

Research Performance and the Organizational Effectiveness of UK Universities

Richard Croucher
Middlesex University Business School, UK

Paul Gooderham*
NHH Bergen, Norway

Marian Rizov
Lincoln International Business School, UK

*Corresponding author:

Richard Croucher,
Middlesex University Business School
The Burroughs
London NW4 4BT
r.croucher@mdx.ac.uk
Tel: +44 -20-8411-5159.

Paul Gooderham
NHH
Helleveien 30
5045 Bergen
Norway
Paul.gooderham@nhh.no
Tel: +47 55 95 90 00

Marian Rizov
Lincoln International Business School
Brayford Pool
Lincoln LN6 7TS
mrizov@lincoln.ac.uk

Research Performance and the Organizational Effectiveness of UK Universities

<i>Purpose of this paper (limit 100 words)</i>	We test Shattock's legacy reputation thesis that non-leading universities in the UK face insuperable resource barriers to entering the leading group.
<i>Design/methodology/approach (limit 100 words)</i>	Employing regression analysis, we examine whether prioritizing research performance is a viable strategy for non-leading UK universities aiming to improve their organizational effectiveness. The dependent variable, organizational effectiveness, is measured by the annual Guardian rankings of universities. The main independent variable, research performance, is measured using "research power" ('RP'). RP is derived from the UK Research Excellence Framework (REF).
<i>Findings (limit 100 words)</i>	For 2008-14 we find that changes in research performance impacted university rankings. However, we also find that changes to the rankings are largely confined to non-leading universities and have not led to these institutions breaking into the group of leading universities. Therefore, Shattock's thesis is supported.
<i>Research implications (limit 100 words)</i>	Future research should use a longer time-span. It should investigate in more detail the mechanisms involved in the research performance-organizational effectiveness relationship.
<i>Implications for Practice (100 words)</i>	Failing to maintain research performance can have significant negative consequences for the rankings of non-leading universities.
<i>Originality/value</i>	This is the first study that examines the relationship between the research performance of universities in the UK with a measure of their overall organizational effectiveness.

Keywords: organizational effectiveness; research performance; UK universities

INTRODUCTION

We test Shattock's (2003, 2010) legacy reputation thesis that non-leading universities face insuperable resource barriers to entering the leading group. More specifically, we examine the role research performance as measured by what is currently known as the Research Excellence Framework (REF) plays in improving the organizational effectiveness of non-leading universities in the UK. Engaging in the REF is a time-consuming and expensive exercise. The question is whether it is a viable means for non-leading universities to improve their organizational effectiveness or whether their managers should use other means.

There is no commonly agreed basis for judging the organizational effectiveness of nonprofit organizations (NPOs) such as universities (Herman and Renz, 2008). Whereas a significant proportion of studies of organizational effectiveness of businesses have used exclusively financial criteria, studies of NPOs are more likely to use a range of nonfinancial criteria such as employee satisfaction, quality, and public image (Baruch and Ramalho, 2006). How best to define organizational effectiveness in the nonprofit sector has been subject to considerable debate (Mitchell, 2013). Part of the challenge is that while over and above the fact that NPOs cannot distribute earnings to anyone and that they receive part of their income from public sources, NPOs are disparate in terms of their activities (Herman and Renz, 2008). Although NPO organizational effectiveness remains a problematic construct, Herman and Renz (2008) suggest a number of "theses" or notions that should guide research. One is that measures should always be multidimensional; NPO effectiveness cannot be assessed with a single indicator. Another is that studies of organizational effectiveness should be comparative; organizations should be compared with similar organizations. Further, while they support and encourage the use of hard evidence to the extent it is

legitimately possible, a measure of NPO effectiveness is not significant “until someone forms judgments of effectiveness from them (and, usually, communicates those judgments) and acts on the judgments” (Herman and Renz, 2008: 404).

We operationalize the organizational effectiveness of universities by using the annual *Guardian University League Table*. As we clarify below, it is multidimensional and uses “hard evidence”. It enables universities to compare themselves with one another, and it is widely recognized by users of higher education. Substantial evidence exists that in countries where universities and other higher education institutions are ranked, they take a significant interest in how they perform (Morgeson and Nahrgang, 2008). Rankings are an important constituent of reputation and therefore have an impact on decision-making by potential students and their families, and by graduate employers (Argenti, 2000). Because they signal which universities are “leading”, rankings are important to consumers of higher education, especially prospective students and their families, who have in recent years had to bear increased costs and risks. They are also significant to the universities themselves as ways of publicly communicating their brands and reputations to other stakeholders. Internally, rankings can elicit powerful reactive responses within universities, stimulating extensive and intensive internal measures to improve their positions (Espeland and Sauder, 2007). Given the importance of rankings to universities, how far they can enhance them is clearly an important issue for all stakeholders and especially for managers.

Our article is structured as follows. We initially discuss what it means to be a leading university in terms of historical advantage. Thereafter, we provide an overview of the key broad university rankings, focusing on the components of the *Guardian* rankings. Next, we discuss the research performance assessment, the REF, showing how it has evolved, how it works and its results. We then address Shattock’s thesis about how the positions or rankings of leading institutions’ appear impregnable. Next, we introduce our research questions. The

aims of our empirical analysis are to assess the stability of these two measures over the period 2008-2014 and to determine whether changes in research performance are associated with changes in the *Guardian* rankings. Finally, we discuss our findings, before drawing conclusions and identifying practical implications for policy makers and managers.

LEADING AND NON-LEADING UNIVERSITIES

UK Universities have expanded throughout the 20th and 21st Centuries. By September 2015, there were 131 in the UK (all figures: HEFCE, 2015). One sizeable group of universities was created post 1992: official lists as of September 2015 show that these numbered 55; a smaller group of 26 had previously been created between 1945 and 1991. Writing recently in the *Financial Times* (2017), Mathew Engel observed that “(UK) universities are no longer judged by vague perceptions, they are ranked ... what was once a vague pecking order (has been transformed) into a set of almost Hindu gradations marking the layers between the Brahmins and untouchables.”

