



Do research performances of universities and disciplines in England converge or diverge? An assessment of the progress between research excellence frameworks in 2014 and 2021

Mehmet Pinar^{1,2} 

Received: 8 February 2023 / Accepted: 20 July 2023 / Published online: 5 August 2023
© The Author(s) 2023

Abstract

Performance-based research funding systems (PBRFSs) have been used in selectively distributing research funding, increasing public money accountability and efficiency. Two recent such evaluations in England were called the Research Excellence Framework (REF), which took place in 2014 and 2021, and the research environment, outputs and impact of the research were evaluated. Even though various aspects of the REF were examined, there has been limited research on how the performance of the universities and disciplines changed between the two evaluation periods. This paper assesses whether there has been convergence or divergence in research quality across universities and subject areas between 2014 and 2021 and found that there was an absolute convergence between universities in all three research elements evaluated, and universities that performed relatively worse in REF in 2014 experienced higher growth in their performance between 2014 and 2021. There was also an absolute convergence in the research environment and impact across different subject areas, but there is no significant convergence in the quality of research outputs across disciplines. Our findings also highlight that there has been an absolute convergence in research quality within the universities (between different disciplines in a given university) and within disciplines (between universities in a given subject).

Keywords Performance-based research funding systems · Research excellence framework · Metrics · Convergence · Divergence · Inequality

✉ Mehmet Pinar
mehmet.pinar@edgehill.ac.uk

¹ Business School, Edge Hill University, Ormskirk L39 4QP, Lancashire, UK

² Departamento de Análisis Económico y Economía Política, Universidad de Sevilla, Avda. Ramón Y Cajal, 1, 41018 Seville, Spain

Introduction

The evaluation of the research performance of the universities and respective funding allocation based on their performance, so-called performance-based research funding systems (PBRFSs), have become widespread worldwide (Dougherty et al., 2016; Hicks, 2012; Pinar & Horne, 2022; Zacharewicz et al., 2019). It has been argued that the PBRFSs emerged based on the new public management reforms to increase public accountability (Hicks, 2012; Leišytė, 2016), and these funding systems aim to allocate resources to more efficient research institutes, provide accountability for public investment and establish reputational yardsticks (REF, 2022; Zacharewicz et al., 2019).

The benefits and negative implications of such systems have been evaluated (see e.g., de Rijcke et al. (2016); Geuna and Piolatto (2016); Watermeyer and Olssen (2016) for a review). One of the benefits of such evaluation systems is found to be the increased quantity and/or quality of the research. For instance, Civera et al. (2020) found that the excellence initiative in Germany led to an increased number of publications, but the quality of such publications declined. On the other hand, Checchi et al. (2019) examined the role of the performance-based funding system in the number and quality of publications in 31 countries over the period 1996–2016, and found that number of publications and citations attracted by these publications increased over time, but the share of articles published in the top journals remained steady. Analyzing the publication practices of Polish researchers between 2009 and 2016, Korytkowski and Kulczycki (2019) demonstrated that both the quality and quantity of publications increased in Poland.

Some of the existing studies found that the PBRFSs in the United Kingdom (UK), Italy and South Korea led to increased performance and funding differences across universities. For instance, the research assessment exercise in 2008 (RAE2008) and research excellence framework in 2014 (REF, 2014) increased the concentration of funding among larger universities in the UK (Marques et al., 2017; Martin, 2011; Pinar, 2020; Torrance, 2020).¹ For instance, Pinar and Unlu (2020a) showed that the inclusion of new criteria to the REF (2014) (i.e., the evaluation of the non-academic impact of the research) led to an increased quality-related research funding gap among universities. In other words, the differences in performance-based funding allocated to universities based on their performances in environment and output categories were lower than the differences in funding allocated based on their performances in the impact category. Furthermore, Jeon and Kim (2018) found that the growth of the number of publications by the top 20% of universities was relatively higher than that by the bottom 20% of universities between 2009 and 2015, which was 7 years after the introduction of the governmental performance-based funding, in South Korea. Grisorio and Prota (2020) found that performance evaluations in Italy led to an increased research gap geographically (i.e., the concentration of high-quality research in northern regions) and could contribute to the already existing regional disparities between the south and north of Italy.

The research quality and inequality changes have also been evaluated using convergence analysis between two research assessment periods in different countries. For instance, there has been a convergence in research quality across Italian universities between two

¹ There have been major structural changes between RAE2008 and REF2014. For instance, assessment procedures were standardized across different units of assessments, and the number panels and units of assessments (subject areas) were decreased in REF2014 compared to the RAE2008 (see e.g., Marques et al., (2017) for a detailed discussion about the structural changes between the RAE2008 and REF2014).

assessment periods (Abramo & D'Angelo, 2021; Checchi et al., 2020). Similarly, analyzing the university and discipline performances in 2003 and 2012 assessments in New Zealand, Buckle et al. (2020) also confirmed convergence in research performance across universities and disciplines (see also Buckle et al. (2022) for convergence analysis in research quality performance across universities in New Zealand between 2003 and 2018). The convergence in research quality across countries and regions has also been examined. For instance, Radosevic and Yoruk (2014) demonstrated that the research production in Central and Eastern Europe, Latin America and Asia Pacific regions caught up with the leading regions between 1981 and 2011. On the other hand, Barrios González et al. (2021) examined research papers published in 121 countries between 2003 and 2016 and found that there was no absolute convergence in terms of the number of papers published across countries, but they found that number of papers published by some countries converged over time.

This paper aims to contribute to the literature that examined the research performances between two research assessment periods and will make use of the last two Research Excellence Frameworks (REF) of the United Kingdom and assess whether there was a convergence or divergence in research quality between and within universities and disciplines between 2014 and 2021. The existing literature found that the research quality converged across universities and disciplines in different countries (see e.g., Buckle et al., (2020, 2022) for the case of New Zealand, Abramo and D'Angelo (2021) and Checchi et al. (2020) for Italian case), but a similar analysis has not been carried out for England. Furthermore, the existing literature argued that there had been a significant concentration of quality-based research funding in fewer universities in earlier assessments in England (e.g., Marques et al., 2017; Torrance, 2020). Therefore, we will examine whether the trend of increased research quality (funding) concentration in fewer universities continued between REF (2014, 2021). Finally, there has been a clear distinction between universities in England (e.g., research-intensive, older established or modern universities), and to test whether there had been differences in convergence (divergence) in research quality between different types of universities, we also carried out our analysis for different samples of universities (i.e., Russell and non-Russell universities, pre-92 and post-92 universities).

The rest of the paper is organized as follows. Section [Literature review](#) provides a literature review on the convergence or divergence of research production and quality. Section [Performance-based research funding in the UK](#) provides the details of the REF in 2014 (REF, 2014) and 2021 (REF, 2021), and how the research quality in England is evaluated in different subject areas and universities. Section [Data](#) provides the details of the data set used, and the convergence concept is introduced in Sect. [Convergence methodology](#). Finally, Sect. [Empirical analysis](#) provides the results, and Sect. [Robustness analyses](#) offers the robustness analyses with alternative measures of composite quality scores. Finally, Sect. [Conclusions](#) concludes and provides policy recommendations.

Literature review

Even though convergence analysis is widely applied to examine the convergence in the standard of living (see e.g., Ganong & Shoag, 2017; Lessmann & Seidel, 2017), poverty (e.g., Crespo Cuaresma et al., 2022) and carbon emissions (e.g., Kounetas, 2018; Li & Lin, 2013), among other areas. The convergence analysis has been recently applied to examine the convergence in research publications across countries and convergence in

research quality between universities within a given country based on the performance-based research funding systems in these countries.

A set of papers examined the quantity and quality of scientific publication differences across countries and highlighted the factors that explain the research quality and publication differences across countries.

It has been found that investment in research and development (R&D) is one of the main factors explaining scientific production and research quality. For instance, Jurajda et al. (2017) examined the number of articles published during 2010–2014 in journals registered by the Web of Science (WoS) in 6 post-communist countries (i.e., Croatia, Czech Republic, Hungary, Poland, Slovakia, Slovenia) and 6 western European countries (Austria, Belgium, Finland, the Netherlands, Portugal, Sweden) and found that the post-communist countries lag behind west European countries in terms of their publications where they argued that the government funding differences between two group of countries might have explained some of the existing differences. Erfanian and Ferreira Neto (2017) showed that the researchers in R&D and R&D expenditures are positively associated with the number of scientific publications after analyzing the number of scientific and technical journal articles published in 31 countries between 2003 and 2011. Similarly, Mueller (2016a) showed that the expenditure on R&D and real gross domestic product (GDP) of the countries are positively associated with the number of citable documents in peer-reviewed sources between 2005 and 2014 in 80 countries. Furthermore, Shelton (2020) examined the number of papers published in 93 countries and also demonstrated that there is a positive association between the number of papers published and gross expenditures on research and development. Xu and Reed (2021) examined the number of papers published across 102 countries and found that internet access, the number of researchers, and research and development investment in these countries are positively associated with the number of papers published. Finally, Lancho-Barrantes and Cantú-Ortiz (2019) found that there was a significant increase in the number of scientific production (i.e., articles, conference proceedings, reviews, chapters of books, books, etc.) in Mexico between 2007 and 2016, and found investment in research and production was highly and positively correlated with the publication counts.

The wealth of the countries also played a significant role in explaining the number of publications and research quality across countries. For instance, Rodríguez-Navarro and Brito (2022) examined the ratio of the number of papers in the global top 10% of most cited articles divided by the total number of articles in 69 countries and found that the GDP per capita is positively correlated with this ratio. Along the same lines, Docampo and Bessoule (2019) also showed that the wealth of countries (proxied by the GDP and GDP per capita) is positively associated with the number of articles and reviews published between 2006 and 2015 in 189 countries.

It has been found that international collaboration also increases the scientific production of the countries. For instance, Nguyen et al. (2021) examined scientific production in Vietnam between 2001 and 2015 and found that the growth of the papers published with international collaborators was larger than the domestically published paper, and the citations received by the papers with international collaborators were larger than the citations received by domestically published papers. Barrios et al. (2021) found that the number of papers published across 121 countries did not converge between 2003 and 2016; however, the number of papers published among the five groups of countries converged over time. They found that countries with more scientific production tend to have a larger economic size and more R&D expenditures, researchers and international collaboration networks.

Overall, several factors explained the scientific production differences across countries, such as cross-country differences in human capital (i.e., researchers), wealth, research and development investments and international collaboration levels. Therefore, convergence and divergence in scientific production across countries could be attributed to the differences in these factors.

Beyond the cross-country analysis of scientific production differences, a set of papers also examined the research quality (production) differences between universities based on the PBRFS in different countries. For instance, Abramo and D'Angelo (2021) analyzed the research performance of universities before and after the introduction of PBRFS in Italy. In their paper, Abramo and D'Angelo (2021) evaluated the bibliometric performance of universities (i.e., the total number of citations divided by the researchers) over the 2007–2011 period (which was the period before the start of the first assessment period in Italy), and the 2013–2017 period (during the initial evaluation period), and they found that the southern universities (i.e., relatively poor performing universities before the assessment period) improved their research quality between 2007–2011 and 2013–2017 more than the central and northern universities (i.e., relatively better-performing universities before the assessment period). Their paper highlighted that the performance-based funding (i.e., financial incentive) served as a carrot for the southern universities to improve their research performance after the introduction of the PBRFS in Italy. On the other hand, Checchi et al. (2020) examined the research quality convergence across Italian universities between two research assessments. They found that the universities that had performed relatively weaker in the first assessment exercise experienced higher growth in their performance between the two assessments, and there was a convergence in research quality across Italian universities between the two assessment periods. In their paper, they demonstrated that the convergence in research quality between the two periods was due to the relatively better performance of researchers that participated in both assessments and the promotion and hiring strategies of the universities that had performed worse in the first assessment period. In particular, they highlighted that researchers who were hired and promoted by the relatively weaker universities performed exceptionally well and contributed the most to the convergence of research quality between the two assessment periods.

There has also been an investigation of the research performance of the universities in New Zealand between research assessment periods based on the PBRFS. Buckle et al. (2020) found that there has been an absolute convergence in research quality performance across universities between the 2003 and 2012 assessment periods. They support their finding by pointing out that universities that performed relatively weaker in the 2003 assessment may have hired new staff with a proven record of high-quality publications and transformed their existing staff. On the other hand, Buckle et al. (2022) examined the convergence in research quality across universities between three research assessments in New Zealand, which took place in 2003, 2012, and 2018, and they found that there was also a significant convergence across universities between 2012 and 2018 assessments. Buckle et al. (2022) also examined the role of separate contributions to the convergence process of staff turnover (i.e., exists of the low-quality researchers and entrants of the high-performing staff members) and quality transformations of existing staff members and found that the exits of the low-performing staff member and transformation of the existing staff members contributed more to the convergence in research quality across universities.

