

# **Business and Management impact assessment in REF2014:**

## **Analysis and reflection \***

Neil Kellard<sup>†</sup> and Martyna Śliwa

University of Essex

June 2016

### **Abstract**

The evaluation of research impact is likely to remain an important element of research quality audits in the UK for the foreseeable future. With this paper, we contribute to debates on impact and relevance of business and management studies (BMS) research through an analysis of REF2014 impact scores within the Business and Management Unit of Assessment. We offer insights into the organizational contexts of UK business schools within which impact is produced, drawing attention to the issues of linkages with research intensity, grant income generation, research team size, career stage and gender of academics, and whether impact activity is focused on private or public sector organizations, and national or international reach. We put forward recommendations for managers responsible for business schools and higher education policymakers regarding management and organizational policies and processes, as well as possible changes to the rules guiding future research excellence audits.

**Keywords:** Impact, business schools, relevance, research excellence framework

---

\* This paper has been accepted for publication in the British Journal of Management.

<sup>†</sup> Corresponding author: Essex Business School, University of Essex, Colchester, UK, CO4 3SQ.  
Email: [nkellard@essex.ac.uk](mailto:nkellard@essex.ac.uk).



## **Introduction**

Since the first Research Assessment Exercise in 1986, the idea of evaluating university research has been well established in the UK system of higher education. The Research Excellence Framework 2014 (REF2014) brought the first exercise in national assessment of research that did not only focus on the evaluation of research quality but also of its impact outside academia. In preparing for REF2014, this newly added emphasis on impact had prompted UK higher education institutions (HEIs) to pay greater attention to evidencing their engagement with external stakeholders and to the ways in which the generation of impact is embedded in the organizational culture and in existing structures for supporting and managing research. At present, British universities have already started to plan for the next assessment of research quality, commonly referred to as REF2020. At this point, it is not certain what form the next assessment is going to take but it would appear that evaluation of research impact is here for the foreseeable future.

In the case of business schools, the topic of research impact has for a number of years now attracted the attention of scholars contributing to the ‘relevance debate’ (e.g. Butler, Delaney and Spoelstra, 2015; Fincham and Clark, 2009; Gulati, 2007; Hodgkinson and Starkey, 2011; Kieser, Nicolai and Seidl, 2015; Pettigrew, 2011). While the debate has addressed a range of issues, from questioning how to define relevance (e.g. Willmott, 2012) to discussing how to combine academic rigour with relevance to practice within business and management studies (BMS) research (e.g. Hodgkinson and Starkey, 2011), little has been said about the organizational conditions and contexts in which BMS research is carried out, and – with the

exception of Butler, Delaney and Spoelstra's (2015) recent piece with a focus on leadership scholars – about the researchers who produce 'impactful research'.

Against this background, our paper contributes to discussions of impact and relevance of business school research. We address the UK higher education sector in which these debates are interwoven with considerations of the periodic audits of university research carried out through REF exercises. We are particularly interested in the insights that REF2014 results provide for our understanding of: 1) what kind of business school contexts are most conducive to generating research impact; 2) who actually produces the type of impact and the underpinning research that would attract high grades from the REF assessors; and 3) what implications this has for business school managers and for policymakers within the UK system of higher education.

In exploring these questions, we have undertaken an empirical analysis of impact scores within the Business and Management Unit of Assessment (UoA) with a focus on comparison between three performance-related clusters of institutions. A cluster-based comparison has allowed us to concentrate on the more general issues of similarities and differences between organizational contexts and the impact activity business schools engage in, rather than on the detail of the content of individual impact case studies. Against the background of extant literature on the impact and relevance of business school research, our paper offers insights into the organizational contexts within which impact is produced, drawing attention to the issues of linkages with research intensity, grant income generation, research team size, career stage and gender of academics, and whether impact activity is focused on private or public sector organizations, and national or international reach. Following from our analysis, we put forward recommendations for managers responsible for business schools and

higher education policymakers regarding management and organizational policies and processes, as well as possible changes to the rules guiding future research excellence audits.

The remainder of this paper is structured as follows. Below, we offer an overview of the literature on impact and relevance of BMS research and point to the importance of the so far under-explored questions about the organizational contexts within which impact is produced. We then provide a background to our analysis explaining key definitions and requirements regarding research impact, as stipulated in the REF guidelines issued in advance of REF2014. This is followed by discussion of the data and methods applied in the analysis. Subsequently, we offer an analysis of impact scores with a focus on three clusters of institutions. Finally, we discuss the findings of our analysis and provide a set of conclusions.

### **Impact, relevance and the organizational context of UK HEIs**

In the Special Issue of this Journal celebrating its 25<sup>th</sup> anniversary, Andrew Pettigrew (2011: 347) observed that ‘it may be appropriate for some members of our local and international communities to raise their aspirations and deliver forms and processes of knowledge which meet the double hurdle of scholarly quality and policy/practice impact’. At the same time, he noted that management research suffers from the lack of ‘a natural focused community’ (Pettigrew, 2011: 349) comprising its recipients. These comments were published soon after Higher Education Funding Council for England’s (HEFCE, 2011) confirmation that in the next national exercise of academic research evaluation (i.e. REF2014), there will be an explicit element to assess the impact arising from excellent research, alongside the outputs and environmental elements.

While being an advocate of business and management ‘research with impact’, back in 2011 Pettigrew pointed out that the effects of successive research evaluation assessments in the UK – which he links to Power’s (1999) notion of the ‘audit society’ – had been both positive and negative. This is because, similarly to all evaluation mechanisms, rather than simply fulfilling the role of measuring and ranking the research carried out by academics, national assessments of research affect in various ways the processes they evaluate, and give rise to variegated outcomes at both the organizational and individual level (Smith, Ward and House, 2011).

An espoused intention behind introducing the impact element into REF2014 was to incentivize behaviours that would shift the notion of academic research away from ‘classical conceptions of knowledge conducted by elites’ (Hazelkorn, 2009: 9) and towards rewarding research that focuses on different stakeholders and benefits wider society. As such, the logic behind placing importance on the impact of research produced within the UK higher education system resonated with calls by scholars in the international arena, including those within the field of BMS (e.g. Aguinis et al., 2014; DeTienne, 2013; Pettigrew, 2011), for the need to broaden the ways in which the impact of research is evaluated.

Despite its underlying socially beneficial objectives, the inclusion of impact in the REF sparked debates and controversy within the academic community from its outset. To some commentators, this was an overwhelmingly negative development: a sign of ‘philistinism and ignorance’ (Ladyman, 2009, quoted in Pettigrew, 2001: 348) of those responsible for higher education policy, and an initiative potentially leading to scholars becoming ““door to door” salesmen for vulgarised versions of their increasingly market orientated products’ (Collini, 2009, quoted in Pettigrew, 2011:

348). Others were critical, if not of the idea of impact assessment *per se*, then of the proposed process for evaluating impact. For example, in their analysis of the responses to the 2009 public consultation on the proposals for a Research Excellence Framework, Smith, Ward and House (2011: 1376) questioned whether the ‘structure of evaluation in the REF really leaves room for research units to describe the variety of interactive processes through which their work can have an impact on numerous different potential users’.

Notwithstanding the above critiques, the need to produce impactful research has, as mentioned above, for a long time been propagated within the academic community, in particular through the ‘relevance debate’ in relation to BMS research (Butler, Delaney and Spoelstra, 2015). In the UK, the debate was sparked by Tranfield and Starkey’s (1998) paper arguing for the development of research in line with a ‘Mode 2’ approach promising to offer ‘a different and potentially more appropriate (useful/relevant) model of the link between theory and practice’, whereby ‘knowledge production and diffusion are interlinked rather than sequentially disaggregated’, so that ‘it becomes more difficult to divide theory and practice’, as ‘mode 2 enables contribution to both simultaneously’ (Tranfield and Starkey, 1998: 351-352). Since Tranfield and Starkey’s seminal piece, discussions about how to produce academically rigorous and practically relevant BMS research have been developed by a number of authors (e.g. Hodgkinson, 2006; Hodgkinson and Starkey, 2011; Starkey, Hatchuel and Tempest, 2009; Starkey and Madan, 2001). In addition to promoting mode 2 research (e.g. Burgoyne and James, 2006; MacLean, MacIntosh and Grant, 2002), scholars have made a case for increasing the relevance of BMS through conceiving management as a design science (e.g. Aken, 2004, 2005; Denyer, Tranfield and van Aken, 2008; Dunbar and Starbuck, 2006; Hatchuel, Starkey and

Tempest, 2010; Hodgkinson and Healey, 2008; Panza and Thorpe, 2010; Romme, 2003) and through furthering evidence-based management in BMS based on systematic review (e.g. Briner and Rousseau, 2011; Briner, Denyer and Rousseau, 2009; Rousseau, Manning and Denyer, 2008).