There are many categorizations of UK universities (see for example Hewitt-Dundas, 2012 and Hicks, 2012). They vary in the terms they use, but they seek to capture widely acknowledged organizational differences and statuses. One such is the distinction between “Russell Group” universities and others. Twenty-four research-intensive institutions have banded together to form the influential Russell Group to advance their interests, lobby government and project their collective identity (Hewitt-Dundas, 2012). Another, similar distinction is that advanced by Hicks (2012) that the UK contains two generic types of university, the ‘elite’ and the ‘non-elite’. We prefer the terms ‘leading’ and ‘non-leading’ as a categorization for our purpose. Yet whatever labels are used, real differences in their nature and status exist. The former raise relatively high proportions of their income from research; the latter derive the bulk of their income from teaching large numbers of students (Hicks,

2012). In 2004-2005, 65 per cent of all research grant and contract income in UK Universities was accounted for by the Russell Group Universities (Hewitt-Dundas, 2012). Fulton (2003: 165) argued that the division between different types of institution would persist and potentially even harden as leading universities build on existing advantages. Thus, research is for them a largely *indirect* way of achieving a better reputation, attracting more good-quality applicants and improving their financial positions.

Shattock (2003) argued that the historic turning point in inter-university differentiation occurred in the mid-1980s with the first REF (then known as the “Research Assessment Exercise”). By 2001 it was clear “that research excellence is stable amongst the top ranked institutions ... high quality research, once established, tends to reinforce itself” (Shattock, 2003:6-7). From 1994, national newspapers began to publish annual league tables that effectively sought to measure the organizational effectiveness of universities in terms of a range of outcomes of “concern to the wider public (as well as students)” (Shattock, 2003:7) such as teaching quality. For the period to 2001, Shattock (2003:15) observes a powerful correlation or “coherence” between research performance and organizational effectiveness as measured by the university league tables. Although universities can compete for both staff and students, Shattock argues that legacy reputation means that leading universities have a powerful, sustainable competitive advantage over non-leading universities in terms of attracting the best faculty and the best students. Similarly, Naidoo et.al (2011) argue that highly ranked UK universities draw on non-replicable, durable historic advantages and cachet. The large body of new universities created after 1992, (commonly known as ‘post-92s’) meant that a large number of universities did not enjoy these advantages and have been labeled non-leading (Hewitt-Dundas, 2012).

Further, drawing on a variety of studies of sustained organizational success (e.g. Ghoshal and Bartlett, 1993; Kay, 1993; Collins and Porras, 1994), Shattock (2003) argues

that these institutions' reputations are underpinned by complementary resources and capabilities developed over time, that are hard to replicate and confer sustainable advantage. These include established bundles of practice, organizational architectures and cultures. Shattock's (2003) argument is consistent with the concept of organizational complementarity (Milgrom and Roberts, 1992; Amable, 2015). As Espeland and Sauder (2007) demonstrate for US law schools, we view the rankings and the REF as complementary since they interact and reinforce each other, as elements within a coherent wider social system. In line with Hewitt-Dundas (2012: 266), we also view them as encompassing the whole range of a university's physical, material and intangible reputational assets and are collectively substantial. Some, such as university longevity, are non-replicable. Others, such as their ability to pay relatively high academic salaries, may be replicable.

Shattock's argument has been buttressed by two factors that emerged since he first wrote. First, the costs to universities of improving research performance have been rising as Geuna and Piolatto (2016) report. However, many costs are hard to capture. For example, full professors' pay has increased through inter-university competition. UK full-professors' salaries (unlike those of other grades) are negotiable, un-capped and have recently been subject to considerable increases as universities have competed for talent (De Fraja et al, 2013). Given the key importance to academics of working in universities employing the best talent in their areas, they are highly motivated to move to them if possible (De Fraja et al, 2013). Thus, financial resources attract further (human) resources (Hicks, 2012). Second, Hunt (2016) has shown through a case study of an attempt to develop a greater research emphasis in a post-92 university that the demands made on both faculty and managers in attempting this are multifarious and complex. They involve intensifying research effort while concurrently dealing with the heavy teaching loads that generate major income streams for these universities.

The significance for non-leading universities of lacking legacy reputation depends on whether the institutional context offers opportunities to challenge established reputations of superior organizational effectiveness in the public arena. Given the apparent powerful correlation between research performance and the annual university league tables, one possible avenue might be to invest in improved research performance. Another might be to distrust this indirect avenue and simply aim directly to improve the organizational effectiveness outcomes captured by the league tables or rankings such as teaching quality. Let us first examine the annual rankings and thereafter research performance.

NATIONAL RANKINGS AND THE SHATTOCK THESIS

Three national rankings of UK universities are published annually – *The Complete University Guide*, the *Guardian* (GRG) and joint *Times* and *Sunday Times* tables. They have established themselves as guides to prospective consumers. The three rankings employ various input and outcome measures (Appendix 1). We have selected the GRG because - unlike the others - it has no research component, allowing a ‘cleaner’ analysis of the impact of changes to research performance. However, the correlation coefficients for the other two rankings (*Complete University Guide* and the joint *Times* and *Sunday Times* ranking) with GRG are 91 percent and 90 percent respectively. Investigating the subsample for the two upper quintiles of universities, the correlations are 92 percent in both cases and for the rest, 88 percent and 87 percent respectively. Thus, little distinguished the three rankings in outcome terms.

As Table 1 shows, GRG uses eight different criteria, each weighted between 5 and 15 percent. Overall, the measure is divided between inputs and outputs. In output terms, it captures student perceptions of feedback, job prospects, overall quality, teaching quality and ‘value added’ (outputs overall = 55 percent). In input measure terms, it includes entry scores,

spending per student and faculty-student ratios (inputs overall = 45 percent). It contains a specific “value-added” factor comparing students’ degree results with their entry qualifications (Heily-Rayner, 2016).

- Table 1 here-

Although the main body of our empirical analysis is not dichotomous, in parts of it we define the ‘leading’ group as approximating to the top quintile of the GRG, most of which are ‘Russell Group’ institutions (see below).¹ We now turn to the specifics of how the UK government assesses research performance through the REF.