Overall, the existing studies that examined the convergence in research quality across universities in Italy and New Zealand found that there has been a significant convergence in research quality across universities based on the PBRFS. These papers argued that the monetary incentives of the performances, strategic investments by the low-performing

universities, and strategic hiring and promotion by these universities allowed them to catch up with high-performing universities in terms of research quality between assessment periods. In other words, in the context of the research assessment, poor performers in the previous assessment could make strategic promotions and hire experienced researchers with a good track record of publications or make research investments (e.g., employment of research office staff helping with the impact case studies, purchase of a better set of technologies allowing the research team to carry out high-quality research, and so on). Furthermore, the concept of “diminishing marginal returns” in research quality is also likely to be observed. For example, a given amount of physical investment to the institutes with high-quality performance could only marginally improve their performance. However, a similar amount of investment to the low-performing institute would lead to a more considerable increase in their performance. Furthermore, the academic job market is increasingly becoming more competitive due to the current economic conditions and increased numbers of Ph.D. graduates (Passaretta et al., 2019; Hancock, 2020; Hnatkova et al., 2022). Therefore, institutes and departments that initially perform worse may attract good Ph.D. graduates who will likely improve their research quality. As the number of Ph.D. completions are submitted as part of the research environment submission of the units to the REF, Pinar and Unlu (2020b) showed that the postgraduate completion numbers were positively associated with the environment performances in most of the units of assessments in REF (2014).

This paper aims to contribute to this stream of literature by examining the research quality performances of universities and disciplines between two REF assessments in 2014 and 2021. Based on the strategic decisions of the universities (i.e., hiring, research and development investment levels, and so on), the performance of the universities in England may converge or diverge, and this paper aims to evaluate this. In the next section, we provide the details of the two recent research assessment exercises in the UK.

Performance-based research funding in the UK

Higher education is provided through colleges, university colleges and universities in the United Kingdom (UK), and there has been a significant expansion of higher education in the UK over the last decades (see e.g., Universities UK, 2018). While there has been increased participation in UK higher education over the last decades, recent reforms to increase tuition and fees and removing the student number cap (see e.g., Hillman (2016) for a detailed discussion on recent reforms) lowered the participation in low-income groups (Callender & Mason, 2017) and students from low-income backgrounds and state schools were found to be less likely to attend so-called prestigious Russell group universities (Boliver, 2013). The increased marketization of higher education in the UK led to an assessment of the teaching quality in higher education providers, known as the Teaching Excellence Framework (TEF), to assess the excellence of the teaching quality and provide quality assurance to the students (Gunn, 2018) and increase employability levels of the graduate students (Matthews & Kotzee, 2021).

While the assessment of the teaching quality of the higher education providers in the UK is a recent phenomenon, the research exercises have been a long tradition in the UK and were carried out first time in 1986 (see Shattock, 2012 for the details of the previous research assessments in the UK). The latest two research assessments in England are called the research excellence framework (REF) and were carried out in 2014 and 2021 (REF,

2014, 2021, respectively), respectively, and were undertaken by the four UK higher education funding bodies: Research England, the Scottish Funding Council, the Higher Education Funding Council for Wales, and the Department for the Economy, Northern Ireland.

The REF is a peer-reviewed research assessment procedure, and the universities were invited to submit submissions in different subject areas, known as units of assessment (UOAs). The requirements changed between REF (2014) and REF (2021); however, the assessment criteria mostly remained the same (see REF (2012) and REF (2019a) for the details of the requirements and assessment criteria of REF (2014, 2021, respectively)). Both REF exercises consisted of three research categories (i.e., environment, outputs, and impact), and the submissions were evaluated in these three categories. The research environment of the units was evaluated based on the research environment template submitted by the units, which described the research environment of the unit and research environment data (i.e., data about research doctoral degrees awarded and research income generated by the unit during the assessment period). The research environment is evaluated based on its ‘vitality and sustainability’, and its contribution to the broader discipline or research base. The units were also required to submit research outputs (i.e., books, journal articles, book chapters, etc.) for evaluation, and these research outputs were evaluated based on their quality. Finally, each unit had to submit an impact template and impact case studies. The impact template provides the submitted unit’s approach to enabling impact from its research, and impact case studies were evaluated based on the ‘reach and significance’ of their impacts on the economy, society and/or culture that were underpinned by excellent research conducted in the submitted unit.

Each unit’s research environment, output and impact submissions were evaluated based on the five-scale star-rating system: four-star (i.e., world-leading quality), three-star (i.e., internationally excellent quality), two-star (i.e., a quality that is internationally recognized), one-star (i.e., a quality that is recognized nationally), and unclassified (i.e., falls below any other quality levels). The overall quality of the submitting units was then calculated based on the weight attached to each element. The weights attached to the research environment, output and impact in REF (2014) were 15%, 65% and 20%, respectively. However, the weight given to the impact element increased from 20 to 25%, the weight given to the output element decreased from 65 to 60%, and the environment element was given the same weight in the REF (2021). Finally, based on these results, universities are allocated quality-based research funding (see e.g., Research England (2022) for the formula used to allocate funding across universities based on the REF results).

It should be noted that there were significant differences distinguishing REF (2014) from REF (2021). Firstly, submitting units were required to return all their academic staff that had significant responsibility for research in REF (2021). However, in the REF (2014), submitting units were able to choose the staff members to be returned. Secondly, the submission requirements of the research outputs were also changed. Up to four research outputs by each member of the submitted staff were required in the REF (2014). However, the submission of the research outputs was universal for the unit and flexible in the REF (2021). The total number of outputs to be submitted to the REF (2021) was required to be equal to 2.5 times the full-time equivalent (FTE) staff returned by the unit. The unit was required to submit a minimum of one output attributed to each staff member returned and no more than five attributed to any staff member. Furthermore, institutions were able to return the outputs of staff that were previously employed, and these outputs were made publicly available during their employment. Thirdly, new measures have been introduced to ensure equality and diversity in research careers, and the submission and assessment of interdisciplinary research were supported in REF (2021) compared to REF (2014).

Fourthly, institutions were also required to submit an institutional-level environment statement in REF (2021), and unit-level environment templates were changed in REF (2021) to reflect the inclusion of the institutional-level environment statement. Fifthly, the importance given to the three elements (i.e., impact, environment and output) was revised (see above for the details). Finally, there was greater consistency in the assessment process in REF (2021) compared to REF (2014). We provided a brief discussion of the changes between REF (2014) and REF (2021); however, the key changes between REF (2014) and REF (2014) are provided in the paragraph in REF (2019b).

Data

The results of the REF (2014) and REF (2021) are obtained from the REF (2014) and REF (2021), respectively. The quality profile for each submission in each research element (i.e., environment, impact and output) is presented as the proportion of the submission judged to meet the quality for each of the starred levels (i.e., 4-star, 3-star, 2-star, 1-star and unclassified). For instance, a submitting unit may have the following profile for their research outputs: 23% (4-star), 30% (3-star), 40% (2-star), 7% (1-star), and 0% (unclassified). This classification suggests that 23%, 30%, 40% and 7% of this submitting unit's research outputs were evaluated as 4-, 3-, 2- and 1-starred quality levels, respectively. The quality profiles of the environment and impact categories are presented in a similar manner. Finally, the percentage of the overall quality that falls into different starred levels is also obtained with the weights attached to three categories. These results are presented for each submitting unit from various universities.

To compare average achievement scores of different subject areas (i.e., units of assessment) and universities, we first obtain average achievement scores of the submitting units in each category by using the grade point average (GPA) as follows:

$$GPA_{ijt}^C = \frac{\%4_{ijt}^* \times 4 + \%3_{ijt}^* \times 3 + \%2_{ijt}^* \times 2 + \%1_{ijt}^* \times 1 + \%U_{ijt}^* \times 0}{100} \quad (1)$$

where GPA_{ijt}^C is the grade point average of the submitting unit in a given research category C that is submitted to the i th unit of assessment by the j th university in a given assessment period (i.e., $t=2014$ and 2021). C is the research category evaluated (i.e., research environment, impact and output). $\%4_{ijt}^*$, $\%3_{ijt}^*$, $\%2_{ijt}^*$, $\%1_{ijt}^*$, and $\%U_{ijt}^*$ are percentages of research classified as 4-star, 3-star, 2-star, 1-star and unclassified by the submitting unit in a given research category.²

Each university submitted different numbers of units to the (REF, 2014, 2021), and 120 English higher education institutes (HEIs) made submissions to both assessments, and the list of the English universities that had submissions to the REF (2014) and/or REF (2021) is provided in the Supplementary Material Table A1. The average GPA scores achieved by a given university j in a given assessment period t in each category C are obtained by the

² It should be noted that the results for the submitting units with a size of 3 FTE or less are not provided. Therefore, these submitting units were not used in the calculation of the overall achievement scores of the universities and disciplines. 14 and 6 units had a size of 3 FTE or less in REF2014 and REF2021, respectively, and their exclusion from the calculation of the overall scores would not have major implications.

full-time equivalent (FTE) weighted arithmetic mean of GPA values of the units of assessment (UOAs) submitted by university j in different units (i) as follows:

$$GPA_{jt}^C = \frac{\sum_{i=1}^{i=n} FTE_{ijt} \times GPA_{ijt}^C}{\sum_{i=1}^{i=n} FTE_{ijt}} \tag{2}$$

Similarly, one could obtain the average GPA scores of different subject areas in REF (2014, 2021) to examine progress in different subjects. REF (2014, 2021) consisted of 36 and 34 units of assessment (UOAs), different fields or subject areas, respectively. There were minor changes in the UOAs in REF (2021). Firstly, the engineering field had four distinctive UOAs in REF (2014): (i) Aeronautical, Mechanical, Chemical and Manufacturing Engineering; (ii) Electrical and Electronic Engineering, Metallurgy and Materials; (iii) Civil and Construction Engineering; and (iv) General Engineering. However, submission to the REF (2021) consisted of only one general engineering subject area. Secondly, the Geography, Environmental Studies and Archaeology unit in REF (2014) was divided into two units in REF (2021): Geography and Environmental Studies; and Archaeology. Whereas the rest of the categories of the UOAs were the same in both assessments. Therefore, to assess the convergence in research quality across different fields, we collated the performances in 4 engineering categories in REF (2014) to compare it with the general engineering field in REF (2021). Similarly, two units in REF (2021) (i.e., Geography and Environmental Studies, and Archaeology) are combined to compare it with the Geography, Environmental Studies and Archaeology unit in REF (2014). The list of the UOAs in REF (2014, 2021) is listed in Supplementary Material Table A2. In summary, we have 33 disciplines to compare achievements and convergence between REF (2014, 2021). To do so, we also obtained the average GPA scores in each research category of a given field i at a given assessment period of t by the full-time equivalent (FTE) weighted arithmetic mean of GPA values of HEIs that submitted to UOA i as follows:

$$GPA_{it}^C = \frac{\sum_{j=1}^{j=m} FTE_{ijt} \times GPA_{ijt}^C}{\sum_{j=1}^{j=m} FTE_{ijt}} \tag{3}$$

In other words, GPA_{it}^C is the GPA of a given i th UOA in each research category (research environment, output and impact) at a given REF period and obtained by the weighted average of the GPAs of the universities that made submissions to the i th UOA.

Table 1 reports the GPA scores obtained by 120 universities in the REF (2014) and (2021), and the growth in GPA scores between REF (2014) and (2021) in different research categories: (i) Environment (E), (ii) Impact (I), and (iii) Output (O). Overall, most universities experienced positive performance improvements in different categories between the two assessment periods. Among 120 HEIs, 94, 87, and 112 institutes improved their environment, impact and output quality between REF (2014) and (2021), respectively. On the other hand, 23, 33, and 8 universities experienced a deterioration in their environment, impact and output quality between REF (2014) and (2021), respectively. Finally, 3 institutes had the same GPA in the environment element in REF, 2014 and REF, 2021.