In arguing for the necessity to build bridges between academics and users of BMS research, the focus of debates addressing the ‘relevance/impact challenge’ (Hodgkinson and Starkey, 2011: 364) has been on the design and content of the research, and the potential to enhance research relevance through transdisciplinarity and a closer connection between research conducted within BMS and other social science disciplines. Critical commentators, on the other hand, have argued for a re-definition of the notion of relevance (Willmott, 2012), have scrutinized the assumptions behind the relevance debate (Bresnen and Burrell, 2013; Knights, 2008; Learmonth, Lockett and Dowd, 2012), and questioned the motivations and compromises made by researchers pursuing activities aimed at becoming relevant to practitioners (Butler, Delaney and Spoelstra, 2015).

To date, however, little has been said about the organizational contexts within which both academically rigorous and policy/practice relevant research is generated. More specifically, questions about the ways in which the organizational contexts of higher education facilitate the development of impactful research, which individuals and groups of researchers produce such research, and what effects this has for organizations and for the UK higher education sector remain unaddressed. These questions are highly relevant both from the point of view of higher education policymakers and university management practitioners.



Insights into the organizational contexts underpinning the generation of academic knowledge in the era of research excellence assessments, however, can be found in other areas of BMS. Here, we note several observations that might contribute to an understanding of the contexts within which and by whom research impact is achieved. As Butler and Spoelstra (2014: 544) argue, both the research agenda of scholars as well as their academic conduct are ‘shaped by the (real or perceived) demand for excellence within the contemporary university’. It has been highlighted that research excellence audits influence decisions about the research questions explored and methodological approaches adopted (McNay, 2007); considerations regarding where, when and how academics publish their research (Pettigrew, 2011); as well as career-related strategies underpinned by the desire to be classified as ‘research-active’, with implications for the self-esteem and work priorities of academics (Lucas, 2006). This suggests that with impact on policy and practice becoming a new point of gravity for evaluating research, those academics wishing to progress in their careers will be inclined to take part in the ‘impact game’ as a route to professional success. In this context, considering the generation of impactful research an important component of an academic career under conditions of evaluation in which full success can only be achieved by a minority (Macdonald and Kam, 2007), gives rise to questions about what effects the assessment of research impact within a national framework of excellence might have on ‘the inevitable divisions that an elite system of evaluation and judgement entails’ (Clarke, Jarvis and Knights, 2012: 14).

These are pertinent issues for both university managers and higher education policymakers, especially in relation to the UK higher education sector, with its long-standing distinction between ‘research intensive’ universities and those with a primary focus on education, and the concomitant differences between the resources

for non-teaching-related activity that different types of institutions have at their disposal. It is often asserted that, compared to post-1992 institutions, research intensive universities have tended to be higher ranked in terms of publication outputs, to attract a greater amount of external research funding and be more likely to employ teams of researchers collaborating within established research centres and groups (e.g. Howell and Annansingh, 2013). It is not clear, however, whether such distinctions translate into varied generation of impact, and in particular whether they manifest in REF2014 impact scores.

As far as individuals and groups of researchers that produce impactful research are concerned, an a priori case might be made for examining the role of career stage and gender. With regard to the former, it is understood that Early Career Academics (ECAs) constitute a group whose members largely take the imperative to produce work in line with the criteria of research excellence audits for granted (Archer, 2008a; 2008b). However, obtaining permanent positions has become increasingly difficult for ECAs, and their ability to attract external research funding – often a precondition to producing research with impact – is below average (Laudel and Gläser, 2008). In order to secure long-term employment under conditions of uncertainty and precarity (Dunn, 2013; Loacker and Śliwa, 2015), ECAs tend to focus their efforts on working towards highly ranked publications (Prasad, 2013), perhaps at the expense of impact. On the other hand, tenured and established academics are more likely to have developed the types of network that are conducive to impact activity.

With regard to gender, studies of academia have documented its salience to the work processes, hierarchies and career progression within universities (e.g. Acker, 2008; Knights and Richards, 2003; Johansson and Śliwa, 2014; van den Brink and

Benschop, 2012). For example, research has shown that the accomplishment of outcomes that are most highly valued in academia is dependent on belonging to networks of influence to which women tend to have weaker access than men (Fletcher et al., 2007). Empirical evidence suggests that men academics are likely to work closely with other men, whereas women, who are also a minority within the higher echelons of academia, are likely to have fewer opportunities to participate in and benefit from informal academic networks through which coaching, mentoring, collaboration-building and information provision take place. Moreover, as Śliwa and Johansson's (2014) study has shown, women academics can self-exclude from taking part in those professional activities that are highly valued in a university setting. Therefore, when it comes to producing impactful research, men academics might also be more likely than women to successfully engage with the impact agenda.

While the relevance debate stresses the need for collaboration between academics and practitioners (e.g. Bartunek and Rynes, 2014; Starkey, Hatchuel and Tempest, 2009), it does not explore whether impactful research is more likely to result from universities' linkages with private or public sector organizations or whether spatially, these ties should be local, national or international. However, for higher education and government policymakers more broadly, this *direction* of impact is of key importance, as it is indicative of the wider role business schools play within the society and economy. Moreover, for university managers, and in particular for those responsible for business schools, knowledge of what type of research is considered of highest relevance and evaluated as the most impactful, can help frame objectives, narratives and future strategy.

On the basis of the discussion presented thus far, in conducting our analysis, we were

interested in finding out about the institutional contexts in which highly evaluated versus less ‘excellent’ impact is produced, and in particular in exploring the links between REF2014 impact scores within the Business and Management UoA, and research intensity, income generation, research team size, career stage and gender of researchers, private or public sector focus, and national or international reach. Before moving on to discussing these issues, below we provide a context to REF2014 impact evaluation, followed by an explanation of our data and methodological approach.

### **Definitions and guidelines for REF2014 impact submissions**

According to the REF guidelines, ‘(f)or the purposes of the REF, impact is defined as an effect on, change or benefit to the economy, society, culture, public policy or services, health, the environment or quality of life, beyond academia’ (REF 02.2011: 26). In using the terms ‘effect on, change or benefit’, the above definition might be considered equivocal regarding the normative aspect of impact. However, the document further specifies the desirability of non-negative effects, by pointing out that ‘(i)mpact includes the reduction or prevention of harm, risk, cost or other negative effects’ (*ibid.*).

To evaluate the quality of impact, the REF panel used the overarching criteria of ‘reach and significance’, with reach defined as ‘the spread or breadth of influence or effect on the relevant constituencies’ and ‘significance’ as ‘the intensity or the influence or effect’ (King’s College London and Digital Science, 2015: 13). These were to be assessed jointly and ranked by a score between ‘four star’ (‘outstanding’) and ‘one star’ (‘recognized but modest’). Impact judged as having ‘little or no reach and significance’, ‘not eligible’ or ‘not underpinned by excellent research produced by the submitted unit’ fell under the category of ‘unclassified’ (REF 02.2011: 44).

HEIs' impact submissions to the REF consisted of two main elements: the 'impact template', stipulating the unit's approach to enabling and supporting impact and 'impact case studies', i.e. descriptions of actual impacts generated within the assessment period on the basis of 'excellent research undertaken' in the unit submitted for assessment. In the overall evaluation of impact, the impact template was given a weighting of 20%, whereas the case studies contributed the remaining 80%. To ensure comparability, each impact case study was supposed to be submitted on a standardized pro-forma.

The number of individual impact case studies submitted by each unit was calculated according to a formula based on the number of staff submitted to the REF by the unit, whereby units submitting up to 14.99 FTE staff members were required to produce two impact case studies, with an additional case study required for each interval of up to 10 FTEs staff submitted to the REF above 14.99 FTEs. HEIs were not allowed to submit more than the required number of case studies whereas submission of fewer than required case studies resulted in awarding a grade of 'unclassified' to each 'missing' case study.

The Guidelines stipulated that, rather than selecting a range of case studies representative of all research impact generated by the unit, HEIs were supposed to submit their 'strongest examples of impact that are underpinned by the submitted unit's excellent research' (REF 02.2011: 28). The definition of excellent research was further qualified as at least 'two star', i.e. 'quality that is recognised internationally in terms of originality, significance and rigour' (REF 02.2011: 29). Where the underpinning research was not considered to be 'predominantly of at least two star quality' (*ibid.*), the case study was to be graded 'unclassified'. The research was

supposed to have occurred between 1<sup>st</sup> January 1993 and 31<sup>st</sup> December 2013, with all outputs attributed to the submitting unit, whereas the impact was expected to have occurred between 1<sup>st</sup> January 2008 and 31<sup>st</sup> July 2013. The link between the underpinning research and the impact was given high importance. Submitting units were required to provide evidence that ‘the impact would not have occurred or would have been significantly reduced without the contribution of [the] research’ (*ibid.*).