RESEARCH PERFORMANCE

For nearly three decades, British universities have had their research performance evaluated and ranked by the British state at roughly five to six year intervals. For the sake of simplicity, we refer to these as the “REF”. The REF of 2014 is the most recent iteration. A government-funded body (The Higher Education Funding Council for England: HEFCE) manages the REF on behalf of counterpart bodies in Wales, Scotland and Northern Ireland. The results of the HEFCE assessments are publicly disseminated and are used to determine the share of substantial government research funding allocated to each university. The purposes of these exercises are multiple and have been discussed by several researchers (see for example Marginson, 2014; Naidoo et al., 2011). However, a central governmental intention has been to raise the overall level of UK research by intensifying competition for state funds (Marginson, 2014). In addition to allocating research-specific, non-project funding support, the REF has sought to benchmark UK research against international

¹ Inevitably, this boundary is somewhat arbitrary. A much narrower view of “leading” would be those sometime called the “G5”: the universities of Oxford and Cambridge, Imperial College London, University College London and the London School of Economics. However, that group is equally arbitrary in that it excludes the 19 other members of the Russell Group universities as well as others widely regarded as prestigious.

standards, while presenting UK research to the world since its detailed assessments are publicly available.

The REF has influenced the academic labor market by encouraging and facilitating universities efforts to seek out and recruit high-performing researchers (Hicks, 2012). It has pushed institutions and researchers to improve economic and social impacts to raise the wider policy, industrial and social relevance of research (Marginson, 2014). However, the REF has also generated considerable controversy. Hicks (2012) argues that, for institutions, the REF has become primarily a competition for prestige rather than just one for funding. Further, Martin and Whitley (2010) even claim that those who conduct the assessments are recruited from leading institutions and may be biased in favor of them. In this way, they argue, the gradings of their own institutions are higher than they otherwise would be, and stratification between departments and institutions is increased.

The REF has changed considerably over time, becoming increasingly complex and demanding for universities since its introduction in 1986 (Martin, 2011). Part of the increasing burden has arisen because the measures used have constantly shifted (Martin, 2011). The costs to institutions of gearing themselves up to perform well affect all university employees and arguably all students. The REF has been argued to militate towards an increasing division of labor between REF researchers and teachers, as faculty are hired on or moved on to teaching-only contracts if they are judged to be unsuccessful researchers (Rowlinson et al., 2015). Internal changes to the division of labor between academics - for example, teaching loads increasing for those not regarded as sufficiently 'research active' while decreasing for others who are provided with time to improve their outputs and career prospects - may affect internal cohesion. From a faculty perspective, these changes could negatively impact individuals' motivation. They could therefore be dysfunctional for the institutions and the academics concerned (Martin and Whitley, 2010: 72).

The REF has evoked considerable reactions within Universities that vary in their nature and intensity (Hewitt-Dundas, 2012; De Fraja et al., 2016). Institutional strategies may, at one end of the spectrum of theoretical possibilities, strongly prioritize REF research performance through faculty incentives, monitoring and recruitment or, at the other, simply collect and transmit whatever its academics produce without seeking to influence that to any degree. More strongly REF-oriented strategies appear more likely to be adopted if institutions perceive them to lead to better university rankings.

We now present the broad outlines of the 2008 and 2014 REFs, as despite their elaboration and increasing complexity their core design was similar. Universities seeking state ‘Quality-Related’ (‘QR’) funding for their general research efforts were required to make entries to the REF though it was and is not compulsory to request QR funds. Crucially, universities may choose which subject areas (‘Units of Assessment’: UoAs) and which members of faculty to submit; there has been no obligation on them to submit all faculty. Universities prepare detailed statements of their research environment (notably their doctoral student completions, the research funding from all sources they have spent and how they nurture research), and list the details of the faculty they are submitting together with a list of selected published outputs. Up to four items may be submitted for each individual faculty member, depending on the proportion of full-time that they work and whether they are defined as ‘early career’ or have personal reasons such as pregnancy or illness for being submitted with less than four items.

From 2008, government research assessment acquired a new and crucial significance as a funding mechanism since the availability of separate sizeable project-based research funding from the research councils was significantly reduced in and around 2008. From 2008, measurement of university research performance was also conducted differently than hitherto; previous exercises were not comparable since earlier criteria had under specified

circumstances allowed academics' outputs to be counted by more than one institution. Simultaneously, for the overwhelming majority of universities that chose to participate, involvement in delivering inputs to the research assessment process became even more costly from 2008. As criticisms were made of assessment methodologies, so requirements became more demanding. The "REF" was experienced as so resource-demanding that one authority described it as a 'Frankenstein monster' (Martin, 2011: 247). Clearly, if changes in research performance do not have consequences for the GRG rankings one could after 2008 question even more than previously the management effort and resources involved in investing in developing and implementing a REF-focused research performance strategy.

In 2014, a requirement introduced in 2008 was extended, specifying that each UoA submission should contain a general statement about efforts to promote the societal and policy impacts of their research and should present a number of 'Impact Case Studies', whose number would vary in proportion to the total number of faculty submitted. These were designed to illustrate actual impacts on policy and practice and represented a considerable additional burden. Consultants confirm that the requirement to produce these studies (7,000 were produced across institutions) meant that the effort required to submit to the REF grew considerably in comparison with previous exercises, costing institutions an extra 150 million pounds in addition to intangible costs (Manville et al., 2014: 10-11). They simultaneously estimated these costs to be high in international terms. The effort could not, as with publications, be reduced by buying impacts generated at other universities simply by hiring their faculty. Potential case studies had to be identified. 'Runners' had to be selected, weaker studies selected out and studies fully documented and supported by named external sources to vouch for the impacts claimed. Impact could (and by implication ought to) be located in research produced up to fifteen years previously, meaning that senior faculty involved at the beginning of that period had frequently left the University (possibly on bad

terms) or retired. The case studies were developed in a context of uncertainty about the precise requirements of what they should consist of and what would be recognized as ‘impact’. Thus, the costs to institutions of submitting to the REF have grown considerably since 2008.

