MINISTRY OF EDUCATION

Te Tāhuhu o te Mātauranga

Research par excellence

The factors associated with higher research quality in New Zealand tertiary education organisations

New Zealand Government

Research and knowledge creation

This report forms part of a series called Research and knowledge creation.

Author:
Warren Smart, Senior Research Analyst
Email: warren.smart@minedu.govt.nz
Telephone: 64-4-463-8035
Fax: 64-4-463-8713

Acknowledgements:
The author gratefully acknowledges comments provided by Roger Smyth (Ministry of Education), Graham Bussell (Ministry of Education), Nuran Çinlar (Tertiary Education Commission), Janet Grice (Tertiary Education Commission), Paula White (Tertiary Education Commission), Professor Charles Crothers (Auckland University of Technology) and Professor Gary Hawke (Victoria University of Wellington) on earlier drafts of this report.

The author also gratefully acknowledges the work of Sean Alexander (Tertiary Education Commission) in constructing the database used in this report.

The author also thanks Virginia Falealili who proofread the report.

All views expressed in this report, and any remaining errors or omissions, remain the responsibility of the author.

Published by:
Tertiary Sector Performance Analysis and Reporting
Strategy and System Performance
Ministry of Education

This report is available from the Ministry of Education's Education Counts website:
http://www.educationcounts.govt.nz

May 2008

ISBN 978-0-478-13853-5 (Print)
ISBN 978-0-478-13854-2 (Web)

Contents

List of tables

Glossary

Research output (RO) score
This is the score between 0 and 7 (with 7 representing the highest performance) awarded to staff by peer-review panels based on the quality of their research outputs.

Peer esteem (PE) score
This is the score between 0 and 7 (with 7 representing the highest performance) awarded to staff by peer-review panels based on the esteem in which they are held by their peers.

Contribution to the research environment (CRE) score
This is the score between 0 and 7 (with 7 representing the highest performance) awarded to staff by the peer-review panels based on their contribution to the research environment.

Weighted research score (WRS)
This is the weighted composite score of the RO, PE and CRE research scores. A weighting of 70/15/15 is applied to each of the scores, respectively, to arrive at the weighted research score. This score has a value between 0 and 700 (with 700 representing the highest performance). These scores are then used as a guide by the peer-review panels when assigning quality categories.

1 Summary

Key findings:

- The quality of research produced by staff is closely aligned with their seniority in the tertiary education sector, with staff of higher academic rank achieving higher research quality.

- Changes made to the Quality Evaluation assessment system to acknowledge the special circumstances of new and emerging researchers would appear to have had the desired effect – new and emerging researchers had an equal likelihood of attracting research funding for their tertiary education organisation compared with their more experienced colleagues.

- Some subject areas that require or engage in professional training and are relatively new research disciplines in the New Zealand tertiary education sector had research performance that was on average lower than other subject areas.

- Although a higher quantity of submitted research outputs was associated with higher research quality, it was not the key factor driving variation in research quality.

The Performance-Based Research Fund (PBRF) has created a strong financial incentive for tertiary education organisations to maximise the quality of their research. In 2007, approximately $130 million was distributed to tertiary education organisations via the Quality Evaluation. In addition, the scrutiny that is placed on the results when they are published also creates an incentive to maximise research quality.

This report used statistical modelling to identify the staff characteristics associated with higher research quality in the 2006 Quality Evaluation. There were two lines of enquiry used in this study. One used generalised ordered logistic regression to analyse the association between the characteristics of around 97 percent of staff participating in the 2006 Quality Evaluation and the likelihood of achieving a higher quality category. The second approach used ordinary least squares to analyse the research quality scores of around 97 percent of staff who had evidence portfolios examined by the peer review panels to identify the factors associated with achieving higher research quality scores.

With multiple factors likely to impact on the quality of research produced by these staff, regression analysis has the advantage of providing estimates of the average association of each characteristic with research quality while holding other confounding factors constant. This statistical modelling approach complements recent analyses of the 2006 Quality Evaluation by the Tertiary Education Commission and contributes to the evaluation of the PBRF.

Overall, the results of this new study confirmed the findings of an earlier study by Smart (2005) of the 2003 Quality Evaluation results. In both studies, the results of the regression modelling showed that the job title of staff was the strongest factor in explaining variation in research quality. Specifically, a higher academic rank was associated with higher research quality.

Changes made to the Quality Evaluation assessment system to acknowledge the special circumstances of new and emerging researchers would appear to have had the desired effect – new and emerging researchers had an equal likelihood of attracting research funding for their tertiary education organisation compared with their more experienced colleagues.

The results showed that a number of subject areas that have elements of professional training performed worse on average than other subject areas. Notable examples of this were ‘Education’ and ‘Nursing’. ‘Nursing’ and to some extent ‘Education’ are also relatively new research disciplines in the New Zealand tertiary education sector which may be a factor in their performance. The analysis showed that subjects in the social sciences/humanities fields or subjects that were identified as involving research that was potentially ‘risky and innovative’ by recent Tertiary Education Commission studies, generally performed better in average than the remaining subjects.

The results showed that although women had a similar likelihood of achieving a higher quality category as men, women who submitted evidence portfolios received slightly lower research output (RO) scores than men who submitted evidence portfolios. However, women who submitted evidence portfolios were given a slightly higher contribution to the research environment (CRE) score on average than men.

Analysis of the ethnic group of staff showed that Māori were less likely than Europeans to achieve more than an R and more than a C quality category. Interestingly, although Māori staff who submitted evidence portfolios received a lower RO score than Europeans, they received higher peer esteem (PE) scores than Europeans.

Not surprisingly, staff at the seven universities with a longer tradition of research achieved higher levels of research performance compared with staff at the remaining tertiary education organisations. The greatest advantage for staff at the older universities is in the CRE score, where the greater opportunity to supervise postgraduate research students may be a factor.

Larger subjects, in terms of full-time equivalent staffing at a tertiary education organisation, were associated with higher research performance. However, the analysis suggests that there are possibly diminishing returns to size. Also, a larger number of submitted research outputs in an evidence portfolio were associated with higher research performance. However it wasn’t the strongest factor affecting research quality reinforcing that the Quality Evaluation is measuring quality rather than quantity of research. The greatest advantage from submitting large numbers of research outputs was in CRE score.

A key limitation of this study is the lack of information on the quality/quantity of the teaching of staff participation in the PBRF Quality Evaluations. Without this information, the impact of the teaching role of researchers on the quality of the research they produce cannot be determined.

2 Introduction

The quality of research produced by staff involved in research in New Zealand's tertiary education organisations is now the biggest determinant of research funding allocated via Vote Education. In 2007, around $130 million was allocated to tertiary education organisations based on the results of the 2006 Performance-Based Research Fund (PBRF) Quality Evaluation. This creates a strong financial incentive for tertiary education organisations participating in the PBRF to maximise the quality of research produced by their staff.[1] In addition, the public scrutiny that follows the release of the results of the Quality Evaluations creates a further incentive for tertiary education organisations to maximise research quality.

This report applies multiple regression analysis to data from the 2006 PBRF Quality Evaluation to identify the demographic and employment-related characteristics of participating staff associated with higher research quality.[2] With multiple factors likely to impact on the quality of research produced by these staff, regression analysis has the advantage of providing estimates of the association of each characteristic with research quality while holding other confounding factors constant. This approach is similar to that used by Smart (2005) in analysing the 2003 Quality Evaluation results and complements recent analyses of the 2006 Quality Evaluation by the Tertiary Education Commission.[3] It also contributes to the evaluation of the PBRF.

One of the key themes examined in this report is whether the PBRF disadvantages certain groups of researchers. Since the PBRF was introduced in 2004, a number of groups of researchers have been identified as possibly being at a disadvantage from the peer assessment process used in the Quality Evaluations. These groups include: new and emerging researchers, women, Māori and Pasifika staff, researchers in professional training subject areas, researchers in risky and innovative subjects, and researchers in the social sciences/humanities.

The performance of these groups in the 2006 Quality Evaluation is analysed to see if their measured research quality is significantly lower than that of staff not considered as being at risk of disadvantage. For new and emerging researchers this analysis is especially pertinent, given that changes made to the 2006 Quality Evaluation were designed take into account their newness to the research environment.

This report has the following structure. Section 3 briefly outlines the process behind the allocation of quality scores and quality categories in the PBRF Quality Evaluation. Section 4 then discusses the dataset and methodology used in the analysis. Section 5 examines the factors associated with achieving a higher PBRF quality category and section 6 examines the factors associated with achieving a higher PBRF quality score. Some final conclusions are presented in section 7.

[1] Although the focus of this report is on the peer-assessed quality of research by New Zealand's tertiary education staff, it is important to acknowledge that research is just one of a number of important tasks performed by academic staff. Tasks such as teaching, acting as critic and conscience of society and community service are also important.
[2] It should be noted that the 2006 Quality Evaluation was a partial round. Staff had the option of submitting a new evidence portfolio or having their result from the 2003 Quality Evaluation carried over.
[3] See Çinlar and Dowse (2008a, 2008b and 2008c) and White and Grice (2008).

3 Measuring research quality[4]

Evidence portfolios prepared by staff were submitted to the Tertiary Education Commission by their tertiary education organisation if it was considered they would be assigned a quality category high enough to attract funding.[5] The submitted evidence portfolios provided evidence of their research performance across three dimensions: the quality of their research outputs (RO), the esteem they are held in by their peers (PE) and their contribution to the research environment (CRE). Examples of what can be provided as evidence of PE include being awarded research-related prizes and invitations to provide conference addresses. Examples of what can be provided as evidence of CRE includes the amount of external research income generated and the supervision of research students.