### **Data and methodology**

Having at our disposal both the results of REF2014 and the actual impact submissions of all units within the Business and Management sub-panel, we identified a number of factors to represent the focal areas previously discussed as related to impact generation. Table 1 shows these factors and describes briefly how they were obtained.

[Insert Table 1 about here]

Next, in terms of methodology, we have followed a three-step approach:

- (i) We have clustered institutions according to their impact GPA score achieved in REF2014 as a measure of impact generation;
- (ii) Within each cluster, we have computed the average value of a number of possible explanatory factors as listed in Table 1. These factors were selected because of their alignment with our research interest in the organizational contexts in which impactful research is generated (e.g., the number of case studies as reflective of research intensity; the length of service; gender of key researchers).
- (iii) We have assessed whether any significant differences in average factor value exist between the clusters.

The primary aim of the clustering in step (i) was to allocate institutions into homogeneous groups tied to their impact performance. A popular approach in the literature is the  $k$ -means clustering algorithm (see, for example, Lopez, Kundu and Ciravegna, 2009) in which  $n$  observations are apportioned to  $k$  clusters and each  $x_i$  observation is allocated to the cluster with the nearest  $\mu_k$  mean. Essentially, the algorithm minimises the following objective function:

$$A = \sum_{j=1}^k \sum_{i=1}^n \|x_i - \mu_j\|^2 \quad (1)$$

To operationalize, the number of clusters are chosen a priori and initial values for the  $\mu_k$  selected. Observations are then allocated to clusters according to (1) and the  $\mu_k$  recalculated. This process continues iteratively until the change in the  $k$ -means is considered small enough.  $K$ -means clustering is typically appropriate when a moderate number of observations are available. Hierarchical clustering (see Burgess and Shaw, 2010) can be useful with smaller numbers of observations.

With regard to steps (ii) and (iii), the purpose was two-fold. Firstly, where average values are significantly different between clusters, this is clearly suggestive of a factor related to impact performance. Secondly, where average values are similar between clusters but significantly different from zero, this is likely to be indicative of a factor required for impact generation. To test differences in average factor values between clusters (i.e.,  $H_0: \bar{f}_{l,j} = \bar{f}_{l,j+m}$ ), we employed a conventional two-sample t-test allowing for unequal variances:

$$t = \frac{\bar{f}_{l,j} - \bar{f}_{l,j+m}}{s} \sim t_z \quad (2)$$

where  $\bar{f}_{l,j}$  is the average value for factor  $l$  in cluster  $j$ . Also  $S = \sqrt{\frac{v_{f_{l,j}}}{n_{f_{l,j}}} + \frac{v_{f_{l,j+m}}}{n_{f_{l,j+m}}}}$  and

the degrees of freedom are represented by:

$$Z = \frac{(v_{f_{l,j}}/n_{f_{l,j}} + v_{f_{l,j+m}}/n_{f_{l,j+m}})^2}{(v_{f_{l,j}}/n_{f_{l,j}})^2 / (n_{f_{l,j}} - 1) + (v_{f_{l,j+m}}/n_{f_{l,j+m}})^2 / (n_{f_{l,j+m}} - 1)}. \quad (3)$$

where  $v_{f_{l,j}}$  and  $n_{f_{l,j}}$  are the variance and number of observations for factor  $l$  in cluster  $j$  respectively.

## Analysis

### *An overview*

We begin the analysis with an overview of the impact performance of UoA 19 Business and Management. Table 2 contains the relevant summary statistics.

[Insert Table 2 around here]

The mean impact GPA score for an institution was 2.99, with the highest score being 3.84 for the University of Cambridge, and the lowest recorded score being 0.3 for the University of York St John. In comparison, the mean score for output GPA was 2.58, with the standard deviation being considerably smaller. This narrower distribution for output GPA relative to impact GPA may well reflect either perception biases or the more established understanding by both universities and the REF panel as to what constitutes high and low quality research output as opposed to the newly introduced process for impact assessment. Given our interest in the institutional contexts that were conducive to generation of excellent REF2014 impact, one intriguing question



refers to the link between output and impact GPA scores: did a high/low output GPA score typically correspond with a high/low impact score? In other words, was it necessary for an institution to be very successful at publication of highly ranked research outputs in order to also be evaluated as a top performer in terms of research impact outside academia? To this end, Figure 1 provides a scatter plot of the impact and output GPA scores for all submitting institutions.

[Insert Figure 1 around here]

As can be seen from the plot, the majority of institutions' scores are concentrated in the upper right quadrant. Fitting a least squares regression results in an upward sloping line, suggesting that at least to some extent those institutions with higher output GPAs were more likely to receive higher grades for impact. On the other hand, the  $R^2$  implies that only 37% of the variation in impact scores can be explained by variation in output GPA. For example, the University of Brighton recorded an output GPA score of 2.57, i.e. very close to the UoA's mean average, but still achieved a very high impact score of 3.73, which placed it within the top five performing institutions by impact GPA. By contrast, the impact score of the University of Surrey was at 2.18 considerably below the mean average for the UoA, locating it within the bottom 10 assessed institutions, whereas its GPA output grade of 2.82 was well above the output mean average.

To place the impact scores of UoA 19 in a more general context, Table 3 and Figure 2 provide an overview of impact scores for all 36 UoAs.

[Insert Table 3 around here]

Table 3 shows that the FTE-adjusted mean average for impact GPA for all units was 3.21. The highest UoA grade, 3.71, was recorded by UoA 1, Clinical Medicine, whilst the lowest was 2.99, recorded by UoA 11, Computer Science.

[Insert Figure 2 around here]

An examination of the histogram in Figure 2 of all UoAs allows for identifying some broader themes. In particular, the sciences (UoAs 1-16) tended to score more highly than the social sciences, humanities and the arts (UoAs 17-36). The score for UoA 19, the Business and Management sub-panel, conforms to this general trend with a mean score of 3.15, which was slightly less than the average for all UoAs. It is also worth noting that the FTE-adjusted mean average output GPA scores for all UoAs were lower than the impact scores. This demonstrates that across institutions submitted to all sub-panels it was common for output scores to correspond with impact scores of at least similar or, indeed, higher value.

#### *Top-, mid- and bottom-ranked impact: similarities and differences*

Using SPSS 21, we applied the *k*-means cluster analysis outlined in our Data and Methodology section to the impact scores for 97 of the 98 institutions in UoA 19. University of York St John, the bottommost node in Figure 1 was removed from the procedure, as this type of clustering is sensitive to outliers. To allocate institutions to either a top, middle or bottom cluster, we chose  $k = 3$ . Convergence was achieved after six iterations allowing for a maximum absolute coordinate change for any centre of 0.000. Table 4 shows the centre and the number of institutions for each cluster.

[Insert Table 4 around here]

It should be noted that although we chose  $k = 3$ , using a two-step cluster approach that does not impose this a priori, also determined that a division into three clusters was the most appropriate way of representing the impact data. Interestingly, Table 4 shows a somewhat uneven distribution of institutions to each of the clusters, with cluster 1 containing the highest number of institutions (i.e., 39) and cluster 3 containing the lowest (i.e., 24). However, in the following statistical analysis, to ensure comparability between clusters we needed to select groupings of equal size. Moreover, to emphasize variability between clusters, we wished to select groups of universities away from the cluster boundaries of 39 and 73. To satisfy these conditions we chose the ten institutions with the highest impact scores to represent cluster one; the ten institutions from position 47 to 56 to represent cluster two; and the ten institutions with the lowest impact scores to represent cluster three. In other words, following from our general overview of impact, we have carried out a comparative analysis of three groups of institutions – the top 10, mid 10 and bottom 10 – in terms of impact assessment within UoA 19. Selecting clusters of this size and position has allowed sufficient variation to enable meaningful observations to be made with regard to any substantive differences across the submissions.

Ranked by impact, the top ten institutions were as follows: University of Cambridge, University of Aberdeen, University of Strathclyde, University of St Andrews, University of Brighton, University of Bristol, Ulster University, London School of Economics, Bournemouth University and University of Reading. The mean average score for impact within the top ten group was 3.71 (see Table 5).

[Insert Table 5 around here]

The scores are tightly clustered with the highest at 3.84 and the lowest 3.60 which shows that the impact scores of each of these institutions were found to be very similar. By contrast, when output GPA is compared, this is less obviously the case: the highest score was 3.29 (LSE) and the lowest 2.37 (Bournemouth University). To illustrate these two characteristics of similar impact scores but more dispersed output scores, Figure 3 plots impact and output GPA in the top ten group.

