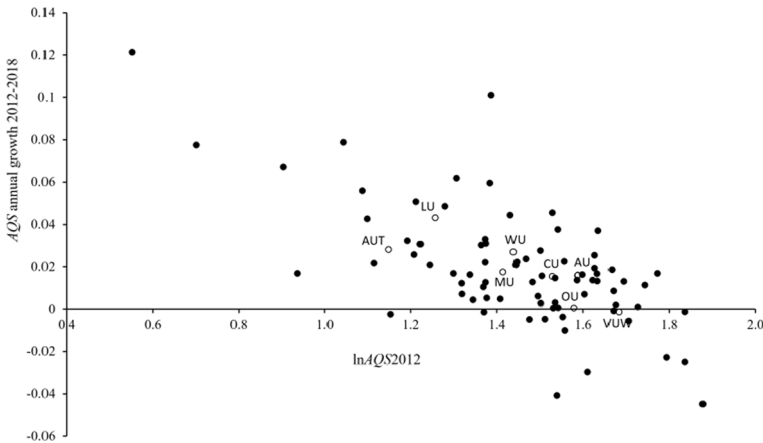
Based on the regression in Table 2, Fig. 2 shows the relationship between  $\log AQS_{2012}$  and expected values of *AQS* growth from the regression result. This is not simply a straight line, since values of  $\log AQS_{2003}$  vary across the observations shown.

**Table 2** Final regression results for *AQS* growth

	Coefficient	Standard error	<i>t</i> -value	95% confidence interval	
$\log AQS_{2012}$	-0.1167	0.0116	-10.09	-0.1396	-0.0937
$\log AQS_{2003}$	0.0220	0.0048	4.55	0.1124	0.0316
$D_{AFE} \times \log AQS_{2012}$	-0.0154	0.0044	-3.48	-0.0242	-0.0066
Constant	0.1699	0.0140	12.13	0.1420	0.1978
Long-run convergence <sup>a</sup>	-0.0947	0.0086	-10.99		
$R^2 = 0.605$	Adj- $R^2 = 0.591$		$F(3, 83) = 42.33$		Obs. = 87

Dependent variable: change in  $\log(AQS)$ , 2012–2018.  $D_{AFE} \times \log AQS_{2003}$  is a slope dummy variable for the Accounting, Finance and Economics discipline (AFE). Adding  $D_{AFE}$  to the regression, to confirm that its prior exclusion was justified, confirmed that the regression is clearly preferred (the *t*-ratios on both AFE dummy variables were less than |2|)

<sup>a</sup>Long run convergence standard errors are obtained by re-parameterising Eq. (3) such that RHS variables include  $\ln AQS_{2012}$  and  $(\ln AQS_{2012} - \ln AQS_{2003})$ , instead of  $\ln AQS_{2012}$  and  $\ln AQS_{2003}$ , giving estimates of  $(\beta_1 + \beta_2)$  and  $\beta_2$  respectively



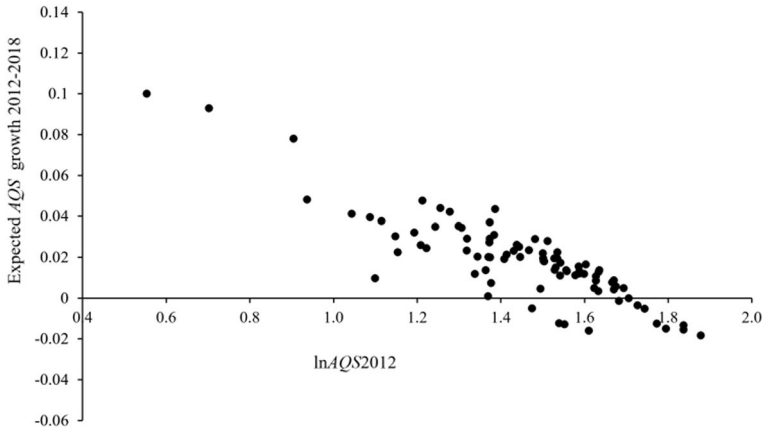
**Fig. 1** Actual *AQS* growth and initial *AQS*

Nevertheless, the dominant convergence relationship between  $\log AQS_{2012}$  and expected *AQS* growth is clear in Fig. 2, despite the small (lagged) divergence tendency contributed by  $\log AQS_{2003}$ . A small additional source of variation observed in Fig. 2 arises from the slightly different convergence parameter applicable to AFE.

Figure 3 compares actual and expected *AQS* growth at various points in the distribution from the 5th to the 95th percentile. This reveals a generally good fit across the distribution of *AQS* growth (2012–2018) observations, especially around the median. Unsurprisingly, expected values fit less well at the 5th and 95th percentiles, but nevertheless perform well, for example capturing around 67% of the large actual *AQS* growth rates at the 95th percentile (0.052/0.078), but under 30% at the 5th percentile.

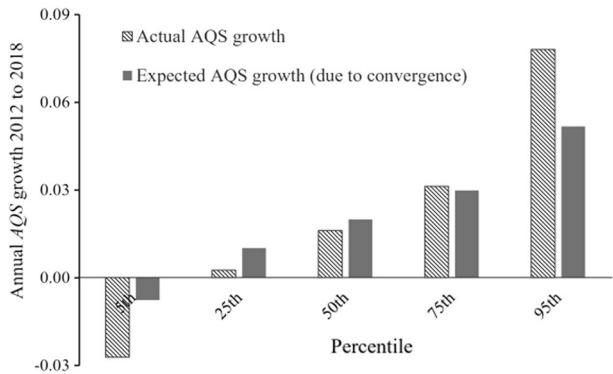
The effects of control variables, including their potential impacts on convergence parameters, were tested by adding three variables to Eq. (3). These were: the initial number of staff FTEs in a university-discipline unit; its gender balance (female staff





**Fig. 2** Expected AQS growth and initial AQS

**Fig. 3** Actual and expected AQS growth at percentiles



share); and staff median age. These variables were added individually and in combination, but none revealed any statistically significant effects, using conventional confidence intervals.

**$\sigma$ -Convergence**

As mentioned above, the overall convergence or divergence in AQS scores is measured by  $\sigma$ -convergence, indicating whether the distribution of AQSs across research units is becoming more or less dispersed. A value of  $\beta < 0$  does not necessarily imply  $\sigma$ -convergence, which also depends on the error variance,  $\sigma^2_\epsilon$  in (3). Table 3 reports the variances of  $\log AQS$  across universities and disciplines separately and combined, and for each discipline within universities. This shows that there were substantial reductions in the overall dispersion of  $\log AQS$  between 2003 and 2018 for all groups: for all universities and disciplines combined and for each discipline within universities.<sup>18</sup>

<sup>18</sup> When AUT (which had a much lower AQS in 2003 than other universities) is excluded the variance still declines but by much less than when AUT is included. Similarly, when Education (which had a much lower

However, the change in variances differs substantially between the earlier period 2003–2012 and the later period, 2012–2018. The largest declines in dispersion occur during 2003–2012, whereas during 2012–2018 changes in variances are much smaller and are not lower in all cases. For all universities combined, the variance is more than halved during 2012–2018, but for disciplines there is little change. For disciplines within universities, the decline of the variance for Massey disciplines is as large as for the earlier period, but for the other universities there is very little change, with half experiencing a small increase and the others a small decrease or no change.

Thus a strong common rate of  $\beta$ -convergence across all university and discipline units during 2012–2018 did not necessarily translate into an overall decline in dispersion across all units, with differences in  $\sigma$ -convergence observed across periods. Whereas substantial  $\sigma$ -convergence is observed for all groups during the initial PBRFS phase in 2003–2012, there is a mix of  $\sigma$ -convergence and  $\sigma$ -divergence during 2012–2018, with a relatively small reduction in variances after 2012. This partly arises from the relatively low variance of  $\log AQS$  achieved by 2012.

