# Determinants of quality of research environment: an assessment of the environment submissions in the UK's Research Excellence Framework in 2014

Mehmet Pinar[*] and Emre Unlu

Business School, Edge Hill University, St Helens Road, Ormskirk, Lancashire L39 4QP, UK

* Corresponding author. Email: mehmet.pinar@edgehill.ac.uk

## Abstract

One of the assessed research elements in the UK's Research Excellence Framework (REF) exercise in 2014 was the research environment. The quality of the research environment was assessed by expert peer reviewers who were given a set of quantitative factors to support their decision making. However, there is no systematic procedure to integrate this quantitative information into the evaluation process. This paper evaluates the relevance of quantitative factors in explaining the assessed quality of the research environment. Findings suggest submitting units with high external research income generation tend to have a better research environment evaluation in almost all the assessed subject areas. The importance given by reviewers to similar quantitative factors was distinctively different in two units of assessment in which the evaluation criteria were the same, which highlights the internal inconsistency of the peer review evaluation. Our findings also confirm the existence of the 'halo effect' in some units of assessment where submitting units that belong to the Russell group and have sub-panel members in the REF exercise obtained higher scores even after controlling for the assessed quantitative factors.

This is the authors' accepted manuscript for the article:

## 1. Introduction

Research assessment exercises have been increasingly used to assess the quality of research produced by higher education institutes (HEIs). These assessment exercises are mainly used to allocate research funds selectively to high performing HEIs, increase accountability for the use of public funds, track the research progress of HEIs and create performance incentives for HEIs and researchers (Hicks, 2012; Rebora and Turri, 2013; Dougherty et al., 2016; Jonkers and Zacharewicz, 2016; Zacharewicz et al., 2019 among many others for a review of research assessment exercises in different countries).

Research performances of HEIs are evaluated in three general ways: evaluation based on the use of quantitative metrics (e.g., quality of journals, citations received, external research income and so on), expert peer review evaluation of research activities, or evaluation based on the use of both quantitative metrics and peer review evaluation (see e.g. Zacharewicz et al., 2019 for detailed evaluation processes used by the European Union member states). Each evaluation method has its strengths and weaknesses. It has been argued that conducting peer review assessments is extremely expensive (e.g., the UK's Research Excellence Framework in 2014 [REF2014 hereafter] costed £246 million; see e.g., Martin, 2011 and Stern, 2016 for further discussion) and that bibliometric evaluation can decrease some of these costs (see e.g., Ancaiani et al., 2015; De Boer et al., 2015; Geuna and Piollato, 2016) because some of the bibliometric factors are correlated with peer review assessments (see e.g., Bertocchi et al., 2015; Wooldridge and King, 2019). On the other hand, it has been also argued that 'novelty' and 'significance' of research activity may not be identified by bibliometric information. 'Novel' research is usually delayed in recognition and published in so-called 'low impact factor journals', but it is more likely to be identified by peer review evaluation (Wang et al., 2017). Given the limitations and advantages of both bibliometric and peer review evaluations, Hicks et al. (2015) argued that some of the useful aspects of quantitative metrics can be used to support peer review evaluation, which could also eliminate some peer review biases (see e.g., Lee et al., 2013 for detailed discussion on peer review biases).

Peer review bias in research evaluation has been discussed in detail. Editors and referees usually judge the quality of the paper by the reputation of the authors or authors' institution (see e.g., Macdonald and Kam, 2007; Bornmann, 2011; Lee et al., 2013; Garcia et al., 2015 among many others). In the context of research assessment exercises, Taylor (2011) finds that the Russell group university submissions obtained relatively higher scores in the Research Assessment Exercise in 2008 (RAE2008 hereafter) after controlling for various quantitative factors. Thorpe et al. (2018a, 2018b) examine whether there were notable

language- related differences in the research environment templates of the high and low scored submissions in the REF2014 and find that submissions from the reputable institutions attempted to take advantage of the 'halo effect' by capitalizing on assessors' existing knowledge of their institutional position. Given the potential bias involved in peer review evaluation, there have been numerous attempts to find relevant bibliometric factors for assessing quality of research outputs and environment (see e.g., Bertocchi et al., 2015; HECFE, 2015; Mryglod et al., 2015; Pidd and Broadbent, 2015; Baccini and De Nicolao, 2016; Bruns and Stern, 2016; Geuna and Piolatto, 2016; UK Forum for Responsible Research Metrics, 2018a among many others).

In this paper, we examine the UK REF2014 research environment evaluations. This is one of the few evaluation systems that use both quantitative data and peer review for evaluating the quality of research environment (see Table 1 of Zacharewicz et al., 2019). The availability of quantitative data in the evaluation process allows us to examine which set of quantitative factors played an important role in peer review evaluations across all units of assessments (UoAs hereafter), which quantitative variables were not important for peer reviewers in their evaluations, and whether the importance given to quantitative variables was different across different subject areas.

Despite some changes, many of the elements for the evaluation of research environments in REF2021[1] are similar to REF2014 (e.g., both REFs consist of the same environment data including postgraduate research degree completions [PGR completions hereafter] and research income generated by submitting units). Hence, although we examine these indicators in the context of REF2014, our results are relevant for REF2021 in three ways. First, our findings could provide a better understanding of how data is used by peer reviewers in different panels and thus provide policy recommendations to the four UK higher education funding bodies (i.e., Research England, the Scottish Funding Council, the Higher Education Funding Council for Wales, and the Department for the Economy, Northern Ireland). Second, policymakers at HEIs that are responsible for returning submissions to the next REF may

---

[1] HEIs are now required to submit institutional-level environment templates, which is aimed to reduce duplication across unit-level submissions and also enable better representation of aspects of an institution's environment that reflects the institutional-level activity. Furthermore, both the impact strategy and contribution to the society and economy are now incorporated into both the unit of assessment and institutional level environment statements which were not part of the unit of assessment level environment statement in the REF2014. Finally, HEIs are required to provide information about their approach to supporting interdisciplinary research in their institutional- and unit-level environment templates (see REF 2019a for detailed changes in the evaluation of the research environment element from REF2014 to REF2021 and see also paragraphs 85-94 in Stern (2016) for recommendations on how research environment evaluations could be improved in the REF2021).

benefit from understanding which quantitative factors play a significant role in the assessed quality of research environment in different fields. Third, the findings of this paper could be used to suggest policies to countries conducting similar research evaluations[2].

The rest of the paper is organised as follows. In section 2, we provide some details on how the research quality and research environment were evaluated in the REF2014 and provide a discussion on how quantitative factors are chosen for the empirical analysis. We then discuss the data used in our analysis in Section 3. Section 4 provides an estimation strategy and empirical findings. Finally, section 5 concludes and provides some policy suggestions.

## 2. The 2014 UK Research Excellence Framework

Research assessment exercises have long been implemented in the UK to inform the allocation of research funding (see e.g., Shattock, 2012 for a thorough discussion on the evolution of research assessment exercises in the UK since the introduction of the Research Selectivity Exercise in 1986). The most recent evaluation of research quality in the UK was the Research Excellece Framework (REF), and was conducted in 2014 to evaluate the quality of research outputs (outputs hereafter), the impact of this research beyond academia (impact hereafter), and the research environment (environment hereafter). All HEIs were invited to make submissions in 36 subject-based UoAs where the quality of research outputs, impact, and research environment were assessed by expert sub-panel members. Evaluations were based on general criteria and level definitions provided for four main panels: medicine, health and life sciences (panel A), physical sciences, engineering and mathematics (panel B), social sciences (panel C), arts and humanities (panel D). Each research element was given quality ratings based on the following five-point scale: world-leading (4-star), internationally-excellent (3-star), internationally recognized (2-star), nationally recognized (1-star), and unclassified (items that fall below the standard of nationally recognized work). The assessed research profiles were then used by the four funding bodies of the UK to allocate mainstream quality-related research (QR hereafter) funding across UoAs and HEIs in the UK from 2015-2016 to the following research funding period when REF2021 results are published. During this period, 65%, 20% and 15% of the total mainstream QR funding were distributed based on the assessed quality of

---

[2] Performance evaluation systems in different countries use similar quantitative data (i.e., PGR completions, external research income generated) in their evaluations and findings in this paper could provide some policy suggestions to the evaluators in these counties. For instance, Austria, Netherlands, Belgium, Denmark, Estonia, Finland, Croatia, Norway, Poland, Slovakia and Italy use the number of PGR completions as performance metrics. Similarly, project funding (external research income generation) is also used as part of metrics in the research performance evaluations in Belgium, Czech Republic, Denmark, Estonia, Finland, Croatia, Norway, Slovakia, France and Italy (see Zacharewicz et al., 2019 for detailed metrics used in different EU countries). Furthermore, France, Italy, Lithuania and Portugal are other EU countries that use peer review evaluation system where these countries also integrated similar indicators to support the peer review evaluations.

research outputs, impact, and environment of submitting unit, respectively, where no QR funding is given to submissions that were classified 2-star or below, and the submissions that were rated as 4-star were given four times as much funding than that rated 3-star (Research England, 2018).

Although the other research elements of the REF have been heavily researched, there is limited research on the environment element, and this limited research primarily focuses on research environment submissions of three UoAs (to our knowledge). Taylor (2011) examined the role of quantitative factors on the quality of research environment in the Business and Management, Economics and Econometrics, and Accounting and Finance UoAs in RAE2008, and Thorpe et al. (2018a, 2018b) analysed the linguistic differences and the relevance of the quantitative data of research environment evaluations in the Business and Management Studies UoA in REF2014. In addition, both Taylor (2011) and Thorpe et al. (2018b) examined whether sub-panel members presented any bias towards units submitted by Russell Group universities, or units that had a panel reviewer in the evaluation from the same HEI. Based on findings for these three UoAs, we will replicate this claim to examine if such effects might exist in other UoAs since previous literature only examined such effects in three UoAs.