[Insert Figure 3 around here]

Looking more closely at the impact case study submissions themselves, Table 6 shows a number of metrics that offer insights into various aspects of impact generating activity among the top ten performing institutions, as stated in the Methodology section.

[Insert Table 6 around here]

As far as key researchers are concerned, it is clear that while in some instances impact case studies were generated by individual researchers, most commonly they were the result of work by small teams. Importantly, key researchers tended to be embedded in the employing university for many years; in the case of eight out of the ten top institutions, the average time in service of the longest employed key researcher was greater than 15 years. Also worth noting is the gender proportion amongst key researchers, whereby in nine out of ten institutions the proportion of women was less than a half. Indeed, in the top performing institution, no women were listed as key researchers behind the four impact case studies.

In terms of the reach of impact generated by the top ten performing institutions, there was a combination of national and international impacts, with an almost equal

distribution of impact generated at the national and international level. In addition, 72% of case studies primarily involved interaction with public sector and non-profit organizations as opposed to the private sphere.

The mid ten institutions ranked by impact by the REF2014 assessors within UoA 19 were: the University of Exeter, Middlesex University, Kingston University, University of Central Lancashire, Robert Gordon University, University of Sussex, London Business School, Royal Holloway (University of London), Loughborough University and Queen's University Belfast. Compared to the top ten, the score for impact (shown in Table 7) was at 3.07 approximately 0.6 lower.

[Insert Table 7 around here]

Again, the mean output GPA score was lower than its impact counterpart at 2.72. The variation in output GPA from 2.19 (University of Central Lancashire) to 3.28 (LBS), shown in Table 7 and Figure 4, confirms that institutions with relatively low output scores could still achieve comparatively higher impact scores.

[Insert Figure 4 around here]

We were interested in whether differences between the quality of research underpinning impact case studies between the top and middle clusters could be found. As a proxy measure for output quality we applied the ABS2010 rankings of journals where they existed. Comparing these data in Tables 6 and 8, we can see no substantial difference in average ABS scores between the two clusters ( $t = -1.12$ ; 10%  $t_{crit} = \pm 1.73$ ). In fact, the mid cluster has a slightly higher mean average for underpinning research quality [Insert Footnote 1].

[Insert Table 8 around here]

Analogously to the top cluster, it would appear that impact case studies in the middle one were typically prepared by small teams of key researchers ( $t = -0.73$ ; 10%  $t_{crit} = \pm 2.92$ ), the majority of whom were men. The reach of the submitted case studies ( $t = 0.09$ ; 10%  $t_{crit} = \pm 1.75$ ) as well as the proportion of impact generated within the public and non-profit sector ( $t = 0.36$ ; 10%  $t_{crit} = \pm 1.77$ ) were also similar to those observed in the top cluster.

Equally, however, it is possible to identify a striking difference between the two clusters. In the middle cluster, the average length of service of the longest employed key researcher was 13.5 years, approximately six years less than the comparative figure in the top cluster ( $t = 1.86$ ; 10%  $t_{crit} = \pm 1.73$ ). Although our data does not allow us to pinpoint the reasons why longer time in post contributed to higher impact scores, it is probable that greater institutional familiarity tends to be coupled with stronger, long-term links with organizations that are direct beneficiaries of impact.

Our analysis leads to further insights as we turn to discussing the REF2014 impact submissions of the bottom ten performing institutions which included the University of the West of Scotland, University of Sunderland, Keele University, University of Surrey, University of Bedfordshire, University of Northampton, University of Worcester, Teesside University, University of Hertfordshire and University of Chester. Within this cluster (see Figure 5), the mean average impact score (shown in Table 9) was at 2.03 considerably lower than their counterparts in the top and middle clusters.

[Insert Figure 5 around here]

[Insert Table 9 around here]

Additionally, only in the bottom cluster was the mean output GPA score (2.11) higher than the mean average impact score. This indicates that while there existed within these institutions the capacity to conduct research and to generate research outputs that were at least of ‘quality that is recognised internationally in terms of originality, significance and rigour’ (REF 02.2011: 29), the capacity for generation of impact, as defined in REF2014, was lower. Moreover, compared to the top cluster, a number of the metrics in Table 10 are considerably different.

[Insert Table 10 around here]

For example, the average number of impact case studies was, at 2.4, noticeably lower ( $t = 2.02$ ; 10%  $t_{crit} = \pm 1.78$ ). This reflects the smaller number of researchers submitted to the REF by most of these institutions – with the University of Surrey as an exception – perhaps indicating a less intensive research environment, a view further supported by the finding that more case studies were based on the work of one key researcher ( $t = 2.10$ ; 10%  $t_{crit} = \pm 1.74$ ). Analogously, the number of underpinning research outputs to support each case study was lower ( $t = 2.77$ ; 10%  $t_{crit} = \pm 1.78$ ). Finally, the amounts of grant funding were lower ( $t = 2.99$ ; 10%  $t_{crit} = \pm 1.77$ ). Our analysis of the content of the case studies, including the evidence for impact provided by the institutions, also demonstrates that their connection to networks of external decision-makers, users and other stakeholder groups tended to be weaker.

At the same time, there are some similarities between the top and bottom cluster. For example, at 17 years, the average length of service of the longest employed key

researcher fell between that of the middle (i.e. 13 years) and top (19 years) cluster ( $t = 0.74$ ; 10%  $t_{crit} = \pm 1.78$ ). As mentioned above, there were fewer impact case studies submitted within the bottom cluster of institutions, indicating that fewer researchers were engaged in impact activity; our findings also suggest that these were those academics who have been with their employing institution for many years. In terms of the gender of key researchers, the percentage of women was higher in the bottom cluster as compared to the top cluster, with nearly half of key researchers being women. However, due to the overall low number of key researchers in the bottom cluster, this finding does not bear statistical significance ( $t = -1.25$ ; 10%  $t_{crit} = \pm 1.78$ ). Finally, as in the case of the other two clusters, the focus of the impact case studies was biased towards engagement with public and non-profit organizations ( $t = 0.82$ ; 10%  $t_{crit} = \pm 1.77$ ), suggesting that across all three clusters, impact submissions within the Business and Management UoA related in the majority of cases to work undertaken with organizations other than private businesses.

## **Discussion and conclusion**

Our analysis triggers a number of responses to the questions we asked at the beginning of this paper. As presented above, with regard to the relationship between institutional output and impact scores in the UoA 19 Business and Management, a least squares estimation gave an  $R^2$  of 0.37, indicating that a high output GPA score was not always a necessary condition for a high impact score. This demonstrates that those organizational contexts that are conducive to generation of 'excellent' impact need not be those that are traditionally viewed as 'elite' HEIs. The presence of post-1992 institutions such as Bournemouth University and the University of Brighton in the top cluster for impact, the highly research intensive London Business School in



the middle cluster and the University of Surrey as well as Keele University in the bottom cluster suggests that if the emphasis on impact remains a permanent element of national research evaluation, and perhaps even increases in importance, there might be a change in the established pattern of distribution of institutions classified within overall REF-based rankings. As a result, the UK business school community might need to develop a new understanding of what it means to be ‘excellent’ and ‘relevant’: an understanding that, indeed, unsettles the traditional privileging of ‘knowledge conducted by elites’ (Hazelkorn, 2009: 9).

From the context of all UoAs, the Business and Management sub-panel’s average grade for impact was slightly lower than the average grade for all 36 panels. In particular, those panels representing science, technology, engineering and mathematics (STEM) disciplines tended to grade more highly than those in the social sciences, humanities and arts. There might be different explanations behind this: i) impact was generally higher in STEM disciplines; ii) impact in non-STEM disciplines was harder to demonstrate through the rules governing the REF2014 process; iii) some panels assessed more harshly than others. We are not in the position to comment on possibilities i) and iii). However, regarding ii), the comparison of the REF2014 impact scores between Business and Management and other units of assessment gives rise to an important observation from the point of view of our interest in the organizational contexts within which ‘excellent’ impact is generated. If the same organizational context, i.e. the same university, is capable of delivering higher impact scores through some departments or research centres than through others, the issue here might not be whether BMS research is more or less relevant to policy and practice than the research produced in other disciplines, but that definitions of impact and relevance might need to be rethought (Willmott, 2012). We would like to propose

that further debate is needed between higher education policymakers and university managers on the extent to which the REF2014 rules for impact assessment succeeded in capturing impact generated by Business and Management unit, especially where it was not easily evidenced or tangible. In contrast to the 'hard sciences', it might be more difficult to determine the impact of the social sciences on the basis of specific pieces of research. As social sciences knowledge is often built through debate that is dispersed and distributed among unconnected scholars from many institutions, linking impact to the work of a single or group of scholars within a particular university becomes problematic (Smith, Ward and House, 2011).