## Contribution of exits, entrants and quality transformations to $AQS$ changes

The previous section estimated the extent of research-quality convergence of universities and discipline groups, although it did not directly consider the precise sources of the  $AQS$  growth in terms of the turnover of staff and their quality transformation. This section examines these sources and their dynamics, revealing how the separate contributions of exits, entrants and quality transformations of incumbent researchers combine to generate overall convergence. The analysis uses a decomposition method to identify the contributions of these three components, for universities and discipline groups.

## Determinants of research quality change

To identify the way in which the component flows combine to affect the  $AQS$  of a university or discipline group, first let  $n_0$  and  $n_1$  denote the number of individuals in a university at times 0 and 1 respectively, and assume for convenience that all are full-time employees. As above,  $G_i$  denotes the score attached to each person,  $i$ , depending on the Quality Category, QC. The four QCs have scores,  $g_k$ , (for  $k=1, \dots, 4$ ), of 10, 6, 2 and 0 (for A, B, C and R researchers respectively). Hence if person  $i$  belongs to category  $k$ , the score is  $G_i = g_k$ . Finally, let  $Q_0$  and  $Q_1$  denote  $AQS$ s at times 0 and 1. By definition, the initial  $AQS$  is:

$$Q_0 = \frac{1}{n_0} \sum_{i=1}^{n_0} G_i \quad (4)$$

---

Footnote 18 (continued)

$AQS$  in 2003 than other disciplines) is excluded the variance still declines but by much less than when Education is included.

**Table 3** Variances of logarithms of AQS, 2003, 2012 and 2018

Groups for which variances are estimated	2003	2012	2018	Obs.
All universities	0.2280	0.0283	0.0121	8
All disciplines	0.0850	0.0135	0.0138	9
All universities and disciplines	0.1541	0.0211	0.0131	17
Across disciplines (within universities)				
AUT	0.1400	0.0161	0.0202	9
Lincoln	0.0935	0.0632	0.0598	7
Massey	0.1584	0.0866	0.0185	9
Auckland	0.0911	0.0084	0.0175	9
Canterbury	0.2358	0.0261	0.0235	9
Otago	0.0843	0.0189	0.0189	9
Waikato	0.0873	0.0071	0.0101	9
VUW	0.1801	0.0295	0.0333	9
Across universities (within disciplines)				
Medicine	0.3625	0.0353	0.0274	8
Engineering	0.2442	0.0243	0.0155	8
Core science	0.3816	0.0259	0.0258	8
Management	0.1690	0.0429	0.0241	8
Acc Fin Eco	0.3203	0.0815	0.0136	8
Humanities	0.2236	0.0401	0.0144	8
Agriculture	0.1009	0.0230	0.0232	8
Law	0.4163	0.1881	0.0360	7
Education	0.1413	0.0265	0.0241	7
All university-discipline units	0.3554	0.0636	0.0346	70

Source Authors’ estimates using data from Appendix Table 7

Suppose  $X_k$  people exit from quality category,  $k$ , and  $E_k$  enter, belonging to quality category,  $k$ . Furthermore, define  $T_k$  as the net transformations of incumbent researchers into quality category,  $k$ . That is,  $T_k$  measures the transfers into  $k$  from all other categories, net of the transfers out of  $k$  into all other categories, so that  $\sum_{k=1}^4 T_k = 0$ . Allowing for all entries, exits and transformations over the period, the proportional change in quality is thus:

$$\frac{\Delta Q}{Q} = \frac{Q_1 - Q_0}{Q_0} = \frac{1}{n_0 + \sum_{k=1}^4 (E_k - X_k)} \left( \frac{1}{Q_0} \sum_{k=1}^4 (E_k - X_k + T_k) g_k - \sum_{k=1}^4 (E_k - X_k) \right) \tag{5}$$

The AQS change therefore depends in a complex way on the flows contained in Eq. (5). However, it is possible, as shown in subsection ‘Decomposition of quality changes’, to decompose the change into the three separate components. First is it useful to consider whether differences in the flows among universities and disciplines is likely to give rise to convergence.

Differential rates of exit, entry and quality transformation can contribute towards convergence for several reasons. For example, there is an upper limit to the extent to which exits can continue to contribute to improving a university or discipline AQS: as the number

of low-quality researchers is reduced, the scope to improve an *AQS* from exits diminishes. Low-quality universities have relatively large numbers of R-rated researchers, encouraging a high exit rate of Rs. Higher-quality universities are able to retain their higher-quality researchers, and at the same time can more readily attract high-quality researchers from other universities in NZ as well as from outside the system. In addition, universities with high *AQS*s may be able to attract strong younger researchers, who enter as C or even R types, but are capable of progressing more rapidly to higher levels.

The PBRFS scoring system implies that, for example, a university with an initial *AQS* that is below 2 can increase its average quality by employing more C-type researchers. But a university with an *AQS* above 2 reduces its average quality by employing more Cs: they will wish to be confident that those who do enter the university are capable of subsequently raising their score. The fact that the attractiveness of C-type researchers varies with the *AQS* means not only that there are differences among universities at any time, but their ability to raise their *AQS* further via recruitment practices is likely to change over time as their *AQS*s change.

A similar argument could be made for the marginal effects of transformations (*T*), and there may also be diminishing benefits from investing in initiatives that improve the research environment and skills of incumbent researchers. For entrants, there is a diminishing effect from a university's budget constraint, as well as from constraints imposed by other university requirements (such as teaching and administration staffing requirements). For these reasons, the effects of exits, entrants and transformations on the rate of improvement of research quality of a university or discipline can be expected to display diminishing marginal productivity analogous to those evident in models of economic growth convergence.<sup>19</sup>

## Decomposition of quality changes

The required decomposition can be obtained by suitably modifying the relevant transition matrix, as follows.<sup>20</sup> Let the initial and final *AQS* for the specified group be denoted by  $Q_1$  and  $Q_2$ . The aim is to decompose  $Q_2 - Q_1$  into components that measure the separate impact on the change in *AQS* of exits, entrants and quality transformations. An *AQS* can be calculated for the final period, using a counterfactual assumption of no entrants into any category, denoted  $Q_3$ .

Furthermore, it is possible to obtain an alternative counterfactual *AQS* in the final period, by setting all exits and entrants to zero: this is denoted  $Q_4$ . The counterfactual of no exits is applied by supposing that those recorded as exiting remain in the quality category in which they were placed in the initial period: that is, the diagonals of the flow matrix are augmented by the number of (FTE) exits. The difference,  $Q_4 - Q_1$ , therefore reflects the effect of quality transformations made by those who remain in the system (since, in calculating  $Q_4$ , those who actually exit are assumed to remain on their respective diagonal). The difference,  $Q_3 - Q_4$ , reflects the separate effect of exits. Finally, the difference,  $Q_2 - Q_3$ , measures the effect of entrants. These are combined to give:

<sup>19</sup> See, for example, Abramovitz (1986), Barro and Sala-i-Martin (1991), Baumol (1986), Dowrick and Gemmell (1991) and Quah (1993).

<sup>20</sup> Examples of transition matrices and comparisons among disciplines and universities can be found in Buckle and Creedy (2019a).

$$Q2 - Q1 = (Q2 - Q3) + (Q3 - Q4) + (Q4 - Q1) \tag{6}$$

These contributions can be derived for the entire university system, and for any unit within the system. Figure 4, derived using the data in Table 7, illustrates the separate average annual contributions to the changes in *AQS* of exits, entrants and quality transformations for all universities combined, each university, and each discipline group. The contributions for each discipline within each university are shown in Fig. 5.

