City Research Online

City, University of London Institutional Repository

Citation: Banal-Estanol, A., Jofre-Bonet, M., Iori, G., Maynou, L., Tumminello, M. & Vassallo, P. (2023). Performance-based research funding: Evidence from the largest natural experiment worldwide. Research Policy, 52(6), 104780. doi: 10.1016/j.respol.2023.104780

This is the accepted version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: https://openaccess.city.ac.uk/id/eprint/30112/

Link to published version: https://doi.org/10.1016/j.respol.2023.104780

Copyright: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

Reuse: Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

City Research Online: http://openaccess.city.ac.uk/ publications@city.ac.uk

Performance-based research funding: Evidence from the largest natural experiment worldwide

Albert Banal-Estanol[1,2,3], Mireia Jofre-Bonet[1,4], Giulia Iori[1], Laia Maynou[5,6,7], Michele Tumminello[8], and Pietro Vassallo*[9]

[1]City, University of London, UK
[2]Universitat Pompeu Fabra, Barcelona, Spain
[3]Barcelona School of Management, Barcelona, Spain
[4]Office of Health Economics, UK
[5]Universitat de Barcelona, Barcelona, Spain
[6]London School of Economics and Political Science, UK
[7]Center for Research in Health and Economics (CRES-UPF), Barcelona, Spain
[8]University of Palermo, Italy
[9]Bank of Italy

March 21, 2023

Abstract

A performance-based research funding system (PRFS) is a nationwide incentive scheme that promotes and rewards university research performance through competition for government funding. The UK's PRFS, currently the Research Excellence Framework (REF), is considered the oldest, largest and most developed payment-by-results system in academia worldwide. Surprisingly, and despite the strong criticisms, little has been done to quantitatively and casually evaluate the intended and unintended effects of the PRFSs. In this paper, we evaluate the incremental impact of the REF 2014 in the fields of Economics and Business. We use a synthetic control method to compare the performance of UK universities with their artificial counterfactual units constructed using data from US universities. Our analysis shows, on the whole, that the introduction of the REF had a significant and positive impact on the quantity and quality of the scientific research produced at UK universities. However, we do not find a significant effect on the per author measures, suggesting that the REF did not result in an increase in research productivity. We also show that the effects are more heterogeneous across universities than across academic disciplines. We do not find evidence of a shift of research focus from Economics to Business topics, as some feared. But our analysis indicates that the REF 2014 may have contributed to the concentration of research excellence in elite institutions.

Keywords: Research Policy; Research Productivity; Research Excellence Framework; Performance-based research funding system; Synthetic Control Method.

* *Contact author:* Pietro Vassallo. Email: pietro.vassallo23@gmail.com.

1 Introduction

The United Kingdom's (UK) public research policy over the last 35 years has revolved around nationwide assessment exercises, firstly known as Research Assessment Exercises (RAEs) and subsequently as Research Excellence Frameworks (REFs). These exercises produce comparable ratings of research performance for all the departments of all the universities and public research institutions in the UK. Based on the results of these assessments, undertaken every three to seven years (1986, 1989, 1992, 1996, 2001, 2008, 2014, 2021), core research funding for the subsequent years is allocated. The resulting *performance-based research funding system* (PRFS) is considered the oldest, largest and most developed payment-by-results incentive scheme in academia worldwide (Zacharewicz et al. 2019), and it has inspired the design and implementation of PRFSs in many other countries, ranging from Italy to Australia (Sivertsen 2017, Geuna & Piolatto 2016).

The primary aim of the REF and of all the other PRFSs is to promote and reward research performance through competition for scarce resources while making universities more accountable. The PRFSs create powerful incentives to increase the quantity and quality of research, both directly, through financial rewards, as well as indirectly, through prestige (Hicks 2012). Indeed, in the UK, while the research councils allocate core government funding on the basis of the REF results, these results are also made public and are used to produce rankings of research performance of university departments. These rankings have a major effect on the fortunes of the universities. Universities, in turn, are induced to incentivise their faculty to publish more and better and/or to recruit more and better academics for their departments.

But, even after several modifications, the existing PRFSs worldwide are still receiving severe criticisms. Some commentators in the UK, for instance, question whether the REF is actually fostering high-quality research (e.g., the University and College Union). A recent mutual learning exercise set up by the European Commission (2020) also acknowledges that "In most countries, PRFS are contentious." As shown by the recent behavioural economics literature (Gneezy et al. 2011), incentives may not work or they may even backfire, for instance, because extrinsic incentives crowd out the researchers' intrinsic motivation to produce high-quality research (Andersen & Pallesen 2008). However, and despite the controversies, there are very few quantitative analyses evaluating the (intended) effects of the PRFSs, the main reason being the lack of good counterfactuals of the targeted universities to reach causality (Gläser et al. 2002, OECD 2010).

Further, critics also complain about the preparation costs and about the possibly more costly side-effects or indirect costs (Harman 2000, Martin & Whitley 2010, Martin 2011, Geuna & Piolatto 2016, Stern 2016). In terms of side-effects, some authors claim, for instance, that the PRFSs, favour the research-intensive, large and/or elite universities, thus increasing, even further, the concentration of research output (Bence & Oppenheim 2005, Martin & Whitley 2010, Clerides et al. 2011, Mingers & White 2015, Jeon & Kim 2018, Guizzo et al. 2021). Commentators in the UK also criticize that the REF allows for strategic behaviour, and for instance allow universities to strategically submit research output to certain discipline panels (e.g. Business) instead of others (e.g. Economics), possibly creating field distortions (Stern 2016, Marques et al. 2017, Johnston & Reeves 2020, Guizzo et al. 2021, Battistin & Ovidi 2022). Surprisingly, and despite the heated debate, little has been done to systematically evaluate the side-effects of the PRFS incentive schemes.

This paper investigates if the REF increases the research performance of UK universities using data from the areas of Economics and Business. We apply a synthetic control method (SCM), using United States (US) economics departments and business schools to construct control units of the UK departments. US universities make use of the same inputs (human input/academics and funding) to produce the same outputs (publications and impact) as their UK counterparts, but they have not been exposed to the REF.[1] In particular, we assess the incremental effect of the REF 2014, which assessed the UK research output between 2009 and 2014 (the "treatment period"), relative to the research output of the 2001-2008 period ("pre-treatment" period). With the replacement of the RAE with the REF in 2014, the PRFS in the UK became much more powerful.[2] The synthetic control matching method allows us to create, for each UK university, a comparable synthetic control unit during the pre-treatment period from US universities. We compare the research performance of each UK university with that of its US synthetic control during the treatment period to causally identify the (incremental) effect of the REF on that university's output. Aggregating over all the UK universities, we obtain the effect of the REF on the whole UK university system.

Our results indicate that the introduction of the REF significantly increased UK universities' research output. Indeed, the number of publications of UK departments grew relative to their control US groups across the whole 2009-2014 treatment pe-

[1]Actually, according to the Hicks definition (Hicks 2012, Checchi et al. 2019, Zacharewicz et al. 2019) the funding of higher education institutions in the US cannot be considered a PRFS because the funding allocation is not based on an ex-post evaluation of research output; instead, it is based on education indicators (Hicks 2012, Jonkers & Zacharewicz 2016).

[2]Compared to its RAE 2008 predecessor, the REF 2014 emphasised the increasing importance of the highest quality research by associating substantially higher rewards for the so-called "world-leading research" (4*) relative to "internationally excellent research" (3*), while it completely eliminated payments for research "recognised internationally" (2*) or "nationally" (1*).

riod, but especially towards the end of that period (2012-2014). Research excellence, measured by the number of publications in top journals also increased, albeit to a lesser extent, again especially towards the end of the assessment period. Still, research productivity, measured by the number of publications per author, and the productivity in research excellence, measured by the number of publications in top journals per author, did not change as the number of authors in UK universities also increased. To test for the effects a PRFS such as the REF may have on the academic disciplines, we analyse the university publications in the fields of Economics and Business. Results show that the proportion of UK research output in Economics did not change as some feared. Moreover, the proportion of research excellence, measured by the fraction of the field's publications in top journals, increased significantly in each of these two fields.

We analyse if the PRFSs produce some of the other alleged distortions by comparing the effects of the REF 2014 on different subsamples of UK universities. To assess if the PRFSs contribute to the concentration of research in fewer universities, we investigate if the gap between the elite, research-intensive group of UK universities that form the Russell Group and the rest of the universities widened after the introduction of the REF. Our findings suggest that, indeed, the research performance of the Russell Group increased more than its counterpart. The REF 2014 increased significantly their number of publications in top journals, for instance, whereas it did not affect that of the remaining UK universities. To further investigate the consequences of strategic behaviour, we analyse the outcomes of the group of universities that submitted to the Economics and Econometrics panel for the RAE 2008 but submitted instead to the Business and Management panel for the REF 2014. Our results suggest that their proportion of Economics publications in top journals decreased relative to their control group, possibly because of a shift in focus and institutional priorities. In contrast, their proportion of Business publications in top journals did not increase.

This paper contributes to two strands of literature. The first strand is on the assessment of the PRFSs (OECD 2010, Hicks 2012, Jonkers & Zacharewicz 2016, Sivertsen 2017, Soderlind et al. 2019, Checchi et al. 2019). We provide the first causal analysis of the intended effects of the REF, as well as the first quantitative analysis of some of the unintended consequences or side-effects. We make use of the Synthetic Control Method (SCM) to construct artificial control units for the UK universities. The second strand is on the methodology of the SCM itself, introduced by Abadie & Gardeazabal (2003) and used widely thereafter (see, e.g., the review in Abadie & Cattaneo (2018)). Abadie & Gardeazabal (2003) used the SCM to quantify the effect of an intervention on one treated unit for one outcome variable.

Acemoglu et al. (2016) extended the SCM to account for the presence of multiple treated units again for one outcome. The method we introduce in this paper allows us to compute the effects of an intervention (the REF 2014) on multiple treated units (all universities in the UK) and on multiple outcomes simultaneously (several measures of quantity and quality of research as well as of academic discipline). Our proposed procedure works by finding a reduced set of (US university) controls that is "robust", i.e., adequate, in constructing the synthetic units for the treatment units and for the outcomes we consider in the analysis.

The paper is organised as follows: Section 2 provides background information about the PRFS systems; Section 3 presents the literature review; Section 4 provides the theoretical background for our hypotheses on the intended and unintended effects the PRFS may have; Section 5 presents our data; Section 6 details our estimation strategy; Section 7 reports the results; Section 8 discusses the results and concludes. Additional tables and results are reported in the Appendices.

2 Background

This section provides, first, a formal definition of what is meant by a Performance-based Research Funding System (PRFS). We also provide a list and the main features of some of the existing PRFS around the world. We then explain in more detail the characteristics of the leading PRFS worldwide (and the subject of our empirical analysis), the UK's REF.

2.1 Performance-based Research Funding Systems (PRFSs)

PRFSs are nationwide incentive policies that use the distribution of research funding to encourage research institutions to improve their research performance (Geuna & Martin 2003, Checchi et al. 2019). The incentives are designed through a competitive game, based on the ex-post assessment of institutions' research performance (Hicks 2012, Zacharewicz et al. 2019). Funding is channelled to the best-performing institutions (Checchi et al. 2019). As shown in Table 1, PRFS have been implemented in many countries during the last decades, representing a major switch from block funding (based on historical data) to the adoption of performance indicators (Checchi et al. 2019).

The characteristics of the existing PRFS worldwide differ widely, not only in terms of the volume of funding allocated but also in terms of the types of assessment that feed into the funding allocation formula (Zacharewicz et al. 2019). In fact, not all research funding schemes can be considered PRFSs. To be considered as such, they need to fulfil a set of characteristics defined by Hicks (2012) and adapted by Zacharewicz et al. (2019). In a PRFS, research must be assessed, the evaluation must be ex-post, the evaluation must be of research outputs and/or impact, part of the funding must depend on the outcome of the evaluation, the assessment and funding allocation should take place at the organisation or sub-organisational level (not at the individual researcher level), and it must be a national or regional system (Hicks 2012, Jonkers & Zacharewicz 2016, Zacharewicz et al. 2019). Following this definition, the US's university funding schemes cannot be considered PRFSs because the funding allocation is not based on an ex-post evaluation of research output (Hicks 2012, Jonkers & Zacharewicz 2016). In most states, the funding allocation formula is based on education indicators (such as course completion, time to degree, transfer rates, or the number of degrees awarded), some of which are tied to strategic objectives of the state government (such as promoting the share of disadvantaged students) (Jonkers & Zacharewicz 2016). In other countries, such as Austria, Germany and the Netherlands, their national funding schemes cannot be considered PRFS because funding is allocated on the basis of teaching indicators (without considering the research output)(Zacharewicz et al. 2019).

Broadly speaking, PRFS can be of two types, depending on how the funding is allocated (Zacharewicz et al. 2019): (i) through a peer-review-based assessment exercise, which can be partly metric-based or "exclusively" peer-review, or (ii) through a quantitative metric-based assessment, i.e., a bibliometric approach (counts of publications, weighted by the quality of the journals or a combination of output and citation-based impact metrics (Hicks 2012)), as reflected in Table 1.

Each of the two types of PRFS has its advantages and disadvantages. The main strengths of peer-review are that it is grounded in specialised knowledge of the field and it helps assess non-quantifiable elements of research (such as novelty) (Jonkers & Zacharewicz 2016). However, it is costly (e.g., in terms of resources and manpower) (Geuna & Piolatto 2016) and complex to implement when the pool of experts is not large (Jonkers & Zacharewicz 2016). Moreover, it is considered to be subjective, conservative, favour mainstream research and disadvantage heterodox approaches (Hicks 2012, Jonkers & Zacharewicz 2016). Bibliometric approaches, on the other hand, have lower costs, are non-intrusive (e.g., it requires less administrative burden to prove the research output), and are perceived to be objective (Jonkers & Zacharewicz 2016). But, the data collection is more challenging (Debackere & Glänzel 2004), particularly when the unit of analysis is the department rather than the university, and the impact measures (such as the impact factor) are considered to disadvantage certain fields, such as the humanities (Hicks & Wang 2009).

While both PRFS's types are currently used, the choice of type of PRFS to assess research performance seems to be related to the unit of analysis (Hicks 2012, Zacharewicz et al. 2019). While peer-review tends to be used for evaluations at the departmental or research group level (Hicks 2012, Zacharewicz et al. 2019), bibliometric approaches tend to be used for organisation-wide evaluations. Zacharewicz et al. (2019) conclude that most countries opt for university or departmental-level evaluations (rather than individual evaluation) to reduce the scope and resource demands of the assessment exercise and that peer-review has a higher degree of acceptance among the academic community. In this line, Hobbs & Roberts (2016) states that funding based on a peer-review assessment at the discipline level is becoming the "gold standard" of the PRFS.

2.2 The UK's Research Excellence Framework (REF)

The UK's current PRFS, the Research Excellence Framework (REF), replaced in 2014 and 2021 the previous scheme, the Research Assessment Exercise (RAE), which was conducted in 1986, 1989, 1992, 1996, 2001 and 2008. The UK's PRFSs are considered the oldest, largest, and most developed PRFSs worldwide (Zacharewicz et al. 2019). The UK scheme falls within the peer-review category of the PRFSs, and in particular within the exclusive peer-review category, as bibliometric indicators are not explicitly used. A panel of experts for each "Unit of Assessment" broadly representing an academic discipline (which we call from now on a "panel"), assesses the quality of the research in that area for all the UK institutions (subdivided into the categories of outputs, impact, and environment). Broadly the level of assessment and funding allocation are, thus, at the university Department level, rather than at the individual or overall University levels. The REF 2014 had 36 panels, and it included Economics and Econometrics and Business and Management, our two subjects of analysis.

Participation in the assessment is voluntary. UK Research institutions decide to which panels they want to submit their research based on the academic discipline of the staff and their research outputs. For instance, a university can submit the research output of the Department of Economics to the Economics and Econometrics panel or to the Business and Management one, possibly jointly with the research output of the Business Department.Over the period 1992 to 2014, the number of UK universities entering the Economics and Econometrics panel declined significantly, from 60 in 1992 to 28 in 2014. Many of the universities that left the Economics and Econometrics panel started submitting to the Business and Management panel or made their existing Business and Management panel submission larger (Johnston & Reeves 2020).

Research institutions can also decide which of their staff to include in the assessment. Eligible staff are those that are employees of the university at the census date (in the case of the REF 2014, October 30th, 2013). Institutions can submit a limited number of research outputs per researcher (usually four) from the publication period (in the case of the REF 2014, between January 1st, 2008 to December 31st, 2013). This should incentivise quality (rather than quantity) of research.

Based on the assessment, the panel of experts of each discipline creates a quality profile for each institution, based on a comparison of its research activity to national and international benchmarks. Results show the proportion of the research activity which can be considered "world-leading research" (4*) relative to "internationally excellent research" (3*), "recognised internationally" (2*) or "nationally" (1*). The quality profile is used by the funding bodies to determine the financial resources that go to each research institution/discipline to fund their future research activities.

The RAE/REFs have changed over time, but it was in the period 2009-2014, when the RAE was replaced by the REF, that the performance-based incentive system became more powerful, giving much steeper incentives to the UK universities than previous assessments. Some of the changes from the RAE 2008 to the REF 2014 included higher rewards for world-leading research (4*), while payments for research "recognised internationally" (2*) or "nationally" (1*) were eliminated (Mingers & White 2015, Geuna & Piolatto 2016, Marques et al. 2017, Chiang 2019).

The REF has inspired, directly or indirectly, the design and implementation of PRFS in many other countries (Sivertsen 2017, Geuna & Piolatto 2016). Many more are considering it (e.g., Hungary). An EU Horizon 2020 project designed a PRFS toolkit, based on mutual learning, to provide recommendations about the design and use of PRFS (European Commission 2020). Table 1 provides a summary of the PRFS worldwide.

Table 1: PRFS worldwide

Country	Year	Name	Type	Level	Assessment
United Kingdom	1986-2008 RAE	RAE	Peer-review	Department	Output
United Kingdom	2009-present	REF	Peer-review	Department	Output, impact and environment
France	2008-present		Peer-review	Department	Quality of research indicators, citations, participation in networks, grants obtained, frontier research and openness to societal changes
Lithuania	2009-present		Peer-review (partly bibliometric)	University	Quality and quantity of research publications
Italy	2001-2010	VTR	Peer-review	Department	Output
Italy	2011-present	VQR	Peer-review (partly bibliometric)	Department	Quality assessment review (publications, citations, external funding, international collaborations, patents registered, the number of doctoral students), recruiting policies and relevance of international teaching activities
Australia	1995-2009	CI	Bibliometric	University	Publications quantity and quality
Australia	2010-present	ERA	Peer-review (partly bibliometric)	Department	Research outputs vis a vis national and international benchmarks.
New Zealand	2003-present	PBRS	Peer-review	University	Quality of research (outputs, contribution to research environment and peer esteem), research degree completions and external research income
Portugal	2015-present	Research Unit Evaluation	Peer-review	Department	Academic performance and the strategic plan submitted
Belgium	2003-present	BOF-key	Bibliometric	University	Master degrees, defended doctorates, gender diversity, publications and citations
Norway	2005-present	Norwegian Performance Based assessment system	Bibliometric	University	Teaching and research indicators (number of PhDs awarded, allocation of EU funding for research, allocation of funding from the Norwegian Research Council and bibliometrics)
Poland	2008-present	Parametric evaluation	Bibliometric	Department	Publications (impact factor), patents, external funding, scientific awards and PhD degrees awarded.
Sweden	2009-present	FOKUS	Bibliometric	University	Publication and citation counts and external funding
Denmark	2009-present	BFI	Bibliometric	University	External research funding, PhD production, student throughput and the Bibliometric Research Indicator
Finland	2010-present		Bibliometric	University	Education performance, research performance (external research funding, PhD production, and a bibliometric indicator) and other considerations
Estonia	2012-present		Bibliometric	University	High level publications, high level research monographs, registered patents, external funding and doctoral graduates
Croatia	2013-present		Bibliometric	University	Scientific production (publications in journals covered by the WOS and SCOPUS, citations and doctoral graduates), national and international competitive research projects and mobility, popularization of science and commercialisation of science.
Slovakia	2013-present		Bibliometric	Department	Scientific papers published/cited in journals listed in WOS, SCOPUS, research monographs published with high-quality publishers and international patents.
Czech Republic	2013-present	NERO	Bibliometric	University	Publications (impact factor), high impact results (like grants) and patents

Source: Hicks (2012), Geuna & Piolatto (2016), Jonkers & Zacharewicz (2016), Zacharewicz et al. (2019), Soderlind et al. (2019), Checchi et al. (2019), European Commission (2020) *Notes:* VTR = Valutazione Triennale della Ricerca, VQR= Valutazione della Qualità della Ricerca, CI=Composite Index, ERA=Excellence in Research for Australia, PBRF=Performance-based research funding, BOF = Bijzondere OnderzoeksFondsen, FOKUS=Forskningskvalitetsutvärdering I Sverige, BFI= Bibliometric Research Indicator, NERO=National Evaluation of Research Organisations.