Table 2 provides the quality of research across 33 UOAs (disciplines) in REF, 2014 and REF, 2021. Firstly, research environment quality in 7 (26) fields decreased (increased) between the two assessment periods. While research environment quality in Law decreased by 6.7%, the research environment quality increased by 10.3% in the Economics and Econometrics field. Secondly, except for three fields (i.e., (i) Psychology,

Table 1 GPA scores of the universities in REF (2014, 2021), growth in GPA scores between REF (2014, 2021)

Institution name	E-14	I-14	O-14	E-21	I-21	O-21	%Δ-E	%Δ-I	%Δ-O
Anglia Ruskin University	2.139	2.363	2.421	2.660	2.852	2.826	24.36	20.69	16.70
Arts University Bournemouth	1.800	2.200	2.461	2.525	2.719	2.611	40.28	23.60	6.08
Aston University	3.291	3.294	2.924	3.098	2.964	3.080	-5.87	-10.04	5.33
Bath Spa University	2.262	2.735	2.518	2.569	2.804	2.702	13.58	2.53	7.30
Birkbeck College	3.161	3.097	2.880	3.215	3.242	3.211	1.70	4.68	11.50
Birmingham City University	2.636	3.098	2.503	2.426	2.824	2.713	-7.96	-8.83	8.38
Bishop Grosseteste University	1.000	1.300	2.334	1.174	1.367	2.324	17.38	5.17	-0.42
Bournemouth University	2.812	2.967	2.621	2.749	3.090	2.647	-2.25	4.15	1.01
Brunel University London	2.877	2.753	2.603	2.923	2.683	2.953	1.61	-2.55	13.44
Buckinghamshire New University	2.003	2.347	2.152	1.646	2.124	2.423	-17.82	-9.48	12.61
Canterbury Christ Church University	2.221	2.507	2.381	2.420	3.069	2.640	8.97	22.43	10.87
City University London	2.987	3.203	2.860	3.200	3.286	3.239	7.11	2.58	13.28
Courtauld Institute of Art	4.000	3.400	3.385	4.000	3.250	3.367	0.00	-4.41	-0.53
Coventry University	2.827	3.144	2.491	2.966	2.952	2.857	4.92	-6.10	14.69
Cranfield University	3.523	3.205	2.919	3.583	3.348	3.013	1.71	4.48	3.25
De Montfort University	2.660	2.994	2.579	2.431	2.695	2.750	-8.63	-10.00	6.64
Edge Hill University	1.994	2.261	2.266	2.105	2.896	2.727	5.60	28.09	20.31
Falmouth University	2.193	2.255	1.946	3.725	3.600	2.845	69.87	59.62	46.19
Goldsmiths' College	3.070	3.105	2.796	3.043	3.221	3.077	-0.89	3.76	10.03
Guildhall School of Music & Drama	2.300	3.600	2.108	3.100	4.000	3.023	34.78	11.11	43.41
Harper Adams University	2.750	2.767	2.611	2.750	2.625	2.547	0.00	-5.13	-2.45
Imperial College London	3.817	3.681	3.157	3.944	3.716	3.510	3.31	0.95	11.19
Institute of Cancer Research	3.666	3.867	3.205	3.914	3.719	3.427	6.77	-3.82	6.93
Institute of Zoology	3.100	3.633	2.969	3.875	3.364	3.372	25.00	-7.40	13.57
Keele University	2.912	3.037	2.824	2.931	3.079	3.103	0.66	1.38	9.88
King's College London	3.624	3.518	3.051	3.747	3.643	3.304	3.38	3.55	8.30
Kingston University	2.653	2.882	2.663	2.741	3.070	2.858	3.30	6.51	7.36
Lancaster University	3.622	3.404	2.963	3.629	3.480	3.254	0.19	2.25	9.82

Table 1 (continued)

Institution name	E-14	I-14	O-14	E-21	I-21	O-21	%Δ-E	%Δ-I	%Δ-O
Leeds Beckett University	1.938	2.405	2.136	2.414	2.795	2.702	24.56	16.22	26.49
Leeds Trinity University	1.306	1.542	2.359	1.699	1.623	2.492	30.07	5.30	5.61
Liverpool Hope University	1.630	1.988	2.405	1.976	2.311	2.752	21.27	16.24	14.42
Liverpool John Moores University	2.733	2.902	2.774	2.458	2.724	2.942	-10.07	-6.13	6.05
Liverpool School of Tropical Medicine	3.425	3.566	2.875	3.585	3.765	3.270	4.65	5.57	13.74
London Business School	3.625	3.064	3.277	3.875	3.188	3.729	6.90	4.05	13.79
London Metropolitan University	2.024	2.448	2.575	2.172	2.880	2.626	7.31	17.68	2.01
London School of Economics & Political Science	3.743	3.504	3.212	3.745	3.465	3.500	0.03	-1.12	8.98
London School of Hygiene & Tropical Medicine	3.910	3.740	2.857	3.923	3.776	3.193	0.34	0.95	11.76
London South Bank University	2.227	2.826	2.492	2.555	3.034	2.736	14.73	7.35	9.80
Loughborough University	3.282	3.175	2.807	3.503	3.496	3.167	6.71	10.10	12.83
Manchester Metropolitan University	2.700	3.082	2.650	3.110	3.311	3.030	15.21	7.44	14.35
Middlesex University	2.751	2.732	2.500	2.671	3.067	2.695	-2.91	12.24	7.79
Newcastle University	3.302	3.367	2.950	3.378	3.392	3.233	2.29	0.77	9.59
Newman University	1.670	1.733	2.537	1.359	1.588	2.371	-18.62	-8.37	-6.55
Norwich University of the Arts	2.300	3.300	2.584	2.825	3.000	2.978	22.83	-9.09	15.25
Nottingham Trent University	2.480	2.774	2.555	2.969	3.233	3.073	19.74	16.52	20.26
Open University	3.035	3.054	2.834	3.129	3.160	2.928	3.10	3.46	3.31
Oxford Brookes University	2.372	2.715	2.717	2.656	2.907	2.918	12.00	7.09	7.39
Queen Mary University of London	3.368	3.338	3.089	3.472	3.463	3.345	3.09	3.74	8.29
Roehampton University	2.565	3.022	2.826	2.740	3.371	3.097	6.84	11.54	9.60
Rose Bruford College	1.700	2.500	2.358	2.100	2.000	2.545	23.53	-20.00	7.93
Royal Academy of Music	2.600	2.900	2.743	1.950	2.000	2.764	-25.00	-31.03	0.77
Royal Agricultural University	1.500	0.800	1.553	2.625	2.000	2.405	75.00	150.00	54.86
Royal Central School of Speech and Drama	3.100	3.867	2.756	4.000	3.833	3.540	29.03	-0.88	28.45
Royal College of Art	3.800	3.543	2.769	3.475	3.055	2.846	-8.55	-13.77	2.78
Royal College of Music	2.800	3.900	2.790	3.550	3.333	2.732	26.79	-14.54	-2.08

Table 1 (continued)

Institution name	E-14	I-14	O-14	E-21	I-21	O-21	%Δ-E	%Δ-I	%Δ-O
Royal Holloway, University of London	3.267	3.252	2.996	3.261	3.318	3.300	-0.19	2.03	10.16
Royal Northern College of Music	2.900	4.000	2.842	3.350	2.250	2.816	15.52	-43.75	-0.91
Royal Veterinary College	3.750	3.455	2.870	3.750	3.833	2.964	0.00	10.94	3.28
School of Oriental and African Studies	2.953	2.828	2.775	3.281	3.308	3.205	11.12	16.98	15.51
Sheffield Hallam University	2.798	2.915	2.708	2.761	3.205	2.861	-1.30	9.94	5.64
Southampton Solent University	1.530	0.961	1.855	1.455	1.926	2.438	-4.96	100.29	31.40
St Mary's University, Twickenham	1.374	1.708	2.206	2.126	2.798	2.851	54.79	63.81	29.24
St.George's, University of London	2.380	3.637	2.936	2.682	3.617	3.211	12.70	-0.54	9.36
Staffordshire University	2.052	2.474	2.140	2.178	3.103	2.796	6.12	25.44	30.67
Teesside University	2.230	2.866	2.565	2.518	2.949	2.749	12.91	2.92	7.16
The University of West London	1.260	1.730	2.094	1.872	2.578	2.898	48.58	49.01	38.43
Trinity Laban Conservatoire of Music and Dance	2.700	3.200	2.676	2.750	3.500	2.720	1.85	9.38	1.64
University College London	3.791	3.542	2.987	3.860	3.652	3.340	1.80	3.10	11.82
University for the Creative Arts	2.700	3.300	2.539	2.750	3.125	3.105	1.85	-5.30	22.29
University of Bath	3.411	3.447	3.036	3.390	3.383	3.266	-0.60	-1.84	7.59
University of Bedfordshire	2.226	2.608	2.565	2.475	2.663	2.784	11.17	2.13	8.55
University of Birmingham	3.315	3.223	2.968	3.656	3.518	3.345	10.28	9.15	12.71
University of Bolton	1.496	1.986	2.179	1.622	1.828	1.984	8.44	-7.95	-8.97
University of Bradford	2.939	3.196	2.868	2.617	2.883	2.895	-10.95	-9.77	0.95
University of Brighton	3.083	3.358	2.623	3.202	3.254	2.920	3.85	-3.10	11.33
University of Bristol	3.528	3.497	2.996	3.675	3.600	3.425	4.17	2.93	14.31
University of Cambridge	3.815	3.488	3.164	3.806	3.612	3.428	-0.23	3.56	8.33
University of Central Lancashire	2.417	2.636	2.496	2.636	2.972	2.776	9.03	12.77	11.20
University of Chester	1.631	2.042	2.187	2.174	2.504	2.582	33.26	22.63	18.06
University of Chichester	2.328	2.807	2.442	2.200	2.475	2.468	-5.49	-11.84	1.10
University of Cumbria	1.492	2.030	2.189	1.945	2.453	2.329	30.35	20.80	6.41
University of Derby	1.643	2.191	2.133	2.196	2.440	2.632	33.61	11.34	23.41

Table 1 (continued)

Institution name	E-14	I-14	O-14	E-21	I-21	O-21	%Δ-E	%Δ-I	%Δ-O
University of Durham	3.430	3.423	2.985	3.512	3.410	3.267	2.39	-0.38	9.43
University of East Anglia	3.181	3.309	3.035	3.360	3.520	3.321	5.62	6.36	9.43
University of East London	2.525	2.948	2.680	2.175	2.620	2.670	-13.87	-11.13	-0.39
University of Essex	3.389	3.124	2.948	3.210	3.349	3.193	-5.28	7.20	8.31
University of Exeter	3.353	3.201	2.976	3.402	3.519	3.264	1.47	9.90	9.68
University of Gloucestershire	1.752	2.638	2.435	1.886	2.152	2.679	7.62	-18.44	10.02
University of Greenwich	1.878	2.495	2.386	2.842	3.064	2.782	51.29	22.78	16.58
University of Hertfordshire	2.556	2.578	2.632	2.760	3.433	2.923	7.96	33.18	11.06
University of Huddersfield	2.720	2.758	2.601	2.685	2.985	2.756	-1.30	8.24	5.95
University of Hull	2.666	2.990	2.617	2.852	3.341	3.059	7.00	11.74	16.91
University of Kent	3.103	3.175	2.848	3.269	3.368	3.215	5.34	6.07	12.87
University of Leeds	3.487	3.466	2.945	3.624	3.437	3.247	3.93	-0.84	10.27
University of Leicester	2.984	3.058	2.871	3.276	3.413	3.260	9.75	11.63	13.54
University of Lincoln	2.223	2.299	2.684	2.701	2.985	3.052	21.47	29.85	13.72
University of Liverpool	3.286	3.280	2.939	3.429	3.467	3.257	4.35	5.72	10.80
University of Manchester	3.637	3.449	2.956	3.809	3.620	3.320	4.74	4.97	12.31
University of Northampton	2.005	2.183	2.083	1.853	2.289	2.286	-7.59	4.88	9.77
University of Northumbria at Newcastle	2.861	2.634	2.698	3.085	3.145	3.019	7.84	19.39	11.87
University of Nottingham	3.534	3.361	2.898	3.553	3.462	3.236	0.55	3.02	11.64
University of Oxford	3.810	3.511	3.172	3.867	3.616	3.344	1.48	3.01	5.43
University of Plymouth	2.699	2.726	2.746	2.842	3.135	3.001	5.31	15.01	9.28
University of Portsmouth	2.630	3.116	2.669	2.933	3.250	2.914	11.50	4.29	9.20
University of Reading	3.265	3.227	2.913	3.475	3.248	3.172	6.43	0.64	8.88
University of Salford	2.741	2.694	2.445	2.926	3.283	2.901	6.77	21.85	18.67
University of Sheffield	3.459	3.432	3.026	3.609	3.585	3.265	4.35	4.45	7.92
University of Southampton	3.408	3.402	3.015	3.547	3.633	3.284	4.07	6.79	8.93
University of Sunderland	2.049	2.141	2.143	2.223	2.748	2.883	8.50	28.34	34.52

Table 1 (continued)

Institution name	E-14	I-14	O-14	E-21	I-21	O-21	%Δ-E	%Δ-I	%Δ-O
University of Surrey	3.081	2.983	2.951	3.464	3.243	3.266	12.43	8.72	10.69
University of Sussex	3.193	3.119	2.927	3.398	3.433	3.204	6.44	10.05	9.47
University of the Arts, London	3.200	3.467	2.996	3.400	3.250	3.092	6.25	- 6.26	3.20
University of the West of England, Bristol	2.738	3.032	2.586	2.759	3.186	2.906	0.76	5.10	12.36
University of Warwick	3.497	3.299	3.131	3.641	3.455	3.354	4.11	4.73	7.13
University of Westminster	2.551	2.906	2.731	2.859	3.267	2.924	12.09	12.44	7.08
University of Winchester	1.969	2.331	2.395	1.939	2.508	2.710	- 1.51	7.60	13.17
University of Wolverhampton	2.226	2.053	2.413	2.254	2.435	2.686	1.28	18.61	11.33
University of Worcester	1.740	2.264	2.099	1.782	2.155	2.524	2.41	- 4.81	20.27
University of York	3.452	3.459	3.011	3.690	3.597	3.329	6.89	4.01	10.54
York St John University	1.249	2.025	2.218	1.949	2.107	2.690	56.05	4.05	21.26

E-14, I-14, and O-14 represent environment, impact, and output, and aggregate GPA scores of the universities in REF (2014), respectively. E-21, I-21, and O-21 represent environment, impact, and output, and aggregate GPA scores of the universities in REF (2021), respectively

%Δ-E, %Δ-I, %Δ-O, and %Δ-A represent environment, impact, and output, and aggregate GPA growth rates (in percentage) between REF (2014, 2021), respectively

Table 2 GPA scores of different UOAs in REF (2014, 2021), growth in GPA scores between REF (2014) and (2021)