Espeland and Sauder (2007) point out that public measures tend to generate widespread ‘gaming’ and this is indeed the case here. Marginson (2014) highlights how institutions used various tactics to attempt to inflate their universities’ research assessments, drawing particular attention to the tactic of submitting a small number of high-performing academics in order to push up the collective quality score for their outputs while excluding their lower-rated colleagues. Marginson argued overall that leading universities were better able to ‘game’ the system in this and other ways than other institutions. It was partly for this reason that educational journalists devised “research power” (‘RP’) ratings that take into account the number of academics involved in achieving quality scores. ‘Research Power’ is not a term or measure used or determined by REF assessors directly but is, rather, an appropriation of particular REF measures into a composite index devised and used by media analysts when they communicate research performance. It is also increasingly used by the academic community (Marginson, 2014). The mid-2016 Stern review of the REF commissioned by the UK government recognized the gaming issue and effectively recommends that RP scores simply replace REF quality (‘GPA’) scores. Another of its key recommendations is that all academics’ outputs be reported by institutions (Department for Business, Energy and Industrial Strategy, 2016). Therefore when gauging research performance, rather than employing REF scores we prefer to employ RP scores.

RESEARCH PERFORMANCE AND NATIONAL RANKINGS

The overall purpose of our study is to examine whether UK universities that managed to improve their research performance between 2008-14 as assessed by the REF also experienced improvements in their national rankings, or whether the rankings are essentially legacy effects and impervious to changes in research performance. If the latter is the case, then it is questionable whether non-leading universities should adopt a resource-intensive strategy geared to improving research performance. In addressing our overall purpose, we examine a number of related issues. How strong are associations between research performance and the annual rankings and over time? Do significant changes to research performance result in similar changes to the rankings? To the extent that this is the case, are significant changes to the research performance of non-leading universities sufficient to propel them into the ranks of leading universities?

The questions we pose are particularly relevant to the viability of non-leading universities intensifying their REF research efforts in order to enhance their rankings. If the substantial investments required to perform well in REF research have no significant consequence for the standing of a university as indicated by rankings then their managers should understand that. The implication would be that non-leading universities might be more effective if they directly targeted selected components of organizational effectiveness contained in the rankings, such as teaching quality. In some UK universities, the size of the task of improving REF research appears considerable. Especially in the post-92s, research was until quite recently a relatively under-emphasized activity and it has in many cases remained a secondary one (Harley, 2002). One of Harley's (2002: 1997) post-92 informants even described it at that point as 'neglected and regarded with suspicion'. Internal critics of research-informed strategies at these universities often regard teaching as the key institutional task and therefore ask whether institutional research investments (frequently presented as cross-subsidies to research) are justified (Fulton, 2003).

Such questioning occurs in a context of broader skepticism voiced by some UK academics about the value and especially the results within academe of state assessments of research performance. In contrast to some other countries where performance-based funding systems have been introduced, a long strand of critical assessments of UK research performance assessments, and their consequences within universities - especially the increasing use of journal rankings lists - now exists (Reborra and Turri, 2013). This forms part of a burgeoning wider literature on growing 'managerialism' in UK universities (see inter alia Puxty et al.; 1994; Humphrey et al., 1995, and more recently Willmott, 2011; Mingers and Willmott, 2013; Tourish and Willmott, 2015; Clarke and Knights, 2015). These criticisms reflect the fundamental ways in which, despite frequent initial skepticism, rankings can come to change how 'attention is redistributed' and affect the way in which academic work, its costs and benefits and even professional identities are understood (Espeland and Sauder, 2009: 605). The critics cited above argue that university managers have taken the opportunities offered by the REF to exercise increasing control over academics and their work, thereby threatening academic freedom. Few defenses have been attempted (for an exception see Rowlinson et al., 2015). If our analysis showed that improved research performance brought no wider benefits, then their managerial use would appear questionable on those pragmatic grounds alone and not only on the principled ones advanced by critics.

We suggest that while the link between research performance and the various GRG organizational effectiveness factors is indirect it is nevertheless likely to be potent as it is part of a virtuous circle. A leading GRG ranking contributes to attracting 'good' students in the sense that they are highly motivated and have high levels of cultural capital (Reay et al., 2009). Universities with 'good' students can better attract leading academics who by definition are relatively successful researchers (De Fraja et al., 2013). The agglomeration of such academics may be configured by management to improve research performance.

However, an alternative model for non-leading universities does exist. The strategy adopted by the post-92 University of Northampton is consistent with Shattock's arguments and offers an alternative approach for non-leading institutions that aligns with their vocational tradition. It sees little point in prioritizing REF performance (University of Northampton, 2015). It argues that alternative or additional funding sources to those offered by the state could be accessed by developing links with business and local communities, by offering practically-oriented research and consultancy. Such universities could thereby build on existing vocational traditions, and simultaneously enhance teaching and overall organizational effectiveness as measured by GRG.

RESEARCH QUESTIONS

To respond to our overall purpose we initially adopt a two-fold explorative approach. One aspect of Shattock's thesis suggests that we should expect little change over time in either RP or GRG terms. Thus, our first explorative research goal involves examining the degrees of association of research performance in 2008 and in 2014 and between GRG in 2008 and 2014. This six-year period is significant because it was in conjunction with the 2008 REF that the research assessment exercise acquired added significance as a determinant of substantial research funding. We then investigate the key issue of whether change in RP between 2008 and 2014 has any impact on change in GRG.

A second aspect of Shattock's thesis is that even if we observed an impact of change in RP and GRG rankings, this will not involve non-leading universities breaking into the ranks of leading universities since the latter's resource complementarities are impervious to any developments by non-leading universities. Thus, our second explorative research task addresses whether any association between changes in RP and GRG rankings between 2008

and 2014 is associated with changes within the rankings of leading universities as opposed to non-leading universities.

OPERATIONALIZATIONS

Dependent Variable

As indicated, we employ the GRG rankings as a measure of the organizational effectiveness of UK universities. As well as employing GRG in 2014 we also employ a measure of GRG change between 2008 and 2014 calculated as rate of change. When calculating changes to GRG between 2008 and 2014, some universities have only data reported by the newspaper for *either* 2008 or 2014 meaning that change cannot be calculated. This reduced the population of universities available for analysis to exactly 100.

Research Performance

In order to operationalize our main independent variable, as we have argued above, we prefer to use RP as our measure of research performance between 2008 and 2014. It measures a combination of overall REF quality gradings and the number of faculty submitted. It is arrived at as follows. All of the REF indicators were measured by expert peer-review panels and were graded 0-4 where 0 suggests that the work is below nationally-recognized standard and 4 represents world-leading research. The composite result in any given unit of assessment (UoA) is referred to as a 'Grade Point Average' (GPA). A further measure is then derived by multiplying the GPA by the full-time equivalent number of researchers submitted, giving a RP ranking between universities in a given UoA. Thus, in a sense RP is a measure of the depth of research quality within a given university's faculty.