## 2.1. Research environment evaluation in REF2014

The main criteria that sub-panel members considered when assessing the quality of research environment submissions was to answer the following question: 'Does this place look like a great place for work where senior and junior researchers will flourish and will move the discipline further?' (Pidd and Broadbent, 2015). Thus, sub-panels assessed the environment submissions according to how the submissions provide evidence of vitality and sustainability in producing an encouraging and facilitating environment for research (see REF, 2011 and REF, 2012a). Environment submissions included two components: environment template and environment data.

### 2.1.1. Environment template

The environment template consisted of the following sub-sections:

- Overview
- Research strategy
- People, including:
    - Staffing strategy and staff development
    - Research students
- Income, infrastructure, and facilities

- Collaboration and contribution to the discipline

The quality of the research environment was assessed in terms of its 'vitality and sustainability' where the 'vitality' aspect considers whether there existed an encouraging and facilitating environment for research, and whether the submitting unit had an effective strategic plan and engaged with the national and international research community. The 'sustainability' of a submitting unit was evaluated based on whether there was a coherent vision for the future and an investment in people and infrastructure. Sub-panel expert reviewers combined the 'overview' and 'research strategy' sections and assessed the following four components of the submissions: i) research strategy, ii) people, iii) income, infrastructure, and facilities, iv) collaboration and contribution to the discipline. Table 1 shows the weights (importance levels) given to each component by expert reviewers (see REF, 2012b, 2012c, 2012d, and 2012e for submission and evaluation criteria for Panels A, B, C, and D, respectively).

<Insert Table 1 approximately here>

The length of the research environment template submissions was based on the number of full-time equivalent (FTE) staff submitted. A submission with a number of FTE staff submitted between 1 and 14.99 was allowed to submit 7 pages with an additional page allowed for every additional 10 FTE staff submitted up to 54.99, and then an additional page for every additional 20 FTE staff submitted for submitting units with more than 55 FTE staff submitted (REF, 2011). In other words, relatively larger units had more space to elaborate on the factors that matter for the quality of their research environment. For the evaluation of the environment templates, the assessors were instructed to apply the criteria of evidence of 'vitality' and 'sustainability' (Pidd and Broadbent, 2015), but language used in the environment templates by submissions from high and low ranked universities were distinctively different and played a major role in the assessors' evaluation. Thorpe et al. (2018a, 2018b) identified this issue by examining the environment template narrative and found that long-established HEIs capitalized on their reputation in the narrative element of the research template to obtain better scores.

**2.1.2. Environment data and the relevance of quantitative factors for the quality of the research environment**

Each submission to the REF2014 provided a detailed set of data which consisted of the number of PGR completions (from 1 August 2008 to 31 July 2013) and total research income generated by the submitting unit (external research income and research income-in-kind from 1 August 2008 to 31 July 2013). Data was provided to the panels to support their evaluation. Therefore, even though peer review evaluation does not follow any specific formula to integrate these data into the evaluation, the quality of the assessed research environment is

expected to be associated with these two factors. We, therefore, hypothesise that submitting units with larger external research income and a higher number of PGR completions obtained higher assessed research environment scores.

*Hypothesis 1*: Submitting units that generated larger external research income received a higher assessed quality of the research environment.

*Hypothesis 2*: Submitting units with more PGR completions received a higher assessed quality of the research environment.

Another relevant quantitative factor was the size of the submitting unit - the total number of FTE staff submitted to REF2014[3]. The size of the submitting unit could have played a role in research environment quality for several reasons. First, the number of FTE staff submitted is a good proxy for the number and quality of research collaborations and research groups within the submitting units (see e.g., Larivière et al., 2015 for the link between team size and scientific collaborations). Second, infrastructure and facilities of larger submissions are likely to be better (e.g., Mryglod et al. [2013] found that bigger research groups have access to more expensive and sophisticated equipment and achieved relatively higher scores in RAE2008). Therefore, we hypothesise that submissions with a higher number of FTE staff submitted were likely to achieve better environment scores.

*Hypothesis 3*: Larger submissions had a higher assessed quality of the research environment.

Another potential determinant of variation in the quality of the research environment could be due to implicit peer review bias (see e.g., Taylor (2011) and Thorpe et al. (2018b)). Panel members received training to overcome such implicit bias, but it is not clear whether this training was useful in overcoming the 'halo effect' in all panels. Therefore, we hypothesise that, if halo effects or implicit biases existed, then submissions from Russell Group Universities and submissions that had sub-panel members in the evaluation process should have achieved higher environment scores.

---

[3] Although we use the number of FTE staff submitted as the size of the submitting unit, it should be noted that the submitting units in HEIs could be selective in deciding which individual researchers that they 'wanted' to submit in relation to having their outputs assessed by a relevant REF panel or researchers from other 'research institutes'. Furthermore, some of the staff members could also be included to any particular UoA rather than having to submit the researcher in their 'home' discipline or UoA, particularly if the university did not make a submission to certain UoAs. Hence, even though submitted FTE in each unit may be considered as a proxy for the size of the submitting unit, this should not be considered as size of the subject area (department) (see Pidd and Boardant, 2015 for a detailed discussion on this issue).

*Hypothesis 4*: There may have been an implicit bias (halo effect) in the REF evaluations if submissions of the Russell Group Universities or submissions that had sub-panel members obtained higher scores in the research environment.

## 3. Data

The data used in this paper comes from the submission data provided on the REF2014 (REF, 2014), including the number of FTE staff submitted, the total number of PGR completions, and the amount of external research income generated by submitting units in the period between 1 August 2008 and 31 July 2013. We controled whether the submitting unit was from a Russell Group university or not[4], and whether the submitting unit had a panel reviewer in the evaluation from the same HEI or not[5]. We constructed our dependent variable by obtaining the grade point average (GPA) of the assessed research quality of each submission by multiplying the proportion of research activity attributed to each quality category with the respective weights given to each quality profile categories: unclassified=0, $1^*$=1, $2^*$=2, $3^*$=3 and $4^*$=4. For instance, if a given submission's research environment were assessed as 50% internationally-excellent (3*) and 50% world-leading (4*), the GPA of the research environment of this submitting unit would be 3.5.

Table 2 includes descriptive statistics across thirty-six UoAs as well as each main panel. On average, all UoAs obtained GPA scores above three except for the UoA11 Computer Science and Informatics. The average quality of the research environment was the lowest for Computer Science and Informatics and was the highest for the UoA2 Public Health, Health Services, and Primary Care. There was also a significant variation across the UoAs in terms of external income generation where, on average, £1.7 million per FTE staff submitted was generated by the UoA1 Clinical Medicine, and £34,613 per FTE staff submitted was generated by UoA29 English Language and Literature. There is a clear pattern in terms of external research income generation across four main panels where medicine, health and life sciences (panel A), physical sciences, engineering and mathematics (panel B), social sciences (panel C), arts and humanities (panel D) generated roughly £940k, 631k, 146k, 64k per FTE staff submitted respectively. There is also a high variation in the size of the submissions across UoAs

---

[4] The Russell Group comprises of the following 24 universities: University of Birmingham, University of Bristol, University of Cambridge, Cardiff University, Durham University, University of Edinburgh, University of Exeter, University of Glasgow, Imperial College London, King's College London University of Leeds, University of Liverpool, London School of Economics and Political Science, University of Manchester, Newcastle University, University of Nottingham, University of Oxford, Queen Mary University of London, Queen's University Belfast, University of Sheffield, University of Southampton, University College London, University of Warwick and University of York.

[5] The list of sub-panel members of each UoA are obtained from https://www.ref.ac.uk/2014/panels/panelmembership/.

where 3571 FTE staff submitted to the UoA1 Clinical Medicine and only 383 FTE staff submitted to the UoA31 Classics. On the other hand, the highest and lowest numbers of PGR completions per FTE staff submitted was in the UoA8 Chemistry (3.85) and UoA36 Communication, Cultural and Media Studies, Library and Information Management (1.03), respectively. Furthermore, the proportion of Russell Group submissions varies across different UoAs (e.g., 21 out of 30 submissions to the UoA1 Clinical Medicine were from the Russell group and only 6 out of 50 submissions to the UoA26 Sport and Exercise Sciences, Leisure and Tourism were from the Russell group). Finally, some UoAs had more of panel members from the same HEI compared to others. For instance, 63% of the assessors of the UoA1 Clinical Medicine were from the HEIs that had submissions in this UoA. Whereas, the UoA26 Sport and Exercise Sciences, Leisure and Tourism had only 18% of the assessors from the HEIs that had submissions to this UoA.

<Insert Table 2 approximately here>

Since each panel has its own criteria to evaluate research environment submissions, we provide correlation matrices between quantitative factors and assessed quality of research scores within the four main panels in Table 3. All quantitative factors and environment GPA scores are positively associated with one another in all panels. The correlation coefficients between the size of the submitting unit (external income generation per FTE staff submitted) and GPA levels are highly and positively correlated, but the correlation coefficients between PGR completions per FTE staff submitted and GPA levels are relatively small. On the other hand, correlation coefficients between external income per FTE staff submitted and size of the submitting unit are higher in panels A and B than those in C and D, suggesting that larger submitting units in panels A and B tend to generate higher external research income. The correlation coefficient between external research income per FTE staff submitted and PGR completions per FTE staff submitted is also higher in panel B, which suggests that submissions with higher external research income tend to have higher PGR completions per FTE staff submitted. This is because receiving larger grants allow the submitting units to allocate more funds for research posts; hence, the number of PGR completions is larger in these submissions.