Unit of assessment name	E-14	I-14	O-14	E-21	I-21	O-21	%Δ-E	%Δ-I	%Δ-O
Clinical medicine	3.596	3.717	2.972	3.643	3.681	3.247	1.32	-0.98	9.25
Public health, health services and primary care	3.684	3.672	2.899	3.653	3.703	3.330	-0.83	0.85	14.84
Allied health professions, dentistry, nursing and pharmacy	3.283	3.305	2.934	3.270	3.211	3.106	-0.39	-2.84	5.89
Psychology, psychiatry and neuroscience	3.428	3.481	2.922	3.249	3.371	3.039	-5.22	-3.15	4.01
Biological sciences	3.494	3.293	3.034	3.467	3.408	3.302	-0.79	3.49	8.84
Agriculture, veterinary and food science	3.424	3.281	2.809	3.450	3.371	3.074	0.78	2.72	9.45
Earth systems and environmental sciences	3.300	3.299	2.975	3.416	3.507	3.382	3.49	6.29	13.67
Chemistry	3.283	3.330	3.149	3.480	3.473	3.424	6.02	4.27	8.72
Physics	3.370	3.173	3.075	3.614	3.433	3.392	7.25	8.17	10.31
Mathematical sciences	3.399	3.159	3.058	3.472	3.480	3.412	2.16	10.16	11.57
Computer science and informatics	2.907	3.030	2.871	2.995	3.224	3.244	3.01	6.40	12.96
Engineering	3.238	3.198	2.988	3.372	3.386	3.230	4.13	5.90	8.09
Architecture, built environment and planning	3.175	3.136	2.813	3.446	3.356	3.081	8.52	7.02	9.55
Geography, environmental studies and archaeology	3.168	3.045	2.822	3.304	3.260	3.173	4.30	7.04	12.45
Economics and econometrics	3.186	3.163	3.013	3.514	3.274	3.272	10.27	3.50	8.59
Business and management studies	3.104	3.116	2.764	3.230	3.246	3.052	4.04	4.19	10.41
Law	3.311	3.098	2.850	3.088	3.285	3.060	-6.74	6.03	7.37
Politics and international studies	3.150	3.231	2.742	3.328	3.314	3.114	5.64	2.56	13.57
Social work and social policy	3.060	3.128	2.760	3.006	3.150	2.977	-1.74	0.68	7.84
Sociology	3.089	3.238	2.850	3.145	3.275	3.137	1.82	1.16	10.06
Anthropology and development studies	3.352	3.166	2.717	3.549	3.342	3.089	5.88	5.56	13.71
Education	3.103	3.083	2.727	3.080	3.262	2.897	-0.75	5.79	6.21
Sport and exercise sciences, leisure and tourism	3.061	3.087	2.718	3.108	3.286	3.122	1.51	6.45	14.84
Area studies	3.174	3.153	2.798	3.420	3.350	3.256	7.76	6.26	16.39
Modern languages and linguistics	3.268	3.074	2.879	3.414	3.319	3.159	4.47	7.98	9.72
English language and literature	3.207	3.124	2.926	3.416	3.372	3.313	6.53	7.95	13.25
History	3.228	3.148	2.973	3.362	3.207	3.195	4.17	1.88	7.48

Table 2 (continued)

Unit of assessment name	E-14	I-14	O-14	E-21	I-21	O-21	% Δ -E	% Δ -I	% Δ -O
Classics	3.345	3.350	2.984	3.418	3.412	3.192	2.20	1.87	6.95
Philosophy	3.287	3.108	2.915	3.396	3.194	3.146	3.32	2.79	7.92
Theology and religious studies	3.088	3.122	2.805	3.267	3.290	3.096	5.77	5.38	10.36
Art and design: history, practice and theory	3.105	3.061	2.705	3.337	3.265	2.964	7.48	6.65	9.60
Music, drama, dance and performing arts	3.048	3.131	2.755	3.261	3.219	3.019	6.99	2.79	9.60
Communication, cultural & media studies, library and information man	3.116	3.043	2.757	3.207	3.350	3.041	2.93	10.09	10.30

E-14, I-14, and O-14 represent environment, impact, and output, and aggregate GPA scores of the universities in REF (2014), respectively

E-21, I-21, and O-21 represent environment, impact, and output, and aggregate GPA scores of the universities in REF (2021), respectively

% Δ -E, % Δ -I, and % Δ -O represent environment, impact, and output, and aggregate GPA growth rates (in percentage) between REF (2014) and (2021), respectively

Psychiatry and Neuroscience, (ii) Allied Health Professions, Dentistry, Nursing and Pharmacy, (iii) Clinical Medicine), impact quality also increased in all other fields, and the impact quality in Mathematical Sciences increased by 11.6%. Finally, the output quality increased in all the fields. The output quality in Area Studies increased the most by 16.39%, and the increase in output quality was more than ten percent in 10 fields.

Overall, most of the universities and disciplines experienced an improvement in their research quality in different research categories, and we will examine whether the changes in research quality between the two periods resulted in convergence or divergence between and within universities and disciplines in the following sections.

Convergence methodology

To examine the convergence or divergence of the research performance across universities and different disciplines between REF (2014) and (2021), we use the β -convergence analysis (see e.g., Barro & Sala-i-Martin, 1991; Mankiw et al., 1992; Quah, 1993 for the technical details). The β -convergence analysis is widely used in the field of economics to examine income per capita convergence across countries and is based on the Solow growth model (Solow, 1956). The β -convergence concept argues that poor countries grow faster than rich ones to catch up with them. The Solow growth model predicts that poor countries will grow faster than rich ones because of the diminishing marginal returns of the capital. In other words, the same amount of capital investment per worker will have a higher return in poorer countries than the rich ones. The convergence in research quality between universities and different subject areas is examined in a similar way by using the β -convergence concept based on the REF, 2014 and REF, 2021 performances.

To test the convergence of research performance in different research categories across institutes and disciplines, we use the following specifications:

$$\log GPA_{j,t}^C - \log GPA_{j,t-1}^C = \alpha + \beta \log GPA_{j,t-1}^C + \varepsilon_{j,t} \tag{4}$$

$$\log GPA_{i,t}^C - \log GPA_{i,t-1}^C = \alpha + \beta \log GPA_{i,t-1}^C + \varepsilon_{i,t} \tag{5}$$

where α is the constant term, and $\varepsilon_{j,t}$ and $\varepsilon_{i,t}$ is the normally distributed error terms. (3) and (4) are used to examine the convergence between universities and disciplines in different research categories (i.e., environment, impact and output), respectively. Along the same lines, we will also analyze the convergence of universities in each discipline by using Eqs. (3) and use (4) to examine whether disciplines within a university also converge.

Overall, when $\beta > 0$, this would suggest that the universities (disciplines) that performed better in REF, 2014 also experienced higher growth in their performance between REF (2014) and (2021), suggesting that there was a divergence across universities (disciplines). On the other hand, $\beta < 0$ suggests that universities (disciplines) that performed poorly in REF, 2014 experienced higher growth in their performance and therefore, there was a convergence across universities (disciplines). On the other hand, $\beta = -1$ reflects full convergence. Hence, the closer the estimated β coefficients to -1 , the faster is the convergence between two assessment periods (i.e., REF, 2014 and REF, 2021). Finally, an insignificant β coefficient or $\beta = 0$ would suggest no significant convergence or divergence.

Table 3 Convergence between universities in different research categories

	(1) Environment	(2) Impact	(3) Output
$\log\text{GPA}_{2014}^{\text{Environment}}$	- 0.225*** (0.0450)		
$\log\text{GPA}_{2014}^{\text{Impact}}$		- 0.410*** (0.0830)	
$\log\text{GPA}_{2014}^{\text{Output}}$			- 0.338*** (0.0679)
Constant	0.293*** (0.0510)	0.485*** (0.0917)	0.435*** (0.0698)
Observations	120	120	120
R-square	0.241	0.401	0.299

Robust standard errors in parentheses

Dependent variable is the growth of the environment, impact and output GPA between two assessment periods in columns 1, 2, and 3, respectively

***, **, and * Represent significance level at the 1%, 5% and 10%, respectively

Empirical analysis

We first start examining the convergence among the universities between two assessment periods in different research quality categories (i.e., environment, impact, output). Table 3 presents the findings. Our findings highlight a strong and significant convergence between universities between two assessment periods in all research categories. In other words, there was an absolute convergence across universities in all research elements. On the other hand, we found that the negative coefficient on impact quality in 2014 was relatively higher than those of coefficients on environment and output quality scores in 2014. In other words, there has been a stronger convergence in impact quality compared to convergence in other research elements. The impact category has been integrated as part of the research assessments in REF, 2014; however, both the research environment and outputs were evaluated in the previous assessments before REF, 2014. For instance, Pinar and Unlu (2020a) showed that the inclusion of the impact element to the REF, 2014 increased the research quality inequality between universities as most academics were not familiar with the impact element in REF, 2014 and had limited public engagement activities such as presentation of their work through social media, taking part in public dialogue, presenting in public exhibitions, making public performances, working with schools/teachers, engaging with the policymakers, and so on (Chikoore et al., 2016; Manville et al., 2015). However, researchers were more familiar with the impact element in the REF, 2021; therefore, there has been a faster convergence in this element.

One of the factors that could have played a role in the absolute convergence in output quality across universities between the two assessment periods may be the changes in the output submission rules between REF (2014) and REF (2021). The submitting units were required to return up to 4 outputs for each member of staff returned in REF (2014) (unless there was a reduction in outputs for specific reasons such as maternity/paternity leave, etc.). However, submission of the outputs in the REF (2021) was universal for the submitting

unit. The units could submit the output of the previous employees and could submit at least a minimum of one output attributed to each staff member and no more than five attributed to any staff member. Therefore, units were more flexible in choosing the strongest set of outputs produced by the units. For example, the strongest outputs of the previous employees could have returned, and up to 5 outputs of a given set of staff members were allowed to be returned if all of their outputs were considered to be strong. Therefore, REF, 2021 had more flexibility to choose the strongest outputs produced by the unit, and therefore, this flexibility could have allowed universities with worse output quality in REF, 2014 to catch up with the universities with better output quality in REF (2021).

There have been distinctive differences among universities in England. The Russell Group of universities is arguably considered to be more established research-intensive universities, and post-92 universities are considered to be modern universities, and they are not established as older institutes that became universities pre-92. For instance, De Fraja et al. (2016) highlighted that the quality of the research produced by the Russell Group universities was higher compared to non-Russell universities in REF (2014). Similarly, Taylor (2011), Thorpe et al. (2018), and Pinar and Unlu (2020b) demonstrated that the Russell Group universities obtained relatively higher scores in the research environment after accounting for various quantitative metrics, suggesting that Russell group universities obtained relatively higher scores compared to non-Russell group universities with similar quantitative achievements. On the other hand, some of the universities are more established and have a long tradition, while some other universities are known as ‘new’ or ‘modern’ universities. These modern universities (i.e., post-92 universities) are evaluated differently in assessing research quality by the existing literature (see e.g., Hewitt-Dundas, 2012; Guerrero et al., 2015). Therefore, we evaluate the convergence between universities using four samples of the universities: Russell and non-Russell group of universities and pre-92 and post-92 universities.³ While all the Russell Group universities were established before 1992, 29 additional universities were also awarded university status before 1992. In other words, the pre-92 sample of universities covers 59 institutions.

Table 4 provides the results when we carry out the convergence analysis for the Russell and non-Russell groups of universities. The findings suggest that there is significant absolute convergence in all categories of research for both Russell and non-Russell groups of universities. The convergence is faster in the non-Russell group of universities in research environment and output compared to the Russell Group of universities. In other words, the size of the negative coefficient on initial environment and output GPA levels was relatively larger for non-Russell group universities than the ones for Russell Group universities. On the other hand, absolute convergence in impact quality in the Russell Group universities is relatively faster compared to the non-Russell group universities (i.e., β coefficient for Russell group universities is closer to -1). We also carry out the convergence analysis when we separate our sample into two groups: pre-92 and post-92 universities, and the results are offered in Table 5. The findings highlight an absolute convergence in research quality between universities when pre- and post-92 universities are considered separately.

³ The Russell group of universities in England are Imperial College London, King’s College London, London School of Economics and Political Science, Newcastle University, Queen Mary University of London, University College London, University of Birmingham, University of Bristol, University of Cambridge, University of Durham, University of Exeter, University of Leeds, University of Liverpool, University of Manchester, University of Nottingham, University of Oxford, University of Sheffield, University of Southampton, University of Warwick, and University of York. Pre-92 and post-92 universities are universities that obtained university status before and after 1992, respectively.