3 Literature overview

We now provide an overview of the existing literature that analyses the consequences of the introduction of the PRFSs incentive schemes on research output.[3] We divide our review of the effects of the PRFSs into intended consequences, that is the increase in quantity and/or quality of research output, and unintended consequences of the PRFSs, in terms of direct costs and distortions.

3.1 Intended consequences

Several studies have evaluated the effect of the PRFSs using descriptive analyses. Butler (2003) argues that the Australian funding allocation scheme increased journal publication production significantly in the last decade, but its impact, in terms of citations, may have declined. Bence & Oppenheim (2005), and Moed (2008) focus on the shift in the evaluation criteria from quantity to quality that occurred in the UK following the introduction of the RAE 1996. Moed (2008) states that authors gradually increased the number of publications in high-impact journals. Auranen & Nieminen (2010) while looking at the PRFSs of eight countries, says that financial incentives do not boost publication production straightforwardly. Hicks (2012) reviews fourteen PRFSs policies in different countries (including the UK) and concludes that the incentive to enhance research comes from the competition for prestige created by the PRFSs. Bloch & Schneider (2016) state that the Norwegian PRFS impacted individual researchers' behaviour by increasing the number of publications per author and the number of co-authors per publication. Carli et al. (2019) observe that the Italian PRFS may favour research excellence among non-outstanding academics but it may slightly reduce the productivity of outstanding academics.

There are very few quantitative studies evaluating the effects of PRFSs. Franzoni et al. (2011), using a before and after approach, provide evidence that country-level incentives in the OECD lead to more submissions and publications in the academic journal Science. Taylor (2011), using a linear regression model, finds that for the RAE 2008, each of the three components of research activity (research output, impact and research environment) are highly correlated with various quantitative indicators (such as, journal quality index, number of research staff, previous RAE outcome, member of the Russell Group, autonomous department of economics or finance). Wang & Hicks (2013) analyses thirty years (1981-2011) of UK aggregate

[3]Other lines of research use the output submitted to the REF to create a ranking of Economics journals (Hole 2017) or to look at the journals that dominate in outputs' submissions (Stockhammer et al. 2021), while other studies predict the results of the next REF (Mryglod et al. 2015, Basso & di Tollo 2022) or evaluate the REF's usefulness as a composite evaluation (Pinar & Horne 2021).

publication data and identify three structural changes at the national level, one of which is related to one RAE (1989) out of six undertaken in that period (1986, 1989, 1992, 1996, 2001, 2008). Checchi et al. (2019) analyse the effect of incentive policies on the number of publications and research quality at the aggregated level for 31 OECD countries using a difference-in-difference approach. Results show that the introduction of a PRFS temporarily increases the aggregate number of publications at the country level and somewhat influences the average research quality measured by the number of citations.

While the literature has already debated the intended consequences of PRFSs, in terms of both quantity and quality of research, very few papers use a quantitative approach and barely any makes causality claims. Perhaps Checchi et al. (2019), who uses a difference-in-difference approach at a very aggregated country level, is the closest. Identifying a good counterfactual is crucial in this setting, as there have been many other changes, such as an increased journal coverage of Scopus, changing publication behaviours (Moed 2008, Franzoni et al. 2011) and more co-authorship.

3.2 Unintended consequences

Regarding the unintended consequences of PRFSs, there is even less quantitative evidence. We now review the extant literature on some of the unintended consequences on direct costs and indirect effects such as those that lead to distortions across fields of research and types of universities.[4]

Some studies have focused on the substantial costs of preparation and submission of the assessment exercises (Harman 2000, Martin & Whitley 2010, Martin 2011, Stern 2016). Geuna & Piolatto (2016) provide a comparative analysis for the UK and Italian PRFS. They show that using metrics in the assessment (Italy) is less costly than exclusive peer-review (UK). In this line, Battistin & Ovidi (2022) states that in PRFSs, peer-reviews and bibliometrics should be viewed as complementary modes of assessment. He concludes that peer-reviews would be more cost-effective if they were to use an automatic classification for the submissions and only perform the costly and time-consuming peer-reviews for publications in low-impact journals and interdisciplinary journals.

The PRFSs may also create strategic behaviour and, thus, distortions. In the UK's REF, for instance, institutions are allowed to choose the panel they submit their research output to. As mentioned before, the number of universities enter-

[4]This review is by no means exhaustive. We are leaving out for instance the distortions produced by scoresetrics. Jonkers & Zacharewicz (2016) claim that while the EU Member States have implemented policies aiming to increase research outputs' quality, they could have generated perverse incentives (gaming the metrics).

ing the Economics and Econometrics panel declined significantly from 1992 to 2014 while the number submitting in the Business and Management panel grew (Johnston & Reeves 2020). Battistin & Ovidi (2022) show that staff of the Economics departments may be submitted to the Business and Management panel, where the subject matter overlaps (Pidd & Broadbent 2015), and claim that this panel would likely rank the research output more highly. Guizzo et al. (2021) argue that the REF incentivised a change in research decisions and publishing strategies of macroeconomic researchers, especially women, junior-level academics and those affiliated with less prestigious institutions, reporting an "unfair playing field."[5]

Some studies claim that the PRFSs may favour the research-intensive, large and/or elite groups of universities (Bence & Oppenheim 2005, Clerides et al. 2011, Mingers & White 2015, Jeon & Kim 2018, Guizzo et al. 2021). In the case of the UK's RAE, Martin & Whitley (2010) raised concerns regarding the unequal distribution of power and the rewarding of academic elites.[6] The PRFS may increase inequality due to the so-called Matthew effect, which refers to the widening gap in the research activities of universities, where inequalities are perpetuated if universities are historically characterised by heterogeneous performance (Jeon & Kim 2018).[7] However, Checchi et al. (2020) and Buckle et al. (2022) argue that PRFSs may not necessarily increase such inequalities. Focusing on the Italian and New Zealand PRFSs, they find that the dispersion in universities' final scores was reduced across exercises because of staff turnover and research quality improvements for the remaining staff, especially at the below-average universities (Buckle et al. 2022).

4 Conceptual framework and hypotheses

As explained in the previous two sections, the PRFSs may create powerful incentives to produce high-quality research, both directly, through financial rewards, as well as indirectly, through prestige (Hicks 2012). But they may also generate undesirable side-effects. This section provides a conceptual framework to analyse the intended

[5]Marques et al. (2017) find that, in the UK, there is also selectivity in the number of staff whose output is submitted to the REF, as a form of reverse engineering. This distortion caused by selectivity was also described in the Stern (2016) review.

[6]Lee et al. (2013) states the RAE and local department decision-making has contributed to the dominance of a select group of departments in Economics by eliminating the heterodox economics Stockhammer et al. (2021). They also claim that the RAEs are exacerbating a monoculture in Economics with no connection to the real world.

[7]It is important to stress that the REF 2014 allocated 50% of the total funding to the top 10 universities and 71% went to the larger Russell Group of research-intensive universities (Davé et al. 2016, Arnold et al. 2018).

and some of the unintended consequences of PRFSs.

4.1 Intended effects: research output and excellence

The PRFSs incentive schemes exist to address a standard moral hazard problem. Indeed, as in the typical moral hazard framework, the principal (in the case of the REF, the government or the research council) encourages the agents (the departments or the universities) to provide privately costly efforts to improve their research performance. As the efforts are imperfectly observed, monetary rewards (the research funding) are based on observed research performance (in the case of the REF, the submitted outputs of the university, mainly).

While most of the economics and management literature would agree that using incentives helps and that steeper incentives produce greater performance, recent research in behavioural economics has shown that incentives may sometimes not work (Gneezy et al. 2011). First, academic extrinsic incentives may crowd out the researchers' intrinsic motivation to produce high-quality research (Andersen & Pallesen 2008). Thus, (steeper) financial incentives may affect behaviour in two ways. Besides the normal economic response to the relative monetary difference, there may be a crowding effect that works through its impact on intrinsic motivation (Frey & Jegen 2001). In addition, the introduction of managerial principles expressing desired outcomes among department members, shifting from self-control to external control and from preferred activities and themes towards productive research subjects might also affect motivation (Carli et al. 2019).

Still, we hypothesise that the positive standard incentive effect outweighs the negative crowding out and motivation effects:

Hypothesis 1: The introduction of the REF increased the "research output" (i.e., the number of publications) of the UK departments.

Many PRFSs, such as the REF, aim to achieve "research excellence" and not just high research output. To that purpose, the theory would say that the incentives need to be even more powerful or "steeper", i.e., one needs to give an even greater reward for the highest-quality outcomes and reduce those of slightly lower quality. This was meant to be one of the changes in the substitution of the RAE by the REF in the UK. It remains to be seen, though, if it works.

Hypothesis 2: The introduction of the REF increased the "research excellence" (i.e., the number of publications in top journals) of the UK departments.

Universities can encourage research excellence by incentivising their researchers to publish more and better, thus improving their individual research productivity, and/or by hiring more (and more productive) members of staff (Marques et al. 2017).

Consistent with the latter, academic staff in the UK increased by 11% during the REF's 2014 assessment period (2008 to 2014) (Higher Education Staff Statistics 2014).[8] Abramo & D'Angelo (2022) argue that individuals and organisations may respond differently to financial and reputational incentives. The individual response depends on the extent that institutions internally deploy the incentives at the individual level and the researchers' sense of belonging to their organisation (Abramo & D'Angelo 2022).

Hypothesis 3: The introduction of the REF increased the "research productivity" (i.e., the number of publications per author) of the UK departments.

Hypothesis 4: The introduction of the REF increased the "productivity in research excellence" (i.e., the number of publications in top journals per author) of the UK departments.

4.2 Unintended consequences: distortions across disciplines

Inevitably, national (one-size fits all) built-in incentive schemes are not going to be discipline-neutral. Even with the same intrinsic motivation, the extrinsic motivation and, in particular, the standards of excellence may be higher in one discipline than in another, especially if the research evaluation is done through peer-review. PRFSs based on peer review, as used in the UK, are becoming more common internationally and becoming the standard model of PRFS (Hobbs & Roberts 2016), despite being slower, costlier, more subjective and less transparent than metric-based PRFSs (Hobbs & Roberts 2016). As recognised by Wilsdon et al. (2015), peer review may be "inconsistent and characterised by a lack of inter-rater reliability". This is a perspective also echoed by the European Commission (2010) expert group on research assessment: "Unintended consequences [. . .] may include over-concentrating on research and favouring particular disciplines or allocating resources and realigning priorities to match indicators".

All these arguments lead us to formulate the following two additional hypotheses.

Hypothesis 5: The introduction of the REF has changed the proportion of "research output across fields" (i.e., the fraction of publications of each field relative to the overall number of publications).

Hypothesis 6: The introduction of the REF has changed the "proportion of research excellence per field" (i.e., the fraction of the publications in that field published in top journals).

[8] Some studies have claimed that the UK's RAE/REF has actually distorted universities' hiring decisions, especially in the years around submission deadlines (Hayri 1997, La Manna 2008, OECD 2010, Stern 2016).

The UK, for instance, has witnessed a heated debate on the presumably higher standards of the Economics and Econometrics RAE/REF panel, especially relative to the sister Business and Management RAE/REF panel. Over the last RAEs/REF, there has been a mass exit of UK institutions from the Economics and Econometrics panel and into the Business and Management panel. As a result of all these moves, researchers may have shifted their research focus from Economics to Business topics. Due to the same potentially higher standards, the quality of the research output in Economics may have declined. At the same time, the quality of the research output in Business may have been raised. Some commentators argue that the UK's PRFS, the REF, has greatly contributed to the shrinking of the Economics discipline's research performance (Lee et al. 2013, Guizzo et al. 2021). Economics and Business are going to be the testbed of our hypotheses.

5 Data

We analyse the effects of the REF 2014, the intervention that (may) have affected the research performance of UK universities between 2009 and 2014. Prior to that, between 2001 and 2008, their research output was assessed by the RAE 2008. We consider the period 2009-2014 as the "treatment years" while the period 2001-2008 is considered the "pre-treatment years". Indeed, as mentioned before, the UK's PRFS became much more powerful for 2009-2014, when the RAE was replaced by the REF. But, more formally, our results should be understood as an incremental effect of the REF relative to the RAE. Thus, our data covers the period from 2001 to 2014.[9]

5.1 Sample

Our initial sample of universities includes all the 103 UK universities, i.e., the ones that submitted their research to the REF 2014 Panels 18 (Economics and Econometrics) and/or 19 (Business and Management).[10] To define a control group of universities not exposed to the REF 2014, or, more generally, to any PRFS, we use the 135 US universities with a top-25% Department of Economics and/or a top-25%

[9] Note that the submission date for the REF 2014 was 29th November 2013 and the assessment period ran from 1st January 2008 to 31st December 2013. Our analysis uses the period that runs from January 1st 2009 to December 31st 2014 to consider forthcoming articles included in the REF submission (and those that just missed the deadline). We also performed a robustness check with the period that runs from January 1st 2008 to December 31st 2013. Both definitions produce very similar results.

[10] https://results.ref.ac.uk/(S(xgg03r2fhnekcwvixcrvmhgf))/Results/ByUoa/18 and https://results.ref.ac.uk/(S(xgg03r2fhnekcwvixcrvmhgf))/Results/ByUoa/19.

Business School according to the December 2018 *RePEc* rating.[11]

We obtained all of the articles from the Scopus database, which (i) include one of these 103 UK institutions or one of the 135 US institutions as an affiliation,[12] (ii) have either "Economics, Econometrics and Finance" or "Business, Management and Accounting" as subject areas, and (iii) were published between 2001 and 2014 (both included). We include articles published in scientific journals identified with an ISSN code and remove publications in books and/or conferences.[13] We further restricted the sample of institutions by dropping those with a per-year average of less than ten papers during the pre-treatment period (2001-2008). Our final data includes 145,536 unique publications, 769 authors, 202 affiliations (121 US and 81 UK) and 975 journals between 2001 and 2014. We aggregated the publication data at the institution and year level.

5.2 Outcome measures and descriptive statistics

We now explain the constructed outcome measures of university research performance.

As a measure of "research output", we count the number of unique publications of each university (affiliation) in each year and name this measure *Number of publications*. To construct an outcome measure that considers quality, we use the 2018 Association of Business Schools' classification of scientific journals, published in the Academic Journal Guide.[14] Journals are classified from 1* (least influential), 2*, 3*, 4* to 4** (most influential). As a proxy of "research excellence," we count the number of publications in categories 3*, 4* and 4** of each university in each year, and name it the *Number of publications in top journals*.

To understand the "research productivity" and the "productivity in research excellence," we compute the *Number of publications per author* and the *Number of publications in top journals per author* by dividing the number of publications in a given year by the number of authors in that institution in that year.[15] The number

[11]See https://ideas.repec.org/top/top.usecondept.html. Throughout the paper, we do not distinguish between the publications of the different departments of the same institution. This is because the division in departments across institutions is highly heterogeneous and difficult to identify in the database.

[12]Universities with several institutions/schools/colleges were manually assigned the same affiliation code. For instance, Cass Business School, which has its affiliation code in Scopus, was assigned to City, University of London's code. Moreover, spelling errors in the universities' names, creating different affiliation codes for the same university, were also addressed.

[13]The number of papers and journals by sub-field and country is shown in Table A1 in Appendix 1.

[14]http://www.CharteredABS.org/academic-Journal-Guide-2018. The classification of journals remains almost invariant over time.

[15]We do not take into account the number of authors, the number of affiliations of each author,

Table 2: Research performance measures: description and summary statistics.

Outcomes	Description	Mean	Min	Median	Max	SD
Research output (Number of publications) (1)	Count of the publications in journals, by institution and year.	68.91	14.28	54.71	250.92	52.62
Research excellence (Number of publications in top journals) (2)	Count of the of publications in journals with an AJG 2018 grade of 3*, 4*, or 4**, by institution and year.	37.51	2.42	26.71	153.00	35.52
Research productivity (Number of publications per author) (3)	Count of the number of publications per author, by institution and year.	1.23	0.95	1.22	1.67	0.14
Productivity in research excellence (Number of publications in top journals per author) (4)	Count of the number of publications per author in journals with an AJG 2018 grade of 3*, 4*, or 4**, by institution and year.	0.61	0.13	0.60	1.10	0.21
Proportion of research output across fields (Proportion of publications in Economics) (5)	Proportion of the number of publications in Economics journals over the total number of publications in journals, by institution and year.	0.22	0.01	0.20	0.54	0.14
Proportion of research excellence per field (Proportion of Economics publications in top journals) (6a)	Proportion of the number of publications in Economics journals with an AJG 2018 grade of 3*, 4*, or 4** over the total number of Economics publications, by institution and year.	0.37	0.00	0.38	0.80	0.21
Proportion of research excellence per field (Proportion of Business publications in top journals) (6b)	Proportion of the number of publications in Business journals with an AJG 2018 grade of 3*, 4*, or 4** over the total number of Business publications, by institution and year.	0.51	0.16	0.51	0.89	0.14

Notes: This table provides the description and the summary statistics (mean, minimum, median, maximum and standard deviation (SD) of the research outcome measures considered in the analysis. AJG stands for Academic Journal Guide.

of authors in a university in a given year is obtained by summing all of the publishing authors with that affiliation in that year. To assign the author to an institution, we use weights to take into account multiple affiliations, with a maximum weight per author per year of one.[16] For years in which an author does not publish, we check if he/she has publications, before and after, in that institution. If he/she does, we consider that this author has been in that year in that institution.

To assess whether the intervention has produced a shift in the fields of research, we constructed the following three variables. Following the classification of the REF 2014 Panels (18 - Economics and Econometrics and 19 - Business and Management),

or the total number of affiliations in the paper. However, we also considered different alternatives: (i) counting the number of publications in each institution, which implies that if two of the paper's authors are from the same institution, the publication counts double; (ii) dividing each publication by the number of co-authors, and attributing to each institution the fraction corresponding to the co-authors in that institution; (iii) similarly, dividing each publication by the number of co-authors and by the number of affiliations of the authors (if an individual has two affiliations, half is attributed to each one); (iv) dividing by the number of affiliations of the authors (but not by the number of co-authors); and (v) dividing by the total number of affiliations in the paper, which implies that if two authors have the same affiliation, it only counts once. The correlations between these measures and our main count measure are all above 0.98, and thus the empirical results are almost identical quantitatively. Authors that only publish one single paper in our study period are not considered.

[16] When authors change affiliations over time, the weight in the original and destination universities is based on the actual papers and is adjusted so that the weight is not more than 1 in any given year.

we separate the publications of each institution by area. We make use of the journal subject categories, in both, the Academic Journal Guide (AJG) journal classification and Scopus to classify the publications into "Economics" and "Business" (see Table A1 in Appendix 1).[17] We classify the publications in the journals of the Scopus ECON subject category as in "Economics" and those in the journals for the other subject categories as in "Business". We compute the *Proportion of publications in Economics*, relative to all publications in Economics and Business in a given year, to measure the "proportion of research output across fields." Moreover, we count which of those are in the area's top journals. We calculate the *Proportion of publications in Economics top journals* relative to the total number of publications in Economics. We do the same exercise and calculate the *Proportion of publications in Business top journals*. These are our two measures of the "proportion of research excellence per field."

