

MEASURING THE IMPACT OF SOCIAL SCIENCE RESEARCH IN UK CENTRAL GOVERNMENT POLICY MAKING

Simon Bastow, Patrick Dunleavy* and Jane Tinkler

LSE Public Policy Group and *Government Department
London School of Economics and Political Science

Assessing the policy impacts of academic research has been plagued by unrealistic Platonic guardian images of the scope for influence. In more recent times case study approaches to assessing impact have been contaminated by the growth of what might be called ‘mini-leaderism’ images, a sort of minor key variant of neo-charismatic/transformational ‘leaderism’. We show that these criteria for assessment are simply disabling and impossibilist, and that no operational criteria can be derived from them. We next present extensive analysis of comparative impacts of different social science and humanities across UK central government. We use a web-based survey of a sample of 370 UK-based academics to operationalize more feasible, comparative metrics of academic influence on policy-making. Thirdly, we analyse extensive data on UK central government spending on commissioning research, to examine the demand and supply sides of the research impact question, demonstrating that the bulk of implementation-specific research goes to consultancies and mediators of various kinds and that university research plays a much more restricted and generally ‘early stages’ role in the social sciences. Finally we draw three key conclusions for the assessment of research influences using metrics. Any effective account is likely to be *relativized*, looking at a lot of data in a comparative way and using internally defined criteria of influence; *robust* (perhaps even *replicable*) in its use of data and development of metrics; and *realistic*, avoiding arbitrarily defined criteria lacking in any demonstrated empirical feasibility.

Assessing the impacts of higher education research has been projected into sharp focus by the bureaucratic organization of the REF 2014 process. HEFCE has collated more than 7,000 ‘impact case studies’ along with the over 200,000 ‘research outputs’ submitted by 52,000 academics. The REF’s obsession with census-like processes has combined with the insistence that these vast quantities of paperwork will be ‘assessed for quality’ by ‘experts’, and not just ‘eyeballed’ by panels and their external members. HEFCE uses no realistic metrics of the external impacts of research, but still deploys in impacts cases the same kind of using radically over-claiming language and categories needed for bureaucratic battles with the Treasury. As a result the piling up of impact case study has encouraged a ‘fairy tales of influence’ approach to assessing the external impacts of academic work.

We seek to counter this un-realism and bureaucratically-licensed hype by pursuing four arguments. First, we critique an ascendant (possibly now predominant) view of academic or research influences on public policy that is radically disabling or ‘impossibilist’, which views the desirable impact to be sought after (and rewarded with 4* grades) as a kind of ‘mini-leaderism’ – a scaled down version of the ‘leaderist’ or ‘transformational leadership’ ideas fashionable across the UK public services and business ideologies. We show that leaderism can only really be assessed in terms of three criteria – welfare gains, counter-factual impact, or leaders ‘dixing’ outcomes precisely. None of these criteria are remotely transferable or applicable to research or academic influences, or to evidence-based policy making.

Second, we use data from our detailed study of how social scientists influence UK policy-making at the central government level to try to map out the contours of what research influence actually looks like. We show how realistic metrics can be defined in a comparative manner. Third, we consider the demand for research from Whitehall departments and other central agencies, utilizing a huge amount of recently released data on how contracting for research actually operates.

Finally, we set out three recommendations for developing data-based metrics on how academics influence policy-making that follow a ‘three Rs’ logic. Effective metrics of academic influence are feasible only so long as they *relativized*, comparing within policy fields and types of academic work, especially within discipline groups; *replicable*, so that assessments of influence are capable of being cross-checked, compared and assessed by different research teams; and *realistic*, taking full account of the strong multiplicity of influences, actors and processes that always (necessarily) come together to shape policy outcomes, and rejecting criteria of influence that are simply disabling and possibilist.

1. Academic influence on public policy is not ‘mini-leadership’

The discussion of how far research evidence can influence public policy-making has long been plagued by absurdly Platonic ‘guardianship’ ideals, which envisage that political decision-makers could somehow be a kind of blank slate on which researchers or scientists could directly inscribe the ‘correct’ course of action to be followed (Bastow et al, 2014, pp. 30). Exorcizing such implausible and non-relevant criteria from public and academic discussions and debates has not been easy, and Platonic criteria still regularly crop up in commentary from research bureaucracies (like HEFCE and the research councils), from STEM scientists and professionals, and from the more naïve politicians or media commentators.

Against such over-simplification, a rich vein of literature illustrates the inherent complexity and ‘messiness’ in the relationship between research and policy (Weiss, 1995; Parsons, 2002; Sanderson, 2009). Concepts such as ‘evidence-based policy making’ (hereafter EBPM) have highlighted both an aspiration towards more systematic and informed policy making, but have also shed light on this inherent complexity and messiness. As Campbell et al. (2007) point out, EBPM has been an artefact of a broad understanding amongst policy makers (and researchers) of the importance of having evidence and knowledge in the mix (Weiss 1999; Nutley et al 2002a). But these authors also point out that even EBPM is often criticised for looking at this complex reality in far too simplified rationalistic (quasi-Platonic) terms (Campbell et al, 2007, p6).