In our analysis, we employ two variants of RP, in levels and in changes: RP either in 2008 or in 2014, or Change in RP between 2008 and 2014.

Control Variables

In our regression analyses, we control for university size, which we operationalize as Number of Students (undergraduate and post-graduate). Further, we control for Faculty Cost per full-time academic, total Capital (non-faculty) Expenditure, and Russell Group membership, all as reported by individual institutions in their annual reports. All control variables are measured at 2008.

DATA ANALYSIS

Correlation Analysis

As an initial step, we employ a correlation analysis to explore the relationship between GRG in 2008 and GRG in 2014 and RP in 2008 and RP in 2014. We report the results in Table 2.

- Table 2 here –

Table 2 indicates that GRG rankings in 2008 and 2014 are highly correlated. In short, there is substantial stability in both GRG and RP over the six-year period. The table also suggests that powerful links exist between GRG rankings and RP. The correlations between RP rank in 2008 and GRG are strong not only for GRG 2008, but are almost unchanged for GRG 2014. The implication is that universities with strong research performances are also well-ranked in the non-research areas of organizational effectiveness that GRG captures.

However, although the association between RP 2008 and GRG in 2014 (or 2008) is strong, it is not perfect.

Table 2 enables us to explore whether changes in RP between 2008 and 2014 have any association with changes in GRG rankings in the same period. When we correlate the rate of change in RP between 2008 and 2014 with the rate of change in GRG for the same

period we find a statistically significant positive correlation. Albeit modest (0.23), the correlation implies that the two rates of change exhibit broadly similar trends.

Regression Analysis

In Table 3, we move beyond correlation to test a model that addresses the issue of the causality between RP in 2008 and GRG in 2014 in a linear regression framework. The purpose of the model is two-fold. First, it is to examine the degree to which RP in 2008 sustains its impact on GRG in 2014, and second to assess the impact of change in RP between 2008 and 2014 on GRG rankings in 2014. Summary statistics for all the regression variables are reported in Appendix 2. In Table 3, we initially introduce RP 2008 as a determinant of GRG 2014. We thereafter extend the estimated specification step-wise and introduce the Change in RP between 2008 and 2014 (model 2), and then the controls - Number of Students, Faculty Cost and Capital Expenditure (model 3) and Russell Group membership (model 4).

- Table 3 here -

In column 2, the explanatory variables are the RP rank in 2008 and RP change between 2008 and 2014. We find that the RP rank 2008 is an important factor affecting the GRG rank 2014 with a marginal effect of 0.76; the marginal effect of the Change in RP rank is similarly high at 0.85. Furthermore, both effects are statistically significant meaning that even though initial conditions matter, RP change between 2008 and 2014 does have a distinct impact on GRG rank 2014. The high measure of explained variance of the regressions is noteworthy; Adjusted R^2 ranges between 0.56 in column 1 and 0.64 in column 2, indicating the strong explanatory power of our simple model.

As a robustness check, we present results from other extended specifications; in column 3 besides the main variables of interest we add controls for number of students, faculty cost per full time and capital (non-faculty) expenditure. Number of students, measuring organization size, has a negative and statistically significant impact while none of

the other control variables achieves statistical significance. In a further step, in column 4 we add to the specification a dummy indicating membership of the Russell group that may capture additional unobserved fixed effects. However, the coefficient is not statistically significant. Further analysis (not reported here) indicates that this is because membership of the Russell Group is highly correlated with research performance. Importantly, the coefficients of the two main variables in the base specification, column 2 remain stable in the extended specifications in columns 3 and 4. In other words, RP in 2008 and Change in RP between 2008 and 2014 are the most important determinants of GRG rank in 2014.

We now turn to the issue of whether changes in RP result in changes to GRG. We estimate a specification where the dependent variable is Change in GRG between 2008 and 2014 and the explanatory variables are the same as in the previous table. The results are displayed in Table 4.

- Table 4 here -

The main finding is that while RP 2008 does not have a statistically significant impact on GRG change, Change in RP ranking does. The marginal effects of RP change are between 1.12 and 1.18 across specifications. The effects of the control variables in columns 3 and 4 are all insignificant. When Change in RP is entered, adjusted R^2 is 0.26 and, in columns 3 and 4, 0.28. While these adjusted R^2 measures are substantially lower than their equivalents in Table 3, they are of importance. While the results in Table 3 indicate that both initial conditions in terms of research status (RP 2008) and changes to RP over the period of analysis play a role in determining GRG ranking in 2014, Table 4 indicates that only change in RP is significantly associated with change in GRG. There is therefore some evidence that those universities improving their RP can expect to change their GRG rank.

In summary, in line with the first facet of Shatttock's thesis that GRG rankings will be largely stable, our correlation analysis suggests considerable stability not only in terms of

GRG rankings but also in RP. However, there is an indication that changes in RP are associated with changes in GRG. The regression analyses add further credibility to this association. The second facet of Shattock's thesis is however that the association between changes in RP and changes in GRG are a feature of non-leading universities. We now turn to this issue.

Location of Variability

As an initial step to investigating where among the population of UK universities the impact of changes in RP on changes to GRG rank is most evident, and to observe how significant it is, we examined the five universities that had had the greatest declines in RP rankings and the five universities with the greatest RP improvements for the period 2008-2014.

Observing the five universities whose RP rankings declined most between 2008 and 2014 in Table 5, we note that none of these was ranked by GRG 2008 as being in the upper quintile of universities.

- Table 5 here -

In that sense, none was a leading institution. The table indicates that all of these universities' GRG rankings also declined. In the case of the University of Bradford its RP declined by 28 places between 2008 and 2014 and its GRG rank by 47 places. We see similar dramatic falls for Edinburgh Napier and Cardiff Metropolitan. While the University of Salford suffered a decline in its RP of 19 places this only had a marginal downward impact on its GRG ranking. It nevertheless appears that in general falls in RP result in falls in GRG rankings, and that these may be sharp.