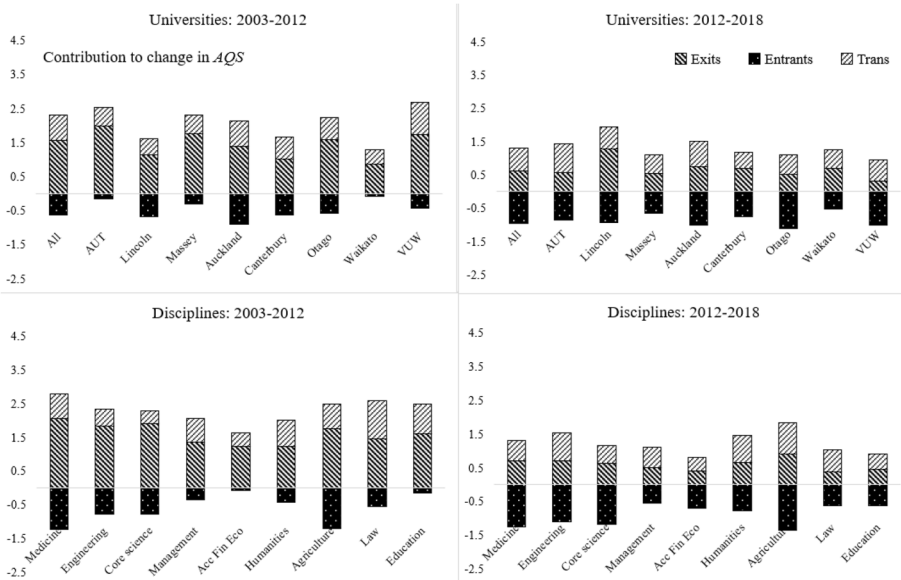
It is evident from Fig. 4 that, for the university system as a whole and for each university and discipline, the net impact of exits is always positive, the net impact of entrants is always negative, and the net impact of quality transformations is always positive. This pattern prevails in both periods, although the sizes of the average annual contributions change. In the second period the positive average annual contribution from quality transformations is larger (0.08 in 2003–2012 and 0.12 in 2012–2018) and from exits is smaller (0.17 in 2003–2012 and 0.11 in 2012–2018). The negative average annual contribution from entrants increases during the second period (−0.07 in 2003–2012 and −0.16 in 2012–2018).

The characteristics for separate disciplines within each university shown Fig. 5, reveals greater diversity in the level of contributions and the changes between the two periods than is evident for the more aggregated university and discipline groups shown in Fig. 4. Although the dominant pattern of these researcher contributions is the same for universities and disciplines as a whole, there are several cases where the impact of exits and quality transformations have a negative effect on *AQS*s, and of entrants having a positive impact. The data for Figures 4 and 5 are provided in Table 7 in the Appendix.

A word of caution is necessary in interpreting the negative contribution of entrants observed above. This contribution,  $Q2 - Q3$ , measures the difference between the final *AQS* and that which would result from having no entrants, and relates only to the *AQS* values at the end of a period. Hence a negative value of  $Q2 - Q3$  suggests that entrants on average, by the end of a period period, are of lower-quality than the incumbents at the end of a period. And of course, incumbents are typically subject to positive net transformations over the period. This does not necessarily mean that the entrants do not contribute positively to a higher overall *AQS*, as their contribution is positive so long as the average quality of entrants is higher than the overall start-of-period *AQS*.

Further details of the contributions are provided in Table 4, for all universities combined. This shows the beginning and end-of-period *AQS*s of various groups, and differences between the *AQS*s of the groups. The table shows how the different groups have contributed to the overall change in the *AQS*. The *AQS* of exits during 2003–2012, of 2.12, was substantially lower than the *AQS* of all researchers, of 2.88, in 2003. In 2003 the *AQS* of those who remained incumbent in the system was 3.68, which was 0.80 higher than the average of all researchers. They improved considerably to 5.18. Interestingly, the entrants during this period had an *AQS* in 2003 of 3.90. Although this was lower than the *AQS* of all researchers in 2012, of 4.55, it was higher than the *AQS* of all researchers in 2003. Therefore, the net effect of entrants was to contribute to an improvement in the score from 2003, of 1.02, but this contribution to the improvement in *AQS* was lower than that of incumbents, which was 2.3.

The second period differs in important ways from the first. The exits during this period had an *AQS* lower than that of all researchers in 2012. However, the difference was about half of the corresponding difference in the earlier period. Entrants, on the other hand, had a much lower score in 2018 than did entrants in the earlier period (by 2012) and was also lower than the *AQS* for all researchers at the start (2012) of this



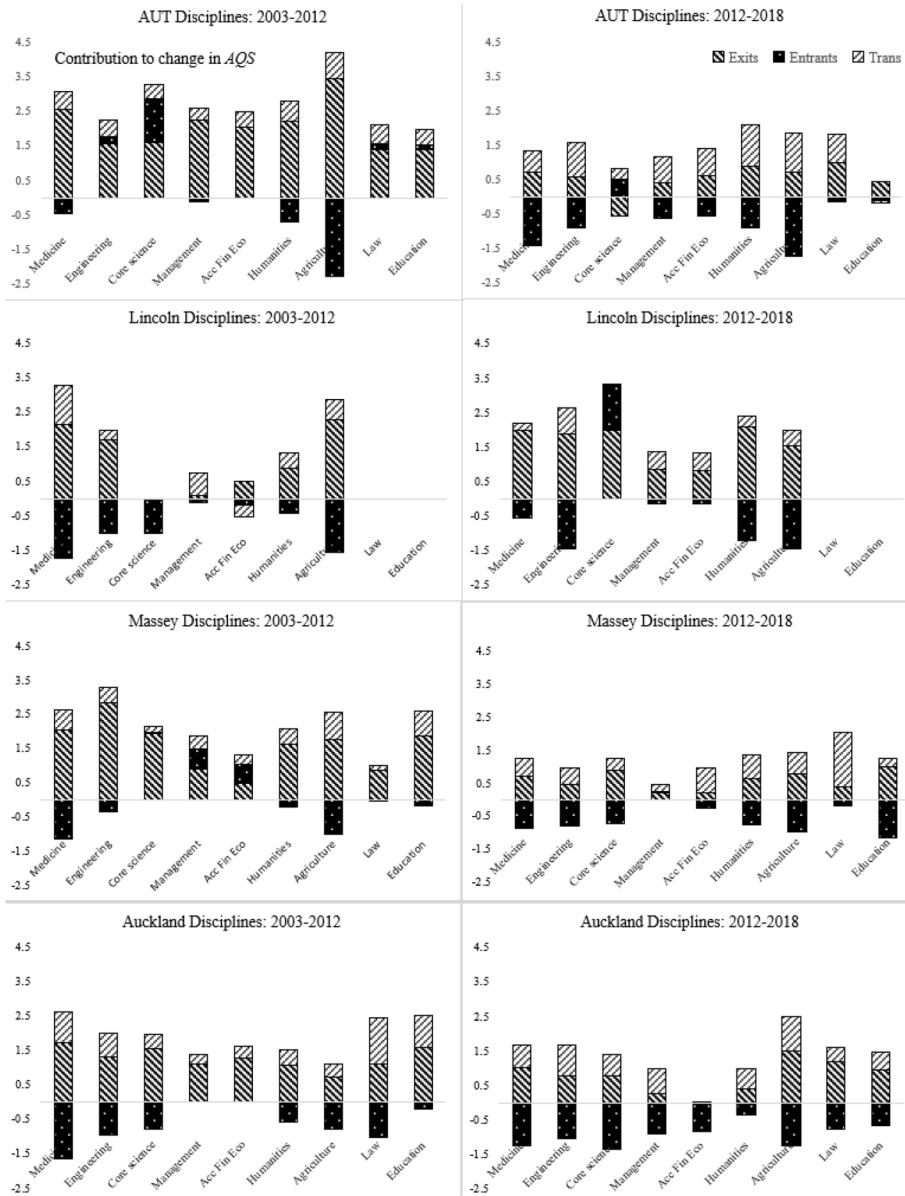
**Fig. 4** Average annual contributions to *AQS* growth of exits, entrants and quality transformations. *Note* Derived from Appendix Table 7