The expert peer-review panels allocated a score from 0 to 7 to each of the RO, PE and CRE categories after evaluating the content of the evidence portfolios, with 7 representing the highest performance and 0 the lowest. A 70/15/15 weighting was then applied to the three component scores and a weighted research score (WRS) between 0 and 700 calculated. The association of these research scores with the demographic and employment-related characteristics of participating staff is examined in Section 6.

For staff not identified as new and emerging a WRS between 600 and 700 was generally assigned an A quality category, a WRS between 400 and 599 was assigned a B, a WRS between 200 and 399 was assigned a C, and a WRS between 0 and 199 an R. The WRS was used only as a guideline by the peer review panels and the final quality categories were assigned on a holistic basis. Nevertheless, only around 1 percent of experienced staff who submitted evidence portfolios received quality categories different from that indicated by their WRS.

Although the WRS awarded to staff ranges between 0 and 700, for the purposes of the ordinary least squares regression analysis in section 6 the WRS is divided by 100, so that it has a comparable base to the three research component scores.

If a staff member was employed for the first time in an academic role after 1 Jan 2000 (either in NZ or overseas) or required to do degree teaching for the first time from this point, they could be nominated by their tertiary education organisation as a 'new and emerging' researcher. New and emerging staff were allocated quality categories using a slightly different approach. Like experienced researchers, new and emerging staff could be assigned an A or a B quality category. However, because new and emerging staff may not have had the opportunity to build up a significant portfolio of research, they could also be assigned a C(NE) or an R(NE) quality category. A C(NE) quality category did not require staff to submit evidence on the PE and CRE components, although they were encouraged to do so.

If new and emerging staff met certain reduced RO requirements they could still attract funding for their institution by being allocated a C(NE) quality category. It was expected that a minimum of two research outputs would be submitted in the evidence portfolio, in addition to a completed doctoral degree thesis.

Staff who received a C(NE) quality category attracted the same level of funding as experienced staff who received a C quality category. Under the 2003 Quality Evaluation assessment system, many of the new and emerging researchers who received a C(NE) quality category may well have received an R quality category.

---

4 The process used by the peer-review panels to allocate quality categories to staff participating in the 2006 Quality Evaluation is briefly outlined in this section. Those seeking a more detailed explanation should refer to Tertiary Education Commission (2007).

5 Where no evidence portfolio was submitted, PBRF-eligible staff were automatically allocated an R or R(NE) quality category.

The association between the demographic and employment related characteristics of staff and the likelihood of being allocated a higher quality category is examined in section 5.

It is important to note that the 2006 Quality Evaluation was a partial round where staff who also participated in the 2003 Quality Evaluation had the option of submitting a new evidence portfolio or having their result from the 2003 Quality Evaluation carried over. Around 35 percent of staff participating in the 2006 Quality Evaluation chose to have their quality category carried over and around 31 percent resubmitted evidence portfolios.

Those staff and tertiary education organisations that resubmitted evidence portfolios were likely to have learned from their experience in the 2003 Quality Evaluation and produced better quality evidence portfolios. Indeed, the Tertiary Education Commission (2007) noted that the peer review panels uniformly commented on the improved quality of presentation of evidence portfolios in the 2006 Quality Evaluation. As a result, the Tertiary Education Commission suggested that the results for these staff in the 2006 Quality Evaluation might be a more accurate reflection of their actual research performance. If this is the case, then this analysis of the 2006 Quality Evaluation results may provide a clearer picture of the factors associated with higher research quality than the previous study by Smart (2005) of the 2003 Quality Evaluation results.

4 Data and method

4.1 Data

The dataset used in this report was supplied by the Tertiary Education Commission following an application to access unit record data from the Performance-Based Research Fund (PBRF) Quality Evaluations.[6] The dataset contained unit record data of the quality scores/categories and demographic and employment-related information for staff who participated in the 2003 and 2006 Quality Evaluations.

4.2 Method

The theoretical framework for this analysis assumes that the quality of research produced by staff participating in the Quality Evaluation is influenced by personal and employment-related factors. Examples of personal factors include demographic characteristics such as age, gender and ethnic group. Examples of employment-related factors of staff include their subject discipline, type of institution employed at and their full-time equivalent status.

This is adapted from the framework used by Johnes *et al*. (1993) and Taylor (1995) who examined the factors associated with variation in performance among United Kingdom university departments in the Research Assessment Exercise.

In this study, PBRF Quality Evaluation data is available at the individual researcher level. Therefore, the quality of research (Q) produced by an individual staff member can be modelled as:

$$Q = f(A, E) \quad (1)$$

where A represents the personal characteristics of a researcher and E represents the employment-related characteristics of a researcher.

Because the research quality of PBRF-eligible staff is measured by a process that involves a mix of numerical scores as well as quality categories, two regression approaches are used to analyse the association of the various staff characteristics with research quality. Ordinary least squares regression is used to analyse the association of the personal and employment-related factors with the numerical research component scores (research output (RO), peer esteem (PE) and contribution to the research environment (CRE)) and generalised ordered logistic regression is used to analyse the association of those factors with the quality categories allocated to participating staff.

As the quality categories assigned to staff have an order to them (from A the highest to R the lowest), generalised ordered logistic regression is applied to examine the association between staff characteristics and the likelihood of staff achieving a higher quality category. For the purposes of this analysis, a C(NE) quality category is treated as a C quality category. Similarly, an R(NE) quality category is treated as an R quality category.

Generalised ordered logistic regression uses a maximum likelihood procedure to generate logit coefficient estimates that can then be expressed as odds ratios. The particular generalised ordered logistic regression model used here relaxes the parallel lines assumption and assumes that the impact of staff characteristics on the likelihood of being in a higher quality category can vary at different levels of research quality.[7]

---

[6] See http://www.tec.govt.nz/upload/downloads/pbrf_dataaccess_final.pdf for more information on the process to access information from the PBRF Quality Evaluations.

[7] For more detail on generalised ordered logistic regression see Williams (2006).

The generalised ordered logistic regression model can be written as:

$$P(Q_i > j) = g(X\beta_j) = \frac{\exp(\alpha_j + X_i\beta_j)}{1+[\exp(\alpha_j + X_i\beta_j)]}, \mathrm{j} = 1,2,...,\mathrm{M}-1 \tag{2}$$

where $P(Q_i > j)$ is the probability that individual i receives a quality category greater than category j and M is the number of categories of the ordinal dependent variable – in this case the number of quality categories.

When M > 2, the model becomes equivalent to a series of binary logistic regressions where the categories of the dependent variable are combined.[8] For example, in this analysis there are four quality categories (A, B C and R). Therefore for j = 1 the R quality category is contrasted with the C, B and A quality categories. For j = 2 the R and C quality categories are contrasted with the B and A quality categories. For j = 3 the R, C and B quality categories are contrasted with the A quality category.

The results of the generalised ordered logistic regression analysis are presented in odds ratio form. An odds ratio is the odds of an event happening divided by the odds of the opposite event happening. For example, suppose that 400 professors achieved a quality category that was higher than an R and 200 did not. The odds of a professor achieving a quality category higher than an R are 400/200 = 2, or 2 to 1. In other words, the chances of a professor being in a quality category higher than an R are reasonably good.

Suppose that 500 senior lecturers achieved a quality category that was higher than an R and 1,000 did not. The odds of a senior lecturer achieving a quality category higher than an R would be 500/1,000 = 0.5, or 1 to 2. The chances of them achieving a quality category higher than an R are therefore significantly lower than for professors.

To analyse the association between the demographic and employment-related characteristics of staff and their research scores (RO, PE, CRE and weighted research score (WRS)), ordinary least squares regression was applied. To make it easier to compare the results of the WRS score analysis with the other three component scores (RO, PE and CRE), the WRS is divided by 100 so that it ranges from 0 to 7, the same as the three component scores.

The dependent variable (Q) is the research score allocated to staff members that submitted evidence portfolios in the Quality Evaluation. This can be modelled as:

$$Q_i = \alpha_i + X_i\beta + \mu \tag{3}$$

where X represents the set of explanatory variables and μ is an error term.

Because staff who had their quality categories rolled over in the 2006 Quality Evaluation did not resubmit evidence portfolios, the scores they achieved in the 2003 Quality Evaluation have been used as a proxy for their performance in the 2006 Quality Evaluation. This of necessity assumes that the research performance of these staff has not changed between Quality Evaluations.

The specific demographic and employment-related characteristics of staff used as explanatory variables in the regression modelling are discussed in the rest of this section. A major limitation to this analysis is that no data is available on the quality or quantity of teaching undertaken by participating staff which is likely to be a key influence on their research performance.

8 This is an advance on the approach used by Smart (2005) where binary logistic regression was used to analyse the association between staff characteristics and the likelihood of achieving more than an R in the 2003 Quality Evaluation.

The age of a staff member is represented by a variable *AGE* which is the age of a participating staff member as at the 30 June 2006. Normally, one might expect the quality of research to be higher for older staff, given that staff gain greater experience over time. However, as the regression model includes explanatory variables that control for the academic rank of staff and whether the staff member was new and emerging, the relationship between age and research quality may be more complex. For example, younger staff in each academic rank may exhibit a higher level of research performance, given that older staff with greater research ability may have been promoted to a higher academic rank.