When we examine those universities with the greatest improvements to their RP ranking in Table 6, we note that none of them had leading GRG rankings in 2008. Nor did they have a leading ranking in 2014.

- Table 6 here -

Northumbria's RP improved by 30 places between 2008 and 2014 while its GRG ranking improved from 87 to 58. We see similar results for Huddersfield and Edge Hill. However, for Worcester and Bedfordshire, paradoxically, improvements in their RP were accompanied by declines in their GRG ranking. The upside possibilities of RP performance therefore appear to be somewhat less predictable than the downside. However, Tables 5 and 6 involve small numbers of universities. Nevertheless, what they do indicate is that the most dramatic changes to RP are a feature of non-leading universities.

Location of Volatility

In Table 7 we explore whether the uppermost quintile of universities together with the next 19 highest ranked universities (i.e. the top 39 institutions) are more or less volatile (heterogeneous) than the rest in terms of their RP and GRG performance. We use a common measure of volatility defined as the standard deviation of the rate of change in RP and GRG rankings. The table indicates a much higher degree of stability (and less heterogeneity) within the top-quintile as well as the top-39 universities compared to the rest. We obtain similar results when we divide universities according to either their 2008 RP or GRG position.

- Table 7 here –

In sum, while both correlation and the regression analyses suggest that GRG rankings and RP are predominantly stable over time, changes to RP are associated with changes in GRG. Our analysis of the location of variability and volatility indicate that radical change to RP occurs outside of the leading (and the lowest-rated) universities and is a feature of middle-ranking universities. While one should exercise extreme caution in extrapolating from our data, given current trends it seems a tall order to expect that improvements or declines in research performance in themselves will result in substantial changes to the

leading GRG rankings. As Shattock argued, reputations at the top of the distribution and those factors that underpin them appear to be very “sticky”.

Our analysis supports the notion that the UK contains two distinct types of universities. On the one hand, there are leading institutions whose RP and GRG did not change to any great degree between 2008 and 2014. On the other, there are non-leading institutions that are significantly more volatile in RP and GRG rankings and indeed, where cases of fairly dramatic variability exist. Our analysis also supports the notion of the difficulty non-leading institutions have, regardless of changes to their research performance, of breaking into the ranks of the leading institutions. Thus, both aspects of Shattock’s thesis are supported.

CONCLUSIONS AND DISCUSSION

Our overall contribution has been to provide empirical support for Shattock’s argument that breaking into the group of leading universities in terms of organizational effectiveness through improving research performance is unlikely for outsiders. We demonstrate that it has not occurred in recent years. We supplement his argument by showing that the RP of non-leading universities is related to their wider GRG rankings. Thus, failing to maintain RP has a negative impact on their GRG rankings that in some cases appears quite dramatic.

Our first research question asked whether change in RP ranking mattered to the GRG ranking. We found a statistically significant relation suggesting that improvements in research performance positively impact changes to GRG rankings. Our second research question asked whether this matters to leading institutions. Our location of variability and volatility analysis indicates that changes to GRG rankings are confined to non-leading universities. Overall, this means that improving research performance does potentially constitute a viable strategy for

non-leading universities to compete with each other, but also that breaking into the ranks of leading institutions is unlikely.

Hence, the rankings of leading institutions are not obviously threatened by the research performance improvements of non-leading institutions. Shattock's arguments about the apparently insuperable difficulties other universities face in trying to break into the leading group therefore attract support from our evidence. For those concerned with state policy, our findings suggest the existence of a firm line between leading universities and non-leading that is reproduced over time. If the reproduction of existing institutional hierarchies is in fact a tacit policy goal and the state is content to allow the regulatory capture that Martin and Whitley (2010) argue is present, then this may be acceptable to policy makers. If on the other hand state management of the system as a whole requires greater inter-institutional parity of esteem in order, for example, to raise the prestige of vocational education, as successive governments have suggested over a long period would be positive, then it is an issue.

Should, then, managers of non-leading universities aiming to increase overall organizational effectiveness abandon trying to improve their universities' REF performance and, like the University of Northampton develop a strategy that is independent of any concern with the REF? Given the path dependency of research performance and the resources involved in preparing for and participating in the REF, this alternative strategy may be tempting. However, there are dangers associated with it. Universities experiencing marked declines in their research performance generally suffered in terms of organizational effectiveness as measured by the GRG rankings. Such falls may not inspire consumer confidence. Equally, although dramatic improvements in research performance do not confer entry to the ranks of the leading universities, they do generally enhance GRG rankings and therefore add to the prestige of these universities in relation to their more immediate competitor institutions. This may be particularly helpful where institutions cluster closely, as in urban contexts. However,

it is conceivable that the demands of the REF are such that for some institutions a ‘Northampton strategy’ may be the only viable strategic option. In this case, a three-level overall structure may emerge among universities: leading institutions, non-leading and ‘REF opt-outs’.

Our analysis cannot provide more fine-grained guidance about the precise level of emphasis that ought to be put on REF-related research by non-leading institutions. This is an effort-reward calculation. On the reward side, there may be different values for institutions depending on their position in the higher education market, the level of local competition and the priorities of the particular segment of consumers that they target. Among consumers targeting non-leading universities, priorities appear to include several factors with no research connection such as buildings and facilities (Price et al., 2003).

Our study has several clear limitations that could be addressed by future research. One is that a six-year period is limited. Future research will depend on whether future REF exercises are comparable. Another limitation is that, despite its virtues as a multidimensional and comparative index, GRG as a measure of the organizational effectiveness of universities could be regarded as contentious. Further, we can only speculate as to whether changes to RP affect certain factors in the GRG ranking more than others. Moreover, the mechanisms through which this occurs are likely complex and our research design does not allow us to touch on them. One possibility is that students perceive improvement in research performance directly through the teaching they receive and respond positively. In broad terms, this would be consistent with Artès et al.’s (2017) study of a medium-sized Spanish university that improved research may contribute to an overall improvement in teaching quality. A second is that students perceive changes to research performance indirectly via enhanced institutional reputation (reflected in their expectations of acquiring what they regard as desirable jobs) and it may be this rather than the learning experience that shifts. A third possibility is that universities that make an effort at improving their research performance are also engaged in

improvements in other areas. This may include their HRM practices including how they select, retain and motivate staff. Indeed all of these possibilities may apply and require further examination.