Table 4 Convergence between universities in different research categories for Russell and non-Russell group universities

	Environment		Impact		Output	
	(1)	(2)	(3)	(4)	(5)	(6)
	Russell	No Russell	Russell	No Russell	Russell	No Russell
$\log\text{GPA}_{2014}^{\text{Environment}}$	- 0.218** (0.0991)	- 0.246*** (0.0511)				
$\log\text{GPA}_{2014}^{\text{Impact}}$			- 0.537*** (0.128)	- 0.445*** (0.0862)		
$\log\text{GPA}_{2014}^{\text{Output}}$					- 0.375*** (0.124)	- 0.405*** (0.0741)
$\log\text{GPA}_{2014}^{\text{Environment}}$ Constant	0.307** (0.127)	0.307*** (0.0543)	0.695*** (0.157)	0.509*** (0.0922)	0.510*** (0.136)	0.491*** (0.0736)
Observations	20	100	20	100	20	100
R-square	0.235	0.233	0.421	0.422	0.341	0.340

Robust standard errors in parentheses

Dependent variable is the growth of the environment, impact and output GPA between two assessment periods in columns 1–2, 3–4, and 5–6, respectively

***, **, and *Represent significance level at the 1%, 5% and 10%, respectively

Table 5 Convergence between universities in different research categories for pre- and post-92 universities

	Environment		Impact		Output	
	(1)	(2)	(3)	(4)	(5)	(6)
	Pre-92	Post-92	Pre-92	Post-92	Pre-92	Post-92
$\log\text{GPA}_{2014}^{\text{Environment}}$	- 0.222** (0.0899)	- 0.322*** (0.0752)				
$\log\text{GPA}_{2014}^{\text{Impact}}$			- 0.292 (0.322)	- 0.480*** (0.102)		
$\log\text{GPA}_{2014}^{\text{Output}}$					- 0.348* (0.175)	- 0.582*** (0.0850)
Constant	0.304*** (0.114)	0.351*** (0.0683)	0.357 (0.386)	0.533*** (0.102)	0.464** (0.190)	0.634*** (0.0799)
Observations	59	61	59	61	59	61
R-square	0.140	0.253	0.051	0.469	0.203	0.456

Robust standard errors in parentheses

Dependent variable is the growth of the environment, impact and output GPA between two assessment periods in columns 1–2, 3–4, and 5–6, respectively

***, **, and *Represent significance level at the 1%, 5% and 10%, respectively

However, the convergence between post-92 universities was relatively faster than the pre-92 universities in all categories of research as the negative coefficients on the initial GPA levels in all research categories were higher for post-92 universities. In other words, the convergence between relatively ‘new’ universities is relatively faster between REF (2014) and REF (2021). The convergence speed between relatively older universities may be lower as some universities had been participating in earlier research assessments, and the convergence between these universities may have taken place in earlier periods. Furthermore, newer institutions may feature a more ‘applied research’ focus and greater investment in co-designed/co-produced research, which may be beneficial to impact generation. Therefore, newer institutions (i.e., post-92 universities) may have experienced a faster convergence in impact quality between two assessment periods. Overall, our findings highlight that irrespective of the sample choice of the universities, there was an absolute convergence between universities between the last two research exercises.

We further examine whether there has been a convergence between universities in each discipline and found that with few exceptions, there was an absolute convergence between universities in each discipline (i.e., UOA) in different research categories, and the results are presented in Table 6.⁴ There is an absolute convergence in research environment quality between universities in 29 subject areas (UOAs) out of 33 fields. However, there was no significant convergence in research environment quality between universities in four UOAs: UOA2: Public Health, Health Services and Primary Care; UOA5: Biological Sciences; UOA10: Mathematical Sciences; UOA20: Sociology. Similarly, except for the four UOAs (i.e., UOA1: Clinical Medicine, UOA21: Anthropology and Development Studies, UOA28: Classics, UOA30: Theology and Religious Studies), there was an absolute convergence in impact quality between universities in 29 UOAs. Finally, we also found an absolute convergence in output quality between universities in 28 UOAs except for five UOAs (i.e., UOA5: Biological Sciences, UOA8: Chemistry, UOA20: Sociology, UOA28: Classics, and UOA29: Philosophy). The convergence between universities was the fastest in environment research quality in two subject areas (i.e., UOA21: Anthropology and Development Studies, and UOA28: Classics) and output quality in four subject areas (i.e., UOA1: Clinical Medicine, UOA23: Sport and Exercise Sciences, Leisure and Tourism, UOA24: Area Studies, and UOA30: Theology and Religious Studies). Finally, the convergence in research quality between universities was the fastest in impact quality in 26 subject areas. The finding suggests that the universities with low impact quality in REF, 2014 were able to evidence the societal and economic impact of their research better in REF, 2021 and experienced relatively faster convergence.

We also carried out our analysis to examine whether there has been a convergence in research quality between 33 subject areas, and the results are presented in Table 7. In line with the findings of the convergence analysis between universities, we also found that the convergence in impact quality between subject areas is relatively faster compared to the convergence in research environment quality. On the other hand, even though the coefficient on the initial output quality is negative, we found no significant convergence in output quality between different subject areas between the two assessment periods.

Finally, we examine whether there has been a convergence in research quality between UOAs in a given university. The universities may have decided to allocate more resources

⁴ Some universities submitted multiple submissions into a given UOA (subject area). We combine these multiple submissions into one submission by using FTE-weighted GPA levels of these multiple submissions to obtain a single quality score for each university in a given UOA (subject area).

Table 6 Convergence between universities in different subject areas

UoA	Environment			Impact			Output			Observations
	$\log E_{i,t}$	SE	R-square	$\log I_{i,t}$	SE	R-square	$\log O_{i,t}$	SE	R-square	
UoA1	-0.334***	(0.0576)	0.397	-0.436	(0.345)	0.145	-1.076***	(0.272)	0.539	24
UoA2	-0.203	(0.220)	0.049	-0.607***	(0.151)	0.308	-0.578***	(0.138)	0.398	24
UoA3	-0.278***	(0.0590)	0.213	-0.560***	(0.0699)	0.442	-0.493***	(0.0547)	0.500	63
UoA4	-0.214**	(0.0871)	0.157	-0.533***	(0.132)	0.348	-0.430***	(0.101)	0.453	61
UoA5	-0.125	(0.144)	0.076	-0.794***	(0.143)	0.520	-0.0903	(0.157)	0.026	33
UoA6	-0.624***	(0.0767)	0.801	-0.724***	(0.117)	0.781	-0.545***	(0.0557)	0.631	17
UoA7	-0.319***	(0.0794)	0.225	-0.930***	(0.0417)	0.848	-0.336***	(0.0689)	0.390	25
UoA8	-0.361***	(0.0834)	0.521	-0.715**	(0.283)	0.354	-0.0885	(0.188)	0.010	28
UoA9	-0.472***	(0.104)	0.448	-0.812***	(0.0723)	0.677	-0.464**	(0.182)	0.227	31
UoA10	-0.116	(0.103)	0.063	-0.481**	(0.223)	0.321	-0.286**	(0.137)	0.177	39
UoA11	-0.217***	(0.0442)	0.266	-0.628***	(0.164)	0.541	-0.602***	(0.0731)	0.689	66
UoA12	-0.164**	(0.0652)	0.113	-0.606***	(0.0951)	0.455	-0.418***	(0.122)	0.411	64
UoA13	-0.497***	(0.108)	0.541	-0.519***	(0.124)	0.406	-0.439***	(0.0960)	0.385	26
UoA14	-0.338***	(0.120)	0.505	-0.431***	(0.0590)	0.487	-0.326***	(0.0749)	0.278	43
UoA15	-0.484***	(0.111)	0.447	-0.489*	(0.243)	0.297	-0.470***	(0.105)	0.709	21
UoA16	-0.379***	(0.0770)	0.354	-0.820***	(0.0349)	0.823	-0.379***	(0.0453)	0.485	78
UoA17	-0.329**	(0.145)	0.209	-0.842***	(0.158)	0.573	-0.351***	(0.106)	0.225	48
UoA18	-0.517***	(0.0346)	0.735	-0.646***	(0.224)	0.463	-0.344***	(0.0768)	0.417	40
UoA19	-0.617***	(0.0374)	0.724	-0.809***	(0.0612)	0.648	-0.523***	(0.0726)	0.450	45
UoA20	-0.109	(0.111)	0.029	-0.671***	(0.0974)	0.521	-0.276	(0.232)	0.049	22
UoA21	-0.600***	(0.0587)	0.738	-0.291	(0.192)	0.161	-0.328**	(0.123)	0.238	14
UoA22	-0.300***	(0.0725)	0.187	-0.305***	(0.111)	0.105	-0.218**	(0.106)	0.062	57
UoA23	-0.225**	(0.0886)	0.117	-0.587***	(0.0751)	0.637	-0.599***	(0.0723)	0.748	38
UoA24	-0.487***	(0.0663)	0.747	-0.838***	(0.195)	0.615	-0.857***	(0.132)	0.732	16

Table 6 (continued)

UoA	Environment			Impact			Output			Observations
	$\log E_{14}$	SE	R-square	$\log I_{14}$	SE	R-square	$\log O_{14}$	SE	R-square	
UoA25	-0.606***	(0.109)	0.604	-0.750***	(0.0990)	0.624	-0.512***	(0.125)	0.323	28
UoA26	-0.594***	(0.0524)	0.726	-0.639***	(0.103)	0.509	-0.532***	(0.0757)	0.479	72
UoA27	-0.511***	(0.107)	0.398	-0.565**	(0.258)	0.229	-0.361***	(0.0826)	0.248	61
UoA28	-0.749***	(0.167)	0.602	-0.946	(0.535)	0.228	-0.582	(0.338)	0.359	15
UoA29	-0.242**	(0.108)	0.144	-0.973***	(0.304)	0.333	-0.470	(0.313)	0.208	26
UoA30	-0.263**	(0.124)	0.235	-0.201	(0.286)	0.040	-0.425***	(0.109)	0.386	23
UoA31	-0.571***	(0.0696)	0.611	-0.759***	(0.0631)	0.606	-0.665***	(0.111)	0.464	62
UoA32	-0.403***	(0.0585)	0.467	-0.631***	(0.182)	0.407	-0.464***	(0.0722)	0.442	57
UoA33	-0.297**	(0.110)	0.149	-0.544***	(0.172)	0.373	-0.518***	(0.106)	0.475	43

$\log E_{14}$, $\log I_{14}$, and $\log O_{14}$ represent logarithm of the environment, impact and output GPA levels in 2014, respectively. The dependent variable is the growth of the environment, impact and output GPA between two assessment periods in each respective case. The constant is included, but not reported. SE stands for robust standard errors

UOAI: Clinical Medicine; UOAI2: Public Health, Health Services and Primary Care; UOAI3: Allied Health Professions, Dentistry, Nursing and Pharmacy; UOAI4: Psychology, Psychiatry and Neuroscience; UOAI5: Biological Sciences; UOAI6: Agriculture, Veterinary and Food Science; UOAI7: Earth Systems and Environmental Sciences; UOAI8: Chemistry; UOAI9: Physics; UOAI10: Mathematical Sciences; UOAI11: Computer Science and Informatics; UOAI12: Engineering; UOAI13: Architecture, Built Environment and Planning; UOAI14: Geography, Environmental Studies and Archaeology; UOAI15: Economics and Econometrics; UOAI16: Business and Management Studies; UOAI17: Law; UOAI18: Politics and International Studies; UOAI19: Social Work and Social Policy; UOAI20: Sociology; UOAI21: Anthropology and Development Studies; UOAI22: Education; UOAI23: Sport and Exercise Sciences, Leisure and Tourism; UOAI24: Area Studies; UOAI25: Modern Languages and Linguistics; UOAI26: English Language and Literature; UOAI27: History; UOAI28: Classics; UOAI29: Philosophy; UOAI30: Theology and Religious Studies; UOAI31: Art and Design; History, Practice and Theory; UOAI32: Music, Drama, Dance and Performing Arts; UOAI33: Communication, Cultural and Media Studies, Library and Information Management

***, **, and *Represent significance level at the 1%, 5% and 10%, respectively

Table 7 Convergence between subject areas in different research categories

	(1) Environment	(2) Impact	(3) Output
$\log\text{GPA}_{2014}^{\text{Environment}}$	- 0.243** (0.102)		
$\log\text{GPA}_{2014}^{\text{Impact}}$		- 0.431*** (0.0826)	
$\log\text{GPA}_{2014}^{\text{Output}}$			- 0.158 (0.0939)
Constant	0.316** (0.120)	0.544*** (0.0954)	0.263** (0.1000)
Observations	33	33	33
R-square	0.109	0.433	0.062

Robust standard errors in parentheses

Dependent variable is the growth of the environment, impact and output GPA between two assessment periods in columns 1, 2, and 3, respectively

***, **, and *Represent significance level at the 1%, 5% and 10%, respectively

between two assessment periods to the subject areas that performed weakly in the REF, 2014 to improve the overall university performance. Furthermore, based on the absolute convergence theory, the same amount of investment in areas with relatively weaker initial performance has a higher return than those with stronger initial performance. Therefore, we conducted our analysis to examine whether there was an absolute convergence across different subject areas in a given university. The results are presented in Table 8 for 40 universities that submitted at least 14 discipline areas to REF, 2014 and REF, 2021.⁵ With few exceptions, we found an absolute convergence in environment, impact and output quality across different disciplines in a given university. In all categories of research (i.e., environment, impact and output), there has been an absolute convergence between disciplines in 32 universities. On the other hand, there was no significant convergence in either category of research quality across different disciplines at Liverpool John Moores University. Furthermore, there was no significant convergence across disciplines in the research environment in 5 additional universities (i.e., University of Bristol, University of East Anglia, University of Durham, Birkbeck College, and University of Essex) and impact quality in 2 additional universities (i.e., Keele University and University of Hull). On the other hand, except for Liverpool John Moores University, the impact quality across disciplines in 39 universities converged between two assessment periods. Finally, the convergence across disciplines was fastest in the environment, impact and output categories in 5, 27 and 7 universities, respectively. Overall, with few exceptions, we found a convergence in research quality across subject areas in universities.