<Insert Table 3 approximately here>

4. **Estimation strategy and empirical analysis**

4.1. **Estimation strategy**

To examine the relationship between environment GPA scores and quantitative factors and to test the hypotheses set in section 2.1.b, we carry out a separate regression analysis for 36 UoA as follows.

$$GPA_i = \alpha + \beta \ln(\text{Income per FTE staff submitted})_i +$$
$$\gamma(\text{PGR per FTE staff submitted})_i + \delta \ln(\text{FTE staff submitted})_i + \theta RM_i +$$
$$\phi PM_i + \varepsilon_i \qquad\qquad (1)$$

where $i$ denotes a submitting unit in a given UoA. $GPA_i$, (Income per FTE staff submitted)$_i$, (PGR per FTE staff submitted)$_i$ and (FTE staff submitted)$_i$ represent the weighted average of research environment quality, research income generated per FTE staff submitted, the total number of PGR completions per FTE staff submitted, and the total number of FTE staff members submitted by unit $i$, respectively. Russell-group membership (RM hereafter) is a dummy variable, and it is equal to one if the submitting unit was from a Russell Group university and 0 otherwise. Panel membership (PM hereafter) is another dummy variable, and it is equal to one if the submitting unit had a panel member in the REF evaluation from the same HEI, and it is equal to zero otherwise.

## 4.2. Determinants of GPA levels across all submissions

Before starting with detailed UoA-level analysis, we first use a pooled regression model concerning the full data set to examine the potential differences between the UoAs and HEIs. We also carry out a 'fixed effects' model, whereby dummy variables are included for each UoA and HEI. The reasons for controlling for the institution and UoA-specific factors is that other sets of factors not considered in our analysis may be correlated with the assessed research environment quality. In REF2014, many submitting units from the same institution provided the same information in their unit-level environment statements, which led to an introduction of an institutional-level environment statement submission in the REF2021 (see paragraph 94 of Stern, 2016 and paragraph 27 of REF, 2019). Institution-specific factors that were used by units in their environment submissions in REF2014 may include the following factors: mission and strategy of the institution, infrastructure and facilities, management style, and other factors that enhance the research environment of the overall institution. UoA-specific factors may include environment narrative used, research culture in different subjects, and panel's attitude towards REF evaluation given the variation in quantitative factors across UoAs among other variables (e.g., some reviewers in one UoA may be biased towards their fields and may have

allocated higher scores compared to other UoAs or reviewers in some panels may give more importance to environment template compared to the data, and so on).

Results are shown in Table 4. In the first column, we use pooled ordinary least squares (OLS) estimation when we consider all submissions without differentiating them across UoAs. We find that PGR per FTE staff submitted, research income per FTE staff submitted, Russell Group membership, and panel membership are positively and significantly associated with the environment GPA scores. In the second column, we control for UoA-specific factors by including UoA dummies, and all the variables in interest remain to be significant and positive. We also find that there are significant differences in GPA scores among UoAs after accounting for quantitative variables (i.e., F-statistics for UoA dummies are significant at the 1% level). In the third column, we control for HEI-specific factors, and in the fourth column, we control for both UoA- and HEI-specific factors by including dummies for UoAs and HEIs, and we find that both UoA- and HEI-specific factors are significant at the 1% level[6]. These findings suggest that quantitative variables played a significant role in GPA scores achieved by submitting units at an aggregate level even after controlling for UoA- and HEI-specific factors. However, the peer review evaluations were mainly at UoA-level, and in the next subsection, we examine in which ways quantitative factors played a role in within-UoA variation of GPA scores for all UoAs, which will also enable us to compare our findings with those of Taylor (2011) and Thorpe et al. (2018b).

<Insert Table 4 approximately here>

### 4.3.  Determinants of GPA levels within the unit of assessments

Table 5 shows regression estimates when submissions are grouped by their corresponding UoAs, which allows investigating the impact of quantitative factors on the environment GPA scores of submitting units in each UoA. Given the relevance of these quantitative factors in the assessment criteria, one should expect the variation among these factors within each UoA to explain the variation in assessed quality by the peer reviewers. However, our analysis shows mixed results, with some quantitative factors playing a significant role in assessed research quality by peer reviewers but others not.

<Insert Table 5 approximately here>

We find that the size of the submission, which is a proxy of the level of collaboration, infrastructure and facilities, and the total research income generated per FTE staff submitted

---

[6]Both Russell Group and panel membership dummies are excluded from the analyses when we include HEI dummies due to multicollinearity problem.

played the most significant role in explaining the assessed quality within each UoA. Size of the submissions affects the assessed quality within 28 UoAs positively and highly significantly (1% level) with the exceptions of eight UoAs (i.e., UoA6 Agriculture, Veterinary and Food Science; UoA13 Electrical and Electronic Engineering, Metallurgy and Materials; UoA14 Civil and Construction Engineering; UoA18 Economics and Econometrics; UoA24 Anthropology and Development Studies; UoA27 Area Studies; UoA28 Modern Languages and Linguistics; UoA31 Classics; UoA36 Communication, Cultural and Media Studies, Library and Information Management). Similarly, research income generated per FTE staff submitted affects the assessed quality in 31 UoAs positively and significantly with the exceptions of five UoAs (i.e., UoA1 Clinical Medicine; UoA8 Chemistry; UoA10 Mathematical Sciences; UoA14 Civil and Construction Engineering; UoA31 Classics). The effect of external research income on assessed quality is not significant in one UoA (i.e., UoA1 Clinical Medicine) because large submitting units generate more income per FTE staff submitted, and when the size of the submissions are controlled for, external research income per FTE staff submitted is not significant for this unit[7].

One of the reasons for the significance of the size of submitting unit may be attributable to the fact that peer reviewers may have given more importance to absolute research income levels and PGR completions rather than research income per FTE staff submitted and PGR completions per FTE staff submitted (i.e., standardised variables). For instance, panel members may have rated environment higher where there were 40 staff and an income of £40 million than they would an environment with 4 staff and a £4 million income (both corresponding to £1 million per FTE staff submitted)[8]. To test this, we re-do the analysis in Table 5 where we replace income per FTE staff submitted and PGR completions per FTE staff submitted variables with the respective total figures (see supplementary file 1 for the detailed results). We find that most of the 'size' effect disappears once the total figures for research income and PGR completions are used, and the FTE staff submitted variable is only positively and highly (1% level) significant for 3 UoAs (i.e., UoA1 Clinical Medicine, UoA19 Business and Management Studies and UoA20 Law). In other words, in most of the UoAs, larger submissions tend to have better scores not because their size but they tend to generate more research income and PGR

---

[7] We carried out regression analysis for the UoA1 Clinical Medicine by excluding the size from the analysis and find that income per FTE staff submitted becomes a significant factor. In other words, submissions in this UoA generated more external research income because of their size. Once the size of the submissions is controlled for external income per FTE staff submitted did not play a significant role in environment scores. The results are available upon request from authors.

[8] The authors thank an anonymous referee for pointing out such a potential issue.

completions. This could be due to the fact that submissions over a certain size may also be indicative of the existence of a critical mass of researchers, which leads to enhanced GPA scores through its effect on the amount of external research income generation and PGR recruitment. For instance, a critical mass of researchers would allow enhanced internal peer review of external grant bids, which would then increase the chances of the submitting unit to obtain such grants. To put it differently, the 'size' of the submitting unit matters through its effect on the total amount of external research income generation and PGR completions. The significance of coefficients of other variables of interest tends to be similar irrespective of standardised or total values of research income and PGR completions used (i.e., significant variables are the same ones for the most of the UoAs) with four exceptions[9]. This is in line with Pidd and Broadbent (2015) who discussed that statistics may be misleading rather than helpful unless peer reviewers are clearly informed about ways in which they should be using quantitative factors in their evaluations.

The effects of other variables on assessed environmental quality show variation across different UoAs. For example, the number of PGR completions per FTE staff submitted plays a highly significant role (i.e., variable is significant at the 1% level) in the assessed quality in only 7 UoAs and has no significant effect on the assessed quality in 16 UoAs (see Table 5 for the details). Furthermore, when total PGR completions are considered in the analysis, PGR completions are highly significant in only 12 UoAs (see supplementary file 1 for details). As shown in Table 3, the correlation between PGR per FTE staff submitted and GPA were relatively weak compared to that of income per FTE staff submitted and size, which is also reflected in the regression analysis. In other words, once other factors are controlled for, the number of PGR completions per FTE staff submitted does not play a significant role in assessed quality in 16 UoAs. This suggests that even though the research income per FTE staff submitted plays a significant role consistently across different UoAs, some of the sub-panel assessors do not place significant importance on the number of PGR completions per FTE staff submitted. This finding suggests that what was found by Thorpe et al. (2018b) and Pidd and Broadbent (2015) with respect to Business and Management submissions bear out on a wider scale. Pidd and Broadbent (2015) argued that the number of FTE staff submitted to the REF does not reflect the actual number of staff members taking supervisory roles and that the Business and

---

[9] For instance, we find that total PGR completions are significant and positive for quality of environment in the UoA6 Agriculture, Veterinary and Food Science, UoA9 Physics, UoA30 History and UoA31 Classics when total research income and PGR completions are used. However, the coefficients of PGR per FTE staff submitted were not significant in these units when standardised variables are used.

Management panel undervalued the number of PGR completions per FTE staff submitted. Our findings show that a similar trend is observed in most of the other UoAs and is not limited to solely Business and Management panel, with panel members in other units do not consider PGR completions per FTE staff submitted as a significant factor in their evaluations either. Another reason why external income generation is significant, but the number of PGR completions component is not significant in some of the UoA evaluations is the intensity of competition for such components. Obtaining a research grant is highly competitive, and there is a relatively fixed sum of grants available. On the other hand, there is a bigger pool of PGR students, and some HEIs recruit lots of weaker PGRs while others are more selective in the admission process. Findings in Table 5 demonstrate that this is something that was taken into account by most peer reviewers while assigning research environment scores.