A third limitation of our study concerns the generalizability of our findings. The REF, a government led system of ranking of research, is unique to the UK. However, if universities in other countries converge around the ABS journal rankings that underpin the REF or similar rankings then within country, and even between-country REF-style assessments may become more likely. Further, as various third party actors now produce rankings of universities across many countries, then one can expect researchers in these countries to ask similar research questions to ours.

REFERENCES

- Amable, B. 2016. Institutional complementarities in the dynamic comparative analysis of capitalism. *Journal of Institutional Economics*, 12(1): 79 – 103.
- Argenti, P. 2000. Branding B-schools: Reputation management for MBA programs. *Corporate Reputation Review*, 3(2): 171-178.
- Artès, J., Pedraja-Chaparro, F., del Mar Salinas-Jimenez, M. 2017. Research performance and teaching quality in the Spanish higher education system: Evidence from a medium-sized university. *Research Policy* 46: 19-29.
- Brynjolfsson, E., Milgrom, P. 2013. Complementarity in organizations. In R. Gibbons & J. Roberts (Eds.), *The Handbook of Organizational Economics*: 11-55. New Jersey, USA: Princeton University Press.
- Clarke, C.A., Knights, D. 2015. Careering through academia: Securing identities or engaging ethical subjectivities? *Human Relations*, 68(12): 1865-1888.
- Collins, J.C., Porras, J.I. 1994. *Built to last: Successful habits of visionary companies*. London, UK: Century.
- De Fraja, G., Facchini, G., Gathergood, J. 2013. How much is that star in the window? Professorial salaries and research performance in United Kingdom universities. University of Nottingham Global Economic Policy Paper 16/13. Retrieved from: <https://www.nottingham.ac.uk/gep/news-events/news/papers/1613.aspx>.
- Department for Business, Energy and Industrial Strategy. 2016. *Research Excellence Framework Review. An independent review of university research funding* by Sir Nicholas Stern. Retrieved from: <https://www.gov.uk/government/publications/research-excellence-framework-review>.
- Espeland, W.N., Sauder, N. 2007. Rankings and reactivity: how public measures recreate social worlds. *American Journal of Sociology*, 113(1): 1-40.

- Espeland, W.N., & Sauder, N. 2009. Rankings and diversity. *Review of Law and Social Justice*, 18(3): 587-608.
- Financial Times* (2017), All work and no play? The new uni experience. May 27, 2017.
- Fulton, O. 2003. Managerialism in British Universities: Unstable hybridity and the complications of implementation. In A. Amaral, O. Fulton & I.M. Larsen (Eds.), *The Higher Education managerial revolution*: 341-346. Abingdon, UK: Routledge.
- Geuna, A., Piolatto, M. 2016. Research assessment in the UK and Italy: Costly and difficult, but probably worth it (at least for a while). *Research Policy* 45: 260-271.
- Ghoshal, S., Bartlett, C. 1993. *The individualized corporation*. London: Heinemann.
- Guardian. 2013. The Guardian University League Table 2014 – Methodology. Retrieved from: <http://www.theguardian.com/education/interactive/2013/jun/04/universityguide-students>.
- Harley, S. 2002. The Impact of Research Selectivity on Academic Work and Identity in UK Universities. *Studies in Higher Education*, 27(2): 187-205.
- HEFCE. 2015. Register of HE providers. London: HEFCE.
- Heily-Rayner, M. 2016. Methodology behind the Guardian University Guide 2017. Retrieved from: <https://www.theguardian.com/education/2016/may/23/methodology-behind-the-guardian-university-guide-2017>.
- Hewitt-Dundas, N. 2012. Research and knowledge transfer activity in United Kingdom Universities. *Research Policy*, 41(2): 262-275.
- Hicks, D. 2012. Performance-based university research funding systems. *Research Policy*, 41(2): 251-261.
- Humphrey, C., Moizer, P., Owen, D. 1995. Questioning the value of the research selectivity process in British University Accounting. *Accounting, Auditing and Accountability*, 8(3): 141-164.

- Hunt, C. 2016. 'Teachers' to 'Academics': The implementation of a modernisation project at one UK post-92 University. *Studies in Higher Education*, 41(7): 1189-1202.
- Kay, J. 1993. *The foundations of corporate success*. Oxford: Oxford University Press.
- Manville, C., Morgan Jones, M., Freason, M., Castle-Clark, S., Hasham, M.L., Gunashekar, S., Grant, J. 2014. *Preparing impact submissions for REF 2014: An evaluation*. Cambridge: Rand Europe.
- Marginson, S. 2014. UK research is getting better all the time—or is it? Retrieved from: <https://www.theguardian.com/higher-education-network/2014/dec/23/ref-uk-research-excellence-getting-better-higher-education>.
- Martin, B.R., Whitley, R. 2010. The UK research assessment exercise. A case of regulatory capture? In Whitley, R., Gläser, J., & Engwall, L. (Eds.), *Reinforcing knowledge production. Changing authority relations in the sciences and their consequences for intellectual innovation*: 51-80. Oxford: Oxford University Press.
- Martin, B.R. 2011. The Research Excellence Framework and the 'Impact Agenda:' Are We Creating a Frankenstein Monster? *Research Evaluation*, 20(3): 247-254.
- Milgrom, P., Roberts, J. 1992. *Economics, Organization and Management*. Englewood Cliffs, NJ: Prentice Hall.
- Mingers, J., Willmott, H.C. 2013. Taylorizing business school research. On the 'one best way' performative effects of journal rankings. *Human Relations*, 66(8): 1051-1073.
- Naidoo, R., Shankar, A., Veer, E. 2011. The consumerist turn in Higher Education: Policy aspirations and outcomes. *Journal of Marketing Management*, 27(11-12): 1142-1162.
- Price, I.F., Matzdorf, F., Smith, L., Agahi, H. 2003. The impact of facilities on students' choice of university. *Facilities*, 21(10): 212-222.
- Puxty, A.G., Sikka, P., Willmott, H.C. 1994. Systems of surveillance and the silencing of UK academic accounting labour. *British Accounting Review*, 26(2): 137-171.