second period. Hence, the net effect of entrants during 2012–2018 was to reduce the *AQS* of all researchers in 2018, from what it would otherwise have been (that is, with no entrants). The *AQS* of entrants in 2018 was 0.97 lower than the score of all researchers in 2012, contrasting with an *AQS* of incumbents which was 1.31 higher than the average of all researchers at the start of the period. In addition, the *AQS* of entrants in 2012 was actually lower than the *AQS* in 2012 of those who exited the system during the second period.

The positive contributions of exits and quality transformations to the growth in *AQS*s are consistent with the new incentives created by the PBRFS. These encouraged universities to remove lower-quality researchers and to retain higher-quality and promising researchers considered more likely to transition to a higher QC over time. The positive net impact of quality transformations reflects decisions to attract and retain good researchers who improve over time. This net positive effect captures the mixture of those who improve and those who decline in quality while remaining within the institution. There were clearly diminishing gains from the positive net transformations of incumbents during the second period compared with the first.

The average annual contribution of exits, although also on average positive in both periods, is smaller in the second period. This reflects the situation where the large gains achieved by the removal of a very high proportion of R researchers during the first period reduced the scope for similar *AQS* gains from this source during the second period. During this second period, the proportion of exits by higher-quality researchers increases substantially.

The contribution of entrants to growth in university *AQS*s clearly deteriorated during the second period compared with the first period: their *AQS* was lower than that of entrants in the first period, and lower than the *AQS* of all researchers at the start of the second period,



**Fig. 5** Contributions to exits, entrants and quality transformations to changes in AQS across disciplines. Notes Derived from Appendix Table 7

and of exits during that period. It clearly became increasingly more difficult to recruit entrants above the average quality.<sup>21</sup>

<sup>21</sup> The lower average quality of entrants appears not to be dominated by hiring younger new PhD graduates; see Buckle et al. (2021).

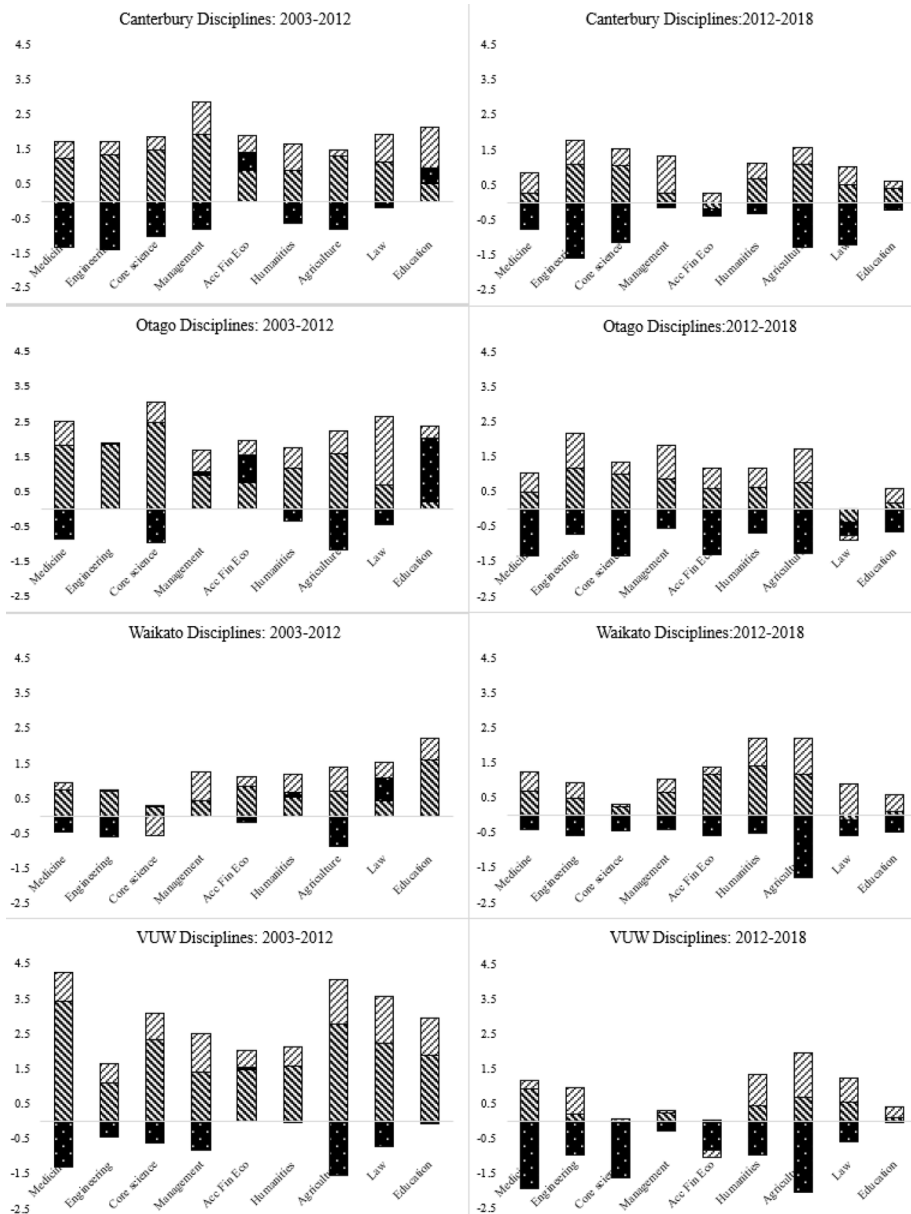


Fig. 5 (continued)

To decompose  $\sigma$ -convergence, Table 5 reports standard deviations of the separate average annual contributions of exits, entrants and quality transformations to changes in AQS. These standard deviations are similar in both periods for universities and disciplines overall. The changes in standard deviations across disciplines within universities (shown in the lower section of the table) are more diverse, as observed in section ‘Growth rates of



**Table 4** AQSs for exiting, entering and incumbent researchers

Period	AQS	Differences in AQS compared to		
		All in 2003	Incumbents in 2003	Incumbents in 2012
<b>2003–2012</b>				
AQS all researchers in 2003	2.88			
AQS of incumbents in 2003	3.68	0.80		
AQS of incumbents in 2012	5.18	2.30	1.50	
2003 AQS of exiting researchers	2.12	−0.76	−1.56	−3.06
2012 AQS of entering researchers	3.90	1.02	0.22	−1.28
AQS of all researchers in 2012	4.55			
<b>2012–2018</b>				
		All in 2012	Incumbents in 2012	Incumbents in 2018
AQS of all researchers in 2012	4.55			
AQS of incumbents in 2012	4.82	0.27		
AQS of incumbents in 2018	5.86	1.31	1.04	
2012 AQS of exiting researchers	4.06	−0.49	−0.76	−1.80
2018 AQS of entering researchers	3.58	−0.97	−1.24	−2.28
AQS all researchers in 2018	4.91			