We also examine the potential 'halo effect' in the peer review evaluation (i.e., whether submitting units that belonged to the Russell Group or had panel members in the evaluation had any impact on the assessed quality of research environment). We find that if submitting units belonged to a Russell Group university, they tended to obtain relatively higher GPA scores in 14 UoAs as the dummy variable on RM is significant at the 5% level (see Table 5). For instance, after controlling for other quantitative factors, assessors that evaluated the UoA28 Modern Languages and Linguistics submissions allocated 0.405 higher environment scores to the submitting units belonging to the Russell Group universities compared to the submitting units from the non-Russell group universities. Similarly, after controlling for observable quantitative factors, we find that submitting units that had sub-panel members in the evaluation tended to have higher GPA scores in four UoAs as the PM dummy variable is significant at the 5% level (i.e., UoA8 Chemistry; UoA19 Business and Management Studies; UoA34 Art and Design: History, Practise and Theory; UoA35 Music, Drama, Dance and Performing Arts). For instance, submitting units that had sub-panel members in the UoA8 Chemistry achieved 0.307 higher GPA scores compared to submitting units that had no panel members in the evaluation. Overall, we find that reviewers in some sub-panels suffered from implicit bias when assessing the research environment. However, there is a good percentage of UoAs that were not affected by implicit bias as both dummy variables on RM and PM are not significant. Hence, results about the significance of Russel Group membership for the Business and Management panel by Taylor (2011) and Thorpe et al. (2018a) cannot be generalized to other panels as our analyses shows that this was not evident everywhere else.

Additional to significant coefficients, it is also essential to evaluate magnitudes of these coefficients because this highlights the relationship between a given quantitative factor and

GPA levels. For instance, a unit increase in PGR completions per FTE staff submitted in the UoA19 Business and Management Studies (UoA24 Anthropology and Development Studies) would increase GPA of that submitting unit by 0.095 (0.287). In other words, reviewers in the UoA19 Business and Management Studies tended to undervalue PGR completions per FTE staff submitted compared to the UoA24 Anthropology and Development Studies ones as a similar increase in PGR completions per FTE staff submitted would lead to different levels of increase in GPAs in these UoAs. On the other hand, doubling the external income generated per FTE staff submitted by submission in the UoA20 Law (UoA22 Social Work and Social Policy) would increase the GPA of that submission by 0.068 (0.345) suggesting that the valuation of the external funding generation by sub-panel members is different in these UoAs. Similarly, implicit bias in the evaluation process is also different across UoAs where reviewers in some UoAs are more biased than the ones in other UoAs (see the coefficients on the RM and PM in Table 5). In other words, the perceived importance of each factor tends to be different across different UoAs.

Overall, we observe that the explanatory power of quantitative factors (i.e., R-squares) varies across UoAs but these factors explain most of the variation in GPA scores within each UoA. One potential explanation for this is that some panel reviewers may have placed more importance on the environment template compared to others. Environment templates included additional information not controlled for in this paper (e.g., the number of research centres within the submitting unit, career development processes offered to staff members, participation and/or organisation of major conferences, major awards won by faculty members, editorial and other collegial services provided by staff members, international collaborations of the submitting unit and so on). These may have played a significant role in GPA scores achieved by submitting units beyond the distinctive language-related differences in environment templates found by Thorpe et al. (2018a, 2018b). We will expand on this discussion in the next subsection where we compare two units that have similar predicted GPA scores with our empirical findings, but reviewers allocated these submitting units different GPA scores.

### 4.4. How good are the quantitative factors in estimating the peer review evaluation?

To show the goodness of fit of the quantitative factors in predicting peer review evaluations, we use the significant factors from Table 5 to predict GPA scores obtained in environment submissions in two UoAs. Figures 1 and 2 provide the scatter plots of the predicted and actual GPAs for the submitting units in UoA12 Aeronautical, Mechanical, Chemical, and Manufacturing Engineering and UoA23 Sociology, respectively. We choose

these two UoAs because the explanatory powers of the quantitative factors in them are relatively high[10]. In these figures, we also integrated the 45-degree line (red line), which highlights the case of equality between predicted and actual GPA values. Results show that quantitative factors can be used to predict environment GPA levels in these two UoAs. Predicted GPA levels were below (above) actual GPA levels, but quantitative factors explain a good portion of the assessed quality of the research environment by peer reviewers in these two UoAs and may be used to decrease the associated workload and implicit bias in the peer review evaluation process.

<Insert Figure 1 appoximately here>

<Insert Figure 2 approximately here>

Quantitative factors examined in this paper explain most of the variation in GPA scores obtained by submitting units in UoA12 Aeronautical, Mechanical, Chemical and Manufacturing Engineering (Figure 1). However, some factors have not been controlled in this paper's empirical analyses, and they may also play a crucial role. We identified two submitting units from the UoA12 where there is a large difference between the predicted and actual scores. For example, our analysis predicts that submitting units of the University of Cambridge and University of Leeds would receive environment scores of 3.24 and 3.23, respectively. However, peer reviewers allocated them environment scores of 3.75 and 3.3, respectively. To shed light on how other factors may have played in the peer review evaluations, we carried out a deeper analysis into the submitted environment templates of both units. Both submitting units generated similar research incomes, and they have submitted similar numbers of FTE to the REF2014[11]. However, there was a large difference between the two in terms of their planned investments, conference organisations, editorial board positions, committee memberships, and international influence. The submitting unit of the University of Cambridge had 900 research papers, 500 presentations (including 10 plenaries, 50 keynotes), organised 12 conferences, chaired 3 international conferences, and the staff members held 30 editorial positions and took part in 21 national and international governance strategy and advisory committee during the period between 1 January 2008 and 31 July 2013. The submitting unit of the University of Leeds included 4 conference organizations, chaired 3 international conferences and the staff

---

[10] The scatter plots of predicted and actual GPAs for other UoAs can be provided by authors upon request. Obviously, the goodness of fit for some other UoAs are relatively lower highlighting the fact that other factors beyond the factors used in the regression analysis also contribute to the variation in GPA levels, and that one would observe relatively more variation across the 45-degree line the lower the R-squares.

[11] University of Cambridge submitted 32.30 FTE staff to this UoA and this submitting unit generated roughly £28 million external research income. Whereas, University of Leeds submitted 31.55 FTE staff and generated roughly £29 million external research income.

members held 20 editorial positions, acted as leading figures in 4 international committees, held 7 leadership positions and took part in 4 committees. Thus, despite roughly similar quantitavie data, the submitting unit of Cambridge University highlighted more activities in their environment template compared to that of the University of Leeds, leading to the allocation of a relatively better environment score to the former compared to the latter one.

Similarly, as displayed in Figure 2, predicted and actual scores are quite close to each other for many submitting units in the UoA23 Sociology. We repeated the analysis into environment templates where there was a large difference between predicted and actual environment GPA scores in two submitting units. Our empirical analysis predicts GPA scores of 3.22 and 3.21 for the units submitted by Lancaster University and the University of Warwick to the UoA23 Sociology; however, peer reviewers allocated GPA scores of 3.5 and 3.125 to these submitting units, respectively. When we look at the environment templates submitted by these two units, we found that Lancaster University reported more research clusters and more editorial positions in their environment templates compared to that by the University of Warwick. The submitting unit of Lancaster University reported 8 research clusters, with staff members who had delivered more than 99 plenary and keynote lectures in international conferences, published 45 monographs and special issues, conducted 22 seminars, edited/co-edited/associate edited 14 journals and books, and held 33 editorial board membership positions. The submitting unit of the University of Warwick reported 5 research clusters, with staff members who had delivered nearly 100 plenary lectures, hosted 4 seminar series and 2 international conferences, edited 8 academic journals and books and held 23 editorial board positions. Although the University of Warwick submitted 34.60 FTE staff members compared to 28.75 of Lancaster University, staff members of from Lancaster University carried out more activities compared to Warwick University.

Overall, even though quantitative data reported as part of environment data explains most of the variation in GPA scores achieved by submitting units, there is still room for submitting units to list other sets of activities in the environment template to obtain higher environment scores. The listing of these quantitative metrics in the environment templates was in the lines of the suggestions of the UK Forum for Responsible Research Metrics (2018b) as most of the HEIs were in favour of the introduction of more quantitative data while evaluating environment submissions to improve comparability and promote a more objective assessment of the research environments of submitting units. However, even though quantitative metrics are used by units in their REF2014 environment templates, it is important that the quality of these activities could also be assessed. For instance, in the case of UoA23 Sociology

submission of Lancaster University, the submitting unit listed the details of the journals in which its staff members were on editorial boards, and also provided examples of keynote lectures given by its staff members. However, the submitting unit of the University of Warwick only referred that they were on the editorial board of 23 journals and were given 100 plenary lectures at international conferences without providing any details of the journals in which they were on editorial board and international conferences in which they gave keynote lectures. Clearly, evaluating the quality of the research activities of the submitting unit of Lancaster University was easier than that of the University of Warwick as they provided more details on the activities that potentially enabled reviewers to evaluate the quality of these activities. In other words, even though providing quantitative metrics are helpful to improve comparability and promote a more objective assessment of the research environments, submitting units should also provide some specific details about these quantitative metrics so the panel members could evaluate the quality of such activities.

Finally, our results are in line with some previous findings on the importance of using key-language in environment submissions. For example, the set of key activities listed by both Lancaster and Warwick's submitting units, and the language used in these two submissions tended to confirm the readers' impression of activities conducted by 'reputable' institutions, and the difference between these and submissions by low-ranked institutions (Thorpe et al., 2018a, 2018b).

## 5. Conclusions and policy implications

In this paper, we evaluated whether the quantitative factors used in peer review evaluations of the research environment submission of the UK's REF2014 were useful in predicting assessed research environment quality. The main finding of this paper is that assessed environment submissions are significantly associated with a set of quantitative indicators such as external income, the size of the submitting unit, and PGR completions by the submitting unit. We find that with few exceptions, submitting units with high external research income generation tend to have better environment scores in all UoAs. We also find that PGR completions per FTE staff submitted played a highly significant role within 7 UoAs but did not play a significant role in explaining the quality of the research environment in 16 UoAs.