In addition, Christenson *et al*. (1997) found that the cognitive ability of academics in higher education diminished somewhat with age. Therefore, as age increases, research quality may well diminish. Also, the performance of staff aged over 65 might be inflated as more able researchers are retained in the tertiary education system while other less able researchers may retire. To capture the potential complexity of the relationship between age and research quality quadratic and cubic functional forms were trialled in the regression models.[9]

Previous studies have found that the gender of participating staff was associated with different levels of measured research quality in the 2003 PBRF Quality Evaluation. Smart (2005) found that women on average received a lower RO and WRS score than men in the 2003 Quality Evaluation after controlling for other factors (such as age and academic rank). Curtis and Phibbs (2006) also analysed the results of the 2003 Quality Evaluation and argued that systemic issues within the academic environment (such as slow promotion for women) disadvantaged the performance of women. Çinlar and Dowse (2008b) analysed the results of the 2006 Quality Evaluation and found that a higher proportion of women in a subject area was associated with lower research performance, but that given other confounding factors, the true underlying relationship was not clear.

In the regression models, the gender of a staff member is represented by a variable (*GENDER*) that takes a value of 1 if they are a woman and 0 if they are a man.

The ethnic group of a staff member may also impact on the quality category assigned to them. For example, Smith and Bruce Ferguson (2006) suggest that using the individual as the unit of assessment in the Quality Evaluation may disadvantage Māori staff as the evidence portfolio process requires an element of self-promotion, which they argue some Māori may find uncomfortable.

The ethnic group of the staff member in the regression models is represented by a variable (*ETHNIC_GROUP*) that has multiple categories. The categories are: 'European', 'Māori', 'Pasifika', 'Asian', 'Other' and 'Unknown'. The reference category in this analysis is European. Where staff reported multiple ethnicities in the staff census, ethnic group was assigned on a prioritised basis. The order of prioritisation was: Māori, Pasifika, Asian, Other and European.

Because a significant proportion of staff did not report an ethnic group in the PBRF Staff Census[10] they have been placed in the 'unknown' category in order to increase the coverage of the dataset used in the regression analysis.[11] However, this approach means that any results that show differences in performance between ethnic groups need to be viewed with caution.

A variable (*LOAD*) is included in the regression model that captures the full-time equivalent status of staff. Although this will not measure the exact amount of time devoted to research by staff that have a mix of teaching and research roles, generally staff with a higher full-time equivalent status should have

[9] For the quadratic functional form the square of the age of staff (AGE^2) is included as a separate variable in the model. For the cubic functional form, the square (AGE^2) and cube (AGE^3) of the age of staff are included as separate variables in the model.
[10] Around 13 percent of staff did not indicate an ethnic group in the staff census.
[11] This is the approach suggested by Fitzgerald and Knuiman (1998) when dealing with missing values for categorical explanatory variables.

greater opportunity to produce research and/or supervise research students and so would be expected to achieve a higher level of research quality.

One of the key changes to the operation of the Quality Evaluation in the 2006 round involved the classification of staff who were identified as new and emerging. It is expected that staff new to the research environment are likely to have a lower level of research quality compared with more experienced researchers because they may still be developing a research programme and not yet have the visibility in their disciplines to accrue peer esteem or make strong contributions to the research environment. A variable is included in the regression model *(EXPERIENCE)* that identifies whether the staff member was a new and emerging researcher. This variable takes a value of 1 if the staff member was 'new and emerging', otherwise 0.[12]

One of the key employment-related variables in the regression models relates to the academic title of participating staff. To achieve promotion within the academic environment and move to a higher position, staff generally have to demonstrate in some form that they have produced a certain standard of performance across a number of dimensions, one of those being research.[13] For example, it would be expected that a professor would have produced research that is of higher quality than staff of lower academic rank.

A number of previous studies have examined the link between academic title and peer-reviewed research performance. Johnes *et al*. (1993) used the proportion of professors in university departments as an explanatory variable in their regression analysis of Research Assessment Exercise scores but found this was not a significant factor. Jayasinghe *et al*. (2001) examined the impact of academic title on the likelihood of success of research grant applications to the Australian Research Council and found that professors were significantly more likely than academics of lesser rank to have successful applications.

For the purposes of this analysis, the academic title of a staff member is represented by a variable (*POSITION*) with multiple categories. The categories used in this analysis are based on broad categories as defined by the Tertiary Education Commission: academic leader, professor, associate professor, senior lecturer, lecturer, assistant lecturer and 'other staff'.[14] The reference category used in the analysis was senior lecturer, the largest group.

The performance of staff may also vary depending on the subject discipline they are associated with. Çinlar and Dowse (2008b) identified three groupings of subjects that *may* potentially be disadvantaged by the Quality Evaluation process. These groupings were: subjects in the social sciences/humanities,[15] subjects that require or engage in professional training[16] (which may impose on time to do research) and subjects that are likely to include higher proportions of 'risky and innovative' research.[17] The individual subjects within each of these groups and the subjects identified as not being at possible disadvantage are listed in Table 1.

---

[12] As the classification of new and emerging researchers was introduced for the 2006 Quality Evaluation, the statistical modelling used by Smart (2005) to analyse the 2003 Quality Evaluation results did not include this variable.
[13] Although caution is advised regarding the usefulness of analysing research performance by the academic rank of staff (see Çinlar and Dowse 2008b), previous statistical studies of peer-reviewed research performance have included this variable in the analysis (see Johnes *et al*, 1993; Jayasinghe *et al*, 2001)
[14] For more detail on how these aggregated positions were determined see Çinlar and Dowse (2008a).
[15] Research in the social sciences/humanities generally has more of a national or local focus than other subject areas and hence may have less potential outlets for publication that are internationally regarded (Çinlar and Dowse, 2008b). For more discussion on this and other concerns for subjects in the social sciences/humanities see Çinlar and Dowse (2008b).
[16] Middleton (2006) also argued that subjects with professional practice requirements, such as education, are at a disadvantage in the Quality Evaluations as the time taken to remain current with professional requirements could reduce the time available to produce research outputs.
[17] For example, intellectual property requirements may impinge on the ability of staff in these areas to produce research outputs (Çinlar and Dowse, 2008b). For more discussion on this and other concerns for risky and innovative research see Çinlar and Dowse (2008b).

Table 1: Subjects identified as having the potential to be disadvantaged by the Quality Evaluation process

Disadvantage less likely	Potentially disadvantaged		
	Professional training	Social sciences/humanities	Risky and innovative
– Accounting and finance – Chemistry – Earth sciences – Ecology, evolution and behaviour – Economics – Law – Management, human resources, industrial relations, international business and other business – Marketing and tourism – Molecular, cellular and whole organism technology – Other health studies (including rehabilitation therapies) – Pharmacy – Physics – Public health – Pure and applied mathematics – Statistics	– Architecture, design, planning, surveying – Clinical medicine – Dentistry – Education – Nursing – Veterinary studies and large animal science	– Anthropology and archaeology – Communications, journalism and media studies – English language and literature; – Foreign languages and linguistics – History, history of art, classics and curatorial studies – Human geography – Māori knowledge and development – Philosophy – Political science, international relations, and public policy – Psychology – Religious studies and theology – Sociology, social policy, social work, criminology and gender studies	– Agriculture and other applied biological sciences – Architecture, design, planning, surveying – Biomedical – Computer science, information technology, information sciences – Design – Engineering and technology – Music, literary arts and other arts – Theatre and dance, film, television and multimedia – Visual arts and crafts

Source: Çinlar and Dowse (2008b)

In the regression models, a variable with multiple categories (*SUBJECT*) captures the subject area of a staff member. Each of the 42 narrow subject areas in the Quality Evaluation is assigned a separate category with the reference category being 'Marketing and tourism'. This is a subject not identified as being potentially disadvantaged by the Quality Evaluation process and had an average quality score (2.84) that was closest of this group to the overall average score (2.96) achieved in the 2006 Quality Evaluation.

The type of tertiary institution a staff member is employed by may influence the research performance of staff. Institutions with a long history of research are likely to be able to provide a better research culture and environment than institutions that may be new to research. Johnes *et al*. (1993) and Taylor (1995) found that departments in older well-established United Kingdom universities achieved higher Research Assessment Exercise scores.

In New Zealand, the seven universities with a longer tradition of research (Auckland, Waikato, Massey, Victoria, Canterbury, Lincoln and Otago) would be expected to have a relatively well-developed research culture and capability, compared with other tertiary education organisations where the move to produce research is more recent.[18] A variable (*TEO_TYPE*) is used in the regression models to capture this effect on research quality. The variable takes a value of 1 if the researcher was employed at one of the following universities: Auckland, Massey, Victoria, Canterbury, Lincoln, Waikato and Otago.[19] If staff were employed by other tertiary education organisations, the variable took a value of 0.

The number of research staff within a subject area has been found to be a significant factor in explaining variations in the research quality in tertiary institutions. Johnes *et al*. (1993) and Taylor (1995) found that the assessed quality of research increased with the number of staff in the academic departments of universities in the Research Assessment Exercise. Taylor (1995) suggested there may

[18] It is also possible that staff at tertiary education organisations new to research may have higher teaching hours as a residual effect from the pre-research environment which may negatively impact on their research performance.
[19] Although Auckland University of Technology is a university, it was previously a polytechnic and was only granted university status in 2000. As such, it is still at a stage of growing its research capacity.

be economies of scale associated with larger numbers of research staff and that larger departments may be more noticeable.

In their analysis of the 2006 PBRF Quality Evaluation results, Çinlar and Dowse (2008b) found a slight negative association between subject size and the average research score of that subject area in the 2006 Quality Evaluation.

In this analysis, a variable (*SIZE*) is included in the regression models to measure the association between the size of a subject within each tertiary education organisation and the quality of research produced by individual staff.[20] This is the size (measured in full-time equivalent staff) of the PBRF subject area within each nominated academic unit (ie school or department) by a tertiary education organisation.

Taylor (1995) identified a quadratic functional form as fitting the data best. In other words the quality of research may eventually decline as the size of the academic unit offering the subject increases. Therefore, a quadratic functional form is also used in this analysis of PBRF performance. This is achieved by including a variable in the regression model ($SIZE^2$) which is the square of *SIZE*.