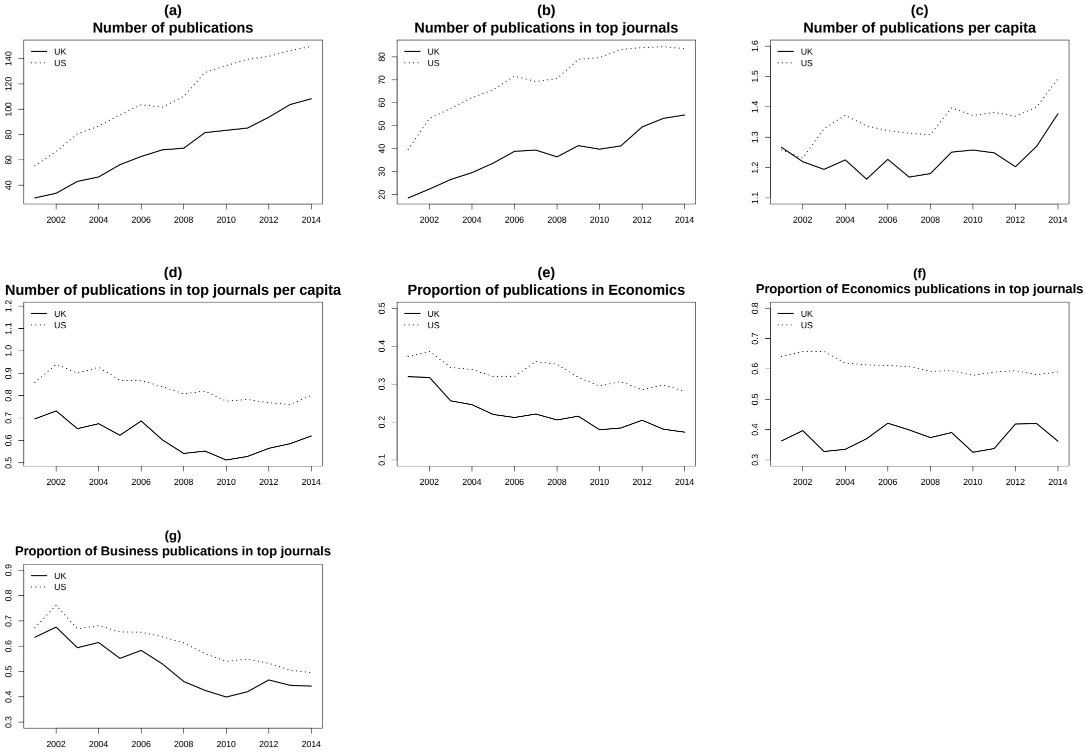
Table 2 contains the summary statistics of the yearly averages of our variables for all universities in both countries. The evolution of the variables yearly averages over time, separated by UK (solid line) and US (dotted line), are described in Figure 1. The number of publications per university per year spans from 14.28 up to 250.92, with an average of 68.91. In top journals, these figures are much lower (i.e., 2.42, 153.00, and 37.51, respectively). The average number of publications per author per year, is 1.23, and its distribution is quite tight around that figure. The average number of publications in top journals per author per year is about half of that figure, but there is a wider variation. On average, 22% of the publications of the universities of our sample are in Economics. On average, about 37% of the university publications in Economics are in top journals, with a minimum level of 0% and a maximum of 80%. For Business, the figures are 51%, 16% and 89%. Tables A2, A3 and A4 in Appendix 1 present the list of universities for the UK and US, and the all-year average of our outcome measures, along with those of the number of affiliated authors. We sort the universities in decreasing order according to their average number of publications.

Figure 1 shows that the average number of publications (panel a) and average number of publications in top journals (panel b) of US and UK universities increased during the REF's 2014 assessment period. The number of publications per author (panel c) remained constant in both countries and only started to raise towards the end of the assessment period. The other variables (panels d, e and g) show a decline

[17]We do not utilize the subject areas of Scopus, (i) Economics, Econometrics and Finance and (ii) Business, Management and Accounting because this system does not match exactly with the classification of the REF panels. Finance, for instance, is included in the "Business and Management" REF Panel, but it is included in the same subject area as Economics and Econometrics in the Scopus database.

over the same period, with the exception of the proportion of Economics publications in top journals (panel f) that remained at about the same level throughout the period.

Figure 1: Evolution of research outcome measures over time.



Notes: This figure shows the evolution over time of the yearly averages of the outcomes of interest, separated by UK (solid line) and US (dotted line).

5.3 Subsamples

To explore whether the REF intervention had a differential impact on different groups of universities, we perform two subsample analyses. First, we compare the impact of the REF 2014 on the universities belonging to the Russell Group (list reported in Table A2), a group of UK universities historically associated with higher research intensity and overall better research performance, to those that do not belong to the Russell Group. While the research intensity is measured as a university aggregate across different fields, our data also shows a higher research performance of this group relative to the rest in our two areas of interest, as shown in Table A2.

Second, we compare the impact of the REF on the universities that submitted to the Economics and Econometrics Panel in both years (Remainers), i.e., to the RAE 2008 and REF 2014, from those universities that submitted to this panel in RAE 2008 but switched to the Business and Management Panel in REF 2014 (Leavers), which are listed in Table A2. This latter analysis aims to test whether the decision of the university on which panel to submit to influences the focus and quality of the departments' research.

6 Empirical Strategy

As described earlier, Figure 1 shows that the average number of publications of UK universities increased during the REF's 2014 assessment period. This may be due to the introduction of the REF, but it may also be due, for instance, to the change in publication practices towards indexed journals (Hammarfelt & de Rijcke 2015) and/or the increase in the number of journals covered by Scopus. Indeed, the number of publications also increased for US universities, despite not being exposed to the REF (or any other PRFS). Similarly, the proportion of publications in Economics journals, out of all publications, decreased in the UK, but it also did for the US, so it is unclear if the reduction of UK Economics publications is due to the introduction of the REF. Therefore, the analysis warrants a method that disentangles what would have happened independently of the introduction of the REF from what the intervention is accountable for. Below, we explain our empirical strategy.

6.1 The Synthetic Control Method

We use the Synthetic Control Method (SCM), introduced by Abadie & Gardeazabal (2003) to casually estimate the effect of a "treatment" (in their case an outbreak of terrorism) on a certain unit (in their case the Basque region in Spain) for a certain outcome of interest (the region's economic performance). They compared the evolution of the outcome of interest for the treated unit, during the treatment period, to the evolution of the same outcome of interest for an artificial unit, created as a convex combination of multiple untreated units (in their case, other Spanish regions). Indeed, the artificial comparator provides information on how the treated unit would have evolved, in the absence of the treatment, in the treatment period. As argued by Abadie & Gardeazabal (2003), the artificial control should be able to reproduce the counterfactual behaviour of the treated one better than if we were to use a single unit as control.

To apply the SCM to our case, we define the pre-treatment period as the years 2001-2008 and the treatment one as 2009-2014. As explained in the Background Section 2, PRFSs' incentives in the UK became much more powerful in the period 2009-2014 than in the previous one (2001-2008). The system gave much higher rewards for outputs of the highest classification (4*) and eliminated the payments of research "recognised internationally" (2*) or "nationally" (1*) entirely. To reflect that change, the name of the intervention changed from "Research Assessment Exercise" to "Research Excellence Framework". Therefore, our study covers the period 2001 to 2014, being 2001 to 2008 the pre-treatment period and 2009 to 2014 the treatment one.

The SCM applies a matching algorithm that follows an iterative, two-step optimization process. The two-stage optimization procedure uses the outcome, and a group of selected covariate variables for the treated and potential control units, to select, for each treated unit, the best weighting of covariates and the best weighting of units from the pool of potential controls to create a synthetic control. In our case, we use the following 13 covariates: the values of the pre-treatment years' outcomes for the seven covariates of interest, as well as the mean and median across these years (two further covariates); and the means across all pre-treatment years of the (i) number of publishing authors, (ii) number of publications, (ii) number of publications in a 3*, 4* or 4** journal and (iv) number of publications in a 4* journal. The optimisation procedure minimises the pre-treatment period difference between the outcome of interest of the treated unit and of the synthetic control. This difference is measured by the root mean square prediction error (Abadie et al. 2010, Bouttell et al. 2018). In Appendix 2.1, we provide the technical details of this process. In Appendix 2.4, we describe how to assess the goodness of fit of our matching process, that is how to assess if the weighted average of controls is able to approximate well the outcome of the treated unit in the pre-treatment period.

For each treated UK university and each treatment year, the SCM computes the treatment effect as the difference between the actual value of the outcome measure for that university and year and that value for the counterfactual university, using the sample of US controls and the weights attributed to each of them. A final step calculates, for each treated UK university, the cumulative Treatment Effect as the sum of the yearly individual effects across all the treatment years (2009 to 2014).

6.2 SCM vs. difference-in-differences

The SCM has advantages and disadvantages relative to the difference-in-differences model, which has been widely used in the treatment effects literature. The SCM allows for the presence of unobserved confounders with time-varying effects, and thus, it does not rely on the *parallel trend assumption* (Abadie et al. 2010). Intuitively, only units that are alike in both observed and unobserved determinants of the outcome variable, and in the effect of those determinants on the outcome variable, should produce similar trajectories of the outcome variable over extended periods of time (Abadie et al. 2015). In our case, there is no single control unit, i.e., a US university, that can be matched to a UK one. The parallel trends assumption, as required for a difference-in-differences estimator of the treatment effect, does not hold for all outcomes of interest in our framework, as suggested by the plots of the average outputs over time in Figure 1. The SCM overcomes this by allowing the creation of controls for the treated units (in our case UK universities) that are a combination of different controls (US ones). In addition, the SCM allows the assessment of the impact of the policy on each individual treated unit. Nevertheless, for comparison reasons, we present the results obtained with a difference-in-differences estimator at the of the results section.

The SCM has a limitation with respect to the difference-in-differences method: traditional statistical inference is inappropriate. The reason is that there are small numbers of treated and control units, and units are not sampled probabilistically (Bouttell et al. 2018). This is addressed by making use of placebo tests, which are obtained by performing the analysis as if the units in the control group were treated units, to generate a distribution of effect estimates under the scenario of no intervention, which can be used to infer the significance of the SCM estimates. In Appendix 2.2, we provide a more technical description of this process.

6.3 SCM for multiple treated units

As the REF 2014 is an intervention that affects all the UK universities, and not just one, we apply a variation of the original SCM designed for the case of multiple treated units (Acemoglu et al. 2016, Kreif et al. 2016). The modified SCM creates first, in a similar way to the original SCM, a matched artificial control university for each UK university. As explained before, the artificial control unit for each UK university is constructed based on a set of potential controls and on a series of covariates with good predictive power over the pre-intervention period. The overall (i.e., all UK universities) average treatment effect is computed by taking into account the quality of the pre-intervention matching for each treated unit. UK universities with a higher-quality matching with its artificial control, derived from US universities, count more towards the all-university weighted average, i.e., their weights are larger. See again Appendix 2.3 for the technical derivation and for the use of placebo tests to infer significance.

6.4 SCM for multiple treated units and multiple outcomes

The approach introduced by Acemoglu et al. (2016) extends the SCM to more than one treated unit. But it is used to study if one and only one outcome has been affected by a specific intervention. In our case we aim to test whether the REF has affected the evolution of *multiple outcome measures* on *multiple treated units* simultaneously. Ideally, we would like to use, for each treated UK university, the same artificial control for all outcomes. Unfortunately, finding, for each treated unit, a unique solution for all outcomes proved to be unstable, particularly because the size of the control set is large and the larger the dimension of the set of controls, the larger the number of possible combinations of controls. Due to the similarity of the chosen controls, mostly around the "average" universities, the algorithm has difficulties finding a unique optimal solution.

For this reason, we introduce an iterative procedure to find a smaller "robust set of controls" for the matching. The procedure starts by performing the SCM for all 81 treated units and for each outcome using all 121 US universities to create control units. Then, we discard from the initial 121 US universities those that were not useful to shape any optimal counterfactual synthetic unit for any of the outcomes, that is, we discard the universities whose optimal coefficients were zero for all treated units and all outcomes. We repeat the procedure until we find a stable set to create control units, which had 19 US universities (this set is marked with a "Yes" in the column labelled *Selected* in Tables A3 and A4). By robust set of controls, we mean that each of these 19 US universities is important to shape the synthetic unit for at least one UK university and for at least one outcome. The composition of optimal weights may change for each specific UK university and each specific outcome, though. Thus, we run the matching process for each outcome variable separately, but all outcomes are considered simultaneously in the selection of the robust control group as explained above. Table A6 in Appendix 2 reports the estimated weights given to the selected 19 US universities to create a control for each UK university for one of the output measures, the number of publications.

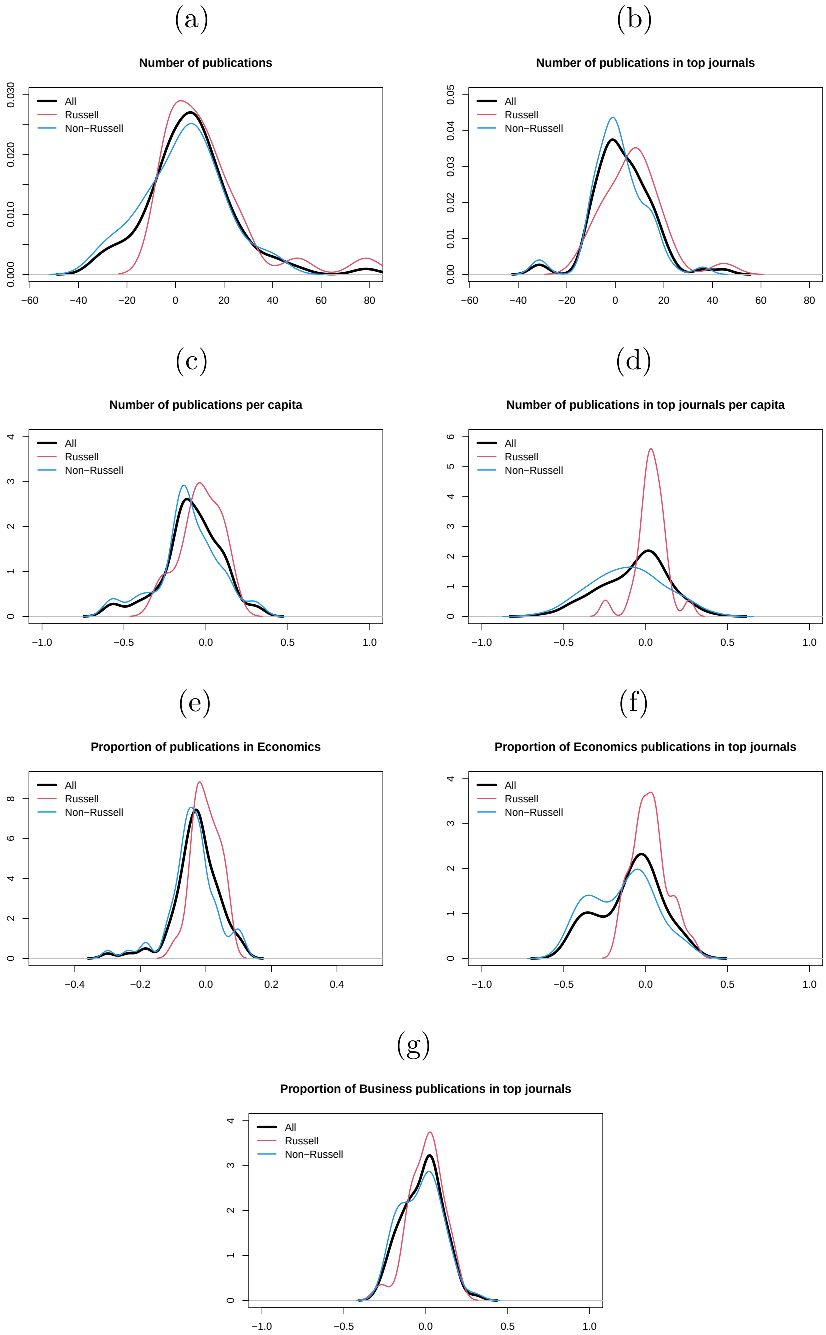
7 Main results

This section presents the SCM results. We first show the SCM results at the individual university level. We then show the Average Treatment Effect (ATT) of the introduction of the REF on all the UK universities. We discuss the intended effects, both at the department level and on a per author basis, as well as the results on the possible field distortions. Finally, as a complement to the SCM results, we present difference-in-differences estimates on the main outcome variables of interest.

7.1 Individual university effects

We start by discussing the effects of the introduction of the REF at the individual university level. Figure 2 presents the distribution of the yearly treatment effects for all the UK universities (in black), as well as that of the subsample of Russell and non-Russell groups of universities (in red and blue, respectively), for all the outcome measures (panels a to g). The figure reveals a high degree of heterogeneity, both overall as well as across and within these two groups of universities. The distributions are indeed broad, with supports spanning from positive to negative values.

Figure 2: Distribution of the yearly treatment effects: All, Russell and non-Russell



Notes: This figure reports the distribution of the yearly treatment effects for all, Russell and Non-Russell groups of universities for: number of publications (panel a), number of publications in top journals (panel b), number of publications per author (panel c), number of publications in top journals per author (panel d), proportion of publications in Economics (panel e), proportion of publications in Economics top journals (panel f), proportion of publications in Business top journals (panel g), as described in Table 2.

The distribution of the effects of the number of publications for all universities (black line in panel a) has a peak at 10 publications. That of the Russell Group (red line) has positive skewness (i.e., a fatter right tail than left tail), indicating that the effect has been large and positive for some units, despite the most frequent yearly effects being close to zero. The Russell Group also has a much lower negative tail compared to the Non-Russell group (blue line) revealing that it was mostly non-Russell universities to be negatively affected by the REF 2014 policy.

The distribution of the effects for the number of publications in top journals has a peak at a much higher value for the Russell Group than for the Non-Russell Group (ten versus zero publications in top journals). But the distribution for the Russell Group has negative skewness, while the distribution for the non-Russell has positive skewness. This is a manifestation of the heterogeneous response to the REF 2014, with some notable strong performances also from within the non-Russell Group of universities and some weak responses even within the Russell group ones.

The Russell Group universities had a much stronger positive impact from REF 2014 in terms of number of publications in top journals per author (panel d), as indicated by the position of the peak (at a positive value for the Russell Group and negative for the non-Russell group) and the shorter and thinner left tail, whereas the distributions of the number of publications per author (panel c) are fairly similar across groups.

Panels f and g show that it is especially the peak of the proportion of Economics publications in top journals, not that of Business, which is markedly higher for the Russell group of universities. In the case of Business, the distribution of the non-Russell group of universities has a fatter negative tail than the Russell Group, but it is relatively smaller than that of the Economics publications. Panel e shows that the distribution of the effects on the proportion of publications in Economics is similar across Russell and non-Russell Groups.

We show the distributions of the subsamples of the Leavers and Remainers in Figure A1 in Appendix 1.[18] Finally, for completeness, Table A5 in Appendix 1

[18]It is interesting to note that the distribution for the number of publications in top journals for the Leavers (panel b) is narrow, showing less heterogeneity in the response of this group, with mostly negative (and not significant) effects. The majority of the Remainers show a strong and positive response, but the negative skewness again reveals heterogeneous behaviour within this group, with some Remainers' outcomes being substantially worse than their placebo counterparts. The distribution of publication per author (panel c) is bi-modal for the Leavers and narrow. This indicates than some of the Remainers had a larger decrease in terms of publication per author than the Leavers group. For all other variables, the Remainers overall had a stronger positive response to the REF 2014 than the Leavers. The majority of Leavers experienced negative effects for the number of publications in top journals per author (panel d), the proportion of publications in Economics (panel e) and the proportion of Economics publications in top journals (panel f). The distributions for the proportion of Business publications in top journals (panel g) are closer

reports the individual treatment effects, cumulated over the treatment period, for all the UK universities (with a codified name) and for all the outcome measures. The first panel shows those of the Russell Group universities and the second panel those of the non-Russell Group.