We expected PGR completions per FTE not playing a significant role in explaining the environment scores in REF2014. Pidd and Broadbent (2015) pointed out that the number of PGR completions within the unit did not reflect the actual number of staff members taking supervisory roles as not all of the staff members were returned to the REF2014. This led the

Business and Management panel to place less importance to PGR completions per FTE data. Furthermore, some institutes may recruit lots of weaker PGRs to increase their PGR completion numbers to improve their environment scores. Our findings are in line with those by Pidd and Broadbent (2015) and suggest that the panel members in other 16 UoAs units did not consider PGR completions per FTE as a significant factor in their environment evaluations (UoA2 Public Health, Health Services and Primary Care, UoA3 Allied Health Professions, Dentistry, Nursing and Pharmacy, UoA5 Biological Sciences, UoA6 Agriculture, Veterinary and Food Science, UoA7 Earth Systems and Environmental Sciences, UoA12 Aeronautical, Mechanical, Chemical and Manufacturing Engineering, UoA15 General Engineering, UoA16 Architecture, Built Environment and Planning, UoA20 Law, UoA22 Social Work and Social Policy, UoA26 Sport and Exercise Sciences, Leisure and Tourism, UoA27 Area Studies, UoA30 History, UoA31 Classics, UoA33 Theology and Religious Studies, and UoA36 Communication, Cultural and Media Studies, Library and Information Management). To overcome some of these problems, the Forum for Responsible Research Metrics provided some advice on the use of indicators in their institutional- and unit-level environment statements in the REF2021 submissions (see REF, 2019b for details). For instance, one suggestion is that submissions should provide the number of eligible or submitted staff, as well as details of percentages of staff members that are on permanent / fixed-term / atypical contracts (UK Forum for Responsible Research Metrics, 2018b). The provision of such data may overcome or minimize the problem discussed by Pidd and Broadbent (2015) as this would enable panel members to interpret the environment data better in the REF2021 evaluations. Additionally, results in this paper allows us to recommend that units should provide details that enable panel members to evaluate the quality of the research environment rather than using quantitative metrics only.

Beyond the relevance of quantitative factors, we also confirm the existence of 'halo effect' (implicit bias) in some UoAs where submitting units that belonged to Russell Group and had panel members in the REF exercise obtained higher scores even after controlling for other quantitative factors. Russell Group membership was significantly (5% level) and positively affected the GPA scores in 14 UoAs (i.e., UoA3 Allied Health Professions, Dentistry, Nursing and Pharmacy, UoA4 Psychology, Psychiatry and Neuroscience, UoA7 Earth Systems and Environmental Sciences, UoA8 Chemistry, UoA11 Computer Science and Informatics, UoA12 Aeronautical, Mechanical, Chemical and Manufacturing Engineering, UoA15 General Engineering, UoA16 Architecture, Built Environment and Planning, UoA26 Sport and Exercise Sciences, Leisure and Tourism, UoA28 Modern Languages and Linguistics, UoA33 Theology and Religious Studies, UoA34 Art and Design: History, Practise and Theory,

UoA35 Music, Drama, Dance and Performing Arts and UoA36 Communication, Cultural and Media Studies, Library and Information Management). Furthermore, HEIs that had panel members in the evaluation process achieved significantly (5% level) higher GPA scores in 4 UoAs (i.e., UoA8 Chemistry, UoA19 Business and Management Studies, UoA34 Art and Design: History, Practise and Theory, UoA35 Music, Drama, Dance and Performing Arts). Our findings suggest that the list of panels above should be conscious about the existence of implicit bias in the REF2021 evaluations. On the other hand, there was no implicit bias (i.e., neither Russell group membership nor panel membership played a significant role in GPA scores) in 18 UoAs (i.e., UoA1 Clinical Medicine, UoA2 Public Health, Health Services and Primary Care, UoA5 Biological Sciences, UoA6 Agriculture, Veterinary and Food Science, UoA9 Physics, UoA10 Mathematical Sciences, UoA14 Civil and Construction Engineering, UoA18 Economics and Econometrics, UoA20 Law, UoA21 Politics and International Studies, UoA23 Sociology, UoA24 Anthropology and Development Studies, UoA25 Education, UoA27 Area Studies, UoA29 English Language and Literature, UoA30 History, UoA31 Classics and UoA32 Philosophy). This suggests that training on implicit biases in the REF2014 paid off in these UoAs. A potential recommendation steaming out of these findings is that funding bodies responsible for REF2021 could invite members from panels with no implicit bias to share their experiences and help training panel members responsible of REF2021 evaluations.

This paper also demonstrates that importance given to similar quantitative factors by reviewers are distinctively different across UoAs. In other words, similar PGR completions per FTE staff submitted, or similar external income generated per FTE staff submitted, increased GPA scores differently across different UoAs. This could be relevant for policymakers in HEIs who are responsible for returning submissions to the REF2021 or the following REF exercises. For instance, based on the findings of this paper, HEIs may choose to prioritise allocation of funds to specific fields (e.g., allocation of scholarship and bursaries for graduate studies, funds for internal and external training of staff members to improve external bid success rates) in which return to such investment is relatively higher in terms of improving research environment scores in the REF exercises. If this type of strategic actions is not preferred option for the funding bodies that are running the future REF exercises, it is essential that they implement some benchmarking of quantitative data for different subjects or panels to avoid such inconsistencies across different UoAs.

Our results can be of interest to countries with performance-based research funding systems. For instance, countries that are using PGR degree completions in their performance evaluation metrics (e.g., Austria, Netherlands, Belgium, Denmark, Estonia, Finland, Croatia,

Norway, Poland, Slovakia and Italy) should be cautious as the number of completions itself does not capture the quality of research environment as some HEIs may recruit PGR students to increase their performance outcomes without necessarily paying attention to the quality of students. Instead, such countries may consider collecting proxies that would enable them to assess the 'quality' of these PGR completions such as postgraduate employment earnings data or the number of PGR students with highly skilled employment after their graduation. On the other hand, countries that use peer review evaluation (e.g., France, Italy, Lithuania and Portugal) should be cautious about the potential 'halo' effect in the evaluation process, and the differences between subject areas.

Finally, some of the submission guidelines of the environment template for REF2021 have changed compared to REF2014. One of the most important changes is that institutions are required to submit an institutional-level environment statement (see paragraphs between 359 and 365 of REF, 2019a). Institutional-level environment templates will consist of information on the size, structure, and mission of the institution, institution's strategy, people, income, infrastructure, and facilities. This statement is planned to be appended to each unit-level template and will be submitted for review but will not be scored separately by the sub-panels. Empirical findings in this paper suggest that the inclusion of such an institutional-level template may further reinforce the already existing 'halo' effect. Institutions with an established reputation may further capitalize on their pre-existing positions to take advantage of the implementation of institutional-level templates (see Thorpe et al., 2018a, 2018b). On the other hand, institutional-level templates may enable reviewers to assess whether the units are able to achieve their strategic missions based on the institutional environment. For instance, institutional-level facilities and institution's strategy would be key for units to achieve interdisciplinary research, diversity and equality, and non-academic impact. Institutions that are relatively large generate more external research income and impact (Pinar and Unlu, 2019), and have relatively more interdisciplinary research centres. Hence being part of relatively large institution may enable units to realize better research environments, which potentially can be reflected better in the unit-level environment templates by referring to the institutional-level templates. In short, our expectation is that the inclusion of institutional-level template may potentially improve the research environment scores of the units that belong to relatively large and reputed institutions.

**References:**

Ancaiani, A., Anfossi, A. F., Barbara, A., Benedetto, S., Blasi, B., Carletti, V., Cicero, T., Ciolfi, A., Costa, F., Colizza, G., Costantini, M., di Cristina, F., Ferrara, A., Lacatena, R. M., Malgarini, M., Mazzotta, I., Nappi, C. A., Romagnosi, S., and Sileoni S. (2015). Evaluating scientific research in Italy: The 2004–10 research evaluation exercise. *Research Evaluation, 24*(3), 242-255

Baccini, A., De Nicolao, G. (2016). Do they agree? Bibliometric evaluation versus informed peer review in the Italian research assessment exercise. *Scientometrics*, 108(3), 1651–1671

Bertocchi, G., Gambardella, A., Jappelli, T., Nappi, C. A., and Peracchi, F. (2015). Bibliometric evaluation vs. informed peer review: evidence from Italy. *Research Policy*, 44(2), 451-466.

Bornmann, L. (2011). Scientific peer review. *Annual Review of Information Science and Technology*, 45, 199-245.

Bruns, S.B., Stern, D.I. (2016). Research assessment using early citation information. *Scientometrics*, 108 (2), 917–935.

De Boer, H., Jongbloed, B., Benneworth, P., Cremonini, L., Kolster, R., Kottmann, A., Lemmens-Krug, K., and Vossensteyn, H. (2015). *Performance-based funding and performance agreements in fourteen higher education systems. Report for the Ministry of Education, Culture and Science.* The Hague: Ministry of Education, Culture and Science.

Dougherty, K. J., Jones, S. M., Lahr, H., Natow, R. S., Pheatt, L., and Reddy, V. (2016). *Performance funding for higher education*. Baltimore: Johns Hopkins University Press.

Garcia, J.A., Rodriguez- Sánchez, R., & Fdez- Valdivia, J. (2015) Bias and effort in peer review. *Journal of the Association of Information Science and Technology*, 66, 2020–2030.

Geuna, A., Piolatto, M. (2016). Research assessment in the UK and Italy: Costly and difficult, but probably worth it (at least for a while). *Research Policy*, 45(1), 260-271.