AQSs derived from data in Table 1. Incumbents refers to researchers who remained within the university system during the period and participated in the beginning and end of period assessment rounds

**Table 5** Standard deviations of the X, E & T average annual contributions to AQS growth

	2003–2012			2012–2018		
	Exits	Entrants	Trans	Exits	Entrants	Trans
Universities	0.045	0.031	0.018	0.047	0.033	0.020
Disciplines	0.036	0.046	0.026	0.029	0.053	0.030
AUT disciplines	0.073	0.105	0.013	0.076	0.175	0.068
Lincoln disciplines	0.106	0.071	0.053	0.092	0.166	0.040
Massey disciplines	0.081	0.066	0.026	0.049	0.067	0.072
Auckland disciplines	0.034	0.060	0.040	0.079	0.053	0.047
Canterbury disciplines	0.045	0.077	0.035	0.074	0.093	0.040
Otago disciplines	0.070	0.106	0.059	0.076	0.064	0.065
Waikato disciplines	0.043	0.050	0.047	0.088	0.073	0.053
VUW disciplines	0.083	0.062	0.036	0.053	0.117	0.080

Source Authors’ calculations using data from Appendix Table 7

The standard deviations are of the average annual contributions of X, E and T to AQS growth

**Table 6** Regression results for *AQS* growth components, *X*, *E* and *T*

	Stage 1 results			Stage 2 results		
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
	Exits	Entrants	Trans	Exits	Entrants	Trans
$\log AQS_{2012}$	-0.0625 (0.0112)**	-0.0847 (0.0591)	-0.0479 (0.0137)**	-0.0433 (0.011)**	-0.1230 (0.0506)**	-0.0424 (0.0119)**
$\log AQS_{2003}$	0.0160 (0.0047)**	0.0650 (0.0273)	-0.0190 (0.0055)**	0.0152 (0.0057)**	0.0584 (0.0233)**	-0.0257 (0.0042)**
$Age_{2012}$	†	0.0068 (0.0030)**	-0.0024 (0.0006)**	†	†	-0.0014 (0.0005)**
$D_{MU}$					-0.0793 (0.0249)**	0.1027 (0.0231)**
$D_{LU}$						-0.0185 (0.0059)**
$D_{AUT}$				0.1802 (0.046)**		
$D_{MU} \times \log AQS_{2012}$						-0.0810 (0.0167)**
$D_{LU} \times \log AQS_{2012}$				0.0185 (0.0048)**		
$D_{AUT} \times \log AQS_{2012}$				-0.1468 (0.0387)**		
$D_{Edu}$						-0.2612 (0.0389)**
$D_{Ag}$					1.4487 (0.4731)**	
$D_{Med}$					-0.1004 (0.0300)**	
$D_{Edu} \times \log AQS_{2012}$						0.1795 (0.0298)**
$D_{Ag} \times \log AQS_{2012}$					-0.9478 (0.2857)**	
$D_{AFE} \times \log AQS_{2012}$						-0.0136 (0.0033)**
Constant	0.1007 (0.0135)**	-0.3608 (0.1894)	0.2355 (0.0410)**	0.0704 (0.0133)**	0.0707 (0.0618)	0.1899 (0.0321)**
Adj- $R^2$	0.262	0.079	0.470	0.434	0.326	0.784
Regression $F$	16.27	3.19	26.24	14.2	7.21	35.65
Obs	87	78 <sup>††</sup>	87	87	78 <sup>††</sup>	87
Long-run convergence	-0.0465 (0.0083)**	-0.0196 (0.0471)	-0.0669 (0.0106)**	-0.0281 (0.0082)**	-0.0646 (0.0384)	-0.0681 (0.0092)**

This table shows the estimated regression coefficients and standard errors are in parentheses

\*\* (\*) = significant at 1% (5%)

† Adding  $Age_{2012}$  is not statistically significant:  $t = -1.04$  in (i);  $t = -1.57$  in (iv); and  $t = 1.36$  in (v)

†† There are fewer observations for entrants due to nine observations where *AQS* fell between 2012 and 2018, such that the log difference is undefined

research quality and convergence’ for changes in variances of logarithms of *AQS*s across disciplines within universities.

## **$\beta$ -Convergence properties of contributions to *AQS* growth**

This section evaluates how the separate components of *AQS* growth (entry, exit and quality transformation) contribute to the  $\beta$ -convergence properties observed for *AQS* growth as a whole. That is, did initially lower-quality units catch-up on higher-quality units largely via improvements to retained staff, or via the removal of low-quality staff and/or recruitment of higher-quality new staff? While results reported in section ‘Contribution of exits, entrants and quality transformations to *AQS* changes’ revealed negative effects on *AQS* growth from new entrants, they could nevertheless have contributed to cross-unit quality convergence. For example, to help catch-up on other units, some universities or disciplines may have used a strategy of quality improvement for retained staff, even if they had limited ability, or made little attempt, to use exits and entrants for this purpose, or vice versa. The previous section revealed substantial variation across universities and discipline groups in the contributions of these different components.

To examine the influences on convergence of exits (*X*), entrants (*E*) and quality transformations (*T*), a two-stage process was followed. First, regressions of Eq. (3) were estimated including control variables, with statistically insignificant controls progressively eliminated. This suggested at most one control variable, each unit’s initial median age of staff,  $Age_{2012}$ , with a significant effect on *AQS* growth, but only for *E* and *T* components. Results are reported in columns (i) to (iii) of Table 6. Second, the previous general-to-specific process was followed, starting with all dummy variables and  $Age_{2012}$  included, and eliminating insignificant variables in turn. This yielded the ‘final’ regressions reported in columns (iv) to (vi) of Table 6.<sup>22</sup>

Allowing for the possibility of different rates of convergence across universities and disciplines, final results in columns (iv) to (vi) suggest that all three components of *AQS* growth contributed towards overall convergence, with  $\beta_1$  negative in all cases, and with additional negative effects from  $\beta_2$  for quality transformations, *T*. Long-run convergence parameters also confirm negative effects in all three cases. In the case of *X* and *T* these are robustly identified (*t*-ratios exceed |3|), while for *E* the estimate (−0.0646) is significantly negative at the 10% level. However, estimated magnitudes suggest that the largest effects on *AQS* growth are from *E* and *T* (around −0.065 to −0.068), with estimated convergence effects for *X* smaller at −0.028.

Results in columns (iv) to (vi) of Table 6 also suggest few statistically significant deviations from a uniform contribution to the convergence rate across universities and disciplines. For exits, *X*, there is some evidence of a larger convergence rate contribution for AUT (which initially had the lowest *AQS* and the fastest growth in the first period), and a slower rate for Lincoln (LU). For *T*, only Massey university (MU) displays a significantly greater rate of convergence. Across disciplines there is some evidence for different convergence rates on *E* or *T* for education, agriculture, and AFE.<sup>23</sup>

<sup>22</sup> Based on regressions (iv) to (vi), re-testing the inclusion of all control variables by adding them to those regressions did not support the inclusion of any.

<sup>23</sup> There are large values in Table 6 for the two  $Ag$  dummies for Entrants, suggesting a dramatic convergence effect and large autonomous *AQS* growth,  $D_{Ag}$ , over 2012 to 2018. However, there are few observations because AUT and LU Agriculture discipline groups units are omitted due to negative *AQS* changes.