7.2 Intended effects: research output/excellence

We now describe the university-aggregate effects to analyse the intended effects of the REF, i.e., the effect on research output (number of publications), research excellence (number of publications in top journals), research productivity (number of publications per author) and productivity in research excellence (number of publications in top journals per author). Table 3 presents the aggregate (all-UK universities) results for these outcomes estimated using the SCM, first at the department level and then on a per author basis. We show the yearly effects as well as the estimated ATT (for the overall period 2009-2014) on the outcomes of interest.

Table 3: REF 2014: yearly effects and ATTs

Outcomes	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**ATT_{09-14}**
Number of publications (p_{match}=0.47)	5.67	0.21	-0.19	6.50*	9.73**	18.96****	41.37**
Number of publications in top journals (p_{match}=0.63)	-0.62	-2.72	-3.95**	4.27	4.01***	8.06**	10.93
Number of publications per author (p_{match}=0.57)	-0.028	-0.022	-0.130***	0.088****	0.020	0.124	0.052
Number of publications in top journals per author (p_{match}=0.63)	0.080	0.046	0.039**	0.164	0.158****	0.076	0.563
Proportion of publications in Economics (p_{match}=0.58)	-0.006	0.031	-0.001	0.054**	0.012	0.064*	0.153
Proportion of publications in Economics top journals (p_{match}=0.47)	0.078***	0.084***	-0.062	0.130****	-0.005	0.026	0.248***
Proportion of publications in Business top journals (p_{match}=0.68)	0.096****	0.096***	-0.065	0.191****	0.242****	0.120****	0.679****

Notes: This table reports the Average Treatments on the Treated (ATTs) of the: (1) number of publications, (2) number of publications in top journals, (3) number of publications per author, (4) number of publications in top journals per author, (5) proportion of publications in Economics, (6) proportion of publications in Economics top journals, (7) proportion of publications in Business top journals, as described in Table 2. Values are marked by *, **, ***, **** if they are significant at a level of 0.10, 0.05, 0.01 or 0.001, respectively.

In terms of research output, the overall ATT reports a positive and significant change in the number of publications, showing an overall increase of 41.37 publications per university department, mostly driven by the increase during the last years of the treatment period (2012-2014). This indicates that the total number of publications grew faster for the UK universities than for their US counterfactuals, as it is also shown in Figure 1 panel a.

Regarding research excellence, the ATT for the total number of top publications is positive but insignificant. Despite being negative in the first three years, it be-

for the two groups.

comes positive and significant in 2013 and 2014. The latter results suggest that the positive effect of the REF on the number of top publications might accrue more slowly. Possibly, this might be due to the long lead publication time for top journals (Hadavand et al. 2022).

While the REF positively affects the number of publications, the ATT for the number of publications per author and publications in top journals per author are not significant, neither overall nor for almost any particular year. This indicates that universities have responded to the policy by hiring more academics, rather than by successfully incentivising and supporting their academics to publish more and better.

Overall our results confirm hypotheses 1 and 2 (the latter only towards the end of the REF cycle), i.e., the introduction of the REF has increased research output and research excellence of UK universities. However, we do not find evidence that the REF has increased per author measures, thus showing that hypotheses 3 and 4 do not hold.

7.3 Unintended effects: distortions across disciplines

To test the effects of the REF on the fields of research, we separate university publications in Economics and Business. While the proportion of publications in Economics journals out of all journals declined in both countries, as visible from Figure 1 panel e, the ATT for the proportion of Economics publications shown in Table 3 is not significant. This indicates that this decline is comparable in the UK and the US.

Table 3 also reports a positive and significant ATT for both, the proportion of Economics publications in top journals and the proportion of Business publications in top journals. In this case, as shown in Figure 1 panel f, the proportion of Economics publications in top journals only marginally changed, but the gap with the US counterfactuals appears to become narrower. For Business, as seen in Figure 1, panel g, the proportion of publications in top journals decreased over time in both countries but at a slower rate for the UK universities than their US counterfactuals.

All in all, the REF does not appear to have favoured one discipline at the expense of the other. Thus hypothesis 5 does not hold. On the contrary, the research excellence was positively affected, relative to the counterfactual, in both areas (even though the impact was stronger in Business). As a result, hypothesis 6 holds (reduction for Economics and increase for Business).

7.4 Differences-in-differences estimates

For comparison purposes, we now present the estimates of the traditional differences-in-differences method. As mentioned before, this method, unlike the SCM, relies on the parallel trends assumption and it does not allow us to identify effects at the individual university level, as it does not match each unit to a control. In this case, we assess the REF 2014 effect directly at the group level.

The differences-in-differences specification is based on two binary variables, the first one *Time* takes value zero for the pre-treatment years (2001-2008) and value one for the treatment ones (2009-2014). The second binary variable, *Treated*, distinguishes the treated units (all 81 UK universities) from the control group units (all 128 US universities) by taking a value equal to one for the former and a value equal to zero for the latter. The average treatment effect on the treated, ATT, is here the coefficient δ associated with the interaction of the two dummy variables, *Time*Treated*, as it also captures the average treatment effect of the policy across universities, accounting for what would have happened independently over time without it. We also include, as two additional covariates, the income and expenditure of each university. We obtain these data for the UK from Higher Education Statistics Agency (HESA) and the data for the US from the National Center for Education Statistics (NCES).

The difference-in-differences equation we estimated is:

$$Y_{it} = \alpha + \beta Treated_{it} + \gamma Time_{it} + \delta Treated_{it} * Time_{it} + \eta X_{it} + \epsilon_{it} \quad (1)$$

where Y_{it} is one of the seven outcome measures of interest (i.e., the number of publications, the number of publications in top journals, the number of publications per author, the number of publications in top journals per author, proportion of publications in Economics, the proportion of publications in Economics top journals, and the proportion of publications in Business top journals, as described in Table 2); X_{it} a set of unit time-variant characteristics (university income and expenditure), ϵ_{it} an idiosyncratic error, and *Time* and *Treated* the aforementioned dummy variables.

Table 4 shows that the effect of the REF 2014 has a positive and significant effect on the average number of publications per university (20.806), the average number of publications in top journals (14.761), and on the proportion of Economics publications in top journals (0.076), which is not in disagreement with the SCM results presented earlier. However, the proportion of Business publications in top journals is not significant in this case.

Table 4: Diff-in-Diff estimates

	Dependent variable: Outcomes						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Constant	27.497***	22.169***	1.314***	0.889***	0.366***	0.613***	0.664***
	(4.032)	(2.808)	(0.017)	(0.017)	(0.012)	(0.017)	(0.014)
Income	0.011***	0.009***	0.00002	0.00000	0.00000	0.00001	−0.00001
	(0.003)	(0.002)	(0.00001)	(0.00001)	(0.00001)	(0.00001)	(0.00001)
Expenditure	0.039***	0.028***	0.00001	0.0001***	0.00001	0.00005***	0.00004***
	(0.004)	(0.003)	(0.00002)	(0.00002)	(0.00001)	(0.00002)	(0.00001)
Time	24.081***	2.476	0.084***	−0.124***	−0.057***	−0.073***	−0.147***
	(5.049)	(3.517)	(0.021)	(0.021)	(0.015)	(0.021)	(0.017)
Treated	21.435***	8.178**	−0.102***	−0.215***	−0.110***	−0.239***	−0.061***
	(4.834)	(3.366)	(0.021)	(0.020)	(0.014)	(0.021)	(0.017)
Time*Treated	20.806***	14.761***	−0.027	0.018	−0.008	0.076***	−0.018
	(6.396)	(4.454)	(0.027)	(0.027)	(0.019)	(0.027)	(0.022)
Observations	1,674	1,674	1,674	1,674	1,674	1,674	1,674
Adjusted R^2	0.396	0.401	0.092	0.271	0.135	0.223	0.157

Notes: This table contains the difference-in-differences estimates of the effect of the REF 2014 policy (captured by the *Time*Trend* coefficient) on: (1) number of publications, (2) number of publications in top journals, (3) number of publications per author, (4) number of publications in top journals per author, (5) proportion of publications in Economics, (6) proportion of publications in Economics top journals, (7) proportion of publications in Business top journals, as described in Table 2. Values are marked by *, **, ***, **** if they are significant at a level of 0.10, 0.05, 0.01 or 0.001, respectively.

8 Elite vs non-elite: exacerbating inequalities?

PRFS, as national (one-size fits all) built-in incentive schemes, may affect different individuals and organisations differently, even within the same field (Carli et al. 2019). The UK higher education system, for instance, is composed of about 130 universities, some created in medieval times and others more recently. A particular set of universities is those that form the prestigious Russell Group of 24 research-intensive universities. As shown before, these universities have a higher research performance, not only across all disciplines but also in our two disciplines of interest.

This section tests whether the UK's PRFS reduces or exacerbates inequalities. As explained before, there is already some evidence that shows that the adoption of performance measures in funding allocation perpetuates inequality (Jeon & Kim 2018). Others, such as Checchi et al. (2020) and Buckle et al. (2022), argue that the PRFSs in Italy and New Zealand reduced inequality. Dispersion in research quality across universities can be reduced if it induces below-average universities to hire and retain higher-quality staff and to induce the remaining to improve their research performance relatively more than above-average universities.

Conceptually, previous literature has shown that individuals and organizations respond differently to financial and reputational incentives (Abramo & D'Angelo 2022). Excellent researchers may already possess the skills to carry out the highest-quality publications and therefore do not benefit from the incentive schemes. Similarly, at the other extreme, researchers that will fail to reach excellent standards, independently of their efforts, may not be affected by the incentive schemes. In addition, the individual response to a PRFS depends not only on the extent that the organization internally deploys incentives at the individual level but also on the researchers' sense of belonging and identification with the cause of their organization.

To assess if the REF has contributed to concentration of research in fewer universities, we analyse the outcomes separately for universities belonging to the elite, research-intensive Russell Group and universities that do not, as a subsampling exercise. Results are displayed in Table 5.

Table 5: REF 2014 ATTs: Russell vs Non-Russell

Outcomes	**ATT_{09-14}**		
	Russell	**Non Russell**	**Diff. Russell-Non Russell**
Number of publications	59.67**	35.98*	23.68
Number of publications in top journals	39.07**	5.78	33.28*
Number of publications per author	0.258	-0.052	0.310
Number of publications in top journals per author	0.577	-0.231	0.808***
Proportion of publications in Economics	0.166	-0.167	0.334*
Proportion of Economics publications in top journals	0.126	0.581****	-0.455***
Proportion of Business publications in top journals	0.858****	0.505****	0.353***

Notes: This table reports the Average Treatments on the Treated (ATTs) of the: (1) number of publications, (2) number of publications in top journals, (3) number of publications per author, (4) number of publications in top journals per author, (5) proportion of publications in Economics, (6) proportion of publications in Economics top journals, (7) proportion of publications in Business top journals, as described in Table 2 for the Russell and Non-Russell universities. Values are marked by *, **, ***, **** if they are significant at a level of 0.10, 0.05, 0.01 or 0.001, respectively.

What stands out in Table 5 is that the number of publications increased for both groups (ATT of 59.67 and 35.98, respectively), and that the difference across groups is not statistically significant. The number of publications in top journals however, increased only for the Russell Group (ATT 39.07), and the difference across groups is significant also (ATT 33.28).

Our results also show that the REF had an insignificant effect on the number of publications per author and the number of publications in top journals per author for both groups. However, the latter measure was significantly higher amongst Russell Group universities.

In terms of the field distortions, we find an insignificant coefficient in the proportion of publications in Economics across the two groups, but the difference is slightly significant as the coefficient is positive for Russell and negative for non-Russell. Finally, we observe a negative and statistically significant difference between the groups for the proportion of publications in top Economic journals (ATT -0.455), showing that the non-Russell Group benefited more from the REF 2014 than the Russell Group for this particular outcome. On the contrary, the Russell Group benefited more from the REF 2014 than the non-Russell Group for the proportion of top journals in Business publications. This may be a reflection of Economics being a more mature field in the UK, limiting the scope for improvement of elite Universities that were already producing top-class Economics research at capacity while further growth was still possible in the area of Business.

In sum, the REF 2014, by widening the gap between the Russell Group relative to the non-Russell Group, in terms of the number of publications in top journals and the number of publications in top journal per author, has contributed to the concentration of research excellence in a few elite universities, exacerbating inequalities.

This is not to say that non-Russell Groups Universities have not benefited from the REF 2014. Notably the gap in the proportion of publications in Economics top journals has narrowed. However, the REF has affected the relative size of the two fields in the two groups, with Business expanding faster than Economics in the non-Russell group. This provides some evidence to claims that the REF may have contributed to the shrinking of the Economics discipline (Lee et al. 2013, Guizzo et al. 2021). However, this was not accompanied by a decline in quality. Rather, the quality of publications has become more homogeneous across the two groups in the area of Economics, but more heterogeneous in the area of Business.

9 Remain vs leave: should I stay or should I go?

The PRFSs may also generate strategic behaviour. As mentioned earlier, universities in the UK decide which REF panel they submit their research to. The number of universities entering the Economics and Econometrics panel declined significantly in recent times, as the number of universities submitting to the Business and Management panel grew (Johnston & Reeves, 2020).

Whatever the reasons for the switching from one panel over the other,[19] this choice may have negatively affected the ability to continue to produce research at similar standards. The loss of professional identity driven by the switching and potential associated managerial decisions to reduce investment in this area may have hampered the recruitment and retention of research-excellent academics, lowering the overall quality of Economic research at these institutions. Some evidence in this direction has emerged from the reports of the Economics and Econometrics sub-panel,[20] which noted how the work cross-referred to the Economics and Econometrics panel from the Business and Management panel was generally of lower quality than

[19] Among the proposed reasons are the increasing number of economics departments being located within business schools, economies of scale associated with entering one rather than two panels, and the level of tolerance of poor performance. We do not know the time at which they made this decision either. Our analysis implicitly assumes that the decision was made at the beginning of the period.

[20] See REF 2014 Panel overview reports, Main Panel C and sub-panels 16-26 https://www.ref.ac.uk/2014/media/ref/content/expanel/member/Main%20Panel%20C%20overview%20report.pdf

that submitted directly to the Economics and Econometrics panel in both 2008 and 2014.[21] If the switching universities felt that the standards are lower in the Business and Management panel, they may have reduced the incentives to produce excellent research in general.

To further investigate this concern, we analyse the outcomes of the group of Universities which submitted to RAE 2008 but not to the REF 2014 (the *Leavers*),[22] separately from those which submitted to the Economics and Econometrics panel in both evaluation rounds (the *Remainers*). This analysis will allow us to test whether Leavers benefited or not from leaving. Results are shown in Table 6.[23]

Table 6: REF 2014 ATTs: Remainers vs Leavers

Outcomes	**ATT$_{09-14}$**		
	Remainers	**Leavers**	**Diff. Remainers-Leavers**
Number of publications	85.45****	-0.57	86.03****
Number of publications in top journals	36.33**	-11.20	47.53***
Number of publications per author	0.028	-0.457	0.485***
Number of publications in top journals per author	0.116	-0.695	0.812***
Proportion of publications in Economics	0.166	-0.391	0.557***
Proportion of Economics publications in top journals	0.117	-0.458*	0.576***
Proportion of Business publications in top journals	0.705****	-0.262	0.968****

Notes: This table reports the Average Treatments on the Treated (ATTs) of the: (1) number of publications, (2) number of publications in top journals, (3) number of publications per author, (4) number of publications in top journals per author, (5) proportion of publications in Economics, (6) proportion of publications in Economics top journals, (7) proportion of publications in Business top journals, as described in Table 2 for the Remainers and Leavers universities. Values are marked by *, **, ***, **** if they are significant at a level of 0.10, 0.05, 0.01 or 0.001, respectively.

We can see that the Remainers experienced a significant increase in the number of articles published (ATT 85.45), and also in the publications in top journals (ATT 36.33), as well as in the proportion of publications in top journals in Business (ATT 0.705). However, the Leavers had a very different fate. The effect of the REF 2014 on their outcomes is insignificant or negative and only significant for one of the outcomes when compared with their placebo counterparts: the proportion of publications in Economics top journals (ATT -0.458).

If we compare the Remainers' ATT effects with those of the Leavers, results show that there is a positive and significant difference between the groups in all outcomes. Thus the REF 2014 has exacerbated inequalities among the Remainer and Leaver groups. In particular, it appears that once universities leave the Economics

[21]The Business and Management panel cross-refers to the Economics and Econometrics panel outputs submitted to them but deemed to be more appropriate for the Economics and Econometrics panel.

[22]The six universities submitted to the Economics and Econometrics panel in RAE2008, but not in REF2014, had the lowest average scores in RAE2008. All six submitted economists in Business and Management panel in REF2014.

[23]Johnston & Reeves (2018) suggest that Russell Group universities and/or Universities based in London and the South of England are more likely to remain. Thus, the REF may have contributed not only to the concentration of research excellence in elite institutions but also to further its geographic fragmentation across the UK.

panel, their ability to produce world-leading and internationally excellent Economics research declines, possibly due to a shift in focus and institutional priorities. Unfortunately, this decline is not accompanied by a rise in research excellence in Business, suggesting that the panel shift was not driven by focused strategic investments in this area.

10 Discussion and Conclusion

The aim of this paper is to causally evaluate the performance of the PRFSs. We have made use of SCM methodology to analyse the intended and some of the unintended consequences of the introduction of the REF, the UK's PRFS, often used as a role model for other PRFS worldwide. Our results indicate that the introduction of a PRFS can improve research performance in terms of, both, research output and research excellence. This is in line with the findings of previous studies (Franzoni et al. 2011, Wang & Hicks 2013, Checchi et al. 2019) and may help convince critics of the potential for PRFS incentive schemes to improve research performance (European Commission 2020).

We show in particular that the REF 2014 had a significant and positive impact on the number of publications as well as the volume (and proportion) of world-leading research during the 2009-2014 assessment period. While publications in top journals did not significantly increase throughout the entire period, they did so in the last two years, 2013 and 2014. We also show that the effects of REF vary more across universities than across academic disciplines. We do not find in particular evidence of a shift in research focus from Economics to Business. We record, in fact, a significant increase in the proportion of publications in top journals in Economics, though this effect is stronger in Business. We do not find, then, support for the argument that potentially higher REF standards in Economics have reduced the quantity and the quality of the Economics research output in the UK (Lee et al. 2013, Guizzo et al. 2021).

Note that our analysis includes, for each university (both treatment and control), all publications from all authors, defined according to the affiliation recorded in their publications. Given this, our findings are less affected by the possible "gaming the system" behaviour the assessed universities may engage in, such as in the self-selection of staff and outputs or in the recruitment of academics shortly before the census date. These tactics, designed to maximise REF performance, were identified as areas of concern by The Stern Independent Review of REF 2014 (Stern 2016). In this respect, our approach allows us to assess more realistically and rigorously the effectiveness of the PRFSs on research performance.

Understanding how PRFSs, through what they measure and reward, shape researchers' and institutions' behaviours and attitudes is crucial from a policy perspective. As recognised by the 2021 Real-Time REF review (Manville et al. 2021), future assessment models should not be limited to evaluating tangible outputs but also the processes that lead to the production of those outputs. An increased focus on recognizing, influencing and incentivising what is valued by a broad and diverse research community, would support a more positive research culture. Our paper offers a rigorous approach to unveil some of the unintended consequences and behaviours driven by REF2014, which can inform the ongoing debates on the design of the future PRFS.

For example, our analysis indicates that, while the effects in the overall quantity and quality of research have been positive, there has been no significant effect on the corresponding per-author measures. This suggests that the REF 2014 did not result in an overall increase in research productivity or in the productivity of research excellence. Universities do not appear to have been successful at incentivising and supporting their academics to publish more or/and better. Rather the increase in output was driven by an increase in the number of active researchers, something influenced by universities' hiring policies. This supports the claim that the RAE/REFs have distorted universities' hiring decisions (Hayri 1997, La Manna 2008, OECD 2010, Stern 2016). After all, the UK's PRFS (and the other PRFSs) target departments/universities, not individuals.