⁵ If a university had more than one submission in a given subject area (UOA), we obtained FTE-weighted GPA levels in different research categories to have a single university performance score in a given subject area.

Table 8 Convergence between subject areas (UOAs) in a given university

Institution name	Environment			Impact			Output			Obs.
	$\log E_{i,t}$	SE	R2	$\log I_{i,t}$	SE	R2	$\log O_{i,t}$	SE	R2	
	University College London	-0.528***	(0.109)	0.438	-0.548***	(0.173)	0.275	-0.412***	(0.109)	
University of Oxford	-0.621***	(0.123)	0.468	-0.445**	(0.178)	0.173	-0.644***	(0.211)	0.342	28
University of Cambridge	-0.821***	(0.220)	0.431	-0.820***	(0.220)	0.364	-0.373**	(0.158)	0.197	28
University of Manchester	-0.409**	(0.162)	0.182	-0.845***	(0.151)	0.522	-0.534***	(0.144)	0.342	28
University of Birmingham	-0.665***	(0.121)	0.533	-0.799***	(0.132)	0.556	-0.647***	(0.117)	0.507	27
University of Nottingham	-0.594***	(0.153)	0.424	-0.540***	(0.133)	0.403	-0.884***	(0.179)	0.587	27
University of Bristol	-0.227	(0.153)	0.072	-0.724***	(0.138)	0.600	-0.536***	(0.0845)	0.516	27
University of Leeds	-0.449***	(0.0935)	0.317	-0.826***	(0.242)	0.261	-0.294*	(0.147)	0.104	26
Newcastle University	-0.425***	(0.0958)	0.408	-0.573***	(0.166)	0.310	-0.459**	(0.193)	0.237	25
King's College London	-0.866***	(0.134)	0.599	-0.721***	(0.201)	0.546	-0.462***	(0.148)	0.253	24
University of Southampton	-0.613***	(0.139)	0.416	-0.911***	(0.135)	0.799	-0.488***	(0.142)	0.315	24
University of Exeter	-0.344**	(0.131)	0.165	-0.576**	(0.239)	0.236	-0.732**	(0.290)	0.255	24
University of Sussex	-0.501***	(0.147)	0.353	-0.962***	(0.101)	0.790	-0.817***	(0.0967)	0.589	24
University of Kent	-0.539**	(0.198)	0.410	-0.605**	(0.232)	0.363	-0.759***	(0.146)	0.595	23
University of Sheffield	-0.535***	(0.115)	0.416	-0.680***	(0.184)	0.333	-0.574***	(0.125)	0.434	23
University of York	-0.655***	(0.111)	0.613	-1.021***	(0.122)	0.764	-0.739***	(0.133)	0.550	23
University of East Anglia	-0.262	(0.219)	0.044	-0.822***	(0.109)	0.519	-0.739***	(0.198)	0.416	23
University of Liverpool	-0.475***	(0.155)	0.306	-0.937***	(0.225)	0.370	-0.514***	(0.0947)	0.444	22
University of Durham	-0.282	(0.181)	0.105	-0.693***	(0.186)	0.367	-0.734***	(0.121)	0.600	22
University of Warwick	-0.551***	(0.179)	0.497	-0.515***	(0.112)	0.238	-0.426*	(0.223)	0.244	22
University of Reading	-0.453**	(0.191)	0.187	-0.685***	(0.190)	0.264	-1.242***	(0.330)	0.393	20
University of Leicester	-0.415***	(0.106)	0.399	-0.573***	(0.0742)	0.405	-0.572**	(0.201)	0.286	19
Queen Mary U. of London	-0.712***	(0.118)	0.710	-0.861***	(0.169)	0.621	-0.519***	(0.162)	0.298	18

Table 8 (continued)

Institution name	Environment		Impact		Output		Obs.			
	$\log E_{i,t}$	SE	R2	$\log I_{i,t}$	SE	R2		$\log O_{i,t}$	SE	R2
University of Lincoln	- 0.810***	(0.176)	0.248	- 1.032***	(0.122)	0.899	- 0.550***	(0.147)	0.452	17
Open University	- 0.390***	(0.105)	0.272	- 0.682**	(0.283)	0.463	- 0.452**	(0.201)	0.195	17
University of Plymouth	- 0.740***	(0.138)	0.579	- 0.917***	(0.201)	0.438	- 0.872***	(0.257)	0.534	17
Brunel University London	- 0.860***	(0.255)	0.471	- 0.807*	(0.410)	0.305	- 0.292**	(0.113)	0.193	16
Keele University	- 0.495***	(0.148)	0.387	- 0.939***	(0.162)	0.596	- 0.247	(0.218)	0.094	16
Lancaster University	- 0.651***	(0.0986)	0.627	- 0.991**	(0.334)	0.461	- 0.958**	(0.391)	0.324	16
Oxford Brookes University	- 0.387*	(0.204)	0.263	- 1.112***	(0.117)	0.713	- 0.616*	(0.285)	0.296	15
Royal Holloway, U. London	- 0.537**	(0.240)	0.322	- 0.815**	(0.357)	0.284	- 0.721**	(0.328)	0.324	15
University of Chester	- 0.443**	(0.168)	0.366	- 0.900***	(0.0598)	0.875	- 0.778***	(0.0924)	0.727	15
Anglia Ruskin University	- 0.755**	(0.318)	0.262	- 0.859***	(0.194)	0.618	- 0.754***	(0.202)	0.553	14
Liverpool John Moores U	- 0.474	(0.321)	0.141	- 1.078	(0.613)	0.196	- 0.254	(0.260)	0.069	14
Loughborough University	- 0.499***	(0.0721)	0.721	- 0.559***	(0.168)	0.470	- 0.775***	(0.156)	0.604	14
Nottingham Trent University	- 0.626**	(0.242)	0.388	- 0.970***	(0.181)	0.593	- 0.857***	(0.0875)	0.827	14
U. of Central Lancashire	- 0.967*	(0.527)	0.288	- 0.580***	(0.161)	0.488	- 0.698**	(0.264)	0.429	14
University of Hull	- 0.658***	(0.156)	0.523	- 0.888***	(0.166)	0.614	- 0.340	(0.211)	0.263	14
Birkbeck College	- 0.258	(0.188)	0.115	- 0.809**	(0.299)	0.291	- 0.562**	(0.239)	0.336	14
University of Essex	0.0641	(0.134)	0.006	- 1.064***	(0.241)	0.635	- 0.545***	(0.0873)	0.654	14

$\log E_{i,t}$, $\log I_{i,t}$, and $\log O_{i,t}$ represent logarithm of the environment, impact and output GPA levels in 2014, respectively
 The dependent variable is the growth of the environment, impact and output GPA between two assessment periods in each respective case. The constant is included, but not reported. SE stands for robust standard errors

***, **, and *Represent significance level at the 1%, 5% and 10%, respectively

Table 9 Convergence in mainstream QR funding between universities

logQRfunding _{2021–2022}	– 0.0865*** (0.0200)
Constant	1.616*** (0.317)
Observations	119
R-square	0.156

Robust standard errors in parentheses

Dependent variable is the growth of the mainstream QR funding between 2021–2022 and 2022–2023 funding periods

***, **, and *Represent significance level at the 1%, 5% and 10%, respectively

Overall, our findings highlight that there has been a convergence between REF (2014) and (2021) across English universities. This is also reflected in the mainstream quality-related research (QR) funding allocations in the 2021–2022 and 2022–2023 periods, where the former funding period used the REF, 2014 results for allocating funds, and the latter funding period was based on the REF, 2021 results (see Research England (2022) for the details of the funding allocation).

The University of Oxford, University College London, University of Cambridge, and Imperial College London obtained the highest QR funding in the 2021–2022 funding period, accounting for 27.13% of the total mainstream QR funding distributed.⁶ However, all these universities experienced a decline in their share in the total mainstream QR funding allocation in the 2022–2023 period, and their total share in the mainstream QR funding declined to 23.99% of the total mainstream QR funding. On the other hand, some of the so-called ‘modern’ universities, such as the University of Northumbria and Manchester Metropolitan University, experienced major increases in their mainstream QR funding between 2021–2022 and 2022–2023. Mainstream QR funding allocated to the University of Northumbria and Manchester Metropolitan University increased roughly by £10.4 million and £5.7 million, and their share in the total mainstream QR funding also increased by 0.74% and 0.38%, respectively. To test whether there has been an absolute convergence in the QR funding received by the universities, we also carried out a convergence analysis. The dependent variable is the growth in the mainstream QR funding between the 2021–2022 and 2022–2023 periods (i.e., the difference between the logarithm of the QR funding in 2022–2023 and 2021–2022 periods), and the independent variable is the logarithm of the initial funding allocation (i.e., the logarithm of the mainstream QR funding in 2021–2022 period), and Table 9 presents the results. The coefficient on the logarithm of the mainstream QR funding in 2021–2022 is negative and significant, highlighting an absolute convergence in the mainstream QR funding. Therefore, the convergence in the research quality between the universities is also reflected in the convergence in QR funding.

It should be noted that the mainstream quality-related research (QR) funding is distributed based on the four-stage formula of Research England (see pages 11 and 16 of Research England (2022) for the detailed stages used by Research England to allocate QR funding to

⁶ The mainstream QR funding allocation data can be obtained here: <https://www.ukri.org/what-we-offer/what-we-have-funded/research-england/>. We used the mainstream QR funding allocations to the institutes in the 2021–2022 and 2022–2023 periods without considering the London weighting on mainstream QR.

higher education providers). In the first stage, 60%, 25% and 15% of the QR funding are allocated to the output, impact and environment pots, respectively. In the second stage, each pot's funding is distributed to each main panel subject based on the proportion of funding allocated to each panel in 2021–2022. Finally, in stages 3 and 4, the total funding in each pain panel is distributed, firstly between its constituent UOAs, and finally between universities, respectively. The QR funding is allocated to the universities proportional to their volume of activity reaching the 3* and 4* quality levels in the REF, multiplied by quality and cost weights. For instance, some subjects are allocated 60% more funding than others (i.e., high-cost laboratory and clinical subjects versus low-cost subjects), which is determined by the cost weights. Quality weights allocated to 3* and 4* research activities were 1 and 4, respectively. Therefore, the funding allocated for the 4* research quality was four times more than those allocated to the 3* research activity, and no funding is provided to the research activities that did not reach the 3* research quality. Finally, universities in inner and outer London are allocated 12 and 8 percent more QR funding than those outside London. Therefore, there is no straightforward relationship between GPA scores and QR funding obtained by universities. However, there is a close correlation between GPA scores and mainstream QR funding.⁷

Robustness analyses

The existing studies obtained the average research quality of the universities by the weighted arithmetic mean of quality scores of different research activities and researchers (see e.g., Buckle et al., 2020, 2022; Pinar and Unlu, 2022). The baseline analyses of this paper rely on the average quality scores of universities (subjects) based on the interval quality scores used. An interval level measure means that each 1-point difference is equivalent to every 1-point difference throughout the scale. In the English system, lengths in feet can be compared because the constituent units of feet (inches) are additive. The GPA in the REF is an aggregation of categorical ratings (i.e., 4-star, 3-star, 2-star, 1-star, and unclassified). Furthermore, the GPA is derived by a linear combination of classification as described in Eq. 1 of the paper. It should be noted that the weights in this formula (4, 3, 2, 1, 0) are at the ordinal level. However, there is no consensus that the difference between “world-leading research” (4-star research activity) and “internationally excellent” (3-star research activity) is the same as the difference between “internationally recognized” (2-star activity) and “nationally recognized” (1-star research activity). Therefore, we conduct a robustness analysis in this section to examine whether alternative construction of average quality scores of universities (subjects) alters or confirms the baseline findings.

Research England relies on the REF results to distribute QR funding, which relies on quality, cost, and London weights (see previous section for details). For instance, QR funding allocated by Research England is based on quality weights. In particular, the quality weights given to 4-star and 3-star research activities are 4 and 1, respectively, and the 2-star, 1-star, and unclassified research activities were given zero weights. In other words, Research England places significantly more importance (weight) on 4-star and 3-star research activities. Hence, we obtained GPAs for universities and subjects with alternative weight allocation across different starred scores in research activities to test

⁷ We regressed the mainstream QR funding received by universities in the 2022–2023 period on the overall GPA scores of universities in REF2021 (both in logs), and found that 65% of the variation in QR funding in the 2022–2023 period is explained by the variation in GPA scores, and 1% increase in GPA would lead to 8.99% increase in QR funding. The results are available upon request.

the convergence across universities and subjects. GPA_1 is obtained by allocating 4 and 1 weights to the 4-star and 3-star research activities, and zero weights to the rest of the quality categories, and GPA_2 obtained by allocating 4, 2 and 1 weights to the 4-star, 3-star and 2-star research activities, and zero weights to the rest of the quality categories. Furthermore, we also obtained the percentage of research activity that is evaluated as 4-star and 3-star by the universities and subject areas as an additional quality measure. Finally, we also obtained quality-weighted volume, which Research England uses to distribute QR funding across subjects and universities. Quality-weighted volume is obtained by the multiplication of the GPA_1 and the size of the universities (i.e., FTE returned by the universities). Finally, we also tested the convergence in the size of the universities and subject areas to examine whether there has been a convergence or divergence in submission sizes of universities and subject areas between the two assessment periods.