HEFCE (2015). The Metric Tide: Correlation analysis of REF2014 scores and metrics (Supplementary Report II to the Independent Review of the Role of Metrics in Research Assessment and Management). HEFCE. DOI: 10.13140/RG.2.1.3362.4162

Hicks, D. (2012). Performance-based university research funding systems. R*esearch Policy*, 41(2), 251-261.

Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., and Rafols, I. (2015). Bibliometrics: The Leiden Manifesto for research metrics. *Nature, 520,*429–431.

Jonkers, K., Zacharewicz, T. (2016). *Research Performance Based Funding Systems: a Comparative Assessment* (JRC Science For Policy Report). European Commission: Joint Research Centre.

Larivière, V., Gingras, Y., Sugimoto, C. R., and Tsou, A. (2015). Team size matters: Collaboration and scientific impact since 1900. *Journal of the Association for Information Science and Technology*, 66(7), 1323-1332.

Lee, C. J., Sugimoto, C. R., Zhang, G., and Cronin, B. (2013). Bias in peer review. *Journal of the American Society for Information Science and Technology*, **64**(1), 2-17.

Macdonald, S., Kam, J. (2007). Ring a ring o'roses: Quality journals and gamesmanship in management studies. *Journal of Management Studies*, 44 (4), 640-655.

Martin, B.R. (2011). The Research Excellence Framework and the 'impact agenda': are we creating a Frankenstein monster? *Research Evaluation*, 20(3), 247-254.

Mryglod, O., Kenna, R., Holovatch, Y., and Berche, B. (2013). Absolute and specific measures of research group excellence. *Scientometrics*, 95(1), 115-127.

Mryglod, O., Kenna, R., Holovatch, Y., and Berche, B. (2015). Predicting results of the research excellence framework using departmental h-index. *Scientometrics*, 102(3), 2165–2180.

Pidd M., and Broadbent J. (2015) Business and management studies in the 2014 Research Excellence Framework. *British Journal of Management*, 26, 569-581

Pinar, M., and Unlu, E. (2020). Evaluating the potential effect of the increased importance of the impact component in the Research Excellence Framework of the UK. British Educational Research Journal, 46(1), 140-160.

Rebora, G., Turri, M., (2013). The UK and Italian research assessment exercises face to face. *Research Policy*, 42(9), 1657-1666.

REF (2011). Assessment framework and guidance on submissions. Available at https://www.ref.ac.uk/2014/media/ref/content/pub/assessmentframeworkandguidanceonsubmissions/GOS%20including%20addendum.pdf

REF (2012a) Panel Criteria and Working Methods. Available at https://www.ref.ac.uk/2014/pubs/2012-01/#d.en.69569

REF (2012b) Part 2A Main Panel A criteria.  Available at https://www.ref.ac.uk/2014/media/ref/content/pub/panelcriteriaandworkingmethods/01_12_2A.pdf

REF (2012c) Part 2B Main Panel B criteria.  Available at https://www.ref.ac.uk/2014/media/ref/content/pub/panelcriteriaandworkingmethods/01_12_2B.pdf

REF (2012d) Part 2C Main Panel C criteria.  Available at
https://www.ref.ac.uk/2014/media/ref/content/pub/panelcriteriaandworkingmethods/01_12_2C.pdf

REF (2012e) Part 2D Main Panel D criteria.  Available at
https://www.ref.ac.uk/2014/media/ref/content/pub/panelcriteriaandworkingmethods/01_12_2D.pdf

REF (2014) Results and submissions database. Available at
https://results.ref.ac.uk/(S(mevbbiplnrg5u1yebd0vb2dr))/

REF (2019a) Guidance on submission. Available at
https://www.ref.ac.uk/media/1092/ref-2019_01-guidance-on-submissions.pdf

REF (2019b) Guidance for institutions on environment indicators. Available at
https://www.ref.ac.uk/media/1019/guidance-on-environment-indicators.pdf (Accessed on
November 15, 2019).

Research England (2018). Guide to research and knowledge exchange funding 2018-19. How
Research England allocates its funds. Available at
https://re.ukri.org/documents/2018/research-england-guide-to-research-and-knowledge-
exchange-funding-2018-19/

Shattock, M. 2012. *Making Policy in British Higher Education 1945-2011*. Berkshire:
McGraw-Hill.

Stern, N. (2016). *Building on Success and Learning from Experience: An Independent
Review of the Research Excellence Framework led by Lord Stern*. London: Department for
Business, Energy and Industrial Strategy. Available at
https://www.gov.uk/government/publications/research-excellence-framework-review
(accessed March 4, 2019)

Taylor, J. (2011). The assessment of research quality in UK universities: peer review or
metrics? *British Journal of Management*, 22(2), 202-217.

Thorpe, A., Craig, R., Hadikin, G., and Batistic, S. (2018a). Semantic tone of research
'environment' submissions in the UK's Research Evaluation Framework 2014. *Research
Evaluation*, 27(2), 53-62.

Thorpe, A., Craig, R., Tourish, D., Hadikin, G., and Batistic, S. (2018b). 'Environment'
Submissions in the UK's Research Excellence Framework 2014. *British Journal of
Management*, 29(3), 571-587.

UK Forum for Responsible Research Metrics (2018a). UK Progress towards the use of
metrics responsibly. Available via: https://www.universitiesuk.ac.uk/policy-and-
analysis/research-policy/open-
science/The%20Forum%20for%20Responsible%20Research%20Metrics/UK%20progress%
20towards%20the%20use%20of%20metrics%20responsibly%2010072018.pdf

UK Forum for Responsible Research Metrics (2018b). FFRRM's advice to the Research Excellence Framework (REF) 2021 panels. Available at https://www.universitiesuk.ac.uk/policy-and-analysis/research-policy/open-science/The%20Forum%20for%20Responsible%20Research%20Metrics/FFRRM%20advice%20to%20REF2021%20panels%20on%20impact%20and%20environment%20indicators%2026July2018.pdf (Accessed on November 15, 2019).

Wang, J., Veugelers, R., Stephan, P. (2017). Bias against novelty in science: A cautionary tale for users of bibliometric indicators. *Research Policy*, 46(8), 1416-1436.

Wooldridge, J., King, M.B. (2019) Altmetric scores: An early indicator of research impact. *Journal of the Association for Information Science and Technology*, 70(3), 271-282.

Zacharewicz, T., Lepori, B., Reale, E., and Jonkers, K. (2019). Performance-based research funding in EU Member States-a comparative assessment. *Science and Public Policy*, 46(1), 105-115.

| Table1: Percentage contributions of each component to the research environment evaluation in each panel | | | | |
|---|---|---|---|---|
| | (i) | (ii) | (iii) | (iv) |
| Panel A | 25 | 25 | 25 | 25 |
| Panel B | 20 | 30 | 30 | 20 |
| Panel C | 25 | 25 | 25 | 25 |
| Panel D | 20 | 40 | 20 | 20 |

Notes: The weight in people component for Panel D is equally divided into staffing strategy and staff development (20%), and research students (20%).