Our results also confirm that the PRFFs may enhance the concentration of research output in elite universities (Bence & Oppenheim 2005, Martin & Whitley 2010, Clerides et al. 2011, Mingers & White 2015, Jeon & Kim 2018, Guizzo et al. 2021). The REF 2014 appears to have reinforced the strong position of the already strong Russell Group universities. This was perhaps to be expected, but it may not have been one of the motivating drivers behind the introduction of the REF. When comparing Russell Group and non-Russell Group universities, our findings show that, while the non-Russell Group displayed a greater increase in the proportion of publications in top journals in Economics, the Russell Group universities experienced a higher increase in the overall numbers of publications and publications in top journals. Nonetheless, we also detected significant within-group heterogeneity in performance: some Russell Group universities under-performing and some non-Russell Group universities over-performing, their respective counterfactuals.

PRFS-related strategic behaviour was also investigated through an analysis of the *Remainers* versus *Leavers* on the Economics and Econometrics panel. Results show that there is a positive and significant difference, for all outcomes, in favour of those who continued to submit to the Economics and Econometrics panel. All outcome measures of the Leavers decreased, though this was only significant for the proportion of publications in top journals in Economics.

A policy recommendation stemming from our analysis is for future assessment exercises (and other PRFS) to encourage and reward a more inclusive culture of collaboration among domestic Higher Education institutions, rather than to make them compete against each other, which increases the performance gap and induces the weak ones to stop submitting altogether. Comparisons of research performance among universities may be more meaningful if carried out against comparable international benchmarks rather than within a heterogeneous Higher Education sector. Our approach provides a feasible method to implement such an innovative model.

The metrics we selected to measure research output and excellence in the current paper are grounded in the reputation of the journals where the research was published. This approach is largely accepted in Economics and Business and in other Social Sciences. But, in fields such as Physics, Mathematics, Chemistry, Biology, and Medicine, greater emphasis is generally placed on the level of attention a study garners, typically measured by its citation rate. Moreover, as compared to REF 2014, REF2021 placed more emphasis on other measures of research performance, such as the inter-disciplinary nature of the research and its socio-economic impact, while reports such as the Harnessing Metric Tide review (Curry et al. 2022) have called for the responsible use of metrics for research assessment. Our approach can support alternative measures of research performance, to be chosen among the ones that may be more relevant to each fields, and provides a robust data-driven methodology that could help streamline research assessment in future exercises.

In fact, the effect of the REF on the mobility of researchers, co-authorships and interdisciplinary collaborations, as well as the gender impact of research excellence evaluation, are important questions (Arnold et al. 2018) that we hope to address with our data and methods at another time. In future analyses, we plan to perform a longitudinal study that includes the period leading to the REF 2021. While the incremental effects of the REF 2021 relative to the REF 2014 may not be as large as those of the REF 2014 relative to the RAE 2008, performing a combined analysis of both schemes might reveal if the effects of a given PRFS are transient– i.e., universities publishing more just before the end of the assessment period–relative to their counterfactual- or permanent.

Acknowledgements

We are indebted to Elsevier for allowing access to the Scopus dataset and to Christian Batchelor for his invaluable technical support. We are also grateful for the comments of the editor, Ben Martin, and two anonymous reviewers. We also acknowledge the feedback received at the Annual Meeting of the Association for Mathematics Applied to Social and Economic Sciences (AMASES), which took place in Perugia in September 2019, at the WEHIA in Japan 2019, at the Internal Seminar at City, University of London in 2019, at the Barcelona School of Economics (BSE) Summer Forum in June 2022, and at the D2 seminar series of the University of Florence Center for Data Science in October 2022.

References

Abadie, A. & Cattaneo, M. D. (2018), 'Econometric methods for program evaluation', *Annual Review of Economics* **10**(1), 465–503.

Abadie, A., Diamond, A. & Hainmueller, J. (2010), 'Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program', *Journal of the American Statistical Association* **105**(490), 493–505.

Abadie, A., Diamond, A. & Hainmueller, J. (2015), 'Comparative politics and the synthetic control method', *Am J Pol Sci* **59**, 495–510.

Abadie, A. & Gardeazabal, J. (2003), 'The economic costs of conflict: A case study of the basque country', *The American Economic Review* **93**(1), 113–132.

Abramo, G. & D'Angelo, C. (2022), 'The different responses of universities to introduction of performance-based research funding.', *Research Evaluation* **30**, 514–528.

Acemoglu, D., Johnson, S., Kermani, A., Kwak, J. & Mitton, T. (2016), 'The value of connections in turbulent times: Evidence from the united states', *Journal of Financial Economics* **121**, 368–391.

Andersen, L. & Pallesen, T. (2008), '"not just for the money?" how financial incentives affect the number of publications at danish research institutions', *International Public Management Journal* **11**(1), 28–47.

Arnold, E., Simmonds, P., Farla, K., Kolarz, P., Mahieu, B. & Nielsen, K. (2018), 'Review of the Research Excellence Framework: Evidence Report', https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/fi
excellence-framework-review-evidence-report.pdf.

Auranen, O. & Nieminen, M. (2010), ‘University research funding and publication performance,Äîan international comparison’, *Research Policy* **39**(6), 822–834.

Basso, A. & di Tollo, G. (2022), ‘Prediction of uk research excellence framework assessment by the departmental h-index’, *European Journal of Operational Research* **296**(3), 1036–1049.

Battistin, E. & Ovidi, M. (2022), ‘Rising stars: Expert reviews and reputational yardsticks in the research excellence framework’, *Economica* **89**(356), 830–848.

Bence, V. & Oppenheim, C. (2005), ‘The evolution of the uk’s research assessment exercise: publications, performance and perceptions’, *Journal of Educational Administration and History* **37**(2), 137–155.

Bloch, C. & Schneider, J. W. (2016), ‘Performance-based funding models and researcher behavior: An analysis of the influence of the norwegian publication indicator at the individual level’, *Research Evaluation* **25**(4), 371–382.

Bouttell, J., Craig, P., Lewsey, J., Robinson, M. & Popham, F. (2018), ‘Synthetic control methodology as a tool for evaluating population-level health interventions’, *J Epidemiol Community Health* **72**, 673–678.

Buckle, R., Creedy, J. & Gemmell, 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.

Butler, L. (2003), ‘Explaining australia’s increased share of isi publications’ effects of a funding formula based on publication counts’, *Research Policy* **32**(1), 143–155.

Carli, G., Tagliaventi, M. R. & Cutolo, D. (2019), ‘One size does not fit all: the influence of individual and contextual factors on research excellence in academia’, *Studies in Higher Education* **44**(11), 1912–1930.

Cavallo, E., Galiani, S., Noy, I. & Pantano, J. (2013), ‘Catastrophic natural disasters and economic growth.’, *Review of Economics and Statistics* **95**(5), 1549–1561.

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. & Malgarini, M. (2020), ‘Convergence or polarisation? the impact of research assessment exercises in the italian case.’, *Scientometrics* **124**, 1439–1455.

Chiang, K.-H. (2019), 'From rae to ref: Trust and atmosphere in uk higher education reform.', *Journal of Education Social Policy* **6**(1), 29–38.

Clerides, S., Pashardes, P. & Polycarpou, A. (2011), 'Peer review vs metric-based assessment: Testing for bias in the rae ratings of uk economics departments', *Economica* **78**(311), 565–583.

Curry, S., Gadd, E. & Wilsdon, J. (2022), 'Harnessing the Metric Tide: indicators, infrastructures & priorities for UK responsible research assessment.', https://doi.org/10.6084/m9.figshare.21701624.v2. [Online; accessed 14-March-2023].

Davé, A., Hopkins, M., Hutton, J., Krcál, A., Kolarz, P., Martin, B., Nielsen, K., Rafols, I., Rotolo, D., Simmonds, P. et al. (2016), 'Landscape review of interdisciplinary research in the uk', *Report to HEFCE and RCUK by Technopolis and the Science Policy Research Unit (SPRU), University of Sussex* **184**.

Debackere, K. & Glänzel, W. (2004), 'Using a bibliometric approach to support research policy making: The case of the flemish bof-key.', *Scientometrics* **59**(2), 253–276.

European Commission (2010), 'Assessing europe's university-based research. expert group on assessment of university-based research', https://op.europa.eu/en/publication-detail/-/publication/93ec2eb0-b614-41df-a894-56895a795a54.

European Commission (2020), 'MLE on Performance-based Funding of Public Research Organisations', https://ec.europa.eu/research-and-innovation/en/statistics/policy-support-facility/mle-performance-based-funding-public-research-organisations.

Franzoni, C., Scellato, G. & Stephan, P. (2011), 'Changing incentives to publish', *Science* **333**(6043), 702–703.

Frey, B. S. & Jegen, R. (2001), 'Motivation crowding theory', *Journal of Economic Surveys* **15**(5), 589–611.

Geuna, A. & Martin, B. (2003), 'University research evaluation and funding: An international comparison.', *Minerva* **41**, 277–304.

Geuna, A. & Piolatto, M. (2016), 'Research assessment in the uk and italy: Costly and difficult,but probably worth it (at least for a while).', *Research Policy* **46**, 260–271.

Gläser, J., Laudel, G., Hinze, S. & Butler, L. (2002), Impact of evaluation-based funding on the production of scientific knowledge: What to worry about, and how to find out (expertise for the german ministry for education and research).

Gneezy, U., Meier, S. & Rey-Biel, P. (2011), 'When and Why Incentives (Don't) Work to Modify Behavior', *Journal of Economic Perspectives* **25**(4), 191–210.

Guizzo, D., Walker, J., Della-Giusta, M. & Fontinha, R. (2021), 'Assessing the impact of the research excellence framework (ref) on uk academic macroeconomists', *Rebuilding Macroeconomics Working Paper Series* **59**.

Hadavand, A., Hamermesh, D. S. & Wilson, W. W. (2022), 'Publishing economics: How slow? why slow? is slow productive? how to fix slow?', *Journal of Economic Literature* (forthcoming).

Hammarfelt, B. & de Rijcke, S. (2015), 'Accountability in context: effects of research evaluation systems on publication practices, disciplinary norms, and individual working routines in the faculty of arts at uppsala university', *Research Evaluation* **24**(1), 63–77.

Harman, G. (2000), 'Allocating research infrastructure grants in post-binary higher education systems: British and australian approaches', *Journal of Higher Education Policy and Management* **22**(2), 111–126.

Hayri, A. (1997), 'The research assessment exercise and transfer of academics among departments', *isk Decision and Policy* **2**(1), 71–86.

Hicks, D. (2012), 'Performance-based university research funding systems', *Research Policy* **41**, 251– 261.

Hicks, D. & Wang, J. (2009), 'Towards a bibliometric database for the social sciences and humanities—a european scoping project', *Final Report on Project for the European Science Foundation* .

Higher Education Staff Statistics (2014), 'Higher Education Staff Statistics: UK', https://www.hesa.ac.uk/data-and-analysis/statistical-first-releases?date$_{f}ilter$

Hobbs, R. & Roberts, L. (2016), 'The stern review of the research excellence framework', *BMJ* **354**.

Hole, A. (2017), 'Ranking economics journals using data from a national research evaluation exercise', *Oxford Bulletin of Economics and Statistics* **79**(5), 0305–9049.

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**(2), 589–606.

Johnston, J. & Reeves, A. (2018), 'An investigation into the role played by research assessment in the socio-geographic fragmentation of undergraduate economics education in the uk', *Higher Education* **76**(4), 589–614.

Johnston, J. & Reeves, A. (2020), 'Research in uk universities: a tale of two subjects ,Äì economics and econometrics; and business and management studies', *European Journal of Higher Education* **10**(2), 164–180.

Jonkers, K. & Zacharewicz, T. (2016), *Research Performance Based Funding Systems: a Comparative Assessment*, Publications Office of the European Union, Luxembourg.

Kreif, N., Grieve, R., Hangartner, D., Turner, A. J., Nikolova, S. & Sutton, M. (2016), 'Examination of the synthetic control method for evaluating health policies with multiple treated units', *Health Economics* **25**(12), 1514–1528.

La Manna, M. (2008), 'Assessing the assessment or, the rae and the optimal organization of university research', *Scottish Journal of Political Economy* **55**(5), 637–653.

Lee, F. S., Pham, X. & Gu, G. (2013), 'The UK Research Assessment Exercise and the narrowing of UK economics', *Cambridge Journal of Economics* **37**(4), 693–717.

Manville, C., d'Angelo, C., Culora, A., Gloinson, E. R., Stevenson, C., Weinstein, N., Wilsdon, J., Haddock, G. & Guthrie, S. (2021), *Understanding perceptions of the Research Excellence Framework among UK re- searchers: The Real-Time REF Review.*, RAND Corporation, Santa Monica, CA.

Marques, M., Powell, J., 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. (2011), 'The research excellence framework and the 'impact agenda':are we creating a frankenstein monster?', *Research Evaluation* **20**(3), 247–254.

Martin, B. & Whitley, R. (2010), 512 The UK Research Assessment Exercise: A Case of Regulatory Capture, *in* 'Reconfiguring Knowledge Production: Changing Authority Relationships in the Sciences and their Consequences for Intellectual Innovation', Oxford University Press.

Mingers, J. & White, L. (2015), 'Throwing out the baby with the bathwater: The undesirable effects of national research assessment exercises on research'.

Moed, H. (2008), 'Uk research assessment exercise: Informed judgement on research quality or quantity?', *Scientometrics* **74**, 153–161.

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

OECD (2010), 'Performance-based Funding for Public Research in Tertiary Education Institutions Workshop Proceedings', https://read.oecd-ilibrary.org/education/performance-based-funding-for-public-research-in-tertiary-education-institutions$_9$789264094611 $- enpage4$. [$Online; accessed5 - December - 2022$].

Pidd, M. & Broadbent, J. (2015), 'Business and management studies in the 2014 research excellence framework', *British Journal of Management* **26**(4), 569–581.

Pinar, M. & Horne, T. J. (2021), 'Assessing research excellence: Evaluating the research excellence framework.', *Research Evaluation* **31**(2), 173–187.

Sivertsen, G. (2017), 'Unique, but still best practice? the research excellence framework (ref) from an international perspective.', *Palgrave Commun* **3**, 17078.

Soderlind, J., Nordstrand, L., Krog Lind, J. & Pulkkinen, K. (2019), National performance-based research funding systems: Constructing local perceptions of research?, *in* R. Pinheiro, L. Geschwind, H. Foss Hansen & K. Pulkkinen, eds, 'Reforms, Organizational Change and Performance in Higher Education: A Comparative Account from the Nordic Countries', Springer International Publishing, Cham, pp. 111–144.

Stern, N. (2016), 'Research excellence framework (ref) review: Building on success and learning from experience.', *Department for Business, Energy Industrial Strategy* **Ref: IND/16/9**.

Stockhammer, E., Dammerer, Q. & Kapur, S. (2021), 'The research excellence framework 2014, journal ratings and the marginalisation of heterodox economics', *Cambridge Journal of Economics* **45**(2), 243–269.

Taylor, J. (2011), 'The assessment of research quality in uk universities: Peer review or metrics?', *British Journal of Management* **22**, 202–217.

Wang, J. & Hicks, D. (2013), 'Detecting structural change in university research systems: A case study of british research policy', *Research Evaluation* **22**, 258–268.

Wilsdon, J., Allen, L., Belfiore, E., Campbell, P., Curry, S., Hill, S., Jones, R., Kain, R., Kerridge, S., Thelwall, M., Tinkler, J., Viney, I., Wouters, P., Hill, J. & Johnson, B. (2015), 'The Metric Tide. Report of the Independent Review of the Role of Metrics in Research Assessment and Management', https://blogs.lse.ac.uk/impactofsocialsciences/files/2015/07/2015$_m$*etrictide.pdf*.

Zacharewicz, T., Lepori, B., Reale, E. & Jonkers, K. (2019), 'Performance-based research funding in eu member states - comparative assessment', *Science and Public Policy* **46**(1), 105–115.

Appendix

Appendix 1: Additional Tables and Figures

Table A1: Number of publications in journals (2001-2014): UK and US.

Field	publications UK	publications US	journals UK	journals US
ACCOUNT	1727	3478	35	24
ECON	13484	37101	174	86
ENT-SBM	975	968	14	6
ETHICS-CSR-MAN	2686	4604	31	13
FINANCE	3208	7986	60	18
ECON HYST	894	710	18	2
HRM&EMP	1849	1828	28	10
IB&AREA	1259	1325	24	8
INFO MAN	885	2977	24	14
INNOV	947	1215	21	2
MDEV&EDU	767	724	10	9
MGDEV&ED	31	7	1	1
MKT	2541	6721	50	14
OPS&TECH	2924	4571	44	10
OR&MANSCI	2966	8691	33	13
ORG STUD	1612	2194	16	9
PLANNING	858	546	7	3
PSYCH (GENERAL)	211	699	2	4
PSYCH (WOP-OB)	335	2339	6	10
PUB SEC	763	2044	11	4
SECTOR	3060	5089	55	23
SOC SCI	2077	2049	17	1
STRAT	552	1059	8	3
Total	46611	98925	688	287

Notes: This table reports the number of publications and journals by journal subject area. The journal subject categories are based on both, the Academic Journal Guide (AJG) journal classification and Scopus. This categories are then used to classify the publications into ”Economics” and ”Business”.