We repeat the same exercise in Table 3 but use different composite quality scores in this case. Table 10 represents the convergence in the environment, impact and output categories across universities when we use different quality measures. The results are presented in columns (1)–(3), (4)–(6), (7)–(9), and (10)–(12) when we use GPA_1 , GPA_2 , percentage of 4-star and 3-star activities, and quality-weighted volume as alternative research quality, respectively. The dependent variable is the growth in alternative quality measures, and finally, in column 13 of Table 10, we use the growth in the size of the submissions (FTE) between two periods. On the other hand, the independent variable is the initial achievement scores in 2014 based on alternative quality scores between columns (1)–(12), and the initial submission size in 2014 in column (13). Overall, we found that irrespective of the quality measures used, we found that there has been a significant convergence in research quality in different research categories across universities. One finding that is worth noting is that the coefficient of the initial quality scores in REF, 2014 shows variation. For instance, the size of the negative coefficient is relatively larger in the environment category when GPA_1 is used as a quality measure. On the other hand, the size of the negative coefficient is relatively larger in impact category when GPA_2 is used as a quality measure. Finally, the size of the negative coefficient is relatively larger in the output category when the percentage of 4-star and 3-star research quality and the quality-weighted volume measures are used. Therefore, the speed of convergence across universities is faster in different research categories depending on the quality measures used. Finally, we also found that there has been a significant convergence in submission sizes across universities between two assessment periods (column 13 of Table 10). In other words, the universities that had smaller submission sizes in REF, 2014 experienced a higher growth in their submission sizes between the two assessment periods. The changes in the submission rules between REF (2014) and (2021) could explain this finding. The universities had to submit all the staff members with a significant research responsibility in REF, 2021. Whereas universities were allowed to choose the staff members that they could submit to the REF, 2014. Therefore, universities that were relatively small in the REF, 2014 submission may have been more selective in choosing the staff members to return in their submission; however, with the implementation of the so-called universal submission in REF, 2021, the size of the universities with smaller submissions in REF, 2014 experienced a larger growth in submission sizes.

We also carried out similar analyses to examine the convergence across subject areas (units of assessment). The analyses are similar to that of Table 7, but instead, we use different quality scores, and the results are presented in Table 11. When we used GPA scores based on new weights given to different quality classifications (i.e., GPA_1 and GPA_2), there is significant convergence in the environment and impact composite scores between different subject areas, but there is no significant convergence in output quality between different

Table 10 Convergence across universities in different research categories with alternative composite quality measures between REF (2014) and (2021)

Dependent variable	(1) Growth in GPA_1		(2) Growth in GPA_2		(3) Growth in percentage of 4-star and 3-star research activity		(4) Growth in quality-weighted volume		(5) FTE size	
	Environment	Impact	Environment	Impact	Environment	Impact	Environment	Impact	Environment	Impact
$\log(\text{Score})_{14}$	-0.319*** (0.0528)	-0.306*** (0.111)	-0.235*** (0.0528)	-0.451*** (0.115)	-0.470*** (0.0559)	-0.315** (0.142)	-0.501*** (0.0787)	-0.185*** (0.0364)	-0.0876** (0.0347)	-0.0894*** (0.0188)
Constant	0.251*** (0.0450)	0.261*** (0.0642)	0.362*** (0.0188)	0.434*** (0.0970)	2.049*** (0.243)	1.416** (0.610)	2.245*** (0.327)	1.716*** (0.230)	1.122*** (0.229)	1.304*** (0.116)
Observations	117	116	120	120	117	116	120	117	116	120
R-square	0.339	0.204	0.195	0.402	0.517	0.247	0.553	0.301	0.078	0.176

The dependent variable is the growth in composite quality scores of universities between REF (2014) and (2021), respectively. Dependent variables in columns (1)–(3), (4)–(6), (7)–(9), and (10)–(12) are the growth in GPA_1 , GPA_2 , percentage of 4-star and 3-star activities, and quality-weighted volume between REF (2014) and (2021), respectively. Dependent variable in column (13) is the growth in submission size of the universities (FTE) between REF (2014) and (2021). The independent variables are the logarithm of the respective scores in 2014. Robust standard errors in parentheses
 ***, **, and *Represent significance level at the 1%, 5% and 10%, respectively

Table 11 Convergence across subject areas in different research categories with alternative composite quality measures between REF (2014) and (2021)

Dependent variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Research category	Growth in GPA_1			Growth in GPA_2			Growth in percentage of 3-star research activity			Growth in quality-weighted volume			Growth in size
	Environ-ment	Impact	Output	Environ-ment	Impact	Output	Environ-ment	Impact	Output	Environ-ment	Impact	Output	FTE
$\log(\text{Score})_{14}$	-0.354*** (0.0944)	-0.496*** (0.0831)	-0.140 (0.115)	-0.297*** (0.0953)	-0.466*** (0.0833)	-0.0893 (0.108)	-0.229* (0.134)	-0.419*** (0.100)	-0.382*** (0.0750)	0.0149 (0.0400)	0.0341 (0.0446)	0.0506 (0.0432)	0.0840** (0.0372)
Constant	0.357*** (0.0739)	0.491*** (0.0579)	0.368*** (0.0385)	0.348*** (0.0942)	0.528*** (0.0786)	0.249*** (0.0807)	1.055* (0.594)	1.900*** (0.442)	1.783*** (0.323)	0.331 (0.308)	0.225 (0.339)	0.308 (0.324)	-0.230 (0.257)
Observations	33	33	33	33	33	33	33	33	33	33	33	33	33
R-square	0.210	0.510	0.027	0.156	0.481	0.015	0.082	0.324	0.376	0.005	0.026	0.043	0.135

The dependent variable is the growth in composite quality scores of subject areas between REF (2014) and (2014) respectively. Dependent variables in columns (1)–(3), (4)–(6), (7)–(9), and (10)–(12) are the growth in GPA_1 , GPA_2 , percentage of 4-star and 3-star activities, and quality-weighted volume between REF (2014) and (2021), respectively. Dependent variable in column (13) is the growth in submission size of the universities (FTE) between REF (2014) and (2021). The independent variables are the logarithm of the respective scores in 2014. Robust standard errors in parentheses

***, **, and *Represent significance level at the 1%, 5% and 10%, respectively

subject areas. These findings align with the results presented in Table 7 when the average composite score is obtained using the GPA formula. When we consider the growth in the percentage of 4-star and 3-star research activity (columns 7–9 of Table 11) as a dependent variable and the percentage of 4-star and 3-star research activity in 2014 as our independent variable, we find that there is significant convergence in all three elements of research activity. In other words, the subject areas with relatively lower percentages of 4-star and 3-star research activity in the environment, impact and output categories in REF, 2014 experienced a higher growth in the percentages of 4-star and 3-star research activity between the two assessment periods. In other words, when there is no distinction between 3-star and 4-star research activity, we also found a significant convergence in output quality between subject areas. On the other hand, when we examined the convergence in size of different subject areas (column 13 of Table 11), we found a significant divergence in sizes between subject areas. In other words, the growth in size of the subject areas tends to be relatively higher for subject areas that had larger submission sizes in REF, 2014. For instance, Theology and Religious Studies, Classics and Area Studies were the subject areas that had the smallest submission sizes in REF, 2014 (i.e., with submission sizes of 322.82, 330.95, and 443.05, respectively), and the growth in sizes of these units between REF (2014) and (2021) was 25.1%, 16.6%, and 22.6%, respectively. On the other hand, Business and Management Studies, Clinical Medicine, and Engineering had the largest submission sizes in REF, 2014 (with submission sizes of 2746.67, 2934.20 and 4158.18, respectively), and the growth in sizes of these units between two assessments was 102.4%, 37.4% and 46.6%, respectively. In other words, the growth of the size of the larger units tend to be higher. Furthermore, beyond these units, there was a tendency for the larger subject areas in REF, 2014 to experience higher growth in their sizes between two assessment periods, and there was an increased divergence between subject areas regarding research active staff members. Finally, columns 9–12 of Table 11 report the results when we use quality-weighted volume (which is obtained by the multiplication of the GPA_1 and the size of the subject areas) as a measure of quality. As there was a significant convergence in quality when the GPA_1 is used and divergence in submission sizes between subject areas, we found no significant convergence or divergence between subject areas when the quality-weighted volume is used as a proxy. Overall, with alternative measures of research quality, we found a convergence in the quality of research between subject areas but a divergence in submission sizes across subject areas.

Conclusions

It has been argued that the research assessment system in the United Kingdom aimed at increasing the concentration of high-quality researchers in a handful number of universities (Buckle et al., 2020) and would lead to a concentration of research funding in a few research-intensive universities (Martin, 2011; Torrance, 2020). In other words, previous research argued that the quality difference between better and worse performers is expected to increase, resulting in increased research-quality funding differences between institutes as these funds are distributed based on the REF results. However, contrary to the expectations of increased divergence between institutes, we found that the research performance of universities and disciplines converged between the two REF exercises. Furthermore, we also found that the convergence in the research quality between universities also resulted in a convergence in mainstream QR funding between universities, and there has been a decline in research funding differences between universities.

We also found that impact quality convergence within universities (i.e., convergence across disciplines in universities) and between universities was the fastest. Even though the inclusion of the impact element in the evaluation of the REF, 2014 resulted in higher performance differences between universities in REF, 2014 (Pinar & Unlu, 2020a), the convergence in impact quality was fastest between REF (2014) and (2021). This is because there has been increased emphasis on the impact element in academia to increase engagement with the societal impact. Between the two assessment periods, further guidance on how quantitative metrics could be used to support impact case studies was provided (Parks et al., 2018), and funding councils carried out workshops to provide additional guidance on the requirements of the impact element (REF, 2018). Therefore, these additional investments and training sessions also resulted in faster convergence in this research area within and between universities and disciplines.

The findings of the paper have various policy implications. Firstly, the financial incentives of the REF led to an increase in the performance of the universities between the two assessment periods, and the increase in the performance of the worse performers in REF, 2014 was relatively higher between REF (2014) and (2021). In other words, the PBRFS in the UK served as an incentive to improve the scientific quality produced in the country. Secondly, significant improvements in research quality between the two assessment periods by the universities that performed poorly in REF, 2014 could also improve their rankings in the league tables because research quality is part of the university rankings. Therefore, these universities are more likely to attract more undergraduate and graduate students in the upcoming years based on their research quality improvements because the research quality affects student university choice (Broecke, 2015; Gibbons et al., 2015). Thirdly, convergence in research quality across universities could also allow students to access better-performing universities when they are less likely to move outside the region for their studies. For instance, Donnelly and Gamsu (2018) demonstrate that state-educated and ethnic minority students were much less likely to leave their areas to study in other regions in the UK. Therefore, students that are less likely to move to other areas are likely to benefit from the increased research quality and convergence in research quality across universities. Fourthly, we found that the economic and societal impact of the universities also improved between the two assessment periods. The increased quality of the impactful research could also lead to increased entrepreneurial activity and economic growth at the regional level based on the university-industry links (Bishop et al., 2011; D'Este et al., 2013; Lehman and Menter, 2016; Mueller, 2016b). Therefore, convergence in research quality across universities could also act as a source in reducing regional disparities in the UK. Finally, one of the significant changes between REF (2014) and (2021) was that the universities were required to return all the staff members with significant research responsibility to REF. Therefore, it is likely that the universities may pursue a different hiring process by hiring new staff members and are more likely to offer research contracts to new staff members with a good record of high-quality publications to improve their research quality scores in the next REF assessment.

Future research could evaluate the strategies used by the universities that experienced relatively higher growth in their research performance between two assessment periods to provide guidance on the practices for improving research quality. Furthermore, this paper examined whether there has been an absolute convergence in research quality between universities and different subject areas between REF (2014) and (2021). However, this paper did not examine the factors that explain the convergence or the lack of convergence across universities and subject areas. Therefore, a future study could conduct a regression analysis

to examine the factors that explain the variation in score changes between two REF exercises to provide further insights about the convergence or lack of convergence.