One of the explanatory variables can only be used in the ordinary least squares regression analysis of research component scores. This variable (*OUTPUT*) measures the number of research outputs[21] submitted by a participating staff member in their evidence portfolios and is therefore not available for those staff that did not submit evidence portfolios.[22] Although this does not capture all research output produced by staff during the assessment window − they were encouraged to submit only outputs that they felt would add to their evidence portfolio − it nevertheless provides a proxy for the amount of research produced during the assessment window that the staff member determined was of sufficient quality to merit inclusion.

OUTPUT is generated by first converting the number of submitted research outputs by staff to a 1 full-time equivalent basis.[23] Then the output figure is allocated to four categories: fewer than 10 outputs, 10 to fewer than 20 outputs, 20 to fewer than 30 outputs, and 30 or more outputs. A categorical variable is used in this case as staff who had their scores carried over from the previous Quality Evaluation do not have information on the number of research outputs in the 2006 Quality Evaluation window of assessment. The number of outputs submitted in 2003 is used as a proxy value for the number of outputs in the 2006 assessment window. As staff could include up to a maximum of 50 outputs in 2003 but only 34 in 2006, the research outputs have been converted to categorical values. The reference category in this analysis is fewer than 10 research outputs.[24]

[20] This is an advance on the approach by Smart (2005) which did not include a variable to capture the size of a subject at a tertiary education organisation.
[21] Note that this analysis does not weight the different types of research output.
[22] Taylor (1995) examined the impact of the number of publications on research performance of United Kingdom university departments in the Research Assessment Exercise.
[23] Taylor (1995) used a similar process in normalising the output of academic departments in his analysis of factors associated with the Research Assessment Exercise.
[24] Smart (2005) did not include a variable to capture the effect of the number of research outputs submitted to the peer review panels on research quality.

5 The factors associated with achieving a higher quality category[25]

Once records with missing information on the age of staff were removed from the dataset, a total of 8,410 staff out of 8,701 participating in the 2006 Quality Evaluation remained in the generalised ordered logistic regression analysis.[26] The results of the generalised ordered logistic regression analysis[27] are presented in Table 2 and represent three scenarios: the likelihood of a staff member achieving more than an R, the likelihood of the staff member achieving more than a C, and the likelihood of the staff member achieving more than a B.[28]

For the continuous explanatory variables in Table 2 (such as *LOAD*), an odds ratio of greater than 1 indicates a higher likelihood of being in a higher quality category as the value of the explanatory variable increases and a value less than 1 indicates a lower likelihood.

For categorical explanatory variables in Table 2 (such as *POSITION*), the odds ratio compares the likelihood of being in a higher quality category compared with the reference category. A value of greater than 1 indicates a higher likelihood of being in a higher quality category compared with the reference group, while a value of less than 1 indicates a lower likelihood.

The statistical significance of the explanatory variables in Table 2 is denoted by a symbol after the odds ratio value. If the odds ratio is followed by an asterisk (*) then the variable is significant at the 5 percent level. If the odds ratio is followed by a two asterisks (**) then the variable is significant at the 1 percent level and if the odds ratio has no asterisk after it then the variable is not statistically significant. For categorical explanatory variables, the statistical significance relates to the difference in likelihood between the category of interest and the reference category.

The odds ratios for *AGE*, AGE^2 and AGE^3 in Table 2 suggest that the association between age and research quality is captured by a cubic functional form. In other words, the likelihood of being in a higher quality category initially increases with age, then decreases, and finally increases again from around the age of 65.[29] This is likely to reflect that more able staff may have been promoted into a higher academic rank, leaving lower performing staff in that academic rank in the older age groups. It may also reflect other factors such as an increased teaching or administrative load. The rise in likelihood of being in a higher quality category for staff aged over 65 is likely to result from the impact of retirements of staff, with more able researchers being retained in the workforce for longer.

Although the odds ratios for the *GENDER* variable in Table 2 suggest that the odds of a woman achieving a higher quality category are slightly lower than the odds of a man, the result is not statistically significant. This implies that women were not disadvantaged in terms of being assigned quality categories once other factors (like academic rank and subject discipline) are controlled for.[30]

There were differences in the likelihood of achieving a higher quality category by ethnic group. Māori staff had a lower likelihood than European staff of achieving more than an R and more than a C quality category. There was no difference in the likelihood of European and Māori staff achieving more than a B quality category. However, it is difficult to be definitive about the impact of ethnic group, given the significant proportion of staff in the unknown category. Hence this finding should be viewed with caution.

---

[25] Summary statistics for the generalised ordered logistic regression model are presented in Appendix A.

[26] Or 97 percent of total participating staff.

[27] A number of models were estimated using a variety of combinations of explanatory variables. The final model presented here is the one that provided the best fit of the data. However, where the results varied significantly to other models these are noted in the text.

[28] The pseudo R^2 value of 0.37 indicates the regression model has a reasonable explanatory power for a cross-sectional study of this nature.

[29] This result is sensitive to the specification of the model. If the *POSITION* variable is excluded from the model, then the nature of the relationship changes and older staff have a higher likelihood of achieving a higher quality category. Without the *POSITION* variable in the model, *AGE* will be capturing the impact of experience and ability on research quality.

[30] This result is sensitive to the specification of the model. If the *POSITION* variable is excluded, then women are less likely than men to be assigned a higher quality category.

The odds ratio value of 1.9857 for the *LOAD* variable indicates that a higher full-time equivalent status resulted in a higher likelihood of achieving more than an R quality category. However, full-time equivalent status was not a statistically significant factor in achieving more than a C or more than a B quality category.

The odds ratio for *EXPERIENCE* indicates that there was no statistically significant difference in new and emerging and experienced researchers achieving more than an R quality category. This suggests that the changes made in the 2006 Quality Evaluation have removed the disadvantage faced by new and emerging researchers in terms of attracting PBRF funding. However, as would be expected, the odds of a new and emerging researcher achieving more than a C and more than a B quality category are significantly less than those of an experienced researcher.[31]

The academic title of staff was one of the key factors associated with variation in research quality in the Quality Evaluation. The odds ratios for the categories of the *POSITION* variable show that staff of higher academic rank had a greater likelihood of being allocated a higher quality category. For example, the odds of a professor achieving more than a B quality category were around 97 times higher than those of a senior lecturer (the reference category). Lecturers, assistant lecturers and ‘other staff’ were less likely than senior lecturers to be in a higher quality category.

The results in Table 2 suggest there is considerable variation across subjects in the likelihood of being in a higher quality category. Subjects where staff exhibited a lower likelihood of being in a higher quality category compared with the reference category of ‘Marketing and tourism’ were: ‘Accounting and finance’, ‘Clinical medicine’, ‘Education’, ‘Law’, ‘Nursing’, and ‘Other health studies’.[32] Three of these subject areas (‘Clinical medicine’, ‘Education’ and ‘Nursing’) have professional training requirements.[33] In addition, ‘Nursing’ and to a lesser extent ‘Education’ are relatively new research disciplines in the New Zealand tertiary education sector which could also contribute to their relatively lower performance.[34]

There were a number of subjects that exhibited a greater likelihood of staff being assigned to a higher quality category, compared with staff in ‘Marketing and tourism’. Some of the best performing subjects in this regard were: ‘Philosophy’, ‘Visual arts and crafts’, ‘Earth sciences’ and ‘Psychology’. Interestingly, three of these four subjects were identified in Table 1 as being at a possible disadvantage in the Quality Evaluation process. This concurs with Çinlar and Dowse (2008b) who found there was no indication that subjects in the social sciences or risky and innovative subject areas performed any worse than the others in the Quality Evaluation and in some cases they performed better than average.

The strong performance of ‘Visual arts and crafts’ indicated in Table 2 illustrates the effect of using regression analysis to hold other factors constant. The actual proportion of staff in the ‘Visual arts and crafts’ area achieving more than an R was 59 percent, compared with 71 percent for staff in the ‘Marketing and tourism’ area. However, the ‘Visual arts and crafts’ area contains few staff of senior academic rank and has a large proportion of staff that are employed at tertiary education organisations other than the 7 older universities. Once these and other factors are controlled for, the performance of staff in the ‘Visual arts and crafts’ improves. For example, the odds of a staff member in this subject area achieving more than an R quality category are three times those of a staff member in ‘Marketing and tourism’, after controlling for other factors.

31 In assigning A and B quality categories experienced and new and emerging researchers were assessed on an equal basis.
32 Specifically, the results in Table 2 show that staff in the ‘Accounting and finance’, ‘Clinical medicine’ and ‘Nursing’ subject areas had a lower likelihood of achieving more than an R and achieving more than a C quality category, compared with staff in ‘Marketing and tourism’. Staff in the ‘Education’, ‘Law’ and ‘Other health’ subject areas had a lower likelihood of achieving more than an R quality category than staff in the ‘Marketing and tourism’ area.
33 The government has already acted in some of these areas by allocating additional funding to nurture researchers in strategically relevant areas - which include the nursing and rehabilitation subject areas. Also, as the colleges of education have been absorbed into the universities, it would be expected that the quality of research in the education area may improve due to the more established research culture that exists at the universities.
34 The Tertiary Education Commission (2007) identified ‘Education’ and ‘Nursing’ as being relatively new disciplines to the New Zealand tertiary education sector.