Table A2: UK universities: outcomes averages (2001-2014)

Universities of UK	(1)	(2)	(3)	(4)	(5)	(6)	(7)	Russell (R)/ Non-Russell(NR)	Remainers (Re)/ Leavers (L)
University of Manchester	250.93	153.00	1.19	0.74	0.24	0.46	0.69	R	Re
University of Oxford	231.86	123.64	1.45	0.81	0.45	0.64	0.52	R	Re
LSE	231.00	143.50	1.39	0.89	0.51	0.71	0.57	R	Re
University of Warwick	208.29	135.14	1.42	0.96	0.30	0.69	0.68	R	Re
University of Cambridge	202.93	117.57	1.27	0.75	0.37	0.60	0.60	R	Re
University of Nottingham	192.86	144.57	1.27	0.96	0.48	0.54	0.89	R	Re
Cardiff University	144.21	99.93	1.24	0.88	0.17	0.37	0.79	R	
University College London	127.14	72.79	1.35	0.83	0.54	0.74	0.49	R	Re
Lancaster University	124.57	89.43	1.18	0.86	0.20	0.42	0.81	NR	
University of Leeds	113.64	69.50	1.17	0.73	0.18	0.48	0.67	R	
Imperial College London	113.57	76.36	1.28	0.85	0.30	0.53	0.74	R	
University of Birmingham	113.36	48.43	1.32	0.58	0.26	0.45	0.44	R	Re
University of Southampton	108.00	68.64	1.20	0.80	0.30	0.63	0.69	R	Re
City University London	108.00	75.07	1.33	0.93	0.22	0.55	0.76	NR	Re
University of Strathclyde	103.36	64.21	1.20	0.74	0.15	0.41	0.66	NR	
University of Sheffield	100.21	48.07	1.22	0.58	0.19	0.31	0.51	R	Re
University of Bath	99.00	58.71	1.28	0.78	0.22	0.33	0.69	NR	
Brunel University London	94.93	45.86	1.23	0.58	0.24	0.38	0.52	NR	Re
University of Reading	87.79	42.57	1.18	0.59	0.28	0.42	0.54	NR	
London Business School	87.21	63.29	1.50	1.10	0.15	0.81	0.74	NR	
University of Edinburgh	85.71	45.00	1.28	0.71	0.17	0.63	0.55	R	
Cranfield University	84.21	48.14	1.03	0.60	0.08	0.12	0.62	NR	
University of York	81.50	40.36	1.32	0.67	0.52	0.57	0.51	R	Re
University of Essex	77.14	55.14	1.29	0.96	0.51	0.78	0.70	NR	Re
Aston University	74.93	50.71	1.44	0.98	0.13	0.24	0.75	NR	
University of Surrey	73.86	40.64	1.23	0.67	0.24	0.46	0.56	NR	Re
University of Glasgow	73.07	38.86	1.11	0.60	0.31	0.51	0.58	R	Re
University of Leicester	72.00	34.07	1.37	0.67	0.38	0.49	0.51	NR	Re
Newcastle University	70.64	34.71	1.12	0.56	0.32	0.52	0.51	R	
University of Exeter	70.64	43.64	1.24	0.77	0.26	0.66	0.62	R	Re
University of East Anglia	70.07	43.21	1.20	0.78	0.54	0.64	0.68	NR	Re
University of Durham	69.71	36.50	1.33	0.69	0.23	0.37	0.57	R	
University of Bristol	69.00	41.07	1.28	0.78	0.38	0.71	0.57	R	Re
University of Kent	64.36	34.29	1.37	0.76	0.38	0.53	0.56	NR	Re
King's College London	62.29	33.93	1.29	0.78	0.10	0.49	0.62	R	
Queen Mary University of London	62.21	35.36	1.53	0.89	0.49	0.66	0.53	R	Re
University of Sussex	61.79	35.93	1.11	0.64	0.26	0.38	0.67	NR	Re
University of Liverpool	60.21	28.21	1.19	0.59	0.12	0.53	0.49	R	
Royal Holloway, University of London	57.00	31.21	1.31	0.74	0.30	0.74	0.49	NR	Re
University of Stirling	55.00	26.71	1.35	0.71	0.28	0.54	0.54	NR	L
Open University	54.71	21.07	1.00	0.39	0.15	0.49	0.39	NR	
Queen's University Belfast	51.00	26.50	1.23	0.63	0.28	0.53	0.50	R	
University of St Andrews	48.00	28.36	1.54	0.91	0.42	0.48	0.69	NR	Re
University of Salford	47.86	17.79	1.05	0.40	0.08	0.16	0.40	NR	
University of Ulster	47.79	16.93	0.98	0.36	0.13	0.09	0.40	NR	
University of Hull	47.29	19.50	1.41	0.57	0.14	0.19	0.45	NR	
Heriot-Watt University	46.71	15.79	1.20	0.40	0.20	0.28	0.33	NR	
University of Aberdeen	45.07	21.86	1.22	0.59	0.40	0.44	0.56	NR	Re
Manchester Metropolitan University	44.00	12.71	1.14	0.35	0.09	0.13	0.32	NR	L
University of the West of England, Bristol	43.07	14.86	1.10	0.42	0.13	0.18	0.43	NR	
Middlesex University	42.93	16.57	1.22	0.43	0.22	0.22	0.46	NR	
University of Bradford	42.79	21.79	1.17	0.58	0.08	0.07	0.54	NR	
Birkbeck College	41.57	22.14	1.42	0.81	0.44	0.59	0.53	NR	Re
Swansea University	41.36	17.00	1.41	0.58	0.45	0.37	0.51	NR	L
University of Portsmouth	39.29	17.14	1.10	0.51	0.26	0.33	0.53	NR	
University of Plymouth	36.21	12.00	1.05	0.38	0.11	0.28	0.39	NR	
Bournemouth University	35.21	12.50	1.14	0.43	0.03	0.11	0.38	NR	
Oxford Brookes University	33.93	10.71	1.31	0.38	0.09	0.11	0.33	NR	
Nottingham Trent University	33.86	11.14	1.27	0.48	0.16	0.34	0.39	NR	
Kingston University	31.50	12.21	1.13	0.41	0.12	0.08	0.41	NR	L
London Metropolitan University	30.14	10.00	1.22	0.41	0.19	0.36	0.34	NR	L
University of Westminster	29.86	9.21	1.09	0.35	0.18	0.13	0.35	NR	
De Montfort University	29.57	15.64	1.18	0.69	0.04	0.17	0.61	NR	
University of Northumbria at Newcastle	28.50	7.29	1.07	0.36	0.03	0.02	0.35	NR	
Leeds Beckett University	27.71	4.36	1.15	0.21	0.03	0.17	0.18	NR	
Sheffield Hallam University	26.64	8.36	1.05	0.35	0.03	0.14	0.36	NR	
Bangor University	25.86	15.29	1.50	0.75	0.30	0.37	0.55	NR	
Glasgow Caledonian University	25.79	8.21	0.95	0.34	0.06	0.14	0.37	NR	
University of Dundee	25.36	15.57	1.07	0.67	0.29	0.40	0.71	NR	L
University of Hertfordshire	23.57	11.00	1.31	0.62	0.33	0.30	0.56	NR	
Coventry University	22.64	7.07	1.03	0.39	0.08	0.07	0.40	NR	
University of South Wales	21.86	4.43	1.12	0.27	0.02	0.00	0.25	NR	
University of Greenwich	21.86	7.86	1.21	0.49	0.09	0.02	0.48	NR	
University of Central Lancashire	21.79	6.14	1.68	0.48	0.04	0.18	0.31	NR	
Edinburgh Napier University	21.21	4.86	1.08	0.26	0.08	0.00	0.27	NR	Re
University of Brighton	20.36	7.43	1.01	0.42	0.04	0.10	0.45	NR	
Keele University	19.79	8.64	1.27	0.58	0.21	0.54	0.44	NR	
Aberystwyth University	18.36	7.21	1.27	0.53	0.19	0.39	0.43	NR	
University of Wolverhampton	17.14	2.43	0.95	0.13	0.03	0.00	0.16	NR	
Robert Gordon University	17.07	6.79	1.13	0.46	0.10	0.18	0.42	NR	
London South Bank University	14.29	4.36	1.28	0.35	0.10	0.15	0.34	NR	

Notes: This table reports the outcomes averages for each UK university for: 1) publications, (2) publications in top journals, (3) publications per author, (4) publications in top journals per author, (5) proportion of publications in Economics, (6) proportion of publications in Economics top journals, (7) proportion of publications in Business top journals, as described in Table 2. Universities are listed in decreasing order according to the number of publications. The last two columns classify the universities in Russell/Non-Russell and in Remainers/Leavers.

Table A3: US universities: outcomes averages (2001-2014)

Universities of US	(1)	(2)	(3)	(4)	(5)	(6)	(7)	Selected
Harvard University	375.93	278.29	1.44	1.08	0.47	0.79	0.74	Yes
University of California-Berkeley	291.57	205.43	1.49	1.07	0.46	0.77	0.68	
University of Michigan	287.64	201.36	1.30	0.92	0.34	0.68	0.74	
University of Pennsylvania	280.79	221.86	1.48	1.19	0.34	0.77	0.81	
Pennsylvania State University	270.71	161.57	1.19	0.73	0.21	0.57	0.63	Yes
Columbia University	267.93	183.43	1.53	1.07	0.39	0.74	0.69	
Texas A&M University	259.14	146.21	1.28	0.76	0.30	0.61	0.60	
Stanford University	257.79	177.57	1.41	0.99	0.43	0.76	0.67	
Cornell University	251.43	158.86	1.36	0.87	0.40	0.63	0.66	
New York University (NYU)	250.64	182.07	1.46	1.08	0.37	0.77	0.73	
University of Illinois at Urbana-Champaign	233.64	143.43	1.29	0.81	0.32	0.61	0.65	
Massachusetts Institute of Technology (MIT)	230.86	179.64	1.47	1.15	0.38	0.83	0.77	
Michigan State University	220.00	155.21	1.29	0.94	0.29	0.58	0.78	
Indiana University	215.86	128.00	1.29	0.79	0.25	0.61	0.63	
University of Maryland	209.71	157.57	1.31	0.98	0.36	0.75	0.76	
Northwestern University	201.71	155.14	1.47	1.15	0.41	0.83	0.75	
Arizona State University	196.43	135.79	1.24	0.87	0.18	0.67	0.71	
Rutgers University-New Brunswick	196.21	116.07	1.25	0.75	0.23	0.50	0.64	
Purdue University	195.00	119.29	1.25	0.79	0.31	0.55	0.68	
Ohio State University	194.21	114.64	1.35	0.81	0.32	0.65	0.57	
University of Chicago	193.07	149.64	1.64	1.29	0.58	0.86	0.72	
University of Texas-Austin	187.29	137.71	1.31	0.99	0.23	0.71	0.77	
University of Wisconsin-Madison	186.64	115.64	1.37	0.86	0.40	0.69	0.60	
University of Florida	185.00	115.29	1.31	0.85	0.22	0.56	0.69	
Duke University	181.29	142.29	1.41	1.12	0.43	0.76	0.81	
University of California-Los Angeles (UCLA)	172.29	114.64	1.48	1.00	0.38	0.79	0.62	
Yale University	166.86	109.50	1.69	1.13	0.55	0.81	0.53	
University of Washington	164.43	98.71	1.31	0.81	0.22	0.61	0.64	
University of Southern California	158.07	109.71	1.31	0.93	0.29	0.72	0.71	
University of Georgia	154.57	88.29	1.26	0.72	0.25	0.50	0.61	
University of Minnesota	149.07	93.86	1.45	0.92	0.32	0.67	0.62	
University of North Carolina-Chapel-Hill	145.86	99.57	1.34	0.93	0.32	0.66	0.71	
Georgia Institute of Technology	144.64	108.79	1.12	0.85	0.15	0.53	0.81	
Georgia State University	137.57	87.86	1.43	0.92	0.26	0.54	0.70	
George Mason University	132.36	60.64	1.34	0.63	0.41	0.46	0.48	
North Carolina State University	129.50	67.36	1.23	0.65	0.34	0.69	0.45	
City University of New York (CUNY)	128.79	63.21	1.28	0.65	0.17	0.48	0.52	
Iowa State University	128.43	77.36	1.17	0.72	0.49	0.61	0.66	
Princeton University	127.86	87.29	1.49	1.05	0.56	0.85	0.53	
Carnegie Mellon University	126.57	104.36	1.08	0.90	0.33	0.82	0.85	
University of California-Davis	126.21	86.36	1.42	0.97	0.56	0.67	0.70	
Florida State University	122.93	74.07	1.34	0.80	0.20	0.56	0.61	
University of Arizona	116.14	77.86	1.37	0.92	0.23	0.68	0.68	
George Washington University	115.57	53.07	1.32	0.61	0.29	0.43	0.49	
University of Connecticut	113.71	75.64	1.32	0.91	0.31	0.51	0.77	
Boston University	105.07	69.29	1.35	0.92	0.37	0.81	0.63	
University of Central Florida	100.93	59.79	1.20	0.74	0.18	0.57	0.63	
University of South Carolina	100.43	63.64	1.30	0.86	0.14	0.48	0.70	
University of California-Irvine	98.07	61.79	1.43	0.92	0.31	0.69	0.63	
Auburn University	97.14	40.43	1.26	0.54	0.22	0.38	0.44	
University of Virginia	97.14	64.36	1.26	0.83	0.31	0.72	0.64	
Temple University	94.50	60.43	1.49	0.97	0.10	0.25	0.69	Yes
University of California-San Diego (UCSD)	93.29	62.07	1.67	1.13	0.58	0.84	0.45	
University of Pittsburgh	92.57	63.21	1.19	0.80	0.27	0.66	0.69	
Syracuse University	92.29	52.86	1.34	0.77	0.32	0.58	0.58	
University of Colorado at Boulder	90.14	52.29	1.45	0.86	0.28	0.71	0.56	
University of Alabama-Tuscaloosa	90.00	45.36	1.27	0.66	0.21	0.48	0.54	
University of Texas-Dallas	89.43	78.36	1.52	1.35	0.23	0.52	0.95	
University of Houston	89.14	56.36	1.31	0.84	0.19	0.73	0.63	
University of Missouri	88.50	45.21	1.32	0.67	0.32	0.49	0.54	
Johns Hopkins University	86.71	45.07	1.31	0.70	0.41	0.70	0.44	

Notes: This table reports the outcomes averages for each UK university for: 1) publications, (2) publications in top journals, (3) publications per author, (4) publications in top journals per author, (5) proportion of publications in Economics, (6) proportion of Economics publications top journals, (7) proportion of publications in Business top journals, as described in Table 2. Universities are listed in decreasing order according to the number of publications. The last column reports the universities that are selected as control units in the SCM.

Table A4: US universities: outcomes averages (2001-2014)

Universities of US (table continued)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	Selected
Louisiana State University	85.79	41.86	1.32	0.66	0.23	0.42	0.54	
Boston College	83.71	63.36	1.47	1.13	0.26	0.78	0.77	
Clemson University	83.64	48.00	1.27	0.77	0.21	0.57	0.63	
University of Iowa	81.93	56.93	1.30	0.92	0.29	0.74	0.71	
University of Illinois at Chicago	80.29	39.64	1.20	0.60	0.27	0.46	0.54	
University of Tennessee-Knoxville	79.79	33.79	1.38	0.60	0.27	0.44	0.45	
University of Kentucky	79.50	41.29	1.20	0.67	0.30	0.53	0.56	
Colorado State University	79.29	32.43	1.29	0.54	0.29	0.44	0.42	
Georgetown University	78.14	49.71	1.34	0.86	0.37	0.71	0.61	
Vanderbilt University	75.86	43.86	1.45	0.85	0.43	0.65	0.54	
Emory University	75.64	54.71	1.35	0.98	0.28	0.73	0.74	
Washington University in St. Louis	74.43	56.64	1.43	1.11	0.35	0.76	0.78	
University of Massachusetts-Amherst	72.43	36.29	1.08	0.55	0.28	0.47	0.54	Yes
University of Oklahoma	72.43	46.86	1.18	0.78	0.14	0.61	0.68	
University of Miami	70.93	46.57	1.42	0.94	0.17	0.57	0.68	
State University of New York-Buffalo (SUNY)	69.00	46.21	1.24	0.86	0.15	0.49	0.75	
University of Notre Dame	66.21	45.00	1.29	0.89	0.24	0.67	0.69	
Drexel University	65.43	37.14	1.24	0.71	0.16	0.48	0.59	
Rice University	64.21	47.14	1.44	1.06	0.26	0.71	0.76	
University of Rochester	63.29	49.14	1.34	1.06	0.53	0.77	0.79	Yes
Dartmouth College	61.64	47.21	1.47	1.13	0.37	0.84	0.73	
Brigham Young University	61.07	39.29	1.04	0.68	0.26	0.53	0.69	
American University	60.64	28.71	1.42	0.68	0.26	0.44	0.50	
Southern Methodist University	60.57	39.07	1.42	0.94	0.41	0.67	0.68	
University of Delaware	60.00	27.64	1.22	0.59	0.29	0.50	0.48	Yes
Oklahoma State University	59.71	29.93	1.45	0.73	0.21	0.38	0.54	Yes
University of California-Santa Barbara (UCSB)	57.43	26.00	1.33	0.63	0.49	0.69	0.29	
University of Kansas	56.79	29.93	1.29	0.70	0.23	0.46	0.57	
Rensselaer Polytechnic Institute	55.29	41.00	1.40	1.04	0.13	0.60	0.79	
University of Hawaii-Manoa	54.71	20.64	1.36	0.54	0.19	0.43	0.41	
West Virginia University	54.36	19.14	1.31	0.50	0.33	0.48	0.33	
University of Oregon	53.36	29.43	1.36	0.78	0.40	0.65	0.52	
Florida Atlantic University	52.79	26.36	1.22	0.63	0.18	0.45	0.57	
Brown University	52.36	33.79	1.57	1.06	0.63	0.82	0.43	
University of California-Riverside	51.21	27.86	1.46	0.79	0.41	0.54	0.57	
Fordham University	50.71	24.21	1.55	0.68	0.25	0.27	0.49	
Virginia Commonwealth University	50.29	23.93	1.23	0.60	0.16	0.54	0.48	Yes
Case Western Reserve University	49.57	33.64	1.36	0.92	0.13	0.64	0.70	
State University of New York-Binghamton (SUNY)	45.57	29.86	1.37	0.89	0.39	0.46	0.77	
DePaul University	43.57	22.14	1.08	0.59	0.10	0.62	0.53	
State University of New York-Albany (SUNY)	42.43	23.36	1.23	0.67	0.23	0.67	0.51	
University of Wyoming	42.07	27.29	1.64	1.09	0.69	0.61	0.72	Yes
Utah State University	41.36	20.57	1.32	0.65	0.38	0.49	0.49	
University of Colorado at Denver	41.14	22.86	1.29	0.75	0.26	0.63	0.58	
University of North Carolina-Greensboro	41.14	17.71	1.25	0.57	0.16	0.44	0.50	Yes
California Institute of Technology	37.21	26.79	1.59	1.15	0.71	0.87	0.36	
Baylor University	36.64	22.71	1.22	0.76	0.15	0.37	0.68	Yes
College of William & Mary	35.64	21.57	1.30	0.81	0.37	0.55	0.69	
University of California-Santa Cruz (UCSC)	35.21	20.07	1.75	1.05	0.57	0.67	0.53	Yes
Santa Clara University	34.57	22.36	1.32	0.85	0.14	0.59	0.66	
Tulane University	33.93	19.71	1.45	0.88	0.26	0.53	0.66	
Tufts University	32.93	17.36	1.41	0.75	0.48	0.69	0.42	
Appalachian State University	31.36	9.71	1.13	0.39	0.35	0.48	0.29	Yes
University of Nevada-Reno	30.79	12.64	1.24	0.51	0.38	0.44	0.45	Yes
Stony Brook University - SUNY	29.36	12.86	1.56	0.73	0.36	0.69	0.35	
University of Maryland-Baltimore County	23.71	10.00	1.09	0.49	0.43	0.38	0.48	Yes
Brandeis University	21.36	12.64	1.38	0.85	0.68	0.47	0.54	Yes
Middlebury College	17.00	6.43	1.57	0.57	0.70	0.36	0.33	Yes
Claremont McKenna College	16.71	10.79	1.39	0.91	0.61	0.60	0.56	Yes
Williams College	12.50	6.93	1.33	0.74	0.68	0.65	0.25	Yes

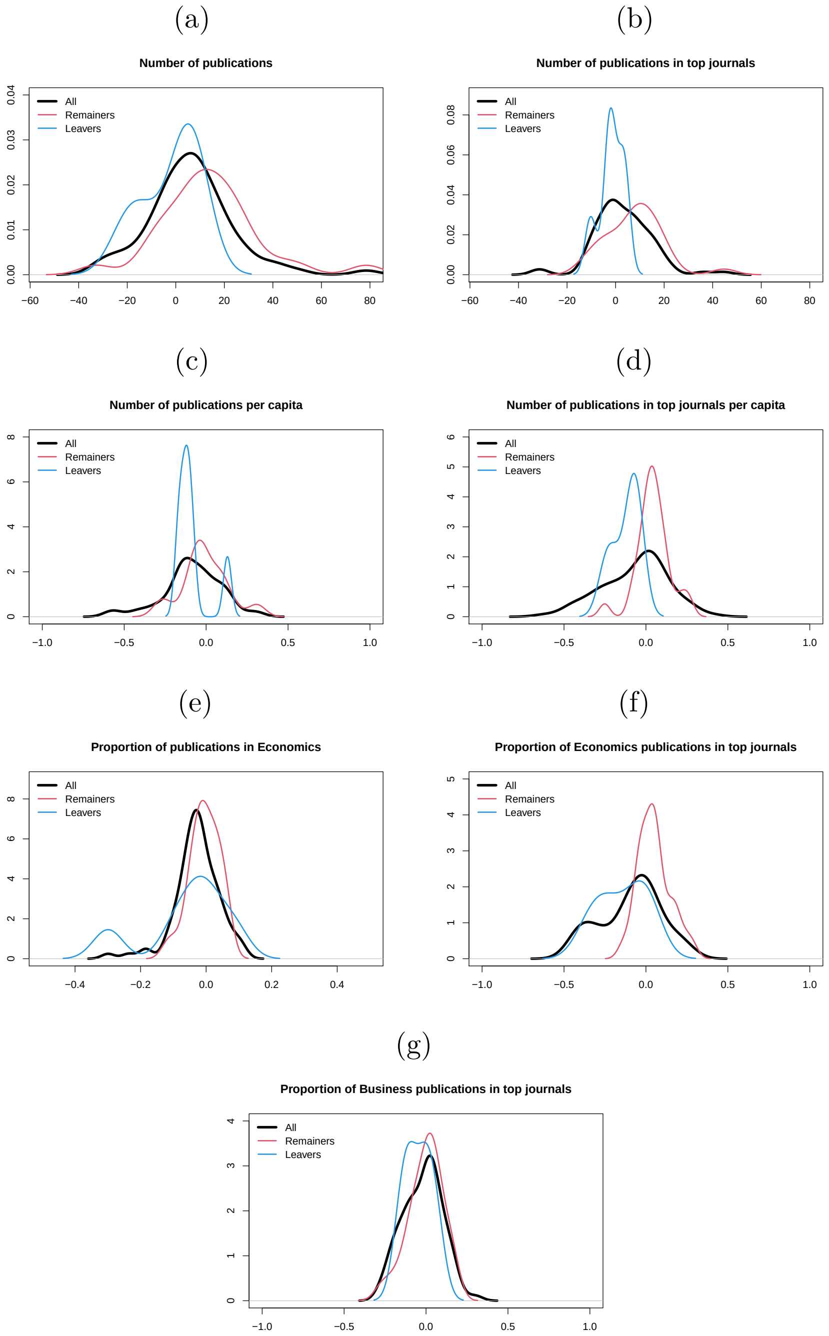
Notes: This table reports the outcomes averages for each US university for: 1) publications, (2) publications in top journals, (3) publications per author, (4) publications in top journals per author, (5) proportion of publications in Economics, (6) proportion of publications in Economics top journals, (7) proportion of publications in Business top journals, as described in Table 2. Universities are listed in a decreasing order according to the number of publications. The last column reports the universities that are selected as control units in the SCM.