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

References

- Abramo, G., & D'Angelo, C. A. (2021). The different responses of universities to introduction of performance-based research funding. *Research Evaluation*, *30*(4), 514–528.
- Barrios González, C., Flores, E., & Martínez, M. A. (2021). Scientific production convergence: An empirical analysis across nations. *Minerva*, *59*, 445–467.
- Barro, R. J., & Sala-i-Martin, X. (1991). Convergence. *Journal of Political Economy*, *100*, 223–251.
- Bishop, K., D'Este, P., & Neely, A. (2011). Gaining from interactions with universities: Multiple methods for nurturing absorptive capacity. *Research Policy*, *40*(1), 30–40.
- Boliver, V. (2013). How fair is access to more prestigious UK universities? *The British Journal of Sociology*, *64*(2), 344–364.
- Broecke, S. (2015). University rankings: Do they matter in the UK? *Education Economics*, *23*(2), 137–161.
- Buckle, R. A., Creedy, J., & Gemmill, N. (2020). Is external research assessment associated with convergence or divergence of research quality across universities and disciplines? Evidence from the PBRF process in New Zealand. *Applied Economics*, *52*(36), 3919–3932.
- Buckle, R. A., Creedy, J., & Gemmill, N. (2022). Sources of convergence and divergence in university research quality: Evidence from the performance-based research funding system in New Zealand. *Scientometrics*, *127*, 3021–3047.
- Callender, C., & Mason, G. (2017). Does student loan debt deter higher education participation? New evidence from England. *Annals of the American Academy of Political & Social Science*, *671*(1), 20–48.
- Checchi, D., Malgarini, M., & Sarlo, S. (2019). Do performance-based research funding systems affect research production and impact? *Higher Education Quarterly*, *73*(1), 45–69.
- Checchi, D., Mazzotta, I., Momigliano, S., & Olivanti, F. (2020). Convergence or polarisation? The impact of research assessment exercises in the Italian case. *Scientometrics*, *124*, 1439–1455.
- Chikoore, L., Proberts, S., Fry, J., & Creaser, C. (2016). How are UK academics engaging the public with their research? *A Cross-Disciplinary Perspective. Higher Education Quarterly*, *70*(2), 145–169.
- Civera, A., Lehmann, E. E., Paleari, S., & Stockinger, S. A. E. (2020). Higher education policy: Why hope for quality when rewarding quantity? *Research Policy*, *49*(8), 104083.
- Crespo Cuaresma, J., Klasen, S., & Wacker, K. M. (2022). When do we see poverty convergence? *Oxford Bulletin of Economics and Statistics*, *84*(6), 1283–1301.
- De Fraja G., Facchini, G., Gathergood, J. (2016). How much is that star in the window? Professorial salaries and research performance in UK Universities. NICEP Working Paper Series 2016–08. ISSN 2397–9771. October 28 2022 Retrieved from <https://www.nottingham.ac.uk/research/groups/nicep/documents/working-papers/2016-08-de-fraja-facchini-gathergood.pdf>
- de Rijcke, S., Wouters, P. F., Rushforth, A. D., Franssen, T. P., & Hammarfelt, B. (2016). Evaluation practices and effects of indicator use—a literature review. *Research Evaluation*, *25*(2), 161–169.
- Docampo, D., & Bessoule, J.-J. (2019). A new approach to the analysis and evaluation of the research output of countries and institutions. *Scientometrics*, *119*, 1207–1225.
- Donnelly, M., & Gamsu, S. (2018). Regional structures of feeling? A spatially and socially differentiated analysis of UK student im/mobility. *British Journal of Sociology of Education*, *39*(7), 961–981.
- Dougherty, K. J., Jones, S. M., Lahr, H., Natow, R. S., Pheatt, L., & Reddy, V. (2016). *Performance funding for higher education*. Johns Hopkins University Press.
- D'Este, P., Guy, F., & Iammarino, S. (2013). Shaping the formation of university–industry research collaborations: What type of proximity does really matter? *Journal of Economic Geography*, *13*(4), 537–558.

- Erfanian, E., & Ferreira Neto, A. B. (2017). Scientific output: Labor or capital intensive? An analysis for selected countries. *Scientometrics*, *112*, 461–482.
- Ganong, P., & Shoag, D. (2017). Why has regional income convergence in the U.S. declined? *Journal of Urban Economics*, *102*, 76–90.
- Gibbons, S., Neumayer, E., & Perkins, R. (2015). Student satisfaction, league tables and university applications: Evidence from Britain. *Economics of Education Review*, *48*, 148–164.
- Grisorio, M. J., & Prota, F. (2020). Italy's national research assessment: Some unpleasant effects. *Studies in Higher Education*, *45*(4), 736–754.
- Guerrero, M., Cunningham, J. A., & Urbano, D. (2015). Economic impact of entrepreneurial universities' activities: An exploratory study of the United Kingdom. *Research Policy*, *44*(3), 748–764.
- Gunn, N. (2018). Metrics and methodologies for measuring teaching quality in higher education: developing the Teaching Excellence Framework (TEF). *Educational Review*, *70*(2), 129–148.
- Hancock, S. (2020). The employment of PhD graduates in the UK: what do we know? Higher Education Policy Institute. October 20 2022. Retrieved from <https://www.hepi.ac.uk/2020/02/17/the-employment-of-phd-graduates-in-the-uk-what-do-we-know/>
- Hewitt-Dundas, N. (2012). Research intensity and knowledge transfer activity in UK universities. *Research Policy*, *41*(2), 262–275.
- Hicks, D. (2012). Performance-based university research funding systems. *Research Policy*, *41*, 251–261.
- Hillman, N. (2016). The Coalition's higher education reforms in England. *Oxford Review of Education*, *42*(3), 330–345.
- Hnatkova, E., Degtyarova, I., Keresshot, M., & Boman, J. (2022). Labour market perspectives for PhD graduates in Europe. *European Journal of Education*, *57*(3), 395–409.
- Jeon, J., & Kim, S. Y. (2018). Is the gap widening among universities? On research output inequality and its measurement in the Korean higher education system. *Quality & Quantity*, *52*, 589–606.
- Jurajda, Š, Kozubek, S., Münich, D., & Škoda, S. (2017). Scientific publication performance in post-communist countries: Still lagging far behind. *Scientometrics*, *112*, 315–328.
- Korytkowski, P., & Kulczycki, E. (2019). Examining how country-level science policy shapes publication patterns: The case of Poland. *Scientometrics*, *119*, 1519–1543.
- Kounetas, K. E. (2018). Energy consumption and CO2 emissions convergence in European Union member countries. A tonneau des Danaïdes? *Energy Economics*, *69*, 111–127.
- Lancho-Barrantes, B. S., & Cantú-Ortiz, F. J. (2019). Science in Mexico: A bibliometric analysis. *Scientometrics*, *118*, 499–517.
- Lehmann, E. E., & Menter, M. (2016). University–industry collaboration and regional wealth. *The Journal of Technology Transfer*, *41*, 1284–1307.
- Leišytė, L. (2016). New public management and research productivity – a precarious state of affairs of academic work in the Netherlands. *Studies in Higher Education*, *41*(5), 828–846.
- Lessmann, C., & Seidel, A. (2017). Regional inequality, convergence, and its determinants—A view from outer space. *European Economic Review*, *92*, 110–132.
- Li, X., & Lin, B. (2013). Global convergence in per capita CO2 emissions. *Renewable and Sustainable Energy Reviews*, *24*, 357–363.
- Mankiw, N. G., Romer, D., & Weil, D. N. (1992). A contribution to the empirics of economic growth. *Quarterly Journal of Economics*, *107*(2), 407–437.
- Manville, C., Guthrie, S., Henham, M.-L., Garrod, B., Sousa, S., Kirtley, A., Castle-Clarke, S., Ling, T. (2015). Assessing impact submissions for REF 2014: An evaluation. October 20 2022. Retrieved from www.rand.org/content/dam/rand/pubs/research_reports/RR1000/RR1032/RAND_RR1032.pdf
- Marques, M., Powell, J. J. W., Zapp, M., & Biesta, G. (2017). How does research evaluation impact educational research? Exploring intended and unintended consequences of research assessment in the United Kingdom, 1986–2014. *European Educational Research Journal*, *16*(6), 820–842.
- Martin, B. R. (2011). The Research Excellence Framework and the 'impact agenda': Are we creating a Frankenstein monster? *Research Evaluation*, *20*(3), 247–254.
- Matthews, A., & Kotzee, B. (2021). The rhetoric of the UK higher education teaching excellence framework: A corpus-assisted discourse analysis of TEF2 provider statements. *Educational Review*, *73*(5), 523–543.
- Mueller, C. E. (2016a). Accurate forecast of countries' research output by macro-level indicators. *Scientometrics*, *109*, 1307–1328.
- Mueller, P. (2016b). Exploring the knowledge filter: How entrepreneurship and university–industry relationships drive economic growth. *Research Policy*, *35*(10), 1499–1508.
- Nguyen, T. V., Ho-Le, T. P., & Le, U. V. (2017). International collaboration in scientific research in Vietnam: An analysis of patterns and impact. *Scientometrics*, *110*, 1035–1051.

- Parks, S., Ioppolo, B., Stepanek, M., Gunashekar, S. (2018). Guidance for standardising quantitative indicators of impact within REF case studies. Published by the RAND Corporation, Santa Monica, Calif., and Cambridge, UK. November 03 2022. Retrieved from <https://www.ref.ac.uk/media/1018/guidance-for-standardising-quantitative-indicators-of-impact.pdf>
- Passaretta, G., Trivellato, P., & Triventi, M. (2019). Between academia and labour market—the occupational outcomes of PhD graduates in a period of academic reforms and economic crisis. *Higher Education*, 77, 541–559.
- Pinar, M. (2020). It is not all about performance: Importance of the funding formula in the allocation of performance-based research funding in England. *Research Evaluation*, 29(1), 100–119.
- Pinar, M., & Horne, T. J. (2022). Assessing research excellence: Evaluating the Research Excellence Framework. *Research Evaluation*, 31(2), 173–187.
- Pinar, M., & Unlu, E. (2020a). 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.
- Pinar, M., & Unlu, E. (2020b). Determinants of quality of research environment: An assessment of the environment submissions in the UK’s research excellence framework in 2014. *Research Evaluation*, 29(3), 231–244.
- Quah, D. T. (1993). Galton’s fallacy and the convergence hypothesis. *Scandinavian Journal of Economics*, 95, 427–443.
- REF (2012). Panel criteria and working methods. REF 01.2012 January 2012. October 17 2022. Retrieved from https://www.ref.ac.uk/2014/media/ref/content/pub/panelcriteriaandworkingmethods/01_12.pdf
- REF (2014). Results and submissions. October 17 2022 Retrieved from [https://results.ref.ac.uk/\(S\(mbfqgnq3acrysgmkgm0ez0py\)\)](https://results.ref.ac.uk/(S(mbfqgnq3acrysgmkgm0ez0py)))
- REF (2018). Notes from impact workshop on continuing case studies and submitting research activities and bodies of work. November 03 2022 Retrieved from <https://www.ref.ac.uk/media/1039/ref-2021-impact-workshop-on-continuing-case-studies-and-research-activities-summary.pdf>
- REF (2019a). Panel criteria and working methods. REF 2019/02 January 2019. October 17 2022 Retrieved from https://ref.ac.uk/media/1450/ref-2019_02-panel-criteria-and-working-methods.pdf
- REF (2019b). Guidance on submissions. May 07 2023 Retrieved from https://www.ref.ac.uk/media/1447/ref-2019_01-guidance-on-submissions.pdf
- REF (2021). Results and submissions: Introduction to the REF results. October 17 2022 Retrieved from <https://results2021.ref.ac.uk/>
- REF (2022). What is the REF? October 04 2022 Retrieved from <https://www.ref.ac.uk/about-the-ref/what-is-the-ref/>
- Radosevic, S., & Yoruk, E. (2014). Are there global shifts in the world science base? Analysing the catching-up and falling behind of world regions. *Scientometrics*, 101, 1897–1924.
- Research England (2022). Research England: How we fund higher education providers. 17 May 17 2023 Retrieved from <https://www.ukri.org/wp-content/uploads/2023/02/RE-150223-HowWeFundHEPs.pdf>
- Rodríguez-Navarro, A., & Brito, R. (2022). The link between countries’ economic and scientific wealth has a complex dependence on technological activity and research policy. *Scientometrics*, 127, 2871–2896.
- Shattock, M. (2012). *Making Policy in British Higher Education 1945–2011*. McGraw-Hill.
- Shelton, R. D. (2020). Scientometric laws connecting publication counts to national research funding. *Scientometrics*, 123, 181–206.
- Solow, R. M. (1956). A contribution to the theory of economic growth. *Quarterly Journal of Economics*, 70(1), 65–94.
- Taylor, J. (2011). The assessment of research quality in UK universities: Peer review or metrics? *British Journal of Management*, 22, 202–217.
- Thorpe, A., Craig, R., Tourish, D., Hadikin, G., & Batistic, S. (2018). ‘Environment’ submissions in the UK’s research excellence framework 2014. *British Journal of Management*, 29, 571–587.
- Torrance, H. (2020). The research excellence framework in the United Kingdom: processes, consequences, and incentives to engage. *Qualitative Inquiry*, 26(7), 771–779.
- Universities UK (2018). Patterns and trends in UK higher education 2018. May 07 2023 Retrieved from <https://www.universitiesuk.ac.uk/sites/default/files/field/downloads/2021-08/patterns-and-trends-in-uk-higher-education-2018.pdf>
- Watermeyer, R., & Olssen, M. (2016). ‘Excellence’ and exclusion: The individual costs of institutional competitiveness. *Minerva*, 54, 201–218.
- Xu, X., & Reed, M. (2021). The impact of internet access on research output - a cross-country study. *Information Economics and Policy*, 56, 100914.
- Zacharewicz, T., Lepori, B., Reale, E., & Jonkers, K. (2019). Performance-based research funding in EU member states—a comparative assessment. *Science and Public Policy*, 46, 105–115.