Overall, these results suggest that the largest components contributing towards  $\beta$ -convergence in *AQS* across universities and disciplines from 2012 to 2018 were via recruitment of entrants and by transforming the quality scores of established and remaining researchers. Exiting staff from those universities and disciplines also facilitated convergence of average quality scores at the next PBRF audit, but made a smaller contribution: just under half of that of each of the other two components. Nevertheless, an important aspect of these results is that entrants can be seen to contribute consistently towards  $\beta$ -convergence yet display a mixed record of  $\sigma$ -convergence and  $\sigma$ -divergence, as shown in Table 5.

## Conclusions

This paper has extended understanding of the convergence properties arising from the New Zealand PBRFS. In particular, it has developed a decomposition method which allows the separate contributions to the convergence process of staff turnover, in terms of exits and entrants, and quality transformations of incumbent staff, to be identified. The approach has also been extended to examine the question of whether changes between two assessment exercises depend on previous researcher quality changes. These contributions can clearly be related to the incentives, and constraints, within the system.

A strong degree of  $\beta$ -convergence was found for the period 2012–2018. Despite the considerable initial variation across universities in their *AQS* scores, a common convergence process was found to operate across all universities, and all discipline groups within and across universities. Furthermore, the results also suggested that universities' growth in research quality from 2012 to 2018 was affected by growth in the prior period. A relatively high rate of improvement in the first period systematically mitigated growth in the second period. There was also some evidence that the *AQS* for one discipline, Accounting, Finance and Economics, may have converged at a slightly faster rate than other disciplines. In addition, it was found that the convergence properties were not affected by differences in gender ratios, median age, and size of universities and their discipline composition.

Strong  $\beta$ -convergence is not necessarily associated with reduced dispersion or  $\sigma$ -convergence. However, the results confirm that a strong  $\sigma$ -convergence process is observed over 2003–2012. Across universities, this was maintained during 2012–2018, despite the fact that the dispersion of *AQS* levels across universities and disciplines was already much reduced by 2012. Across disciplines,  $\sigma$ -convergence was substantially reduced during 2012–2018, with  $\sigma$ -divergence in some cases.

When considering the contributions of researcher exits, entry and quality transformation, two outcomes are especially noteworthy. First, exits and quality transformations both contributed, as expected, to improvements in *AQS*s on average across universities and disciplines. New entrants (whether from outside the NZ system or cross-university transfers within it) raised *AQS* levels over the period 2003–2012, as they had an average quality in excess of the initial *AQS*, although their contribution to growth was less than that of incumbents. Over the period 2012–2018, entrants on average actually served to reduce *AQS* levels as their average quality was below the average in 2012. The combination of substantial quality improvement over the first period, combined with a difficulty of recruiting (and retaining) higher-quality researchers in the second period, seems to have contributed to significant 'decreasing returns' from the PBRFS.

Second, the overall decreasing returns combined with similar effects within universities and disciplines, such that all three components of change contributed to a process of  $\beta$ -convergence, rather than divergence, in *AQS* levels across universities and disciplines. The rate of *AQS* convergence was around  $-0.06$ , or 6%, per year for entrants and transformations, and somewhat slower, at around  $-0.03$  for exits.

Further, these convergence rates were relatively uniform across universities and disciplines with only a few exceptions. These exceptions are: a slower rate of convergence via exits for Lincoln University; a faster rate of convergence via exits for AUT; and a faster rate of convergence via quality transformations for Massey University. Among disciplines, the exceptions include: faster convergence via quality transformations for AFE, and via entrants to Agriculture; but a slower rate of convergence via quality transformation in Education.

In the New Zealand case, the PBRFS funding formula was not designed explicitly to favour increased concentration of researchers, nor to favour low-quality units. Rather, funding allocation was both scale-neutral and neutral with respect to the location of individuals with different quality scores. This level playing field in financial PBRF incentives across research units may have encouraged initially low-rated universities and disciplines to aim for substantial improvement over 2003–2018, and avoided discouraging a process of knowledge transfer across universities. These aspects may have facilitated the convergence outcome observed here.

The results found for the NZ PBRFS do not necessarily carry over to assessment exercises in other countries. This is because the process depends on the particular incentive structure created by a PBRFS, and these differ among countries. The basic methods used to investigate whether there is  $\beta$ - and  $\sigma$ -convergence can be applied to other countries for which their assessment process assigns some kind of quality score to universities and discipline groups, and do not require longitudinal information about individual researchers. However, the present analysis, identifying separate contributions to change, has been possible due to the New Zealand PBRFS's design, which requires all qualifying academics to be assessed in all rounds. Such data are not generally collected as a matter of course for other countries. Nevertheless, the same fundamental processes are involved in generating research quality improvements in any system. Hence, it would be useful for countries to collect, where possible, longitudinal information about individual researchers. Crucially, the incentives created by a PBRFS matter, and it would be useful for the evaluation process to be designed such that the information necessary to investigate responses to incentives is collected.

## Appendix

### Contributions of exits, entrants and quality transformations to *AQS* changes

**Table 7** Contributions of exits, entrants and quality transformations to changes in AQS: 2003–2012 and 2012–2018

	AQS 2003			AQS 2012			AQS 2018		
	Exits	Entrants	Trans	Exits	Entrants	Trans	Exits	Entrants	Trans
University	All universities	2.88	1.57	0.73	4.55	0.63	0.69	4.91	0.69
	AUT	0.77	2.00	0.54	3.15	0.57	0.87	3.73	0.87
	Lincoln	2.56	1.15	0.47	3.51	1.29	0.66	4.54	0.66
	Massey	2.10	1.76	0.55	4.11	0.54	0.56	4.56	0.56
	Auckland	3.67	1.40	0.73	4.89	0.75	0.75	5.38	0.75
	Canterbury	3.55	1.02	0.65	4.61	0.69	0.50	5.05	0.50
	Otago	3.18	1.60	0.64	4.85	0.52	0.59	4.86	0.59
	Waikato	2.98	0.87	0.43	4.21	0.69	0.58	4.95	0.58
	VUW	3.11	1.75	0.93	5.38	0.32	0.64	5.34	0.64
	Medicine	2.93	2.06	0.74	4.49	0.70	0.62	4.55	0.62
Discipline	Engineering	3.00	1.83	0.51	4.57	0.70	0.84	5.01	0.84
	Core science	3.79	1.91	0.39	5.31	0.64	0.51	5.28	0.51
	Management	2.21	1.35	0.72	3.95	0.50	0.61	4.51	0.61
	Acc Fin Eco	2.41	1.24	0.39	3.96	0.40	0.42	4.09	0.42
	Humanities	3.17	1.24	0.76	4.74	0.65	0.81	5.43	0.81
	Agriculture	3.65	1.75	0.75	4.95	0.92	0.93	5.45	0.93
	Law	3.04	1.46	1.12	5.07	0.38	0.66	5.5	0.66
	Education	1.38	1.61	0.89	3.74	0.46	0.44	4.02	0.44
	Medicine	0.56	2.54	0.53	3.17	0.71	0.65	3.12	0.65
	Engineering	1.06	1.57	0.47	3.30	0.60	0.98	4.00	0.98
AUT	Core science	0.66	1.61	0.40	3.93	-0.56	0.29	4.19	0.29
	Management	0.87	2.25	0.35	3.35	0.40	0.79	3.91	0.79
	Acc Fin Eco	0.50	2.05	0.45	3.00	0.63	0.79	3.88	0.79
	Humanities	0.85	2.22	0.58	2.96	0.89	1.21	4.15	1.21
	Agriculture	1.81	3.43	0.76	3.74	0.72	1.16	3.90	1.16
	Law	0.73	1.39	0.55	2.84	1.02	0.81	4.55	0.81
	Education	0.57	1.40	0.45	2.55	0.44	-0.06	2.82	-0.06