Table A5: Cumulated treatment effects per university

University (anonymised)	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Russell 1	40.38	109.38*	1.56****	0.42	0.42	0.48	1.02****
Russell 2	-25.26	-21.18	-0.24	0.02	0.42	-0.06	0.90****
Russell 3	173.1*	118.44	0.72****	0.30	0.24	-0.12	0.18
Russell 4	-29.82	-32.10	-0.30	-0.02	0.02	-0.24	-0.06
Russell 5	47.88	30.72	-0.96****	0.06	-0.30	-0.24	0.06
Russell 6	-26.76	-64.74	1.74****	-1.50	0.06	-0.84	-0.72
Russell 7	44.34	-8.88	-0.48	0.66	-0.06	0.18	0.12
Russell 8	-2.10	4.80	-1.50****	0.66	0.06	0.30	0.90*
Russell 9	13.86	130.02*	0.48	1.62	-0.24	1.20****	0.54
Russell 10	-18.00	-36.30	-0.24	0.24	-0.54	0.30	-0.48
Russell 11	155.70****	53.05	0.42	0.06	-0.06	0.24	-0.48****
Russell 12	106.56	51.78	1.02	0.24	0.36	1.08	-0.78
Russell 13	97.14	84.30	0.60	0.60	-0.18	-0.78	-0.42
Russell 14	61.62	33.72	0.30	0.30	0.12	0.18	0.30
Russell 15	48.06	-10.44	-0.84	-0.54	-0.12	-0.30	0.18
Russell 16	93.72	54.30	-0.60	-0.12	0.24	1.68****	-1.62
Russell 17	-3.84	53.70	0.78	0.18	-0.12	-0.66	0.42
Russell 18	-17.52	37.26	-0.06	0.54	-0.18	-0.78	0.24
Russell 19	146.82*	81.30	-0.06	0.30	0.06	-0.24	0.02
Russell 20	93.24	42.72	0.60	0.54	-0.24	0.42	0.24
Russell 21	300.84****	268.08****	-0.30	-0.30	0.30	0.36	-0.48
Russell 22	58.44	-77.04	-0.54	-0.12	-0.12	0.72	-0.12
Russell 23	-3.90	48.12	0.18	0.90	-0.18	-0.12	0.66****
Russell 24	470.58****	88.62	-0.30	0.12	0.18	1.08****	-0.48
Total Russell	59.67*	39.07*	0.26	0.58	0.17	0.12	0.86****
Non-Russell 25	93.96	48.00	0.66	1.14****	-0.06	-0.36	-0.12
Non-Russell 26	56.40	18.18	-1.14	-1.02	0.60*	-0.66	-0.06
Non-Russell 27	141.66	69.72	0.06	0.48	-0.06	0.18	0.06
Non-Russell-Leaver 28	6.42	-18.12	0.78****	-1.50	-0.12	0.12	-0.66
Non-Russell 29	-122.04	-186.78*	-1.98****	-1.38****	-0.12	-1.62****	-0.06
Non-Russell-Leaver 30	43.38	-5.40	-0.90	-0.48	-1.80****	-0.42	-0.48
Non-Russell 31	115.68	-5.34	-0.78	-1.32	-0.42	-2.52****	-1.44****
Non-Russell 32	102.42	6.90	-0.42	-1.14	-0.24	-0.06	-0.72****
Non-Russell 33	-51.66	-37.98	-3.42****	-3.72****	-0.36	-2.52****	-0.54
Non-Russell 34	-165.72*	-43.02	-2.34****	-2.64****	-0.24	-1.32****	-1.20****
Non-Russell 35	51.18	-3.48	-1.26	-1.02	-0.12	-1.74****	-0.78
Non-Russell 36	-15.90	24.84	-0.78	-1.86	-0.42	-1.62****	-0.06
Non-Russell 37	-152.88*	-7.74	-2.52****	0.06	-0.12	-2.40****	1.20****
Non-Russell-Leaver 38	-123.90****	-61.62	-0.72****	-0.30	-0.24	-1.14*	0.42
Non-Russell 39	-76.02	-40.56	-1.74	-1.50	-0.36	-0.84	-0.90*
Non-Russell 40	-182.28*	-189.42****	1.44****	1.68****	-0.30	1.20	0.84
Non-Russell 41	177.72***	136.38***	-0.12	0.30	-0.30	0.48	0.36
Non-Russell 42	-5.16	-24.18	-0.84	-1.38	0.24	-0.48	-1.14**
Non-Russell 43	168.66*	86.28	-0.78	-2.10*	0.12	1.56****	-0.24
Non-Russell 44	-58.02	-49.80	-0.96****	-0.30	0.06	0.12	0.78***
Non-Russell 45	-26.04	47.76	-1.14****	0.24	-1.08*	-0.36	0.24
Non-Russell 46	13.50	-3.00	0.84****	1.56	-0.18	-0.60	0.78
Non-Russell 47	31.86	35.10	-2.10*	-0.90	-0.24	-1.80****	0.18
Non-Russell 48	-57.54	-70.56	0.54	-0.42	-0.30	-0.06	-1.26****
Non-Russell 49	33.72	-8.40	-1.02	-1.80****	-0.54	-2.64****	-0.66****
Non-Russell 50	116.52	95.82	-0.48	0.60	-0.18	0.06	-0.06
Non-Russell 51	-5.94	-36.60	-2.70****	-1.98****	-0.72	-2.10****	-1.32****
Non-Russell 52	10.86	-49.14	-0.78	-1.86	-0.48	-2.52****	-1.62****
Non-Russell 53	-88.68	-26.16	-1.08	-2.58****	0.42	-2.52****	0.24
Non-Russell 54	65.64	34.98	-0.90****	-0.30	-0.18	-1.80****	-0.42
Non-Russell 55	-20.22	40.74	-0.90****	-0.84	-0.42	-0.90	0.36
Non-Russell 56	33.90	43.14	-0.24	-0.06	-0.30	-0.06	0.30
Non-Russell 57	39.36	30.30	-0.01	0.18	0.12	-0.48	0.54
Non-Russell 58	-125.82****	-5.70	-3.30****	-2.52****	-0.06	-2.28****	0.00
Non-Russell 59	-197.76*	90.30	-0.54	0.78	0.18	1.08	0.72
Non-Russell 60	21.60	3.42	-1.50	-0.12	-0.42	-2.16****	-0.48
Non-Russell-Leaver 61	-77.52	-15.36	-0.66	-0.60	-0.60*	-2.10****	-0.96****
Non-Russell 62	112.32	105.96*	0.36	1.38	-0.12	0.78	0.42
Non-Russell 63	213.48*	215.82****	0.12	0.90	0.24	0.72	0.90****
Non-Russell 64	65.64	63.66	-0.18	1.14****	0.12	-0.12	1.02****
Non-Russell 65	98.46	4.92	-0.72	-0.72	-0.06	-0.12	-1.08****
Non-Russell-Leaver 66	24.60	17.94	-1.02	-0.30	0.54	-0.06	0.06
Non-Russell 67	-63.78	-9.30	-1.02****	-0.18	0.66*	-1.80****	0.30
Non-Russell 68	240.66*	66.66	1.86****	1.62	-0.72*	1.38****	0.54
Non-Russell 69	123.18****	103.08*	0.84	2.40	-0.42	0.12	1.80****
Non-Russell 70	250.38****	86.64	-0.90	-1.02	-0.30	-1.02	0.12
Non-Russell 71	-40.02	-58.08	0.00	-1.74	-0.72	-2.28****	-1.20****
Non-Russell 72	54.18	84.42****	0.90	0.18	-0.36	-0.60	0.48
Non-Russell 73	93.48	-59.52	0.18	0.06	0.60	-0.24	0.06
Non-Russell 74	-9.24	-25.56	0.42	-0.60	-0.60*	-0.72	-1.14****
Non-Russell 75	-67.92	-59.46	-0.78	-1.86*	-1.08*	-2.94****	-0.84
Non-Russell 76	50.94	-8.76	-3.54****	-2.40****	-0.30	-0.66	-0.66
Non-Russell 77	9.84	-48.96	-0.60	-2.82****	-0.48	-2.40****	-0.90*
Non-Russell 78	70.50	10.86	0.06	-0.84	-1.44*	-0.42	-0.36
Non-Russell 79	75.30	-19.68	-0.24	-2.94****	-0.36	-1.14*	-1.14****
Non-Russell 80	45.54	7.38	1.80*	0.78	-0.06	0.30	-0.36
Non-Russell-Leaver 81	52.62	24.42	-0.48	-1.20	0.12	-1.62****	-0.06
Total Non-Russell	35.98*	5.78	-0.052	-0.231	-0.167	0.581****	0.505****

Notes: This table provides the cumulated treatment effects per university (separating the universities by Russell and Non-Russell and Remainers and Leavers) for the: (1) number of publications, (2) number of publications in top journals, (3) number of publications per author, (4) number of publications in top journals per author, (5) proportion of publications in Economics, (6) proportion of publications in Economics top journals, (7) proportion of publications in Business top journals, as described in Table 2. Values are marked by *, **, ***, **** if they are significant at a level of, 0.10, 0.05, 0.01 or 0.001, respectively.

Figure A1: Distribution of the yearly treatment effects: All, Remainers and Leavers



Notes: This figure reports the distribution of the yearly treatment effects for all, Remainers and Leavers groups of universities for: number of publications (panel a), number of publications in top journals (panel b), number of publications per author (panel c), number of publications in top journals per author (panel d), proportion of publications in Economics (panel e), proportion of publications in Economics top journals (panel f), proportion of publications in Business top journals (panel g), as described in Table 2.

Appendix 2: Further Details on the Synthetic Control Method

2.1. Single treated unit: treatment effects

The SCM creates first the artificial matching unit, for each treated unit for each outcome measure, making use of the pre-treatment information of the treated unit and the set of available untreated units, or the so-called control pool. The matching algorithm follows an iterative two-step optimisation process:

(i) The **inner optimization** estimates the weights that minimize the distance between treated and untreated units' covariates over the pre-treatment period

$$\mathbf{w}(\mathbf{V}) = \arg_{\mathbf{w}} \min ||\mathbf{X}_1 - \mathbf{X}_0\mathbf{w}||_{\mathbf{V}} = \arg_{\mathbf{w}} \min \sqrt{(\mathbf{X}_1 - \mathbf{X}_0\mathbf{w})'\mathbf{V}(\mathbf{X}_1 - \mathbf{X}_0\mathbf{w})} \quad (1)$$

where $\mathbf{X}_1$ is the 13×1-matrix containing the values of the covariates over the pre-treatment period for the treated unit; $\mathbf{X}_0$ is the 13×121-matrix containing the values for the untreated units; $\mathbf{w}$ is the 121×1-vector of optimal weights to create a convex combination of untreated units. $\mathbf{V}$ is a positive-definite and diagonal 13×13-matrix, which assigns weights to the variables used in the optimization process. This matrix is initialized at the identity matrix at the beginning of the iterative algorithm.

(ii) The **outer optimization** serves to improve the result by estimating $\mathbf{V}$. Specifically, $\mathbf{V}$ is chosen such that the solution to the $||\mathbf{X}_1 - \mathbf{X}_0\mathbf{w}||_{\mathbf{V}}$ optimization problem minimizes the (pre-intervention) mean square prediction error (MSPE) (Abadie et al. 2010) for the outcome measure over the pre-treatment period,

$$\text{MSPE} = \frac{1}{8}\sum_{t=2001}^{2008}(Y_t - Y_t^*)^2 \quad \text{where} \quad Y_t^* = w(\mathbf{V})'Y_{jt} \quad (2)$$

where Y_{jt} is the 121×1-vector containing the values of the outcome variable for the 121 US universities at time t.

Steps **(i)** and **(ii)** are repeated iteratively until convergence. We use the R packages *Synth* and *improveSynth* to perform the analysis. As iterative algorithm, R uses both Nelder-Mead and BFGS methods and then chooses the most performing one. The weights determine the artificial control unit. The estimated coefficients, $\mathbf{w}$, for each UK university, for one outcome measure, the number of publications, are reported in Table A6.

Then, the SCM computes the difference, $\hat{\alpha}_t$, between the actual values of the outcome measure with those Y^* of the artificial university during the intervention years t (2009-2014). The cumulated Treatment Effect (cTE) for each UK university

i is calculated as the sum of the yearly treatment effects across all treatment years:

$$\text{c}\hat{\text{TE}}_i = \sum_{t=2009}^{2014} \hat{\alpha}_{it} \qquad \text{where} \quad \hat{\alpha}_{it} = Y_{it} - Y_{it}^{*} \tag{3}$$

2.2. Assessing the significance of the treatment effects

As mentioned above, the SCM does not generate standard p values that can be used to test the significance of the treatment effects. To overcome this issue, Abadie et al. (2010) proposes to run the so-called placebo tests. Placebo analysis involves performing SCM for each unit in the controls' pool as if they were treated, using the rest of controls as their pool. This process generates a *distribution of placebo effect estimates*. The placebo tests yield null distributions (i.e., distributions under the null hypothesis of no effect due to intervention) for both the yearly differences as well as for the cumulated treatment effect, against which one can compare the original effect estimates. In fact, by being non-parametric, the placebo test approach has the advantage of not imposing any distribution on the errors. If the intervention is the cause of the observed effect, then the gap between the treated units and its synthetic control should be largest for the actual treated unit than for the placebo units (Bouttell et al. 2018). Otherwise, it is reasonable to think that the estimated effects are observed by chance.

The idea of the placebo tests proposed by Abadie et al. (2010) is akin to the classic framework for permutation inference. As in permutation tests, we apply the SCM to every potential control in our sample. This will assess whether the effects estimated by the SCM for the universities affected by the REF 2014 are large relative to the effect estimated for a control university chosen at random. This inferential exercise is exact in the sense that, regardless of the number of available comparison control universities, time periods, and whether the data are individual or aggregate, it is always possible to calculate the exact distribution of the estimated effect of the placebo interventions.

We consider each of the control US universities in the control set as if they were treated. Thus, we apply the SCM to create the best synthetic counterfactual for each US university in the control set using a combination of the remaining universities in the control set. This yields a group of yearly placebo treatment effects, α_{jt}^{PL}, as well as a cumulative placebo treatment effect cTE_j^{PL} for each university j in this potential control set. Following Abadie et al. (2010), we drop the yearly effects and the cumulative treatment effects of the US universities that have a pre-treatment Mean Squared Predictive Error (MPSR) greater or equal to twice that of the treated unit so that we only retain the N_{PL} placebos that are comparable to the treated unit.

Having a distribution of placebos allows us to conduct a two-sided hypothesis. If the REF 2014 did not have any effect, we would expect the effects on the treated UK universities to be similar to the ones computed for the untreated US universities. The p-values for a treated unit i associated to the yearly treatment effect t can be calculated as:

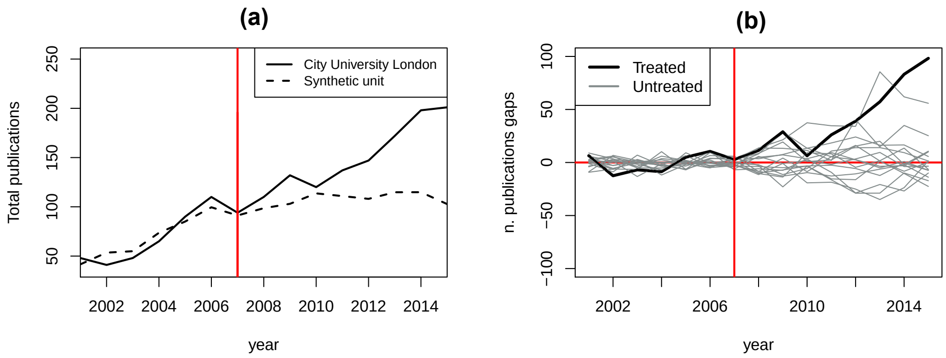
$$p_{it} = \frac{1}{N_{PL}} \sum_{j} \mathbb{1}\{|\alpha_{jt}^{PL}| \geq |\hat{\alpha}_{it}|\} \text{ for each } t, \tag{4}$$

where $\mathbb{1}\{\cdot\}$ is the indicator function that takes value one if the argument in parentheses is true and zero otherwise, and N_{PL} is the number of universities in the potential control set. Thus, the p_{it} is the proportion of universities in the placebo group for which the treatment for that year is larger than that of i. Similarly, the corresponding statistic when taking into account the aggregated effect for all years for unit i, cTE_i, is:

$$p_{cTE_i} = \frac{1}{N_{PL}} \sum_{j} \mathbb{1}\{|cTE_j^{PL}| \geq |c\hat{T}E_i|\}. \tag{5}$$

Panel (a) in Figure A2 shows, as an example, the evolution of the number of publications for City, University of London, and that of its artificial counterfactual university. In panel b, we show the treatment effects $\hat{\alpha}_{it}$ of City (bold line) and those of the control units for placebo tests.

Figure A2: Example: Number of Publications of City, University of London, and its synthetic control unit



2.3. Multiple treated units: average treatment effects and significance

Acemoglu et al. (2016) calculates the Average Treatment Effect on the Treated (ATT) for a group of treated units by computing a fit-weighted ATT. The fit-weighted ATT is computed as

$$\hat{\text{ATT}} = \frac{\sum_{i \in Treat} \frac{c\hat{TE}_i}{\hat{\sigma}_i}}{\sum_{i \in Treat} \frac{1}{\hat{\sigma}_i}}; \quad \text{where } \hat{\sigma}_i = \sqrt{\frac{\sum_{t=2001}^{t=2008} \hat{\alpha}_{it}^2}{8}} \tag{6}$$

where $\hat{\sigma}_i$ is the Root of the Mean Squared Predictive Error of the estimated effects over the pre-treatment period (RMSPE) for treated unit i, that is, the discrepancy between the actual and counterfactual patterns before the intervention. The $\hat{\text{ATT}}$ describes a weighted average of the cumulative effects over the intervention period, using the inverse of the RMSPE, $\frac{1}{\hat{\sigma}_i}$, as weights. This implies that universities with a better matching have a higher impact on the estimate of the ATT which provides an unbiased estimate of ATT.