Another insight from our analysis refers to who in the institution actually produced impact as evidenced in the REF2014 submissions of UoA 19. It is clear from the findings that those HEIs that submitted impact case studies highly evaluated by the REF2014 assessors were able to showcase the work of small teams with predominantly established researchers, some of whom had been in post for 20 years or longer, had developed a solid research base and had been successful in attracting external funding. In the first place, this demonstrates that impactful research tends to be linked to the researcher's long-term employment in the institution and result from collective effort; effort that involves not only engagement with external stakeholders but also joint work on research outputs underpinning impact and securing funding for impact-related activity. The submission of a post-1992 institution, i.e. the University of Brighton, which was able to achieve a position in the top cluster for impact evaluation, amongst mainly traditional research orientated universities, was similarly characterised by case studies prepared by research teams, with a high average longest serving time of key researchers and a high amount of grant funding. On the other hand, case studies submitted by institutions in the lower clusters were to a greater

extent likely to be based on the work of one key researcher, with fewer research outputs underpinning the case studies and a lower amount of funding. For example, both the universities of Keele and Surrey were in the bottom ten cluster, and commensurately had relatively low levels of reported grants and numbers of key researchers. For university managers this implies the need to support the development of a ‘critical mass’ of groups of researchers in their institutions, collaborating on generation of impactful research.

At the same time, it is important to draw management and policy implications from the finding that the category of Early Career Academics was under-represented as key researchers in impact submissions. In the context of the REF2014 rules this is unsurprising, since generating impact takes time and, as explicitly acknowledged by the rules, sometimes a long period will pass before the underpinning research influences practice and/or policy. This, however, also means that ECAs are to a much lesser extent in the position to produce impact than those with long careers in academia. Therefore, with the assessment of impact featuring and perhaps even increasing in importance in subsequent research excellent audits, the already often precarious position of ECAs within the higher education context (e.g. Archer, 2008; Prasad, 2013) might become weaker still: viewed through the lens of REF impact criteria, established researchers become more ‘valuable’ for universities than relatively new scholars. Such an evaluation system – especially as it is currently based on the assumption that impact is attributed to an institution and, unlike research outputs, does not ‘move’ with the researcher should she or he decide to change employers – thus presents an incentive for university managers to place greater emphasis on reward and retention of long-serving impact-generating scholars, and less on improving conditions for ECAs’ development. We would urge managers to

introduce policies and processes counter-acting the possible long-term consequences of this system-based incentive, since these consequences are negative for both the universities, whose future success depends on how well they cultivate their current early career staff, and for the careers and livelihoods of individual ECAs. To higher education policymakers, we would advise a revision of REF-impact rules so that they do not disadvantage early career academics.

Further, with regard to who produces impactful BMS research, our findings draw attention to issues of gender. While it had been intended for REF2014 to strengthen the measures promoting equality and diversity (REF, 2015), the results point to an under-representation of women amongst the key researchers behind the generation of impact. Indeed, while it was rather common for impact case studies built around the work of a team of three or more, there was only one case study amongst those submitted by the 30 institutions within the three analysed clusters where all three key researchers on the team were women; the focus of the case study was women entrepreneurship.

While our methodological approach does not allow us to answer why women were under-represented in the impact submissions, some intuitive explanations can be put forward based on previous literature on gender in academia. For example, it might be that – similar to difficulties with accessing academic networks (Fletcher et al., 2007) – women find it more difficult than men to access national and international networks of influence outside the workplace, and hence have lower chances than men to make impact on government policies or international regulation. It is also possible that some women academics self-excluded (Śliwa and Johansson, 2014) from participation in impact activity for REF2014. Irrespective of the actual reasons, if the REF definition

of impact as well as the submission requirements remain the same (e.g. regarding submission of only the ‘strongest’ case studies), the organizational hierarchies within HEIs might become more gender-imbalanced than they are at present. As with the case of ECAs discussed above, this poses a challenge for policymakers and university managers to design such policies and processes so as to enable more women to be included as key researchers.

Another issue that our analysis sheds light on relates to the majority of submitted impact having been generated through projects based on engagement with public sector and non-profit organizations rather than commercial entities, even in cases where impact claims referred to changes in business practice. Across the three clusters analysed, there were only two institutions – University of Bournemouth and University of Brighton – that had all their case studies based on direct engagement with private organizations. The underpinning reasons and the implications of this situation require more in-depth exploration. It might be that the current definition of research impact, with its explicit focus on ‘the economy, society, culture, public policy or services, health, the environment or quality of life’ (REF 02.2011: 26), does not obviously direct the attention of HEIs, including business schools, towards engagement with businesses as ‘relevant constituencies’ as far as impact is concerned. It might also be that impact on business was more difficult to evidence than, for example, impact on policy and regulation, or that university departments, including business schools, tend to co-operate with other public sector organizations and that these kinds of organizations are much more likely than businesses to commission work by universities.

It might also be that the REF2014 rules, which required the research underpinning

impact to be of at least ‘two star’ quality, made it difficult for many business schools, especially those based in post-1992 universities, to showcase their long established engagement with industry and their influence on management and organizational practice through the work of those staff who are active in applied research and knowledge transfer activity but do not focus on publication of research outputs. This separation of ‘impact’ from ‘engagement’ might have distorted the picture of relevance to practice by excluding the type of impact traditionally generated by business school academics engaging with businesses. Further, it might have also discouraged scholars from dedicating time to liaising with industry and pursuing knowledge transfer activities as these are considered of low priority in the context of the REF. We would therefore call for higher education policymakers and managers responsible for business schools to reflect on the ways in which the current regime of ‘excellence’, as defined by the REF criteria, shapes the decisions and actions of institutions and academics with regard to impact activity, and consider the long-term desirability of the possible shift away from seeing a broader societal value of business schools in cultivating links with and influencing the practices of business organizations.

As far as enriching the academic literature is concerned, our research makes an important contribution to the relevance debate in relation to BMS research. While extant literature has focussed predominantly on issues of research design and content (e.g. Tranfield and Starkey, 1998; Hodgkinson and Starkey, 2011; Briner and Rousseau, 2011), our paper contributes with insights into key questions about the organizational contexts within which impactful research is generated. Future studies could explore further the kind of institutional contexts, conditions and actors – for example, not only in terms of career stage and gender but also ethnicity – that produce

impactful research.

### **Footnote 1**

Our analysis focuses on comparing two cluster means with each other i.e., cluster one versus cluster two and cluster one versus cluster three. As a robustness check, we carried out the ANOVA analysis in SPSS and found that all empirical results were qualitatively similar. These results are available on request.

### **References**

Acker, J. (2008) 'Helpful men and feminist support: more than double strangeness', *Gender, Work and Organization*, 15, pp. 288-293.

Aguinis, H., D. Shapiro, E. Antanacopoulou and T. Cummings (2014). 'Scholarly impact: a pluralist conceptualization', *The Academy of Management Learning and Education*, 13, pp. 623-639.

van Aken, J. E. (2004). 'Management research based on the paradigm of the design sciences: the quest for field-tested and grounded technological rules', *Journal of Management Studies*, 41, pp. 219–246.

van Aken, J. E. (2005). 'Management research as a design science: articulating the research products of Mode 2 knowledge production in management', *British Journal of Management*, 19, pp. 19–36.

Archer, L. (2008a). 'The new neoliberal subjects? Young/er academics' constructions of professional identity', *Journal of Education Policy*, 23, pp. 265-285.

Archer, L. (2008b). 'Younger academics' constructions of authenticity, success and professional identify', *Studies in Higher Education*, 33, pp. 385-403.

Bartunek, J. M. and S. L. Rynes (2014). 'Academics and practitioners are alike and unlike: the paradoxes of academic – practitioner relationships', *Journal of Management*, 40, pp. 1181–1201.

Bresnen, M. and G. Burrell (2013). 'Journals `a la mode? Twenty years of living alongside Mode 2 and the new production of knowledge', *Organization*, 20, pp. 25–37.

Briner, R. B., D. Denyer and D. M. Rousseau (2009). 'Evidence-based management: concept cleanup time?', *Academy of Management Perspectives*, 23, pp. 19–32.

Briner, R. B. and D. M. Rousseau (2011). 'Evidence-based I–O psychology: not there yet', *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 4, pp. 3–22.

Burgess, T. and N. Shaw (2010). 'Editorial board membership of management and business journals: a social network analysis study of the Financial Times 40', *British Journal of Management*, 21, pp. 627-648.