**Table 7** (continued)

	AQS 2003	Exits	Entrants	Trans	AQS 2012	Exits	Entrants	Trans	AQS 2018	
Lincoln <sup>a</sup>	Medicine	2.15	2.13	-1.73	1.14	3.69	2.00	-0.54	0.2	5.35
	Engineering	2.40	1.69	-1.00	0.27	3.36	1.90	-1.45	0.74	4.55
	Core science	5.00	0.00	-1.00	0.00	4.00	2.00	1.33	0.00	7.33
	Management	1.86	0.08	-0.12	0.65	2.47	0.86	-0.15	0.52	3.70
	Acc Fin Eco	2.05	0.5	-0.18	-0.35	2.02	0.82	-0.14	0.51	3.21
	Humanities	2.66	0.89	-0.41	0.45	3.59	2.09	-1.20	0.32	4.80
	Agriculture	2.94	2.28	-1.56	0.58	4.24	1.55	-1.44	0.45	4.80
	Medicine	2.15	2.06	-1.14	0.59	3.66	0.73	-0.87	0.53	4.05
	Engineering	1.51	2.84	-0.35	0.46	4.46	0.45	-0.78	0.50	4.63
	Core science	2.98	1.95	0.03	0.16	5.12	0.90	-0.71	0.35	5.66
Massey	Management	1.58	0.91	0.58	0.39	3.46	0.20	0.05	0.22	3.93
	Acc Fin Eco	2.07	0.48	0.56	0.28	3.39	0.20	-0.26	0.75	4.08
	Humanities	2.37	1.62	-0.22	0.47	4.24	0.65	-0.75	0.71	4.85
	Agriculture	2.92	1.78	-0.99	0.79	4.50	0.78	-0.98	0.64	4.94
	Law	0.78	0.87	-0.04	0.13	1.74	0.39	-0.18	1.65	3.60
	Education	1.39	1.89	-0.17	0.72	3.83	1.01	-1.13	0.23	3.94
	Medicine	3.66	1.73	-1.65	0.90	4.64	1.04	-1.24	0.63	5.07
	Engineering	4.04	1.32	-0.95	0.68	5.09	0.78	-1.04	0.88	5.71
	Core science	4.16	1.55	-0.79	0.42	5.34	0.79	-1.33	0.60	5.40
	Management	3.09	1.11	0.01	0.28	4.49	0.27	-0.91	0.71	4.56
Auckland	Acc Fin Eco	3.39	1.28	0.01	0.33	5.01	0.02	-0.84	0.00	4.19
	Humanities	4.33	1.09	-0.57	0.44	5.29	0.40	-0.35	0.58	5.92
	Agriculture	4.81	0.75	-0.79	0.36	5.13	1.52	-1.23	0.98	6.40
	Law	3.65	1.11	-1.01	1.34	5.09	1.19	-0.76	0.41	5.93
	Education	1.61	1.58	-0.19	0.95	3.95	0.95	-0.67	0.53	4.76

Table 7 (continued)

		AQS 2003	Exits	Entrants	Trans	AQS 2012	Exits	Entrants	Trans	AQS 2018	
Canterbury	Medicine	4.23	1.25	-1.31	0.47	4.64	0.28	-0.75	0.56	4.73	
	Engineering	4.64	1.34	-1.39	0.38	4.97	1.09	-1.58	0.70	5.18	
	Core science	4.04	1.49	-1.02	0.38	4.89	1.05	-1.13	0.49	5.30	
	Management	2.60	1.95	-0.79	0.91	4.67	0.25	-0.15	1.08	5.85	
	Acc Fin Eco	2.46	0.91	0.51	0.49	4.37	-0.19	-0.20	0.26	4.24	
	Humanities	3.45	0.90	-0.63	0.77	4.49	0.69	-0.31	0.43	5.30	
	Agriculture	4.63	1.31	-0.80	0.17	5.31	1.08	-1.29	0.50	5.60	
	Law	3.73	1.15	-0.18	0.80	5.50	0.52	-1.20	0.50	5.32	
	Education	0.91	0.52	0.44	1.18	3.05	0.42	-0.20	0.20	3.47	
	Medicine	3.12	1.81	-0.87	0.69	4.75	0.50	-1.33	0.55	4.47	
	Engineering	2.76	1.83	0.02	0.00	4.61	1.16	-0.73	1.02	6.06	
	Core science	3.57	2.48	-0.98	0.55	5.62	1.01	-1.33	0.35	5.65	
	Management	2.50	0.95	0.11	0.62	4.18	0.86	-0.55	0.97	5.46	
	Acc Fin Eco	2.77	0.74	0.79	0.42	4.72	0.60	-1.28	0.58	4.62	
Otago	Humanities	3.48	1.16	-0.34	0.59	4.89	0.62	-0.69	0.57	5.39	
	Agriculture	4.42	1.57	-1.19	0.64	5.44	0.75	-1.26	0.96	5.89	
	Law	4.10	0.68	-0.45	1.95	6.28	-0.36	-0.38	-0.14	5.40	
	Education	1.58	0.19	1.83	0.34	3.94	0.18	-0.64	0.42	3.90	
	Medicine	3.48	0.76	-0.47	0.18	3.95	0.71	-0.39	0.54	4.81	
	Engineering	4.26	0.72	-0.60	0.02	4.40	0.49	-0.58	0.44	4.75	
	Core science	4.79	0.27	0.03	-0.56	4.53	0.26	-0.43	0.05	4.41	
	Management	3.08	0.43	0.01	0.82	4.34	0.65	-0.39	0.40	5.00	
	Acc Fin Eco	2.98	0.86	-0.19	0.26	3.91	1.17	-0.59	0.20	4.69	
	Humanities	2.78	0.54	0.13	0.54	3.99	1.43	-0.51	0.79	5.70	
	Agriculture	4.59	0.70	-0.88	0.71	5.12	1.17	-1.78	1.03	5.54	
	Waikato	Medicine	4.23	1.25	-1.31	0.47	4.64	0.28	-0.75	0.56	4.73
		Engineering	4.64	1.34	-1.39	0.38	4.97	1.09	-1.58	0.70	5.18
		Core science	4.04	1.49	-1.02	0.38	4.89	1.05	-1.13	0.49	5.30
Management		2.60	1.95	-0.79	0.91	4.67	0.25	-0.15	1.08	5.85	
Acc Fin Eco		2.46	0.91	0.51	0.49	4.37	-0.19	-0.20	0.26	4.24	
Humanities		3.45	0.90	-0.63	0.77	4.49	0.69	-0.31	0.43	5.30	
Agriculture		4.63	1.31	-0.80	0.17	5.31	1.08	-1.29	0.50	5.60	
Law		3.73	1.15	-0.18	0.80	5.50	0.52	-1.20	0.50	5.32	
Education		0.91	0.52	0.44	1.18	3.05	0.42	-0.20	0.20	3.47	
Medicine		3.12	1.81	-0.87	0.69	4.75	0.50	-1.33	0.55	4.47	
Engineering		2.76	1.83	0.02	0.00	4.61	1.16	-0.73	1.02	6.06	
Core science		3.57	2.48	-0.98	0.55	5.62	1.01	-1.33	0.35	5.65	
Management		2.50	0.95	0.11	0.62	4.18	0.86	-0.55	0.97	5.46	
Acc Fin Eco		2.77	0.74	0.79	0.42	4.72	0.60	-1.28	0.58	4.62	