Table 2: Generalised ordered logistic regression results
Dependent variable = allocated quality category (A, B, C, R)

Explanatory variables	Categories (where applicable)	Odds ratios (likelihood of achieving more than indicated quality category) R	C	B
AGE		1.8671**	2.0360**	3.4498*
AGE²		0.9864**	0.9850**	0.9737*
AGE³		1.0001**	1.0001**	1.0002***
GENDER	Male		Reference category	
	Female	0.98636	0.9097	0.8928
ETHNIC_GROUP	European		Reference category	
	Māori	0.6082**	0.5342**	1.3056
	Pasifika	0.9016	0.7165	3.2044
	Asian	0.9793	0.8779	1.0496
	Other	1.1195	0.9591	1.0461
	Unknown	1.2787	1.2099	1.3700*
LOAD		1.9857**	0.9914	0.8633
EXPERIENCE	Experienced		Reference category	
	New and emerging	0.8778	0.3430**	0.2686*
POSITION	Other staff	0.2317**	0.2730**	2.7012*
	Assistant lecturer	0.2173**	0.1262**	1.0505
	Lecturer	0.1978**	0.1495**	0.2515**
	Senior lecturer		Reference category	
	Associate professor	11.6764**	11.2386**	12.6804**
	Professor	22.3759**	36.5568**	97.0338**
	Academic leader	1.0996	2.0666	21.8136**
SUBJECT	Accounting & finance	0.2861**	0.3909**	0.8520
	Agriculture & other applied biological sciences	1.1175	1.1972	1.1767
	Anthropology & archaeology	1.9957	3.0337**	2.4699
	Architecture, design, planning, surveying	0.8825	1.5267	1.6807
	Biomedical	2.2434*	3.0285**	3.4943**
	Chemistry	1.3103	2.1113*	3.2927**
	Clinical medicine	0.3066**	0.4706*	0.8482
	Communications, journalism & media studies	0.9036	1.9785*	0.3789
	Computer science, IT, information sciences	1.0398	1.5633	2.2705
	Dentistry	0.7042	0.4951	3.2712
	Design	0.4977	0.7089	1.4911
	Earth sciences	3.4707**	4.7778**	2.0018
	Ecology, evolution & behaviour	3.5621**	3.4296**	3.8553**
	Economics	0.9535	1.6166	0.6432
	Education	0.4131**	1.2151	2.5176*
	Engineering & technology	1.5226	2.4878**	5.9009**
	English language & literature	1.9173*	1.5674	3.7988*
	Foreign Languages & Linguistics	0.5729*	1.1070	3.8282*
	History, history of art, classics & curatorial studies	2.6858**	2.8171**	2.0605
	Human geography	3.0945*	2.3138*	2.5769
	Law	0.3853**	1.4029	1.1408
	Management, human resources, industrial relations, international business & other business	0.7128	0.8121	0.9339
	Māori knowledge & development	0.8822	2.8393**	0.7519
	Marketing & tourism		Reference category	
	Molecular, cellular & whole organism biology	1.1978	3.0190**	2.4778*
	Music, literary arts & other arts	2.0714**	4.8615**	3.3206*
	Nursing	0.1490**	0.1471**	0.3494
	Other health studies (including rehabilitation therapies)	0.5518**	1.2182	2.0972

Table 2: Generalised ordered logistic regression results (continued)

Explanatory variables	Categories (where applicable)	Odds ratios – (likelihood of achieving more than indicated quality category) R	C	B
SUBJECT cont…	Pharmacy	0.5621	6.0063**	2.2715
	Philosophy	2.4282	5.2628**	11.1042**
	Physics	2.7221*	3.5806**	3.2029*
	Political science, international relations & public policy	1.4448	2.3304*	6.2190**
	Psychology	0.9574	2.3073**	7.7600**
	Public health	1.8351*	1.3074	3.0861*
	Pure and applied mathematics	0.8445	2.0264*	5.6444**
	Religious studies & theology	1.3665	4.7749**	1.2516
	Sociology, social policy, social work, criminology & gender studies	1.0393	1.4765	2.2590
	Sport & exercise science	0.7861	0.6875	0.6583
	Statistics	1.1533	1.3439	3.5886*
	Theatre & dance, film, television & multimedia	1.3537	3.4205**	14.5631**
	Veterinary studies & large animal science	0.5227	0.6928	2.3549
	Visual arts & crafts	3.0815**	3.4737**	6.0591**
TEO_TYPE	University (excluding Auckland University of Technology)	8.2996**	8.4854**	16.4631**
	Other tertiary education organisation	Reference category		
SIZE		1.0251**	1.0295**	1.0158
$SIZE^2$		0.9997**	0.9996**	0.9997*
Log likelihood	-6,770			
Pseudo R^2	0.37			
N	8,410			

Notes:
1. **, * indicates significant at the 1 percent and 5 percent level, respectively.
2. Robust standard errors were used to indicate statistical significance.
3. A C(NE) is treated as a C quality category in this analysis. Similarly, an R(NE) is treated as an R quality category.

The *TEO_TYPE* variable has an odds ratio greater than 1 and is statistically significant for all three scenarios. This indicates that being employed in an institution with a long history of research and well-established research culture resulted in a higher likelihood of achieving a higher quality category. For example, the odds of a staff member from one of the seven older universities achieving more than a B was around 16 times the odds of a staff member from one of the other tertiary education organisations in the 2006 Quality Evaluation.

The size of a subject area (in terms of full-time equivalent staff) at a tertiary education organisation had a positive relationship with research quality in all three scenarios, but the odds ratio of less than 1 on the square of the *SIZE* variable suggests that there may be diseconomies of scale. In other words, there may be an optimal size for a subject area within a tertiary education organisation.

6 The factors associated with achieving higher research quality component scores[35]

Having examined the factors that are associated with achieving a higher quality category in section 5, the focus in this section is on the factors associated with achieving higher research component quality scores. This allows for the impact of the demographic and employment-related characteristics of staff on each of the component scores (research output (RO), peer esteem (PE) and contribution to the research environment (CRE)) to be assessed separately. As previously mentioned in section 4, the WRS score used in this analysis has been divided by 100 so that it ranged from 0 to 7, the same as the other three research component scores. This makes comparison of the results in Table 3 easier.

It is important to remember that the dataset used in this section is restricted to those staff that had evidence portfolios assessed by the peer review panels. Most of the staff that received an R or R(NE) quality category are unlikely to have had their evidence portfolios examined by the peer-review panels so the results in this section should not be seen as representative of all staff participating in the 2006 Quality Evaluation.

It should also be remembered that for staff that had their quality category rolled over from the 2003 Quality Evaluation, the research component scores they received in the earlier evaluation have been used as a proxy for their performance in the 2006 Quality Evaluation.

A total of 6,801 staff in the 2006 Quality Evaluation submitted evidence portfolios to a peer review panel in either 2003 or 2006. After removing records that contained missing values for age, 6,612 records remained in the analysis.[36]

The results of the ordinary least squares regression analysis of the preferred model[37] are presented in Table 3.[38] The table includes the regression coefficients for the analysis of the three research component scores and the composite weighted research score (WRS). For the continuous variables, *AGE, LOAD* and *SIZE* the coefficients can be interpreted as the change in score associated with a 1 unit increase in the independent variable, after controlling for other factors. For categorical variables (such as *GENDER*), the coefficient indicates how the research score varies from the reference category, controlling for other factors.

The statistical significance of the explanatory variables in Table 3 is denoted by a symbol after the coefficient value. If the coefficient is followed by an asterisk (*) then the variable is significant at the 5 percent level. If the coefficient is followed by two asterisks (**) then the variable is significant at the 1 percent level and if the coefficient has no symbol after it then the variable is not statistically significant. For categorical explanatory variables, the statistical significance relates to the difference in average research component score between the category of interest and the reference category.

The regression coefficients for *AGE*, AGE^2 and AGE^3 are statistically significant for all of the research component scores. The positive sign for *AGE* and AGE^3 and the negative sign for AGE^2 suggests that the research scores of staff initially rise with age, then fall and then rise again. This mirrors the relationship between age and the likelihood of receiving a higher quality category described in section 5 and reflects the impact of some staff being promoted and the effect of some staff retiring from age 65 onwards.

---

35 Summary statistics for the dataset used in the ordinary least squares analysis are presented in Appendix B.
36 Or 97 percent of panel-assessed staff.
37 A number of models were estimated using a variety of combinations of explanatory variables. The final model presented here is the one that provided the best fit of the data. However, where the results varied significantly from other models this is noted in the text.
38 The explanatory power of the ordinary least squares regression models was reasonable for a cross-sectional analysis, with between 59 to 66 percent of the variation in research scores explained by the variation in the explanatory variables.

The ordinary least squares regression results showed there were differences in some of the research scores achieved by men and women. On average, women received an RO score that was 0.14 points lower than men, controlling for other factors. However, women achieved a CRE score 0.06 points higher than men.[39] Overall, women received a slightly lower weighted research score (WRS) than men.[40] The small scale of this difference is reflected in the finding in section 5 that gender did not have a statistically significant effect on the likelihood of staff being in a higher quality category.

The research scores of staff exhibited some variation among ethnic groups. The coefficients for the categories of the *ETHNIC_GROUP* variable show that Māori staff received a lower RO score on average than Europeans but received a higher PE score.[41] Also, Asian staff received a slightly lower PE and CRE score than Europeans.[42]

The coefficients for *LOAD* show that the full-time equivalent status of staff had a positive association with research performance. The strongest relationship was with CRE score, followed by the PE score and RO score.[43] An increase of 0.1 in full-time equivalent status was associated with an increase of 0.08 in CRE score, an increase of 0.06 in PE score and an increase of 0.04 in RO score. The greater impact on CRE score is not surprising, given that a higher full-time equivalent status may allow for more time to be devoted to activities such as supervision of postgraduate students.

The ordinary least squares regression results show that new and emerging researchers received lower RO, PE and CRE scores than experienced researchers. The biggest disadvantage for new and emerging researchers was in RO score, where they received a score 0.43 points lower on average than experienced researchers. This compares with PE and CRE scores that were 0.33 and 0.31 points lower, respectively.