Burgoyne, J. G. and K. T. James (2006). 'Towards best or better practice in corporate leadership development: issues in Mode 2 and design science research', *British Journal of Management*, 17, pp. 303–316.

Butler, N., H. Delaney and S. Spoelstra (2015). 'Problematizing “relevance” in the business school: The case of leadership studies', *British Journal of Management*, 26, pp. 731-744.



Butler, N. and S. Spoelstra (2014). 'The regime of excellence and the erosion of ethos in critical management studies', *British Journal of Management*, 25, pp. 538–550.

Clarke, C., D. Knights and C. Jarvis (2012). 'A labour of love? Academics in business schools', *Scandinavian Journal of Management*, 28, pp. 5-15.

Collini, S. (2009). 'Commentary: impact on humanities', *The Times Literary Supplement*, 13 November, pp. 18–19.

Denyer, D., D. Tranfield and J. E. van Aken (2008). 'Developing design propositions through research synthesis', *Organization Studies*, 29, pp. 393–413.

DeTienne, D. R. (2013). 'Assessing scholarly impact in the world of Google', *The Academy of Management Learning and Education*, 12, pp. 1–3.

Dunbar, R. L. M. and W. H. Starbuck (2006). 'Learning to design organizations and learning from designing them', *Organization Science*, 17, pp. 171–178.

Dunn, C. (2013). 'More universities use zero-hours contracts than research shows', *The Guardian*, 5<sup>th</sup> September 2013, <http://www.theguardian.com/higher-education-network/blog/2013/sep/05/university-lecturers-zero-hours-contracts> accessed on 20th November 2015.

Fincham, R. and T. Clark (2009). 'Introduction: can we bridge the rigour–relevance gap?', *Journal of Management Studies*, 46, pp. 510–515.

Fletcher, C., R. Boden, J. Kent and J. Tinson (2007). 'Performing women: the gendered dimensions of the UK new research economy', *Gender, Work and Organization*, 14, pp. 433-453.

Gulati, R. (2007). 'Tent poles, tribalism and boundary spanning: the rigor–relevance debate in management research', *Academy of Management Journal*, 50, pp. 775–782.

Hatchuel, A., K. Starkey and S. Tempest (2010). 'Strategy as innovative design: an emerging perspective'. In J. A. C. Baum and J. Lampel (eds), *The Globalization of Strategy Research. Advances in Strategic Management*, 27, pp. 3–28. Bingley: Emerald Group.

Hazelkorn, E. (2009). 'Impact of global rankings on Higher Education research and the production of knowledge', UNESCO Forum on Higher Education, Research and Knowledge Occasional Paper no. 15.

HEFCE (2011). *Decisions on Assessing Research Impact*, REF01.2011. Bristol: Higher Education Funding Council.

Hodgkinson, G. P. and M. P. Healey (2008). 'Toward a (pragmatic) science of strategic intervention: design propositions for scenario planning', *Organization Studies*, 29, pp. 435–457.

Hodgkinson, G. P. and K. Starkey (2011). 'Not simply returning to the same answer over and over again: reframing relevance', *British Journal of Management*, 22, pp. 355–369.

Howell, K. E. and F. Annansingh (2013). 'Knowledge generation and sharing in UK universities: A tale of two cultures?', *International Journal of Information Management*, 33, pp. 32–39.

Kieser, A., A. Nicolai and D. Seidl (2015). 'The practical relevance of management research: turning the debate on relevance into a rigorous scientific research program', *The Academy of Management Annals*, 9, pp. 143–233.

King's College London and Digital Science (2015). *The nature, scale and beneficiaries of research impact: An initial analysis of Research Excellence Framework (REF) 2014 impact case studies*. Bristol: HEFCE.

Knights, D. (2008). 'Myopic rhetorics: reflecting epistemologically and ethically on the demand for relevance in organizational and management research', *Academy of Management Learning and Education*, 7, pp. 537–552.

Knights, D. and W. Richards (2003). 'Sex discrimination in UK academia', *Gender, Work and Organization*, 10, pp. 213-238.

Ladyman, J. (2009). 'Against impact', *Oxford Magazine*, No. 294, pp. 4–5.

Laudel, G. and J. Gläser (2008). 'From apprentice to colleague: The metamorphosis of early career researchers', *Higher Education*, 55, pp. 387–406.

Learmonth, M., A. Lockett and K. Dowd (2012). 'Promoting scholarship that matters: the uselessness of useful research and the usefulness of useless research', *British Journal of Management*, 23, pp. 35–44.

Loacker, B. and M. Śliwa (2015). 'Moving to stay in the same place?' Academics and theatrical artists as exemplars of the 'mobile middle', *Organization*, DOI: 1350508415598247.

Lucas, L. (2006). *The research game in academic life*, Maidenhead: Open University Press.

Lopez, L., S. Kundu and L. Ciravegna (2009). 'Born global or born regional? Evidence from an exploratory study in the Costa Rican software industry', *Journal of International Business Studies*, 40, 1228-1238.

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

Macdonald, S. and J. Kam (2011). 'The skewed few: people and papers of quality in management studies', *Organization*, 18, pp. 467–475.

MacLean, D., R. MacIntosh and S. Grant (2002). 'Mode 2 management research', *British Journal of Management*, 13, pp. 189–201.

McNay, I. (2007). 'Research assessment: researcher autonomy'. In: C. Kayrooz, G. Akerlind, M. Tight (eds), *Autonomy in social science research: The view from United Kingdom and Australian universities. International Perspectives on Higher Education Research*, 4, Elsevier: New York, pp. 183–216.

Panza, K. and R. Thorpe (2010). 'Management as design, but what kind of design? An appraisal of the design science analogy for management', *British Journal of Management*, 21, pp. 171–186.

Pettigrew, A. (2011). 'Research with impact', *British Journal of Management*, 22, pp. 347-354.

Power, M. (1999). *The Audit Society: Rituals of Verification*. Oxford: Oxford University Press.

Prasad, A. (2013). 'Playing the game and trying not to lose myself: a doctoral student's perspective on the institutional pressures for research output', *Organization*, 20, pp. 936-948.

REF (2011). 'REF 2014: assessment framework and guidance on submissions', <http://www.ref.ac.uk/pubs/2011-02/> accessed on 20<sup>th</sup> November 2015.

REF (2015). 'Equality and diversity in the 2014 Research Excellence Framework. A report by the Equality and Diversity Advisory Panel', <http://www.ref.ac.uk/media/ref/content/equal/EDAP%20final%20report.pdf> accessed on 20<sup>th</sup> November 2015.

Romme, A. G. L. (2003). 'Making a difference: organization as design', *Organization Science*, 14, pp. 558–573.

Rousseau, D. M., J. Manning and D. Denyer (2008). 'Evidence in management and organizational science: assembling the field's full weight of scientific knowledge through reflective reviews', *Academy of Management Annals*, 2, pp. 475–515.

Śliwa, M. and M. Johansson (2014). 'The discourse of meritocracy contested/reproduced: foreign women academics in UK business schools', *Organization*, 21, pp. 821-843.

Smith, S., V. Ward and A. House (2011). "'Impact" in the proposals for the UK's Research Excellence Framework: Shifting the boundaries of academic autonomy', *Research Policy*, 40, pp. 1369-1379.

Starkey, K., A. Hatchuel and S. Tempest (2009). 'Management research and the new logics of discovery and engagement', *Journal of Management Studies*, 46, pp. 547–558.

Starkey, K. and P. Madan (2001). 'Bridging the relevance gaps: aligning stakeholders in the future of management research', *British Journal of Management*, 12 Special Issue, pp. S3–S26.

THE (2014). 'Research Excellence Framework: Institutions ranked by subject', 18<sup>th</sup> December 2014,

<https://www.timeshighereducation.co.uk/sites/default/files/Attachments/2014/12/17/g/o/1/sub-14-01.pdf>, accessed on 23rd June 2015.

Tranfield, D. and K. Starkey (1998). 'The nature, social organization and promotion of management research: towards policy', *British Journal of Management*, 9, pp. 341–353.

Van den Brink, M. and Y. Benschop (2012). 'Slaying the seven-headed dragon: the quest for gender change in academia', *Gender, Work and Organization*, 19, pp. 71–92.

Willmott, H. (2012). 'Reframing relevance as "social usefulness": a comment on Hodgkinson and Starkey's "Not simply returning to the same answer over and over again"', *British Journal of Management*, 23, pp. 598–604.