Recently Cairney (2014) has argued that we need to modify EBPM into a more realistic idea of ‘bounded-EBPM’ (invoking the relationship of ‘bounded rationality’ to that of ‘instrumental rationality’). This involves building a much clearer understanding of the ‘science of policy making’ into our appreciation of how science and research filters into policy. ‘In different ways’, writes Cairney, ‘policy theories can be used to conceptualize the use of information in a large, messy policy process, far removed from that idea that “the evidence” has a direct input to a clearly definable policy process’ (2014, p10). The more we come to understand this complexity and variation in policymaking systems, the more uncertainty we have about controlling or predicting their behaviour (p13). In such circumstances it is ‘naïve’ and ‘inappropriate’ to persist with an attachment to rational policy making that implies top-down or central government control of the system (p13). So as our uncertainty in the policy process itself increases, so too the constraints on our ability to predict and isolate the effects of research and knowledge within.

Yet in defending realism against the scholastic over-reach of Platonic idealism, it is important that we also guard against a radically different kind of over-demand for research influence, one powerfully fuelled by ideologies imported into policy-making from the study of

business organizations. Another equally potent and far more fashionable source of misplaced criteria for assessing academic impacts stems from the out-growth of ‘leaderism’ as the highest form of managerialism. Managerialism prioritizes managerial control, (typically as offsetting professionalism, with its emphasis on professional expertise and client-centred autonomy). It denotes ‘the belief that all aspects of organizational life can and should be managed according to rational structures, procedures, and modes of accountability in the pursuit of goals defined by policymakers and senior managers’ (Wallace and Pocklington 2002, p. 68). Yet in business a cult of neo-charismatic or ‘transformational leadership’ has also grown up since the 1990s (Bass, 1991; Bass and Riggio, 2005) that seeks to supercede managerialist with a view of leadership as going further towards ‘inspiring others in collaborative endeavours’ (O’Reilly and Reed, 2011).

The same ideas have been extensively transferred to the assessment of leadership in government by influential commentators, such as Joseph Nye. In his 2008 book, *Powers to Lead*, we learn that:

Such leaders are called transformational in the sense of changing what would otherwise be the course of history. They raise new issues and new questions.... Transformational leaders empower and elevate their followers... Transformational leaders induce followers to transcend their personal interest for the sake of the higher purposes of the group that provides the context of the relationship... Transformational leadership also includes an element of “intellectual stimulation” (broadening followers’ awareness of situations and new perspectives’) (p.7 and 62)

Explicitly citing Bernard Bass, Nye stresses that ‘transformational leaders appeal to the collective interests of a group or organization ... using primarily soft power’ (p.62), as opposed to merely ‘transactional leaders’ who rely on interest-based responses to incentives or efforts to coerce behaviours. And Nye stresses too that ‘transformational leaders’ create counter-factual effects, that would not have existed without their interventions and actions. Their approach is thus to be contrasted with merely ‘eventful’ leadership that works by capitalizing on what would have happened anyway.

These background ideas have had a profound impact in the UK public services sector recently, triggering the rapid growth of a heavily reworked version of managerialism called ‘leaderism’ by O’Reilly and Reed (2008, 2010, 2011; 2012). They show that in the ten years to 2008 more than 13 times as many articles involving ‘leadership’ in their titles were published about public services and public management as in the ten years leading up to 1997 (1,428 articles as opposed to 124 in the earlier decade) (O’Reilly and Reed, 2010, p. 963). They summarize the core beliefs of ‘leaderism’ thus:

- in an endemic situation of competition, survival and progress require social

co-ordination;

- social co-ordination is best achieved through single or small groups of specially gifted and/or positioned individuals who lead;
- individuals that lead use particular moral, intellectual, interpersonal, cognitive, material, or politico-cultural, resources in order to achieve social co-ordination;
- such social co-ordination by those who lead places them in a pre-eminent role;
- to perform this role leaders must be empowered by giving them sufficient room to manoeuvre - the 'right', or authority, to lead (adapting Pollitt 1993, pp. 2–3);
- those who lead require effort and commitment from those being led;
- such social co-ordination leads to progress which benefits all those involved (O'Reilly and Reed, 2010, p. xx).

The view of research influences on public policy-making implied by the official discourse of the UK research councils and crystallized in stark form in HEFCE's REF criteria shows a strong influence of leaderist ideas, only somewhat reined back so as to apply to the academic context of influence. We can sum up this notion in simple terms as the view that public policy influences from research should show *mini-leaderism*, the same kind of 'transformational', collective self-interest inducing effects as organizational or political leaders, only scaled down somewhat to fit with academics' lowlier status in the scheme of things.