**Table 7** (continued)

	AQS 2003	Exits	Entrants	Trans	AQS 2012	Exits	Entrants	Trans	AQS 2018
Law	2.42	0.45	0.65	0.43	3.95	-0.12	-0.47	0.90	4.26
Education	1.85	1.61	0.01	0.62	4.09	0.10	-0.48	0.50	4.21
Medicine	3.04	3.44	-1.28	0.82	6.02	0.92	-1.93	0.24	5.25
Engineering	3.48	1.10	-0.45	0.55	4.68	0.21	-0.95	0.75	4.69
Core science	4.08	2.34	-0.62	0.74	6.54	0.01	-1.62	0.07	5.00
Management	2.94	1.42	-0.82	1.08	4.62	0.26	-0.29	0.04	4.63
Acc Fin Eco	2.62	1.47	0.09	0.48	4.66	0.02	-0.81	-0.22	3.65
Humanities	3.63	1.57	-0.03	0.55	5.72	0.44	-0.95	0.91	6.12
Agriculture	3.74	2.77	-1.53	1.30	6.28	0.69	-2.02	1.27	6.22
Law	3.05	2.23	-0.72	1.33	5.89	0.56	-0.60	0.67	6.52
Education	0.91	1.89	-0.06	1.07	3.81	0.11	-0.04	0.32	4.20

<sup>a</sup>There were no observations for Lincoln University in the 2012 and 2018 assessment rounds for Law and Education discipline groups

**Acknowledgements** We are grateful to the New Zealand Tertiary Education Commission (TEC) for providing the data, and to Ashley Ball for research assistance in compiling quality scores for universities and disciplines, supported by a VUW research grant.

**Funding** Open Access funding enabled and organized by CAUL and its Member Institutions.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Abramo, G., & D'Angelo, C. A. (2022). The different responses of universities to introduction of performance-based research funding. *Research Evaluation*, 30, 514–528.
- Abramovitz, M. (1986). Catching up, forging ahead, and falling behind. *Journal of Economic History*, 46, 385–406.
- Adams, J. (2008). *Strategic review of the performance-based research fund: The assessment process*. Report Prepared for the Tertiary Education Commission, Wellington.
- Adams, J., & Gurney, K. (2010). *Funding selectivity, concentration and excellence—How good is the UK's research?* Higher Education Research Institute.
- Barker, K. (2007). The UK Research Assessment Exercise: The evolution of a national research evaluation system. *Research Evaluation*, 16, 3–12.
- Barro, R. J., & Sala-i-Martin, X. (1991). Convergence. *Journal of Political Economy*, 100, 223–251.
- Baumol, W. J. (1986). Productivity growth, convergence, and welfare: What the long-run data show. *American Economic Review*, 75, 1072–1085.
- Boyle, G. (2008). Pay peanuts and get monkeys? Evidence from academia. *The B.E. Journal of Economic Analysis and Policy*. <https://doi.org/10.2202/1935-1682.1976>
- Broadbent, A. (2010). The UK Research Assessment Exercise: Performance measurement and resource allocation. *Australian Accounting Review*, 52, 14–23.
- Buckle, R. A., & Creedy, J. (2019a). The evolution of research quality in New Zealand universities as measured by the performance-based research fund process. *New Zealand Economic Papers*, 53, 144–165.
- Buckle, R. A., & Creedy, J. (2019b). An evaluation of metrics used by the performance-based research fund process in New Zealand. *New Zealand Economic Papers*, 53, 270–287.
- Buckle, R. A., & Creedy, J. (2020). The 'disciplinary effect' of the performance-based research fund process in New Zealand. *New Zealand Economic Papers*, 54, 107–126.
- Buckle, R. A., Creedy, J., & Gemmell, N. (2020). Is external research assessment associated with convergence or divergence of research quality across universities and disciplines? Evidence from the PBRF process in New Zealand. *Applied Economics*, 52, 3919–3932.
- Buckle, R.A., Creedy, J., & Ball, A. (2021). Fifteen years of a PBRFS in New Zealand: incentives and outcomes. *Australian Economic Review*, 54, 208–230.
- Campos, J., Ericsson, N. R., & Hendry, D. F. (Eds.). (2005). *General-to-specific modelling*. Edward Elgar Publishing.
- Castle, J. L., Doornik, J. A., & Hendry, D. F. (2011). Evaluating automatic model selection. *Journal of Time Series Econometrics*, 3, 1–31.
- Castle, J. L., & Hendry, D. F. (2014). 'Data mining' with more variables than observations. *VoxEU.org* (CEPR's Policy Portal), 13 August 2014.
- Checchi, D., Malgarini, M., & Sarlo, S. (2019). Do performance-based research funding systems affect research production and impact? *Higher Education Quarterly*, 73, 45–69.
- Checchi, D., Mazzotta, I., & Malgarini, M. (2020). Convergence or polarisation? The impact of research assessment exercises in the Italian case. *Scientometrics*, 124, 1439–1455.
- de Boer, H., Jongbloed, B., Bennenworth, P., Cremonini, L., Kolster, R., Kottmann, A., Lemmens-Krug, K., & Vossensteyn, H. (2015). *Performance-based funding and performance agreements in fourteen higher education systems*. Universiteit Twente.

- Dowrick, S., & Gemmell, N. (1991). Industrialisation, catching up and economic growth: A comparative study across the world's capitalist economies. *Economic Journal*, *101*, 263–275.
- Ehrenberg, R., Kasper, H., & Rees, D. (1991). Faculty turnover at American colleges and universities: Analyses of AAUP data. *Economics of Education Review*, *10*, 99–111.
- Hare, P. G. (2003). The United Kingdom's Research Assessment Exercise: Impact on institutions, departments, individuals. *Higher Education Management and Policy*, *15*, 43–62.
- Hendry, D. F., & Doornik, J. A. (2014). *Empirical model discovery and theory evaluation: Automatic model selection methods in econometrics*. MIT Press.
- Hicks, D. (2012). Performance-based university research funding schemes. *Research Policy*, *41*, 251–261.
- Kolarz, P., Dijkstal, E., Nielsen, A. F., & Farla, K. (2019). *International landscape study of research and innovation systems*. Technopolis Group.
- Martin, B. (2011). The research excellence framework and the impact agenda, are we creating a Frankenstein Monster? *Research Evaluation*, *20*, 247–254.
- New Zealand Tertiary Education Commission (2019) *Improving research quality: The results of the PBRF 2018 quality evaluation*. Tertiary Education Commission.
- OECD. (2010) *Performance-based funding for public research in tertiary education institutions: Workshop proceedings*. OECD.
- Payne, A. A., & Roberts, J. (2010). Government oversight of public universities: Are centralized performance schemes related to increased quantity or quality? *Review of Economics and Statistics*, *92*, 207–212.
- Quah, D. T. (1993). Galton's fallacy and the convergence hypothesis. *Scandinavian Journal of Economics*, *95*, 427–443.
- Shin, J. C., & Cummings, W. K. (2010). Multilevel analysis of academic publishing across disciplines: Research preference, collaboration, and time on research. *Scientometrics*, *85*, 581–594.
- Wilsdon, J., Allen, L., Belfiore, E., Campbell, P., Curry, S., Hill, S., Jones, R., Kain, R., Kerridge, S., Thelwall, M., Tinkler, J., & Johnson, B. (2015). The metric tide: Report of the independent review of the role of metrics in research assessment and management. *HEFCE*. <https://doi.org/10.13140/RG.2.1.4929.1363>
- Woelert, P., & McKenzie, L. (2018). Follow the money? How Australian universities replicate national performancebased funding mechanisms. *Research Evaluation*, *27*, 184–195.
- Xu, Y. J. (2008). Faculty turnover: Discipline-specific attention is warranted. *Research in Higher Education*, *49*, 40–61.