| UoA No | UoA name | Weighted average GPA | Total income per FTE staff submitted | FTE staff submitted | PGR completions per FTE submitted | Number of submissions | Number of RM | Number of PM |
|---|---|---|---|---|---|---|---|---|
| | **Table 2. Descriptive statistics for each UoA and Panel** | | | | | | | |
| 1 | Clinical Medicine | 3.55 | £1,670,735 | 3,570.9 | 2.19 | 30 | 21 (70%) | 19 (63%) |
| 2 | Public Health, Health Services and Primary Care | 3.66 | £1,136,220 | 1,354.3 | 1.24 | 31 | 21 (68%) | 19 (61%) |
| 3 | Allied Health Professions, Dentistry, Nursing and Pharmacy | 3.35 | £334,208 | 2,744.7 | 1.80 | 90 | 20 (22%) | 36 (40%) |
| 4 | Psychology, Psychiatry and Neuroscience | 3.43 | £521,039 | 2,516.8 | 2.24 | 81 | 22 (27%) | 20 (25%) |
| 5 | Biological Sciences | 3.50 | £1,005,414 | 2,373.3 | 2.76 | 43 | 22 (51%) | 16 (37%) |
| 6 | Agriculture, Veterinary and Food Science | 3.57 | £642,545 | 1,042.2 | 1.69 | 27 | 10 (37%) | 16 (59%) |
| | Panel A | 3.49 | £940,238 | 13,602.3 | 2.09 | 310 | 116 (37%) | 126 (41%) |
| 7 | Earth Systems and Environmental Sciences | 3.22 | £589,267 | 1,380.1 | 1.79 | 43 | 18 (41%) | 15 (34%) |
| 8 | Chemistry | 3.24 | £823,378 | 1,229.1 | 3.85 | 35 | 21 (57%) | 15 (41%) |
| 9 | Physics | 3.36 | £1,410,791 | 1,703.6 | 2.10 | 40 | 22 (55%) | 15 (38%) |
| 10 | Mathematical Sciences | 3.35 | £177,194 | 1,930.3 | 1.30 | 52 | 24 (46%) | 23 (44%) |
| 11 | Computer Science and Informatics | 2.91 | £373,398 | 2,044.2 | 2.04 | 89 | 23 (26%) | 20 (22%) |
| 12 | Aeronautical, Mechanical, Chemical and Manufacturing Engineering | 3.27 | £728,508 | 1,152.0 | 3.07 | 25 | 12 (48%) | 14 (56%) |
| 13 | Electrical and Electronic Engineering, Metallurgy and Materials | 3.13 | £761,716 | 1,070.8 | 3.49 | 37 | 19 (51%) | 16 (43%) |
| 14 | Civil and Construction Engineering | 3.26 | £430,013 | 390.5 | 2.12 | 14 | 9 (64%) | 5 (36%) |
| 15 | General Engineering | 3.26 | £515,893 | 2,446.9 | 2.23 | 61 | 18 (29%) | 17 (27%) |
| | Panel B | 3.22 | £630,771 | 13,347.4 | 2.32 | 402 | 166(41%) | 140 (35%) |
| 16 | Architecture, Built Environment and Planning | 3.19 | £216,962 | 1,024.8 | 1.38 | 44 | 12 (27%) | 15 (33%) |
| 17 | Geography, Environmental Studies and Archaeology | 3.16 | £226,279 | 1,685.6 | 1.48 | 74 | 35 (47%) | 34 (46%) |
| 18 | Economics and Econometrics | 3.15 | £165,617 | 755.7 | 1.44 | 28 | 16 (57%) | 11 (39%) |
| 19 | Business and Management Studies | 3.11 | £102,246 | 3,311.9 | 1.44 | 98 | 24 (24%) | 24 (24%) |
| 20 | Law | 3.29 | £48,216 | 1,551.4 | 1.02 | 66 | 23 (35%) | 25 (38%) |
| 21 | Politics and International Studies | 3.17 | £109,243 | 1,271.7 | 1.73 | 55 | 23 (42%) | 17 (31%) |
| 22 | Social Work and Social Policy | 3.09 | £187,557 | 1,301.9 | 1.19 | 62 | 15 (24%) | 21 (34%) |
| 23 | Sociology | 3.14 | £223,711 | 703.6 | 1.71 | 29 | 13 (45%) | 13 (45%) |
| 24 | Anthropology and Development Studies | 3.36 | £229,850 | 561.6 | 2.01 | 25 | 11 (44%) | 13 (52%) |
| 25 | Education | 3.13 | £202,187 | 1,441.8 | 2.51 | 76 | 21 (28%) | 22 (29%) |
| 26 | Sport and Exercise Sciences, Leisure and Tourism | 3.06 | £102,004 | 786.7 | 1.17 | 49 | 6 (12%) | 9 (18%) |
| | Panel C | 3.16 | £146,107 | 14,396.6 | 1.53 | 608 | 199 (33%) | 204 (34%) |
| 27 | Area Studies | 3.17 | £94,784 | 483.1 | 1.45 | 23 | 11 (48%) | 10 (43%) |
| 28 | Modern Languages and Linguistics | 3.23 | £71,994 | 1,379.6 | 1.23 | 54 | 26 (48%) | 24 (44%) |
| 29 | English Language and Literature | 3.23 | £34,613 | 1,969.3 | 1.35 | 88 | 22 (25%) | 22 (25%) |
| 30 | History | 3.23 | £65,015 | 1,780.5 | 1.27 | 81 | 23 (28%) | 22 (27%) |
| 31 | Classics | 3.28 | £78,856 | 382.6 | 1.20 | 22 | 16 (73%) | 12 (55%) |
| 32 | Philosophy | 3.27 | £68,085 | 590.6 | 1.21 | 40 | 21 (53%) | 14 (35%) |
| 33 | Theology and Religious Studies | 3.12 | £66,537 | 410.7 | 3.19 | 32 | 15 (47%) | 10 (31%) |
| 34 | Art and Design: History, Practise and Theory | 3.18 | £89,434 | 1,603.8 | 1.04 | 82 | 15 (18%) | 23 (28%) |
| 35 | Music, Drama, Dance and Performing Arts | 3.09 | £46,731 | 1,140.3 | 1.15 | 82 | 27 (33%) | 25 (30%) |
| 36 | Communication, Cultural and Media Studies, Library and Information Management | 3.11 | £68,680 | 932.1 | 1.03 | 67 | 10 (15%) | 17 (25%) |
| | Panel D | 3.20 | £64,415 | 10,672.4 | 1.29 | 574 | 186 (32%) | 179 (31%) |

Notes: UoAs between 1-6 are included in Main Panel A, 7-15 are in Main Panel B, 16-26 are in Main Panel C, and 27-36 are in Main Panel D. Average GPA scores for each UoA are obtained by using the number of FTE staff submitted for a given UoA as weights. Total income per FTE staff submitted and PGR per FTE staff submitted are obtained by summing all research income and doctoral degrees within each UoA and standardizing these with the total number of FTE staff members submitted in each UoA. The number of submissions provides the total number of submissions in each UoA where joint submissions are considered as one submission. The number of Russell Group membership (RM) submissions provides the number of submissions in each UoA by the Russell Group universities. Finally, the number of panel membership (PM) provides the number of submissions that had panel members in the evaluation.

| Table 3. Correlation matrix of quantitative factors and environment GPA scores by each panel | | | | |
|---|---|---|---|---|
| **Panel A: Medicine, health and life sciences** | | | | |
| | GPA | PGR/FTE staff submitted | ln(Income/FTE staff submitted) | ln(FTE staff submitted) |
| GPA | 1.00 | | | |
| PGR/FTE staff submitted | 0.35 | 1.00 | | |
| ln (Income/FTE staff submitted) | 0.67 | 0.34 | 1.00 | |
| ln (FTE staff submitted) | 0.71 | 0.25 | 0.58 | 1.00 |
| **Panel B: Physical sciences, engineering, and mathematics** | | | | |
| | GPA | PGR/FTE staff submitted | ln(Income/FTE staff submitted) | ln(FTE staff submitted) |
| GPA | 1.00 | | | |
| PGR/FTE staff submitted | 0.43 | 1.00 | | |
| ln (Income/FTE staff submitted) | 0.62 | 0.50 | 1.00 | |
| ln (FTE staff submitted) | 0.78 | 0.26 | 0.51 | 1.00 |
| **Panel C: Social sciences** | | | | |
| | GPA | PGR/FTE staff submitted | ln(Income/FTE staff submitted) | ln(FTE staff submitted) |
| GPA | 1.00 | | | |
| PGR/FTE staff submitted | 0.35 | 1.00 | | |
| ln (Income/FTE staff submitted) | 0.53 | 0.34 | 1.00 | |
| ln (FTE staff submitted) | 0.73 | 0.13 | 0.35 | 1.00 |
| **Panel D: Arts and humanities** | | | | |
| | GPA | PGR/FTE staff submitted | ln(Income/FTE staff submitted) | ln(FTE staff submitted) |
| GPA | 1.00 | | | |
| PGR/FTE staff submitted | 0.33 | 1.00 | | |
| ln (Income/FTE staff submitted) | 0.63 | 0.23 | 1.00 | |
| ln (FTE staff submitted) | 0.59 | 0.09 | 0.37 | 1.00 |
| Notes: Panels A, B, C, and D consist of 302, 396, 606 and 571 submissions, respectively. | | | | |

| Table 4. Determinants of quality of research environment with pooled estimations | | | | |
|---|---|---|---|---|
| VARIABLES | (1) | (2) | (3) | (4) |
| PGR per FTE staff submitted | 0.0605*** | 0.102*** | 0.0312*** | 0.0661*** |
| | (0.00975) | (0.0132) | (0.00861) | (0.0111) |
| ln(Income per FTE staff submitted) | 0.0613*** | 0.138*** | 0.0230*** | 0.0855*** |
| | (0.00719) | (0.0123) | (0.00691) | (0.0115) |
| ln(FTE staff submitted) | 0.359*** | 0.436*** | 0.275*** | 0.342*** |
| | (0.0205) | (0.0196) | (0.0191) | (0.0188) |
| Russell Group member (RM) | 0.342*** | 0.259*** | | |
| | (0.0270) | (0.0235) | | |
| Panel member (PM) | 0.230*** | 0.151*** | | |
| | (0.0266) | (0.0227) | | |
| Observations | 1,790 | 1,790 | 1,790 | 1,790 |
| UoA dummies | NO | YES | NO | YES |
| HEI dummies | NO | NO | YES | YES |
| F-statistic for UoA dummies | | 20.67*** | | 21.60*** |
| F-statistic for HEI dummies | | | 8.91*** | 6.95*** |
| R-square | 0.550 | 0.688 | 0.673 | 0.771 |
| Notes: Robust standard errors in parentheses. Constant, UoA and HEI dummies are included but not reported. HEIs that had less than 6 UoA submissions are excluded from the analysis. ***, ** and * represent significance level at the 1%, 5% and 10%, respectively. UoA 36 and the University of Leeds were treated as reference points when UoA and HEI dummies are included, respectively. | | | | |