The results in Table 3 show that researchers of higher academic rank received higher component scores.[44] For example, professors received an RO score 1.47 points higher than senior lecturers, controlling for other factors. Similarly, lecturers received an RO score 0.53 points lower than senior lecturers. The biggest advantage for those with a higher academic rank was in PE score, where the score for a professor was 1.91 points higher than for a senior lecturer. The *POSITION* variable explained about 10 percent of the variation in the research component scores, the highest explanatory power of any of the independent variables in the model.

One of the advantages of examining the impact of the explanatory variables on each research component score separately is that it can help to identify the exact area which is problematic for staff in lower-performing subject areas. For example, staff in the 'Nursing' area perform best on the CRE score in that there is no statistically significant difference between their score and staff in the 'Marketing and tourism' subject area. However, the RO score of staff in the 'Nursing' area was 0.83 points lower than that achieved by staff in 'Marketing and tourism'.

'Education' was another of the lower performing subject areas in the 2006 Quality Evaluation. The results in Table 3 show that staff in 'Education' received a PE score that was 0.26 points lower than the reference subject 'Marketing and tourism'. However, there was no statistically significant difference in RO and CRE scores.

---

39 The analysis of the 2003 Quality Evaluation results by Smart (2005) found that women received a lower RO score than men, but did not receive a higher CRE score.

40 This result is sensitive to the inclusion of the *POSITION* variable in the model. If this variable is excluded women have a significantly lower score than men across all four of the measures in Table 3.

41 This compares with Smart (2005) who found Māori staff received a higher PE score than European staff but received a similar RO score.

42 As was mentioned in the previous section, because of the high number of staff in the 'unknown' category, these results should be viewed with caution.

43 This mirrors the findings of Smart (2005) in his analysis of the 2003 Quality Evaluation results.

44 Smart (2005) found a similar relationship between academic rank and research quality in his analysis of the 2003 Quality Evaluation results.

Table 3: Ordinary least squares regression results
Dependent variable = research score RO (0-7), PE (0-7), CRE (0-7) and WRS (0-7)

Explanatory variables	Categories (where applicable)	Coefficients			
		RO	PE	CRE	WRS
AGE		0.2957**	0.2611**	0.2547**	0.2813**
AGE^2		-0.0066**	-0.0056**	-0.0049**	-0.0061**
AGE^3		0.0001**	< 0.0001**	< 0.0001**	< 0.0001**
GENDER	Male	Reference category			
	Female	-0.1399**	0.0039	0.0608*	-0.0895**
ETHNIC_GROUP	European	Reference category			
	Māori	-0.1733*	0.2056**	0.1543	-0.0756
	Pasifika	-0.2048	-0.1068	-0.1002	-0.1721
	Asian	0.0249	-0.1124*	-0.1889**	-0.0290
	Other	0.0334	-0.0121	-0.0396	0.0184
	Unknown	0.1142**	0.0991*	-0.0016	0.0935**
LOAD		0.4409**	0.6192**	0.8216**	0.5137**
EXPERIENCE	Experienced	Reference category			
	New and emerging	-0.4395**	-0.3314**	-0.3108**	-0.3554**
POSITION	Other staff	-0.4003**	-0.4447**	-0.4609**	-0.4158**
	Assistant lecturer	-0.4932**	-0.6482**	-0.7845**	-0.5486**
	Lecturer	-0.5275**	-0.6300**	-0.6366**	-0.5619**
	Senior lecturer	Reference category			
	Associate professor	0.8341**	1.0012**	0.9510**	0.8793**
	Professor	1.4714**	1.9100**	1.8209**	1.5930**
	Academic leader	0.3718**	0.5678**	0.5619**	0.4336**
SUBJECT	Accounting & finance	-0.1073	-0.2889*	-0.2118	-0.14.8
	Agriculture & other applied biological sciences	-0.0845	0.0523	0.4819**	0.0212
	Anthropology & archaeology	0.5737**	0.6687**	0.8879**	0.6355**
	Architecture, design, planning, surveying	0.4039**	0.1438	0.1710	0.3539**
	Biomedical	0.5917**	0.3572**	0.7954**	0.5857**
	Chemistry	0.4642**	0.4615**	0.7211**	0.5052**
	Clinical medicine	-0.1381	-0.1718	-0.0328	-0.1306
	Communications, journalism & media studies	0.5531**	0.2481	0.3765**	0.4748**
	Computer science, IT, information sciences	0.3991**	0.1678	0.3197**	0.3638**
	Dentistry	0.3891	0.2645	0.4418*	0.3736
	Design	0.0636	-0.1220	-0.1860	0.0071
	Earth sciences	0.4917**	0.4753**	0.9313**	0.5550**
	Ecology, evolution & behaviour	0.4952**	0.4845**	0.7972**	0.5398**
	Economics	0.3434**	0.1117	0.1694	0.2984**
	Education	-0.0307	-0.2554*	-0.0043	-0.0602
	Engineering & technology	0.7233**	0.4955**	0.6901**	0.6900**
	English language & literature	0.7176**	0.4260**	0.3557*	0.6349**
	Foreign Languages & Linguistics	0.2993*	0.0851	0.0370	0.2343*
	History, history of art, classics & curatorial studies	0.7363**	0.6091**	0.4750**	0.6816**
	Human geography	0.3564**	0.4451**	0.5827**	0.4014**
	Law	0.0511	0.3199*	0.1318	0.1091
	Management, human resources, industrial relations, international business & other business	0.0795	0.1013	0.4039**	0.1304
	Māori knowledge & development	0.2971	0.1338	0.1703	0.2699*
	Marketing & tourism	Reference category			
	Molecular, cellular & whole organism biology	0.6319**	0.4179**	0.8033**	0.6275**
	Music, literary arts & other arts	0.5519**	0.4513**	-0.0433	0.4542**
	Nursing	-0.8300**	-0.4465**	-0.0112	-0.6483**
	Other health studies (including rehabilitation therapies)	0.0386	0.2216	0.4727**	0.1347

Table 3: Ordinary least squares regression results (continued)

Explanatory variables	Categories (where applicable)	Coefficients			
		RO	PE	CRE	WRS
SUBJECT cont…	Pharmacy	0.3596	0.2324	0.6887**	0.3714
	Philosophy	1.0637**	1.2041**	1.2407**	1.1076**
	Physics	0.3839**	0.3465*	0.5654**	0.4201**
	Political science, international relations & public policy	0.8606**	0.6844**	0.5363**	0.7866**
	Psychology	0.4446**	0.5080**	0.7063**	0.5005**
	Public health	0.2716*	0.2682*	0.5257**	0.3054**
	Pure and applied mathematics	0.9922**	0.6617**	0.7019**	0.9003**
	Religious studies & theology	0.5499**	0.4157*	0.5538**	0.5356**
	Sociology, social policy, social work, criminology & gender studies	0.1301	0.3533**	0.5385**	0.2267*
	Sport & exercise science	-0.3101*	0.0964	0.3068*	-0.1560
	Statistics	0.4446**	-0.1160	0.0473	0.3169*
	Theatre & dance, film, television & multimedia	0.5929**	0.3299	0.1118	0.4895**
	Veterinary studies & large animal science	-0.2376	-0.2938	-0.0437	-0.2202
	Visual arts & crafts	0.6285**	0.0606	-0.4185**	0.4053
TEO_ TYPE	University (excluding Auckland University of Technology)	0.7514**	0.7286**	0.8569**	0.7578**
	Other tertiary education organisation	Reference category			
SIZE		0.0095**	0.0102**	0.0093**	0.0097**
SIZE^2		-0.0001**	-0.0001**	-0.0001**	-0.0001**
OUTPUT	<10	Reference category			
	≥ 10 & <20	0.7020**	0.6762**	0.6466**	0.6819**
	≥ 20 & < 30	1.1643**	1.1347**	1.1013**	1.1398**
	≥ 30	1.5586**	1.6737**	1.6599**	1.5806**
CONSTANT		-2.7778**	-3.3576**	-4.1376**	-2.9911**
R^2		0.59	0.63	0.64	0.66
N		6,612	6,612	6,612	6,612

Notes:
1. **, * indicates significant at the 1 percent and 5 percent level, respectively.
2. Robust standard errors were used to indicate statistical significance.
3. In this analysis the WRS score was divided by 100 so that it ranged from 0 to 7, the same as the other three research component scores. This makes comparison of the ordinary least squares regression results easier.

The largest variation in performance was exhibited by staff in the ‘Visual arts and crafts’ area. They achieved an RO score 0.63 points higher than that achieved by ‘marketing and tourism’ staff, but their CRE score was 0.42 points lower than that achieved by staff in ‘marketing and tourism’.

This result illustrates how using regression analysis to hold other factors constant can provide a different insight into subject performance. The raw average RO score for ‘Visual arts and crafts’ (3.1) is below that of ‘Marketing and tourism’ (3.4), but the subject area of ‘Visual arts and crafts’ contains few academics of senior rank and also has a large proportion of staff employed outside of the seven older universities. After controlling for these and the remaining factors, the subject performance of ‘Visual arts and crafts’ in terms of RO score improves markedly.

Staff at the seven older universities had higher research component scores on average than staff at the remaining tertiary education organisations. The greatest advantage from being at an older university was in CRE score, which was 0.86 points higher than achieved by staff at other tertiary education organisations. It is likely that the greater opportunity to supervise students at the postgraduate level at older established universities is a factor in this result.

The size of a subject (in terms of full-time equivalent staff) within a tertiary education organisation had a positive relationship with the research component scores. However, the statistical significance of the square of the *SIZE* variable, which has a negative sign, indicates that eventually the larger size of a subject (in terms of full-time equivalent staff) may reduce the research quality of an individual researcher.