To create our (pseudo) placebo tests, we follow the generalization proposed by Cavallo et al. (2013) of Abadie et al. (2010)'s placebo approach to do inferences about the average effect estimated across multiple treated units. A null distribution of placebo ATT effects is again needed. Following Acemoglu et al. (2016), we create 5,000 placebo treatment groups of the same size as the number of treated units, in our case the 81 UK universities, which are extracted with replacement through bootstrap from the set of control units (even if the control group is of smaller size than 81). We again index all these placebo ATTs over j. The p-values for the overall ATT are given by:

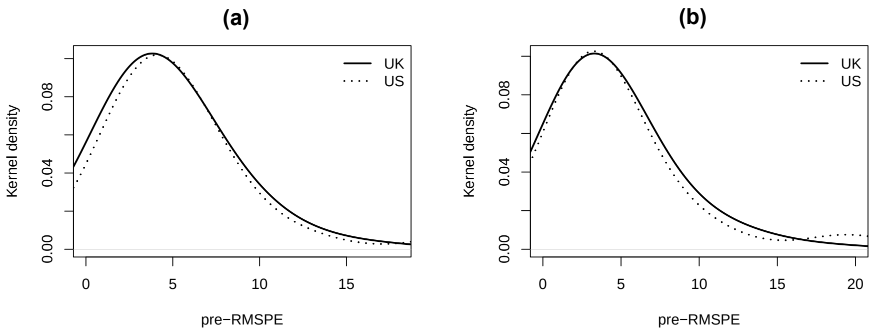
$$\hat{p}_{ATT} = \frac{1}{5000} \sum_{j} \mathbb{1}\{|ATT_j^{PL}| \geq |\hat{ATT}|\}. \quad (7)$$

We base our placebo tests on the $\hat{p}_{ATT}$ above and consider that, if it is less than 5%, the average effect of the REF 2014 is significant, i.e., not random or an artefact of the method.

2.4. Assessing the goodness of fit of the matching

There are several ways to check if a weighted average of controls is able to approximate the outcome of the treated unit in the pre-treatment period. One could visually compare the difference in pre-treatment outcomes between the treatment unit and its synthetic control. Or, one could look at the distributions of the root of the mean square predictive error over the pre-treatment (RMSPE). Nevertheless, although there is currently no consensus on what constitutes a 'good fit' or how to judge the similarity between treated and control units (Bouttell et al. 2018), most of the works making use of the SCM compare the *distance* between the treatment and synthetic control unit/s during the pre-treatment period, i.e., the RMSPE, for the treated and for the placebo control (US) units. Figure A3 below shows the distributions of the pre-treatment RMSPEs for the treated UK universities and for the placebo US universities for the first two outcome measures of interest, the number of publications in journals and of the number of publications in top journals. As can be appreciated, the distributions of UK universities and US controls match well in both cases.

Figure A3: Kernel estimate distributions of pre-treatment RMSPE for control (US) and UK universities



Notes: It is used for the assessment of the quality of SCM matching for two of the outcomes: number of publications in journals (Panel a) and top journals (Panel b).

Following the placebo approach of Cavallo et al. (2013), we look at the proportion of placebos that have pre-treatment RMSPEs at least as large as the average

RMSPEs of the treated units. But, instead of using the average of the pre-treatment treated unit RMSPE distribution as a reference for comparison, we consider the *median*, which is more robust to potential outliers or, in our case, UK universities that show a poor matching with their counterfactuals. These universities would be the ones discarded for the computation of the corrected ATTs, following the approach used by Acemoglu et al. (2016). We call the proportion of placebo p_{match}:

$$p_{match} = \frac{1}{N_{PL}} \sum_{j=1}^{N_{PL}} \mathbb{1}\{\text{RMSPE}_j^{PL} \geq \text{RMSPE}_{Median}^{UK}\}. \tag{8}$$

Note that p_{match} gives the proportion of the considered placebos that have RMSPEs above the median of the associated to the treated. If placebo RMSPEs are very frequently smaller than those of the treated, then the control group is not able to properly replicate the patterns of the treated units. Thus, the larger the p_{match}, the better the quality of the matching. However, the control units are somehow similar in that we should not expect their RMSPEs to be too high. Therefore, if the control group can reasonably reproduce the treated units, we expect their pre-treatment RMSPE distributions to be close to one another but not too often be the placebo ones smaller than the treated. Thus, we consider that a p_{match} of about 50% indicates an acceptable match.

For all the outcomes, the p_{match} values are reported in the first column of Table 3. For all outcomes, p_{match} is about 50% or higher, and, thus, we consider the quality of the match acceptable. Since the matching is done for each outcome variable, as explained in section 6, the set of matching coefficients is allowed to be different. As an example, Table A6 shows matching coefficients for the outcome 'Number of Publications'.

Table A6: SCM estimated coefficients: Number of Publications

Treated	Synthetic control composition
Aberystwyth University	Brandeis University (0.452), Claremont McKenna College (0.432), University of Maryland-Baltimore (0.104), Baylor University (0.013)
Aston University	Florida Atlantic University (0.791), University of Georgia (0.209)
Bangor University	Middlebury College (0.330), Claremont McKenna College (0.294), University of Maryland-Baltimore (0.231), Williams College (0.144)
Birkbeck College	Middlebury College (0.393), West Virginia University (0.269), Syracuse University (0.199), University of Rochester (0.056), University of Massachusetts-Amherst (0.046), Brandeis University (0.036)
Bournemouth University	Williams College (0.430), University of Maryland-Baltimore (0.240), Middlebury College (0.225), Appalachian State University (0.105)
Brunel University London	University of Texas-Dallas (0.777), Arizona State University (0.088), State University of New York-Buffalo (0.061), Purdue University (0.044), University of North Carolina-Greensboro (0.023), Florida Atlantic University (0.007)
Cardiff University	University of Chicago (0.274), Washington University in St. Louis (0.216), Vanderbilt University (0.197), New York University (0.158), Oklahoma State University (0.093), University of Illinois at Urbana-Champaign (0.038), University of Texas-Dallas (0.025)
City University London	University of Delaware (0.381), Florida Atlantic University (0.292), University of Georgia (0.254), Northwestern University (0.073)
Coventry University	Claremont McKenna College (0.423), Middlebury College (0.198), Fordham University (0.197), Appalachian State University (0.148), University of Maryland-Baltimore (0.024), Brandeis University (0.010)
Cranfield University	Florida Atlantic University (0.652), University of Georgia (0.288), Texas A&M University (0.040), University of Arizona (0.021)
De Montfort University	University of Nevada-Reno (0.423), Claremont McKenna College (0.352), University of North Carolina-Greensboro (0.133), College of William & Mary (0.091)
Edinburgh Napier University	Appalachian State University (0.631), University of Maryland-Baltimore (0.139), Claremont McKenna College (0.118), Baylor University (0.112)
Glasgow Caledonian University	Appalachian State University (0.562), Florida Atlantic University (0.293), Baylor University (0.116), West Virginia University (0.029)
Heriot-Watt University	Baylor University (0.410), West Virginia University (0.334), Florida Atlantic University (0.103), University of North Carolina-Greensboro (0.092), University of Alabama-Tuscaloosa (0.060)
Imperial College London	University of California-Santa Barbara (0.278), University of North Carolina-Greensboro (0.247), Stanford University (0.187), Georgia State University (0.180), University of California-Davis (0.068), University of California-Los Angeles (0.039)
Keele University	Brandeis University (0.548), University of Maryland-Baltimore (0.345), Fordham University (0.106)
King's College London	Baylor University (0.304), University of North Carolina-Greensboro (0.304), Syracuse University (0.193), Williams College (0.100), Temple University (0.070), Harvard University (0.017), Florida Atlantic University (0.011)
Kingston University	Appalachian State University (0.271), Williams College (0.270), University of Maryland-Baltimore (0.250), Florida Atlantic University (0.105), Baylor University (0.057), Oklahoma State University (0.047)
Lancaster University	University of Texas-Dallas (0.626), University of Georgia (0.274), Texas A&M University (0.066), Florida Atlantic University (0.034)
Leeds Beckett University	Brandeis University (0.558), University of Maryland-Baltimore (0.242), Baylor University (0.106), Claremont McKenna College (0.095)
London Business School	State University of New York-Buffalo (0.755), Boston College (0.152), Harvard University (0.087), University of Oklahoma (0.006)
London Metropolitan University	University of Nevada-Reno (0.590), Middlebury College (0.204), Baylor University (0.148), Brandeis University (0.053)
LSE	Harvard University (0.336), University of Georgia (0.296), University of Connecticut (0.213), MIT (0.150)
London South Bank University	Williams College (0.504), Middlebury College (0.201), Fordham University (0.198), Brandeis University (0.091), Claremont McKenna College (0.007)
Manchester Metropolitan University	Middlebury College (0.264), Boston College (0.233), Baylor University (0.170), University of North Carolina-Greensboro (0.153), Florida Atlantic University (0.100), Syracuse University (0.078)
Middlesex University	Florida Atlantic University (0.275), University of Maryland-Baltimore (0.243), University of Nevada-Reno (0.168), Claremont McKenna College (0.135), University of Alabama-Tuscaloosa (0.131), Baylor University (0.048)
Newcastle University	University of Nevada-Reno (0.499), Baylor University (0.150), Tufts University (0.085), Princeton University (0.064), Rutgers University-New Brunswick (0.064), Auburn University (0.061), Syracuse University (0.040), Harvard University (0.037)
Nottingham Trent University	University of Colorado at Denver (0.345), Claremont McKenna College (0.253), Middlebury College (0.235), University of North Carolina-Greensboro (0.089), Williams College (0.052), University of Maryland-Baltimore (0.025)
Open University	Brandeis University (0.322), University of California-Riverside (0.248), University of Oklahoma (0.184), Baylor University (0.113), University of Iowa (0.098), University of Maryland-Baltimore (0.035)
Oxford Brookes University	Middlebury College (0.483), Baylor University (0.180), Florida Atlantic University (0.144), Oklahoma State University (0.109), University of Maryland-Baltimore (0.084)

Treated (Table continued)	Synthetic control composition
Queen Mary University of London	University of Maryland-Baltimore (0.293), Florida Atlantic University (0.282), University of Tennessee-Knoxville (0.252), University of North Carolina-Greensboro (0.089), University of Alabama-Tuscaloosa (0.070), University of Georgia (0.013)
Queen's University Belfast	University of North Carolina-Greensboro (0.588), University of California-Santa Barbara (0.227), University of Maryland-Baltimore (0.095), University of Texas-Dallas (0.043), Purdue University (0.036), Florida Atlantic University (0.010)
Robert Gordon University	Middlebury College (0.516), Claremont McKenna College (0.433), University of North Carolina-Greensboro (0.033), Williams College (0.018)
Royal Holloway, University of London	University of California-Santa Cruz (0.533), Florida Atlantic University (0.215), University of California-Santa Barbara (0.137), City University of New York (0.078), University of Texas-Dallas (0.029), Georgia State University (0.009)
Sheffield Hallam University	Brandeis University (0.437), University of Maryland-Baltimore (0.301), Baylor University (0.177), West Virginia University (0.085)
Staffordshire University	Claremont McKenna College (0.786), Williams College (0.214)
Swansea University	Fordham University (0.622), Appalachian State University (0.205), University of Texas-Dallas (0.102), University of Maryland-Baltimore (0.041), West Virginia University (0.030)
University College London	University of Chicago (0.315), Rice University (0.273), City University of New York (0.273), Fordham University (0.093), University of California-Santa Barbara (0.046)
University of Aberdeen	Brigham Young University (0.350), State University of New York-Albany (0.258), Stony Brook University (0.140), Washington University in St. Louis (0.091), Fordham University (0.070), University of Iowa (0.060), University of Minnesota (0.030)
University of Bath	Florida Atlantic University (0.488), University of Georgia (0.284), University of Alabama-Tuscaloosa (0.210), University of Michigan (0.010), West Virginia University (0.008)
University of Bedfordshire	Claremont McKenna College (0.702), Middlebury College (0.298)
University of Birmingham	Stanford Uni (0.237), Rensselaer Polytechnic Institute (0.196), Uni of California-Santa Cruz (0.186), Uni of Colorado at Denver (0.180), Uni of Rochester (0.116), Georgia State Uni (0.050), Temple Uni (0.034)
University of Bradford	University of Tennessee-Knoxville (0.473), Fordham University (0.204), Baylor University (0.152), University of Maryland-Baltimore (0.086), Claremont McKenna College (0.085)
University of Brighton	Williams College (0.596), University of Maryland-Baltimore (0.388), Florida Atlantic University (0.016)
University of Bristol	West Virginia University (0.393), University of Delaware (0.231), Brandeis University (0.170), Iowa State University (0.123), Syracuse University (0.053), Boston College (0.031)
University of Cambridge	University of California-Santa Barbara (0.433), Harvard University (0.261), Rensselaer Polytechnic Institute (0.150), MIT (0.069), Georgia State University (0.045), University of California-Los Angeles (0.042)
University of Central Lancashire	Claremont McKenna College (0.521), University of Maryland-Baltimore (0.433), Appalachian State University (0.045)
University of Dundee	West Virginia University (0.493), Middlebury College (0.353), University of Maryland-Baltimore (0.152)
University of Durham	University of Tennessee-Knoxville (0.533), University of Alabama-Tuscaloosa (0.237), Fordham University (0.117), Baylor University (0.059), University of Maryland-Baltimore (0.054)
University of East Anglia	Appalachian State University (0.337), Syracuse University (0.255), University of North Carolina-Greensboro (0.155), Oklahoma State University (0.132), University of Texas-Dallas (0.061), Iowa State University (0.046), University of Rochester (0.013)
University of East London	Claremont McKenna College (0.674), Middlebury College (0.315), University of Maryland-Baltimore (0.011)
University of Edinburgh	Claremont McKenna College (0.323), University of Texas-Dallas (0.294), Georgia State University (0.225), University of California-Santa Barbara (0.108), MIT (0.050)
University of Essex	University of Maryland-Baltimore (0.288), Georgia Institute of Technology (0.181), Oklahoma State University (0.164), University of Wyoming (0.158), University of Rochester (0.077), University of California-Santa Barbara (0.070), Iowa State University (0.056), University of Massachusetts-Amherst (0.006)
University of Exeter	University of Maryland-Baltimore (0.263), University of Delaware (0.250), University of California-Riverside (0.163), Arizona State University (0.157), Baylor University (0.115), University of Iowa (0.039), North Carolina State University (0.014)
University of Glasgow	University of Maryland-Baltimore (0.322), University of Massachusetts-Amherst (0.257), Iowa State University (0.169), University of North Carolina-Greensboro (0.100), George Washington University (0.086), Oklahoma State University (0.038), University of Tennessee-Knoxville (0.025)
University of Greenwich	Claremont McKenna College (0.747), University of North Carolina-Greensboro (0.126), Appalachian State University (0.069), University of Maryland-Baltimore (0.059)
University of Hertfordshire	Claremont McKenna College (0.737), University of Maryland-Baltimore (0.232), Baylor University (0.017), Fordham University (0.014)
University of Hull	University of Maryland-Baltimore (0.428), Oklahoma State University (0.375), Florida Atlantic University (0.193)
University of Kent	University of California-Santa Cruz (0.483), Florida Atlantic University (0.214), University of California-Santa Barbara (0.148), University of Texas-Dallas (0.134), University of Maryland-Baltimore (0.021)
University of Leeds	University of North Carolina-Greensboro (0.290), University of Georgia (0.264), University of Florida (0.226), University of Texas-Dallas (0.210), Arizona State University (0.010)

Treated (Table continued)	Synthetic control composition
University of Leicester	University of North Carolina-Greensboro (0.364), Williams College (0.225), Dartmouth College (0.195), Boston College (0.095), Harvard University (0.066), Rensselaer Polytechnic Institute (0.055)
University of Liverpool	University of California-Santa Cruz (0.405), College of William & Mary (0.290), University of Texas-Dallas (0.171), Claremont McKenna College (0.068), City University of New York (0.066)
University of Manchester	Pennsylvania State University (0.565), Texas A&M University (0.205), Purdue University (0.165), Northwestern University (0.064)
University of Northumbria at Newcastle	Middlebury College (0.691), Baylor University (0.174), Brandeis University (0.135)
University of Nottingham	Texas A&M University (0.739), Florida Atlantic University (0.124), Syracuse University (0.066), Columbia University (0.048), City University of New York (0.023)
University of Oxford	Harvard University (0.477), University of Chicago (0.231), Georgia State University (0.141), City University of New York (0.119), Stanford University (0.032)
University of Plymouth	University of Maryland-Baltimore (0.460), University of Nevada-Reno (0.377), Claremont McKenna College (0.074), University of Alabama-Tuscaloosa (0.061), Baylor University (0.028)
University of Portsmouth	University of North Carolina-Greensboro (0.586), Florida Atlantic University (0.193), University of California-Santa Barbara (0.110), Stony Brook University (0.100), University of Maryland-Baltimore (0.011)
University of Reading	Florida Atlantic University (0.298), Oklahoma State University (0.286), University of Florida (0.225), Syracuse University (0.190)
University of Salford	University of Maryland-Baltimore (0.552), Stony Brook University - SUNY (0.181), University of Chicago (0.102), Oklahoma State University (0.088), University of Texas-Dallas (0.077)
University of Sheffield	Oklahoma State University (0.547), University of Georgia (0.453)
University of Southampton	University of Maryland-Baltimore (0.625), Iowa State University (0.208), University of California-Berkeley (0.157), University of Illinois at Urbana-Champaign (0.011)
University of South Wales	Claremont McKenna College (0.289), University of Maryland-Baltimore (0.243), Stony Brook University (0.184), Brandeis University (0.177), Middlebury College (0.078), Fordham University (0.029)
University of St Andrews	University of Maryland-Baltimore (0.521), Colorado State University (0.135), Baylor University (0.125), California Institute of Technology (0.115), University of Texas-Dallas (0.074), Stony Brook University (0.030)
University of Stirling	University of California-Santa Barbara (0.391), University of North Carolina-Greensboro (0.361), Florida Atlantic University (0.229), City University of New York (0.009), University of Texas-Dallas (0.007)
University of Strathclyde	University of California-Santa Barbara (0.399), University of Virginia (0.357), University of California-Los Angeles (0.160), University of Minnesota (0.030), University of Pittsburgh (0.023), Stanford University (0.019), University of Illinois at Urbana-Champaign (0.011)
University of Sunderland	Claremont McKenna College (0.834), Middlebury College (0.166)
University of Surrey	Temple University (0.485), Syracuse University (0.245), West Virginia University (0.189), University of North Carolina-Greensboro (0.080)
University of Sussex	Fordham University (0.555), University of Texas-Dallas (0.259), University of Maryland-Baltimore (0.098), Iowa State University (0.049), Appalachian State University (0.026), University of Rochester (0.016)
University of the West of England, Bristol	Appalachian State University (0.311), University of Maryland-Baltimore (0.308), Florida Atlantic University (0.216), University of California-Santa Barbara (0.097), Oklahoma State University (0.069)
University of Ulster	Virginia Commonwealth University (0.434), Middlebury College (0.278), Boston College (0.144), Baylor University (0.054), Harvard University (0.033), Syracuse University (0.032), Florida Atlantic University (0.025)
University of Warwick	Pennsylvania State University (0.297), Yale University (0.257), Purdue University (0.231), University of Georgia (0.120), Florida State University (0.058), University of Chicago (0.037)
University of Westminster	University of Maryland-Baltimore (0.426), Middlebury College (0.248), Appalachian State University (0.183), University of North Carolina-Greensboro (0.105), Florida Atlantic University (0.038)
University of Wolverhampton	Middlebury College (0.535), Appalachian State University (0.304), University of Maryland-Baltimore (0.081), Williams College (0.080)
University of York	Dartmouth College (0.643), Princeton University (0.287), Boston College (0.059), University of North Carolina-Greensboro (0.011)

Notes:. This table provides the Synthetic control method (SCM) estimated coefficients (weights) for the Number of Publications in each UK university. The results for all other variables are available upon request.