## Tables and Figures

**Table 1: Data from REF 2014 and UoA 19 Case Studies**

Variable name	Description	Source
Impact GPA	Institutional impact GPA score	Times Higher Education
Output GPA	Institutional output GPA score	Times Higher Education
No. of case studies	Number of impact case studies submitted by each institution	REF 2014 Results and Submissions (see <a href="http://results.ref.ac.uk/">http://results.ref.ac.uk/</a> )
Av. no. of outputs	Average number of underpinning research outputs listed to support each institutional impact case study	REF 2014 Results and Submissions (see <a href="http://results.ref.ac.uk/">http://results.ref.ac.uk/</a> )
Percent of journals	Percentage of the underpinning research outputs represented by journal articles	REF 2014 Results and Submissions (see <a href="http://results.ref.ac.uk/">http://results.ref.ac.uk/</a> )
Av. ABS score	Average Association of Business School 2010 (ABS2010) Quality Guide ranking of each journal in which underpinning research was published	REF 2014 Results and Submissions (see <a href="http://results.ref.ac.uk/">http://results.ref.ac.uk/</a> ) and Association of Business School 2010 (ABS2010) Quality Guide
Av. listed grant (£k)	Average listed grant amounts for each institutional impact case study	REF 2014 Results and Submissions (see <a href="http://results.ref.ac.uk/">http://results.ref.ac.uk/</a> )
Av. no. of key researchers	Average number of key researchers for each institutional impact case study	REF 2014 Results and Submissions (see <a href="http://results.ref.ac.uk/">http://results.ref.ac.uk/</a> )
Av. longest time in post (yrs)	Average length of service for longest serving key researcher for each institutional impact case study (i.e., the length of time a key researcher has been working in the institution prior to REF2014)	REF 2014 Results and Submissions (see <a href="http://results.ref.ac.uk/">http://results.ref.ac.uk/</a> ) and key researcher websites and CVs.
% of women key researchers	Percentage of women key researchers for each institutional impact case study	REF 2014 Results and Submissions (see <a href="http://results.ref.ac.uk/">http://results.ref.ac.uk/</a> )
Av. percent public	Percentage of an institution's case studies stemming from primary interaction with public and non-profit organizations	REF 2014 Results and Submissions (see <a href="http://results.ref.ac.uk/">http://results.ref.ac.uk/</a> ). Obtained by a thorough reading of case studies submitted to REF2014.
Av. percent national reach	Percentage of national reach which represents the proportion of an institution's case studies where the impact was generated within one country only	REF 2014 Results and Submissions (see <a href="http://results.ref.ac.uk/">http://results.ref.ac.uk/</a> ). Obtained by a thorough reading of case studies submitted to REF2014

**Table 2: Summary Statistics for UoA 19**

	Mean	Standard Deviation	Maximum	Minimum
Impact GPA	2.99	0.57	3.84	0.30
Output GPA	2.58	0.40	3.35	1.13

**Table 3: FTE-adjusted Impact GPA for all Units**

	Mean	Standard Deviation	Maximum	Minimum
Impact GPA	3.21	0.15	3.71	2.99

**Table 4: *k*-means Clustering for Impact GPA**

Cluster	1	2	3
Centre	3.49	2.99	2.31
No. of Institutions	39	34	24



**Table 5: Summary Statistics for Top 10 Institutions**

	Mean	Standard Deviation	Maximum	Minimum
Impact GPA	3.71	0.08	3.84	3.60
Output GPA	2.84	0.26	3.29	2.37

**Table 6: Impact Case Study Metrics for Top 10 Institutions**

	Cambridge	Aberdeen	Strathclyde	St. Andrews	Brighton	Bristol	Ulster	LSE	B'mouth	Reading	Institutional average
No. of case studies	4.00	2.00	8.00	3.00	2.00	3.00	3.00	8.00	2.00	5.00	4.00
Av. no. of outputs	5.25	6.00	6.00	6.00	5.00	6.00	6.00	5.38	6.00	5.80	5.74
Percent of journals	0.52	0.75	0.71	0.61	1.00	0.61	1.00	0.47	0.5	0.76	0.69
Av. ABS score	2.43	2.29	2.73	2.88	2.43	2.27	2.14	3.47	2.33	3.11	2.61
Av. listed grant (£k)	931.75	335.00	1218.38	577.33	550.00	155.90	398.33	569.68	118.12	0.00	485.45
Av. no. of key researchers	1.75	1.50	2.63	2.00	3.50	1.33	2.67	1.75	3.50	2.20	2.28
Av. longest time in post (yrs)	18.50	18.50	16.63	11.33	18.00	20.67	21.67	15.75	14.50	38.20	19.37
% of women key researchers	0.00	33.33	33.33	66.67	14.29	25.00	37.50	14.29	42.86	9.09	27.64
Av. percent public	100.00	100.00	75.00	100.00	0.00	100.00	50.00	100.00	0.00	90.00	71.50
Av. percent national reach	75.00	100.00	75.00	33.33	0.00	66.67	0.00	43.75	50.00	40.00	48.38

Note: The 'average percentage public' represents the proportion of case studies stemming from primary interaction with public and non-profit organizations. The 'average percentage national reach' represents the proportion of case studies where the impact was generated within one country only.

**Table 7: Summary Statistics for Middle 10 Institutions**

	Mean	Standard Deviation	Maximum	Minimum
Impact GPA	3.07	0.03	3.10	3.00
Output GPA	2.72	0.30	3.28	2.19

**Table 8: Impact Case Study Metrics for Middle 10 Institutions**

	Exeter	Middlesex	Kingston	UCLan	Robert Gordon	Sussex	LBS	Royal Holloway	L'boro	Queen's Belfast	Institutional average
No. of case studies	4.00	5.00	3.00	2.00	2.00	5.00	11.00	5.00	7.00	6.00	5.00
Av. no. of outputs	5.75	5.80	5.67	6.00	5.50	5.4	5.45	5.8	4.86	4.67	5.49
Percent of journals	73.91	68.97	58.82	83.33	72.73	51.85	76.67	82.76	70.59	82.14	72.18
Av. ABS score	3.09	2.53	2.80	2.88	2.50	3.23	3.45	2.62	2.22	2.76	2.81
Av. listed grant (£k)	954.85	609.54	19.80	87.39	22.50	916.42	132.74	879.92	81.94	266.59	397.17
Av. no. of key researchers	1.75	2.80	4.67	2.50	6.50	1.80	1.55	1.60	2.14	1.67	2.70
Av. longest time in post (yrs)	8.00	21.20	9.67	5.00	10.00	11.20	24.60	9.00	23.00	13.17	13.48
% of women key researchers	14.29	0.00	50.00	40.00	38.46	22.22	29.41	37.50	40.00	20.00	29.19
Av. percent public	75.00	90.00	50.00	75.00	75.00	80.00	40.91	50.00	35.71	91.67	66.33
Av. percent national reach	50.00	80.00	66.67	50.00	50.00	60.00	0.00	40.00	42.86	33.33	47.29

Note: The 'average percentage public' represents the proportion of case studies stemming from primary interaction with public and non-profit organisations. The 'average percentage national reach' represents the proportion of case studies where the impact was generated within one country only.

**Table 9: Summary Statistics for Bottom 10 Institutions**

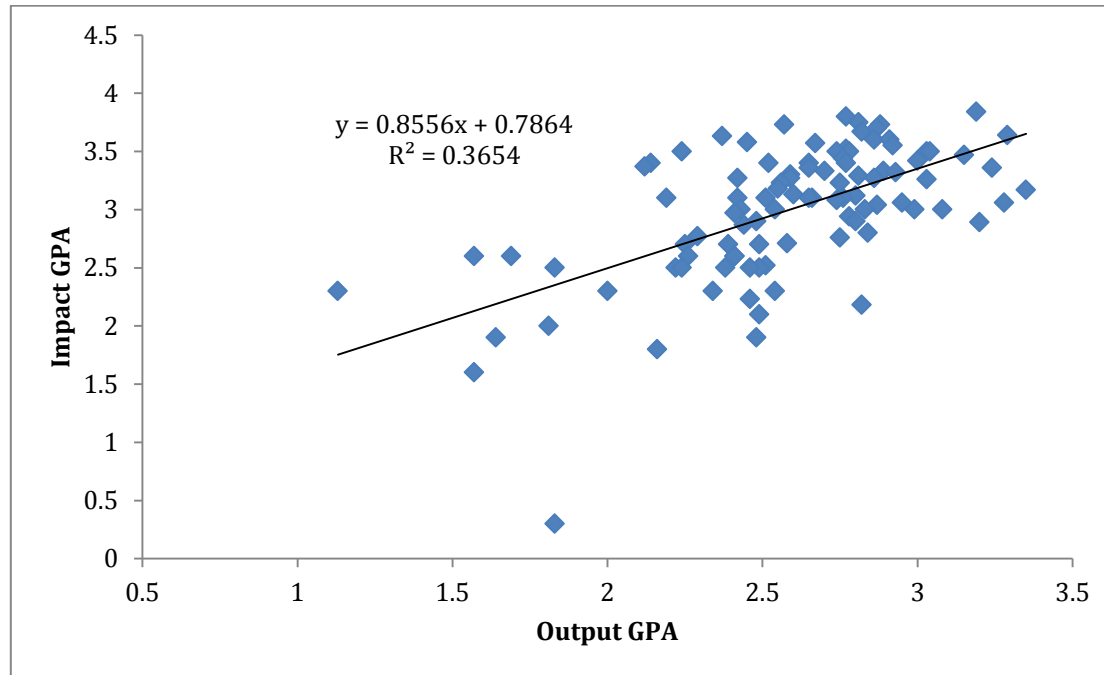
	Mean	Standard Deviation	Maximum	Minimum
Impact GPA	2.03	0.23	2.30	1.60
Output GPA	2.11	0.54	1.13	2.82