The number of submitted research outputs had a positive relationship with the research component scores. However, this variable only explained around 9 percent of the variation in the research component scores, indicating that quantity is not the key factor driving the allocated quality scores.

There are signs of diminishing returns to output, with this most evident in the RO score. For example, staff who submitted between 10 and 20 outputs received a RO score 0.70 points higher than staff who submitted less than 10 outputs. Staff that submitted between 20 and 30 research outputs received an RO score 1.16 points higher than staff who submitted less than 10 outputs. For staff that submitted more than 30 outputs their average RO score was 1.56 points higher than staff who submitted less than 10 outputs. The advantage from submitting more than 30 outputs compared with less than 10 outputs was smallest for the RO score and largest for the PE score.

7 Conclusion

This study of the 2006 Performance-Based Research Fund (PBRF) Quality Evaluation results using regression analysis showed that the job title of staff had one of the strongest associations with research quality, with staff of higher rank achieving higher research quality − a result mirrored in an earlier analysis of the 2003 Quality Evaluation (Smart, 2005). This suggests there is a strong alignment between the quality of research produced by staff and their seniority in the tertiary education sector.

Of those groups of researchers identified as being at a possible disadvantage in the Quality Evaluation process, staff in some of the subjects that require professional training exhibited the lowest relative performance in the 2006 Quality Evaluation − especially staff in the areas of 'Education' and 'Nursing'. 'Nursing' and to a lesser extent 'Education' are also newer research disciplines within the New Zealand tertiary education sector which may be a factor in their performance. Further analysis suggested that it was the research output (RO) score that was most problematic for staff in the 'Nursing' area while the peer esteem (PE) score was most problematic for staff in the 'Education' area.

The government has moved to assist some of these areas in developing research capability through the Building Research Capability in Strategically Relevant Areas Fund. Also, the merger of the remaining colleges of education into the universities may help to improve research performance, given that staff employed at the seven oldest universities performed better than staff employed at other tertiary education organisations.

There was evidence that changes made to the 2006 Quality Evaluation to assess new and emerging researchers differently from more experienced staff had the desired effect. The analysis showed that new and emerging researchers had an equal likelihood of attracting funding for their tertiary education organisation under the new rules in the Quality Evaluation compared with more experienced staff.

The analysis of the impact of gender and ethnic group on research quality showed a mixed picture. Women who submitted evidence portfolios received slightly lower RO scores than men but received slightly higher CRE scores. However, analysis of all the staff participating in the Quality Evaluation showed that women and men had an equal likelihood of achieving a higher quality category. Similarly, although Māori staff who submitted evidence portfolios received lower RO scores than European staff, they received higher PE scores. However, Māori staff had a lower likelihood than European staff of achieving more than an R and more than a C quality category.

A larger number of researchers in a subject area at a tertiary education organisation was associated with higher research performance, although the regression modelling suggests that there is an eventual optimum size to the number of staff in a subject area. There was also a positive association between the quantity of research outputs submitted in evidence portfolios and the research component scores, however it wasn't the strongest factor associated with research quality, reinforcing that the Quality Evaluation is measuring quality rather than quantity of research.

Currently, one of the biggest limitations to any analysis of the research performance of staff in the PBRF Quality Evaluations is the lack of data on the quantity and/or quality of teaching by participating staff. Without this information, the impact of the teaching role on research performance cannot be evaluated. From an analytical viewpoint, if information on the teaching commitments of staff could be collected in future Quality Evaluations, it would enable the research performance of staff to be placed in context and identify any possible tradeoffs between teaching and research performance.

Appendix A: Summary statistics from the generalised ordered logistic regression model

Table 4: Summary statistics from the generalised ordered logistic regression model – continuous variables

Variable	Mean	Std dev
AGE	47.7	10.0
LOAD	0.93	0.17
SIZE	23.6	24.9

Note: These statistics refer to the dataset used in the generalised ordered logistic regression analysis. This represents around 97 percent of staff that participated in the 2006 Quality Evaluation. N = 8,410.

Table 5: Summary statistics from the generalised ordered logistic regression model – categorical variables

				% with more than indicated quality category		
Variable	Category	Number of staff	Proportion of sample	R	C	B
All		8,410	100%	66.8%	32.5%	7.4%
GENDER	Male	4,921	58.5%	74.6%	40.7%	10.4%
	Female	3,489	41.5%	55.7%	21.0%	3.2%
ETHNIC_GROUP	European	5,454	64.9%	65.8%	32.8%	7.3%
	Māori	480	5.7%	42.7%	14.2%	2.7%
	Pasifika	87	1.0%	50.6%	11.5%	3.4%
	Asian	507	6.0%	69.0%	26.6%	4.5%
	Other	893	10.6%	67.4%	31.1%	7.2%
	Unknown	989	11.8%	83.3%	46.0%	12.1%
EXPERIENCE	Experienced	6,577	78.2%	71.1%	39.9%	9.4%
	New and emerging	1,833	21.8%	51.3%	6.3%	0.3%
POSITION	Other staff	258	3.1%	45.0%	11.6%	3.1%
	Assistant lecturer	403	4.8%	51.1%	6.0%	0.7%
	Lecturer	2,856	34.0%	45.4%	6.9%	0.3%
	Senior lecturer	2,930	34.8%	75.2%	33.2%	1.6%
	Associate professor	843	10.0%	97.9%	82.3%	16.0%
	Professor	807	9.6%	98.9%	92.4%	49.2%
	Academic leader	313	3.7%	54.0%	22.4%	7.7%
SUBJECT	Accounting & finance	242	2.9%	52.1%	20.2%	4.5%
	Agriculture & other applied biological sciences	180	2.1%	78.3%	35.6%	5.6%
	Anthropology & archaeology	76	0.9%	89.5%	52.6%	10.5%
	Architecture, design, planning, surveying	177	2.1%	65.0%	29.4%	4.5%
	Biomedical	236	2.8%	90.7%	53.8%	15.7%
	Chemistry	183	2.2%	87.4%	50.8%	15.3%
	Clinical medicine	295	3.5%	73.2%	37.6%	6.8%
	Communications, journalism & media studies	135	1.6%	55.6%	20.7%	0.7%
	Computer science, IT, information sciences	425	5.1%	66.4%	29.4%	5.6%
	Dentistry	39	0.5%	84.6%	35.9%	17.9%
	Design	86	1.0%	40.7%	9.3%	1.2%
	Earth sciences	146	1.7%	93.8%	63.0%	11.6%
	Ecology, evolution & behaviour	200	2.4%	93.0%	52.5%	14.5%
	Economics	151	1.8%	82.8%	45.7%	6.0%
	Education	1,036	12.3%	34.2%	12.4%	2.7%
	Engineering & technology	447	5.3%	78.5%	42.3%	13.0%
	English language & literature	119	1.4%	81.5%	37.8%	9.2%
	Foreign Languages & Linguistics	204	2.4%	62.3%	25.5%	6.9%
	History, history of art, classics & curatorial studies	198	2.4%	89.9%	49.5%	8.6%
	Human geography	67	0.8%	94.0%	50.7%	13.4%

Table 5: Summary statistics from the generalised ordered logistic regression model – categorical variables – continued

Variable	Category	Number of staff	Proportion of sample	% with more than indicated quality category R	C	B
SUBJECT cont…	Law	214	2.5%	73.4%	49.5%	9.3%
	Management, human resources, industrial relations, international business & other business	404	4.8%	65.8%	27.0%	4.5%
	Māori knowledge & development	187	2.2%	46.0%	20.3%	2.1%
	Marketing & tourism	187	2.2%	71.1%	29.9%	5.3%
	Molecular, cellular & whole organism biology	369	4.4%	82.7%	45.8%	8.1%
	Music, literary arts & other arts	159	1.9%	76.7%	42.1%	6.3%
	Nursing	249	3.0%	18.9%	3.2%	0.4%
	Other health studies (including rehabilitation therapies)	208	2.5%	50.0%	20.2%	4.3%
	Pharmacy	20	0.2%	70.0%	45.0%	5.0%
	Philosophy	71	0.8%	88.7%	60.6%	23.9%
	Physics	107	1.3%	92.5%	55.1%	14.0%
	Political science, international relations & public policy	108	1.3%	86.1%	46.3%	13.0%
	Psychology	245	2.9%	79.2%	45.7%	17.1%
	Public health	192	2.3%	83.9%	34.4%	8.9%
	Pure and applied mathematics	132	1.6%	81.8%	50.8%	18.9%
	Religious studies & theology	63	0.7%	47.6%	25.4%	3.2%
	Sociology, social policy, social work, criminology & gender studies	241	2.9%	67.2%	26.6%	4.1%
	Sport & exercise science	108	1.3%	56.5%	14.8%	0.9%
	Statistics	98	1.2%	83.7%	40.8%	11.2%
	Theatre & dance, film, television & multimedia	91	1.1%	51.6%	15.4%	3.3%
	Veterinary studies & large animal science	71	0.8%	74.6%	33.8%	9.9%
	Visual arts & crafts	244	2.9%	59.4%	15.6%	2.9%
TEO_TYPE	University (excluding Auckland University of Technology)	6,262	74.5%	80.3%	41.9%	9.8%
	Other tertiary education organisation	2,148	25.5%	27.4%	5.3%	0.4%

Note: These statistics refer to the dataset used in the generalised ordered logistic regression analysis. This represents around 97 percent of staff that participated in the 2006 Quality Evaluation.