| UoA no | UoA name | PGR completions per FTE staff submitted | Income per FTE staff submitted | FTE staff submitted | RM | PM | Obs. | R-square |
|---|---|---|---|---|---|---|---|---|
| | Table 5. Determinants of quality of research environment within each UoA | | | | | | | |
| 1 | Clinical Medicine | -0.250** | 0.134 | 0.509*** | 0.061 | -0.136 | 30 | 0.717 |
| 2 | Public Health, Health Services and Primary Care | 0.008 | 0.351* | 0.378*** | 0.147 | 0.115 | 31 | 0.661 |
| 3 | Allied Health Professions, Dentistry, Nursing and Pharmacy | 0.006 | 0.323*** | 0.498*** | 0.344*** | 0.086 | 90 | 0.704 |
| 4 | Psychology, Psychiatry and Neuroscience | 0.055** | 0.113** | 0.458*** | 0.366*** | 0.007 | 81 | 0.786 |
| 5 | Biological Sciences | 0.069 | 0.251*** | 0.448*** | 0.140 | -0.046 | 43 | 0.848 |
| 6 | Agriculture, Veterinary and Food Science | 0.124 | 0.327*** | 0.375* | -0.049 | 0.445 | 27 | 0.806 |
| 7 | Earth Systems and Environmental Sciences | -0.051 | 0.253*** | 0.383*** | 0.150** | 0.073 | 43 | 0.807 |
| 8 | Chemistry | 0.093*** | 0.110 | 0.360*** | 0.181** | 0.307*** | 35 | 0.869 |
| 9 | Physics | 0.182*** | 0.158** | 0.332*** | 0.010 | 0.036 | 40 | 0.865 |
| 10 | Mathematical Sciences | 0.200** | 0.043 | 0.508*** | -0.009 | 0.131 | 52 | 0.868 |
| 11 | Computer Science and Informatics | 0.089* | 0.272*** | 0.490*** | 0.304** | 0.115 | 89 | 0.778 |
| 12 | Aeronautical, Mechanical, Chemical and Manufacturing Engineering | -0.018 | 0.412*** | 0.391*** | 0.243** | 0.146* | 25 | 0.913 |
| 13 | Electrical and Electronic Engineering, Metallurgy and Materials | 0.130** | 0.351** | 0.217** | 0.224* | 0.165 | 37 | 0.847 |
| 14 | Civil and Construction Engineering | 0.333* | 0.245 | 0.746* | -0.218 | 0.129 | 14 | 0.593 |
| 15 | General Engineering | 0.089 | 0.370*** | 0.377*** | 0.323** | -0.037 | 61 | 0.867 |
| 16 | Architecture, Built Environment and Planning | 0.108 | 0.302*** | 0.499*** | 0.320*** | 0.006 | 44 | 0.867 |
| 17 | Geography, Environmental Studies and Archaeology | 0.134** | 0.331*** | 0.630*** | 0.242* | 0.010 | 74 | 0.876 |
| 18 | Economics and Econometrics | -0.169** | 0.342*** | 0.294** | 0.202 | 0.098 | 28 | 0.751 |
| 19 | Business and Management Studies | 0.095* | 0.145*** | 0.612*** | 0.183* | 0.272*** | 98 | 0.771 |
| 20 | Law | 0.065 | 0.068** | 0.562*** | 0.165 | 0.026 | 66 | 0.717 |
| 21 | Politics and International Studies | 0.133** | 0.245*** | 0.546*** | 0.127 | 0.152 | 55 | 0.745 |
| 22 | Social Work and Social Policy | 0.074 | 0.345*** | 0.629*** | 0.265* | 0.184 | 62 | 0.776 |
| 23 | Sociology | 0.254*** | 0.191*** | 0.463*** | 0.196 | 0.050 | 29 | 0.909 |
| 24 | Anthropology and Development Studies | 0.287** | 0.144** | 0.203 | -0.150 | 0.310 | 25 | 0.648 |
| 25 | Education | 0.099*** | 0.322*** | 0.614*** | 0.175 | 0.174 | 76 | 0.821 |
| 26 | Sport and Exercise Sciences, Leisure and Tourism | 0.154 | 0.269*** | 0.562*** | 0.299** | -0.118 | 49 | 0.819 |
| 27 | Area Studies | -0.043 | 0.287*** | 0.210* | 0.039 | 0.225 | 23 | 0.693 |
| 28 | Modern Languages and Linguistics | 0.101** | 0.249*** | 0.265*** | 0.405*** | -0.027 | 54 | 0.579 |
| 29 | English Language and Literature | 0.287*** | 0.097*** | 0.582*** | 0.127 | 0.100 | 88 | 0.814 |
| 30 | History | 0.232 | 0.108** | 0.332*** | 0.113 | 0.073 | 81 | 0.723 |
| 31 | Classics | 0.265 | 0.083 | 0.261** | 0.227 | 0.188 | 22 | 0.701 |
| 32 | Philosophy | 0.230** | 0.111** | 0.639*** | 0.030 | -0.032 | 40 | 0.777 |
| 33 | Theology and Religious Studies | 0.010 | 0.080* | 0.627*** | 0.443** | -0.030 | 32 | 0.656 |
| 34 | Art and Design: History, Practise and Theory | 0.301*** | 0.135*** | 0.403*** | 0.359*** | 0.220** | 82 | 0.701 |
| 35 | Music, Drama, Dance and Performing Arts | 0.267*** | 0.095*** | 0.378*** | 0.344*** | 0.218** | 82 | 0.721 |
| 36 | Communication, Cultural and Media Studies, Library and Information Management | 0.108 | 0.090*** | 0.495** | 0.518*** | 0.107 | 67 | 0.589 |

Notes: ***, ** and * represent significance level at the 1%, 5% and 10%, respectively.

| UoA | UoA name | PGR completions | Income | FTE staff submitted | RM | PM | Obs. | R-square |
|---|---|---|---|---|---|---|---|---|
| | Supplementary file 1. Determinants of quality of research environment within each UoA when total PGR completions and research income used in analysis | | | | | | | |
| 1 | Clinical Medicine | -0.574** | 0.186 | 0.892*** | 0.099 | -0.149 | 30 | 0.734 |
| 2 | Public Health, Health Services and Primary Care | -0.018 | 0.353* | 0.037 | 0.146 | 0.124 | 31 | 0.662 |
| 3 | Allied Health Professions, Dentistry, Nursing and Pharmacy | 0.064 | 0.307*** | 0.127 | 0.341*** | 0.079 | 90 | 0.706 |
| 4 | Psychology, Psychiatry and Neuroscience | 0.182*** | 0.094** | 0.153 | 0.343*** | 0.018 | 81 | 0.800 |
| 5 | Biological Sciences | 0.028 | 0.272** | 0.161 | 0.153 | -0.070 | 43 | 0.843 |
| 6 | Agriculture, Veterinary and Food Science | 0.417*** | 0.227** | -0.255 | -0.084 | 0.457* | 27 | 0.859 |
| 7 | Earth Systems and Environmental Sciences | -0.044 | 0.246** | 0.187* | 0.137* | 0.080 | 43 | 0.804 |
| 8 | Chemistry | -0.029 | 0.191 | 0.223 | 0.181* | 0.353*** | 35 | 0.843 |
| 9 | Physics | 0.201 | 0.162** | -0.043 | 0.047 | 0.063 | 40 | 0.840 |
| 10 | Mathematical Sciences | 0.200** | 0.046 | 0.268** | -0.012 | 0.138* | 52 | 0.865 |
| 11 | Computer Science and Informatics | 0.147 | 0.267*** | 0.062 | 0.306** | 0.125 | 89 | 0.776 |
| 12 | Aeronautical, Mechanical, Chemical and Manufacturing Engineering | -0.027 | 0.407*** | 0.013 | 0.237** | 0.150 | 25 | 0.912 |
| 13 | Electrical and Electronic Engineering, Metallurgy and Materials | 0.372*** | 0.358*** | -0.521** | 0.220* | 0.143 | 37 | 0.861 |
| 14 | Civil and Construction Engineering | 0.904** | 0.023 | -0.170 | -0.087 | 0.065 | 14 | 0.628 |
| 15 | General Engineering | 0.102 | 0.385*** | -0.125 | 0.361** | -0.039 | 61 | 0.864 |
| 16 | Architecture, Built Environment and Planning | 0.116 | 0.293*** | 0.096 | 0.326*** | 0.012 | 44 | 0.869 |
| 17 | Geography, Environmental Studies and Archaeology | 0.207** | 0.324*** | 0.083 | 0.237* | 0.000 | 74 | 0.877 |
| 18 | Economics and Econometrics | -0.245* | 0.346*** | 0.188 | 0.193 | 0.094 | 28 | 0.749 |
| 19 | Business and Management Studies | 0.132 | 0.146*** | 0.313*** | 0.199* | 0.271*** | 98 | 0.769 |
| 20 | Law | 0.129 | 0.062** | 0.376*** | 0.135 | -0.001 | 66 | 0.725 |
| 21 | Politics and International Studies | 0.287*** | 0.176** | 0.103 | 0.107 | 0.159 | 55 | 0.769 |
| 22 | Social Work and Social Policy | 0.036 | 0.353*** | 0.246 | 0.273* | 0.181 | 62 | 0.774 |
| 23 | Sociology | 0.246** | 0.196** | -0.035 | 0.171 | 0.148 | 29 | 0.890 |
| 24 | Anthropology and Development Studies | 0.558*** | 0.133** | -0.510* | -0.126 | 0.302 | 25 | 0.652 |
| 25 | Education | 0.279*** | 0.270*** | -0.013 | 0.253* | 0.205** | 76 | 0.837 |
| 26 | Sport and Exercise Sciences, Leisure and Tourism | 0.130 | 0.264*** | 0.162 | 0.286** | -0.077 | 49 | 0.816 |
| 27 | Area Studies | 0.007 | 0.277** | -0.078 | 0.021 | 0.205 | 23 | 0.690 |
| 28 | Modern Languages and Linguistics | 0.190* | 0.239*** | -0.179 | 0.367*** | -0.023 | 54 | 0.596 |
| 29 | English Language and Literature | 0.309*** | 0.095*** | 0.166 | 0.125 | 0.085 | 88 | 0.821 |
| 30 | History | 0.343*** | 0.085** | -0.087 | 0.085 | 0.049 | 81 | 0.758 |
| 31 | Classics | 0.353*** | 0.009 | -0.073 | 0.199 | 0.165 | 22 | 0.748 |
| 32 | Philosophy | 0.215** | 0.119** | 0.293* | 0.000 | -0.027 | 40 | 0.770 |
| 33 | Theology and Religious Studies | 0.175 | 0.064 | 0.348 | 0.391** | -0.018 | 32 | 0.685 |
| 34 | Art and Design: History, Practise and Theory | 0.416*** | 0.126*** | -0.152 | 0.278*** | 0.132 | 82 | 0.735 |
| 35 | Music, Drama, Dance and Performing Arts | 0.323*** | 0.076*** | -0.031 | 0.311*** | 0.174* | 82 | 0.752 |
| 36 | Communication, Cultural and Media Studies, Library and Information Management | 0.107 | 0.089*** | 0.290 | 0.526*** | 0.124 | 67 | 0.589 |

Notes: ***, ** and * represent significance level at the 1%, 5% and 10%, respectively.