- Reay, D., Crozier, G., Clayton, J. 2009. 'Strangers in paradise'? Working-Class students in elite universities. *Sociology*, 43(6): 1103-1121.
- Reborra, G., Turri, M. 2013. The UK and Italian Research assessment exercises face to face *Research Policy*, 42: 1657-1666
- Rowlinson, M., Harvey, C., Kelly, A., Morris, H. 2015. Accounting for research quality: research audits and the journal rankings debate. *Critical Perspectives in Accounting*, 26: 2-22.
- Shattock, M. 2003. *Managing Successful Universities*. Maidenhead: McGraw-Hill Education.
- Shattock, M. 2010. *Managing Successful Universities*. Second edition. Maidenhead: McGraw-Hill Education.
- Tourish, D., Willmott, H.C. 2015. Journal rankings, mindless measures and the ABS guide. *Critical Perspectives in Accounting*, 26: 37-46.
- University of Northampton. 2015. *Transforming Lives. Inspiring Change: Strategic Plan, 2015-2020*. Northampton: University of Northampton.
- Willmott, H.C. 2011. Journal list fetishism and the perversion of scholarship: reactivity and the ABS list. *Organization*, 18(4): 429-442.

TABLES**Table 1 The weighting of GRG factors**

GRG element	Proportion as percentages
Entry score	15
Feedback as rated by graduates	10
Job prospects	15
Final-year students opinions about the overall quality of their course	5
Spending per student	15
Faculty-student ratio	15
Teaching quality as rated by graduates of the course	10
Value added: entry qualifications related to final degree classification	15

Source: Guardian (2013)

Table 2 Correlation analysis, RP and GRG (n=100)

	RP 2008	RP 2014	Change RP	GRG 2008	GRG 2014	Change GRG
RP 2008	1	0.96 (0.00)	0.22 (0.02)	0.82 (0.00)	0.77 (0.00)	0.24 (0.02)
RP 2014		1	0.04 (0.61)	0.81 (0.00)	0.80 (0.00)	0.18 (0.06)
Change RP			1	0.22 (0.02)	0.03 (0.76)	0.23 (0.02)
GRG 2008				1	0.82 (0.00)	0.36 (0.00)
GRG 2014					1	-0.07 (0.48)
Change GRG						1

Note: The pairwise correlation coefficients and their p-values in brackets are reported.

Change denotes rate of change. The correlation coefficients shown in **bold** indicate five percent level of significance or better.

Table 3 Regression analysis: GRG rank in 2014 (n=100)

Dependent variable: GRG rank in 2014				
	1	2	3	4
RP 2008 rank	0.71 (0.06)	0.76 (0.05)	0.82 (0.07)	0.82 (0.08)
Change RP rank		0.85 (0.21)	1.14 (0.22)	1.13 (0.22)
Number of students			-1.15 (0.30)	-1.14 (0.30)
Faculty cost			0.06 (0.14)	0.06 (0.15)
Capital expenditure			0.04 (0.03)	0.03 (0.04)
Russell group				1.27 (7.17)
Adj. R sq.	0.56	0.64	0.69	0.69

Note: Marginal effects and standard errors (in brackets) are reported. The marginal effects shown in **bold** indicate five percent level of significance or higher.

Table 4 Regression analysis: Change in GRG rank between 2008 and 2014 (n=100)

Dependent variable: Change in GRG rank between 2008 and 2014				
	1	2	3	4
RP 2008 rank	-0.12 (0.08)	-0.04 (0.05)	-0.003 (0.07)	0.01 (0.08)
Change RP rank		1.12 (0.21)	1.16 (0.22)	1.18 (0.22)
Number of students			-0.13 (0.28)	-0.14 (0.28)
Faculty cost			0.06 (0.15)	0.05 (0.15)
Capital expenditure			-0.01 (0.03)	-0.004 (0.03)
Russell group				-3.02 (6.55)
Adj. R sq.	0.04	0.26	0.28	0.28

Note: Marginal effects and standard errors (in brackets) are reported. The marginal effects shown in **bold** indicate five percent level of significance or higher.

Table 5 Universities with the greatest decline in RP rank 2008-2014

Institution	RP rank decline	GRG 2008 rank	GRG 2014 rank
Edinburgh Napier University	17	55	85
University of Salford	19	84	85
Cardiff Metropolitan University	20	76	105
Robert Gordon University	20	40	55
University of Bradford	28	43	90

Table 6 Universities with the greatest improvement in RP rank 2008-2014

Institution	RP rank improvement	GRG 2008 rank	GRG 2014 rank
Northumbria University	30	87	58
University of Worcester	26	54	95
University of Huddersfield	25	95	42
University of Bedfordshire	25	94	110
Edge Hill University	21	109	70

Table 7 Volatility in performance comparisons (n=100)

Performance measure	Top-20	Top-39	Outside Top-39
RP	0.10	0.12	0.23
GRG	0.21	0.36	0.68

Appendix 1

Measures used in the Guardian Guide, The Complete University Guide, The Times

Measures	The Guardian	The Complete University Guide	The Times
Student satisfaction	✓	✓	✓
Research assessment		✓	✓
Entry standards	✓	✓	✓
Student/staff ratio	✓	✓	✓
Spend on academic services	✓	✓	✓
Spend on student facilities		✓	✓
Good honours degrees		✓	✓
Completion rates		✓	✓
Graduate prospects	✓	✓	✓
Value added (qualifications upon entry are compared with the degree award that a student receives at the end of their studies)	✓		

Sources: <http://www.thecompleteuniversityguide.co.uk>;

https://en.wikipedia.org/wiki/Rankings_of_universities_in_the_United_Kingdom

Appendix 2

Summary statistics for regression variables (n=100)

Variable	Definition	Mean (S.D.)
GRG 2014	Guardian university ranking guide 2014 rank	56.38 (34.06)
GRG Change	Change in rank, 2008-2014	0.07 (19.69)
RP 2008	Research Assessment Exercise 2008 rank	57.17 (35.87)
RP change	Change in rank, 2008-2014	0.80 (9.29)
Number of students	Total number of students in 2007, thousands	17.71 (7.28)
Staff cost	Total staff expenditure per full-time member of academic staff, thousands	104.02 (13.58)
Capital expenditure	Total capital expenditure, millions	76.63 (77.62)
Russell group	Membership in the Russell group of universities	0.23 (0.42)