Appendix B: Summary statistics from the ordinary least squares regression model

Table 6: Summary statistics from the ordinary least squares regression model – continuous variables

Variable	Mean	Std dev
AGE	47.7	10.0
LOAD	0.95	0.16
SIZE	23.0	22.3

Note: These statistics refer to the dataset used in the ordinary least squares regression analysis. This represents around 97 percent of staff that participated in the 2006 Quality Evaluation and had their evidence portfolios assessed in either 2003 and/or 2006. N = 6,612.

Table 7: Summary statistics from the ordinary least squares regression model – categorical variables

				Average score			
Variable	Category	Number of staff	Proportion of sample	RO	PE	CRE	WRS
All		6,612	100.0%	3.6	3.2	3.0	3.5
GENDER	Male	4,127	62.4%	3.9	3.4	3.3	3.8
	Female	2,485	37.6%	3.1	2.7	2.6	3.0
ETHNIC_GROUP	European	4,254	64.3%	3.6	3.2	3.1	3.5
	Māori	286	4.3%	2.9	2.7	2.5	2.8
	Pasifika	59	0.9%	2.9	2.5	2.4	2.7
	Asian	424	6.4%	3.4	2.7	2.6	3.2
	Other	684	10.3%	3.6	3.1	3.0	3.5
	Unknown	905	13.7%	4.0	3.5	3.3	3.8
EXPERIENCE	Experienced	5,339	80.7%	3.9	3.5	3.4	3.8
	New and emerging	1,273	19.3%	2.5	1.9	1.7	2.3
POSITION	Other staff	165	2.5%	2.9	2.3	2.2	2.7
	Assistant lecturer	263	4.0%	2.7	1.9	1.7	2.5
	Lecturer	1,790	27.1%	2.7	2.0	1.8	2.5
	Senior lecturer	2,536	38.4%	3.5	3.0	2.9	3.3
	Associate professor	834	12.6%	4.8	4.5	4.4	4.7
	Professor	803	12.1%	5.5	5.5	5.4	5.5
	Academic leader	221	3.3%	3.3	3.0	2.9	3.2
SUBJECT	Accounting & finance	176	2.7%	3.1	2.5	2.3	2.9
	Agriculture & other applied biological sciences	156	2.4%	3.5	3.3	3.5	3.5
	Anthropology & archaeology	74	1.1%	4.0	3.8	3.8	4.0
	Architecture, design, planning, surveying	125	1.9%	3.7	3.1	2.9	3.6
	Biomedical	224	3.4%	4.3	3.7	3.9	4.1
	Chemistry	170	2.6%	4.2	3.8	3.8	4.1
	Clinical medicine	240	3.6%	3.8	3.5	3.3	3.7
	Communications, journalism & media studies	92	1.4%	3.4	2.6	2.5	3.2
	Computer science, IT, information sciences	315	4.8%	3.7	3.1	2.9	3.5
	Dentistry	36	0.5%	3.9	3.4	3.3	3.7
	Design	60	0.9%	2.5	1.8	1.5	2.2
	Earth sciences	141	2.1%	4.3	4.0	4.1	4.2
	Ecology, evolution & behaviour	192	2.9%	4.1	3.7	3.8	4.0
	Economics	139	2.1%	3.9	3.3	3.1	3.7
	Education	586	8.9%	2.7	2.2	2.2	2.5
	Engineering & technology	362	5.5%	4.3	3.6	3.5	4.1
	English language & literature	103	1.6%	4.0	3.4	3.1	3.8
	Foreign Languages & Linguistics	163	2.5%	3.3	2.7	2.4	3.1
	History, history of art, classics & curatorial studies	190	2.9%	4.1	3.6	3.2	3.8

Table 7: Summary statistics from the ordinary least squares regression model – categorical variables – continued

Variable	Category	Number of staff	Proportion of sample	Average score RO	Average score PE	Average score CRE	Average score WRS
SUBJECT cont…	Human geography	66	1.0%	4.1	3.8	3.7	4.0
	Law	187	2.8%	3.8	3.8	3.3	3.7
	Management, human resources, industrial relations, international business & other business	319	4.8%	3.4	3.0	3.1	3.3
	Māori knowledge & development	123	1.9%	3.0	2.7	2.5	2.9
	Marketing & tourism	160	2.4%	3.4	3.0	2.7	3.2
	Molecular, cellular & whole organism biology	336	5.1%	3.9	3.3	3.4	3.7
	Music, literary arts & other arts	137	2.1%	3.9	3.4	2.7	3.7
	Nursing	94	1.4%	2.0	2.1	2.3	2.1
	Other health studies (including rehabilitation therapies)	133	2.0%	3.2	3.0	3.0	3.1
	Pharmacy	20	0.3%	3.5	2.9	3.1	3.3
	Philosophy	67	1.0%	4.5	4.3	4.0	4.4
	Physics	104	1.6%	4.1	3.6	3.5	4.0
	Political science, international relations & public policy	96	1.5%	4.3	3.7	3.3	4.1
	Psychology	219	3.3%	3.9	3.6	3.6	3.9
	Public health	173	2.6%	3.7	3.4	3.4	3.6
	Pure and applied mathematics	117	1.8%	4.5	3.7	3.5	4.2
	Religious studies & theology	40	0.6%	3.4	2.8	2.7	3.2
	Sociology, social policy, social work, criminology & gender studies	193	2.9%	3.2	3.1	3.1	3.2
	Sport & exercise science	74	1.1%	2.8	2.7	2.7	2.8
	Statistics	90	1.4%	3.9	2.9	2.8	3.6
	Theatre & dance, film, television & multimedia	64	1.0%	3.0	2.3	1.8	2.8
	Veterinary studies & large animal science	62	0.9%	3.4	3.0	3.0	3.3
	Visual arts & crafts	194	2.9%	3.1	2.1	1.3	2.7
TEO_TYPE	University (excluding Auckland University of Technology)	5,625	85.1%	3.8	3.4	3.3	3.7
	Other tertiary education organisation	987	14.9%	2.4	1.9	1.6	2.2
OUTPUT	<10	1,269	19.2%	2.2	1.6	1.5	2.0
	≥ 10 & <20	1,912	28.9%	3.2	2.6	2.5	3.0
	≥ 20 & < 30	1,313	19.9%	4.0	3.4	3.3	3.8
	≥ 30	2,118	32.0%	4.7	4.4	4.3	4.6

Notes:
1. These statistics refer to the dataset used in the ordinary least squares regression analysis. This represents around 97 percent of staff that participated in the 2006 Quality Evaluation and had their evidence portfolios assessed in either 2003 and/or 2006.
2. The WRS score presented in this table has been divided by 100. This makes comparison between the WRS score and the other three research component scores easier.

References

- Christensen, H., Henderson, A., Griffiths, K. & Levings, C. (1997) Does aging inevitably lead to declines in cognitive performance? A longitudinal study of elite academics, *Journal of Personal Individual Differences*, 23(1), pp 67-78.
- Çinlar, N. & Dowse, J. (2008a) *Human resource trends in the tertiary academic workforce*, Wellington: Tertiary Education Commission.
- Çinlar, N. & Dowse, J. (2008b) *Staffing and performance trends in research subject areas*, Wellington: Tertiary Education Commission.
- Çinlar, N. & Dowse, J. (2008c) *Baseline trends in research in the tertiary education sector*, Wellington: Tertiary Education Commission.
- Curtis, B. & Phibbs, S. (2006) Body politics within the academy: gender and the Performance-Based Research Fund in Bakker, L., Boston, J., Campbell, L. & Smyth, R. (eds) *Evaluating the Performance-Based Research Fund*, pp 401-416, Wellington: Institute of Policy Studies.
- Fitzgerald, P. & Knuiman, M. (1998) Estimation of regressive logistic regression analyses of familial data with missing outcomes, *Australian and New Zealand Journal of Statistics*, 40(3), pp-305-316.
- Jayasinghe, U., Marsh, H. & Bond, N. (2001) Peer review in the funding of research in higher education: the Australian experience, *Educational Evaluation and Policy Analysis*, 23(4), pp 343-364.
- Johnes, J., Taylor, J. & Francis, B. (1992) The research performance of UK universities: a statistical analysis of the results of the 1989 Research selectivity exercise, *Journal of the Royal Statistical Society, Series A (Statistics in Society)*, 156(2), pp 271-286.
- Middleton, S. (2006) Researching identities: impact of the Performance-Based Research Fund on the subject(s) of education in Bakker, L., Boston, J., Campbell, L. & Smyth, R. (eds) *Evaluating the Performance-Based Research Fund*, pp 477-500, Wellington: Institute of Policy Studies.
- Smart, W. (2005) *What determines the research performance of staff in New Zealand's tertiary education sector?* Wellington: Ministry of Education.
- Smith, A. & Bruce Ferguson, P. (2006) Impact of the Performance-Based Research Fund on Māori research and researchers in Bakker, L., Boston, J., Campbell, L. & Smyth, R. (eds) *Evaluating the Performance-Based Research Fund*, pp 379-400, Wellington: Institute of Policy Studies.
- Williams, R. (2006) Generalised ordered logit/partial proportional odds models for ordinal dependent variables, *The STATA Journal*, 6(1), pp 58-82.
- White, P. & Grice, J. (2008) *Participation and performance by Māori and Pasifika researchers in the Performance-Based Research Fund, 2003-2006*, Wellington: Tertiary Education Commission.
- Taylor, J. (1995) A statistical analysis of the 1992 Research Assessment Exercise, *Journal of the Royal Statistical Society, Series A (Statistics in Society)*, 158(2), pp 241-261.
- Tertiary Education Commission (2007) *Performance-Based Research Fund: evaluating research excellence – the 2006 assessment,* Wellington: Tertiary Education Commission.

MINISTRY OF EDUCATION
Te Tāhuhu o te Mātauranga

newzealand.govt.nz