**Table 10: Impact Case Study Metrics for Bottom 10 Institutions**

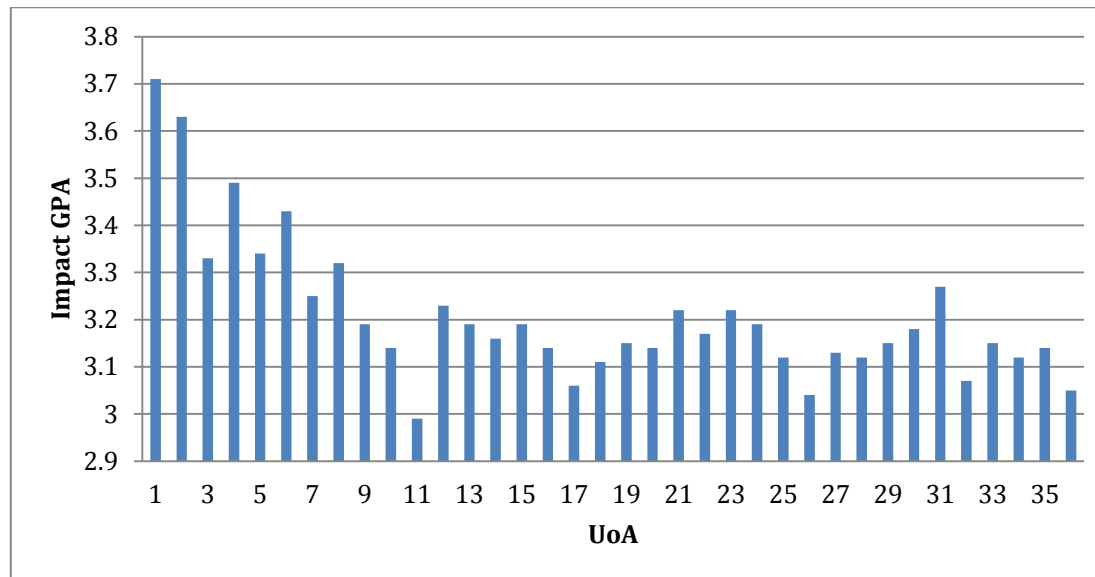
	UWS	Sunderland	Keele	Surrey	Bedfordshire	Northampton	Worcester	Teeside	Hertfordshire	Chester	Institutional average
No. of case studies	2.00	2.00	3.00	5.00	2.00	2.00	2.00	2.00	2.00	2.00	2.40
Av. no. of outputs	5.00	5.00	5.33	4.80	5.00	6.00	2.50	4.50	5.00	5.50	4.86
Percent of journals	70.00	70.00	50.00	62.50	50.00	91.67	80.00	33.33	80.00	72.73	66.02
Av. ABS score	2.67	1.25	2.67	3.00	1.80	1.67	3.00	3.00	2.83	2.50	2.44
Av. listed grant (£k)	100.36	19.50	41.11	27.60	609.42	0.00	46.00	0.00	62.39	0.00	90.64
Av. no. of key researchers	2.5	1.50	1.00	1.40	1.50	2.50	1.00	2.00	1.00	2.00	1.64
Av. longest time in post (yrs)	24.00	22.00	20.00	11.60	17.00	17.50	23.00	5.00	14.00	18.00	17.21
% of women key researchers	20.00	100.00	33.33	28.57	100.00	0.00	100.00	0.00	100.00	0.00	48.19
Av. percent public	100.00	75.00	83.33	50.00	50.00	100.00	75.00	100.00	100.00	100.00	83.33
Av. percent national reach	50.00	100.00	66.67	60.00	50.00	100.00	100.00	50.00	0.00	100.00	67.67

Note: The ‘average percentage public’ represents the proportion of case studies stemming from primary interaction with public and non-profit organizations. The ‘average percentage national reach’ represents the proportion of case studies where the impact was generated within one country only.

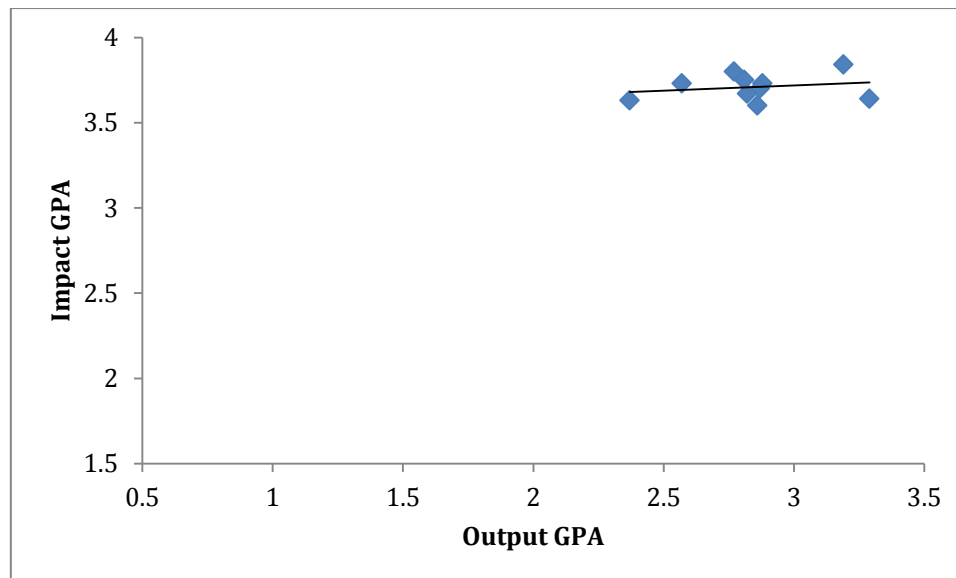
**Figure 1: Impact and Output GPA for UoA 19**



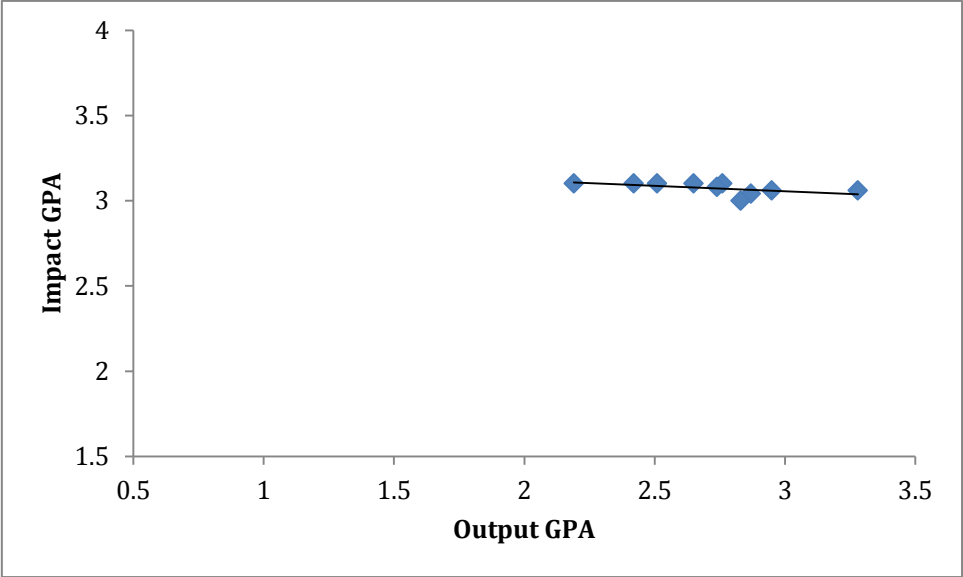
**Figure 2: FTE-adjusted Impact GPA by UoA**



**Figure 3: Impact and Output GPA for Top 10 Institutions**



**Figure 4: Impact and Output GPA for Middle 10 Institutions**



**Figure 5: Impact and Output GPA for Bottom 10 Institutions**