Thus HEFCE can claim to assess academic or research influence primarily in terms of the 'reach' and 'significance' of concretely observable and testified-to impact. We can trace mini-leaderism in the shifting of public policy into a configuration that would not otherwise have been achieved (Nye's counter-factual criterion of transformation leadership), in research 'making a difference'. According to HEFCE, the research councils and just about every Russell group university's PR machine we can also trace research influences of a neo-charismatic/transformational kind in the achievement of unambiguously valued social welfare improvements, in the attainment of improved efficiency, or more 'bangs per buck', greater or more improved outcomes that only the most churlish radicals would deny constitute welfare gains for society as a whole. Academia helps the whole of government and society to achieve results that go beyond the self-interest of participants, in the form of evidence that induces us all to pull together for the common weal. Thus research impact implies 'mini-leaderism', and this is the goal in pursuit of which 7,200 impact case studies have been written, re-written, finessed and extensively re-styled.

The problems of this approach are those that vitiate the parent concept of 'leaderism' itself. In the form that business personell find so plausible, and that a thousand airport books reiterate, leaderism is just an ideology, a cult or a quasi-religion. It perhaps has some therapeutic value in

inducing more ‘democratic’ or humanistic striving behaviours, or an effort to avoid the naked instrumentalism of ‘transactional’ leaders. But leaderism cults largely or completely lack any evidential basis – still less any worthwhile methodology by which evidence could even be sought, let alone gathered in. Leaderism thrives on parables and fables, and needs no more than the right emotional colouration to secure its adherents’ fervent advocacy, as Nye’s influential (2008) book neatly demonstrates.

In a critique of overweening leaderism, and a commentary on the ‘elusive quality’ of leadership, Dunleavy and Steunenberg (2006) set out three potentially operationalizable tests that might be used to try and tie down or measure a leader’s influence on policy-making:

A. *The welfare or proximity test* – a leader exercises influence when a before/after comparison shows that the policy outcomes has moved closer to the leader’s optimum position (her ‘bliss point’). Thus a successful leader is one who gets more of what she wants as a result of her interventions and actions.

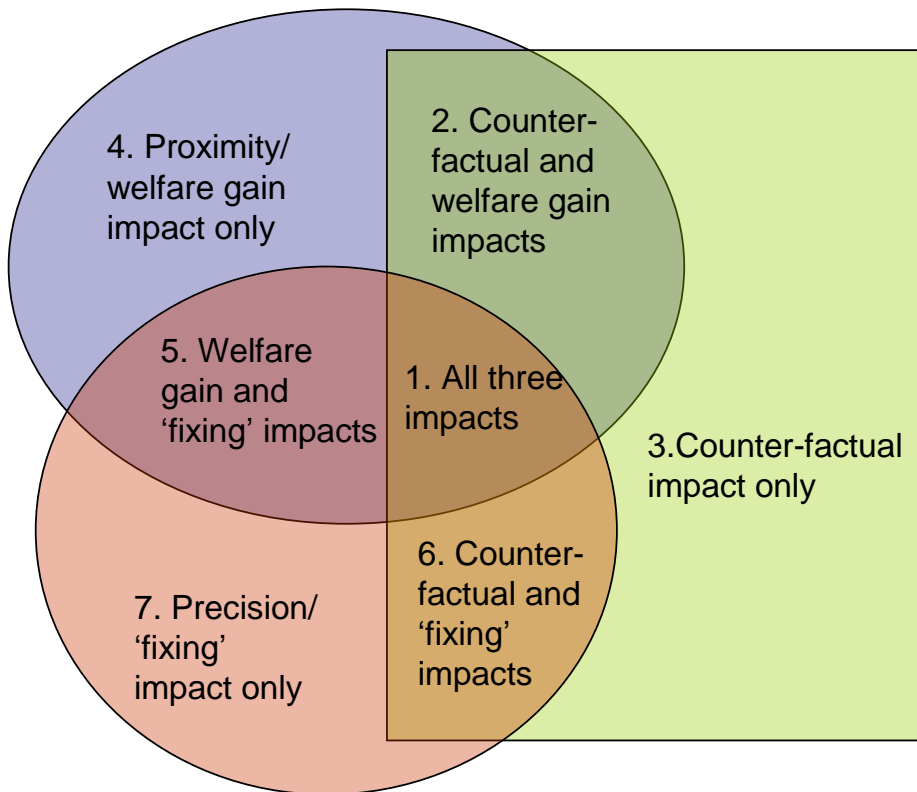
B. *The counterfactual test* – a leader exercises influence when her involvement makes a difference to what would have happened anyway? Here a successful leader is one who does not necessarily get more of what she wanted – for the environment may be changing in an adverse direction, or the whole context of the situation may be moving against her. However, we can chart a welfare improvement of some kind for the leader in terms of how the actual end point stands in relation to what would have happened without the leader’s actions.

C. *The ‘fixing or precision test* – a leader exercises influence to the extent that they are able to directly fix or finalize the outcome, limiting what happens to a defined location or zone within their preference space.

Figure 1 shows that there are seven possible combinations of these three criteria. Area 1 here occurs only if a leader can make plausible claims for influence under all the criteria, matched against the claims of all the other influences on policy making, including the influence of their own delegates or sub-leaders; of opponents; and of the environmental influence (Dunleavy and Steunenberg, 2006, p.26). It is very difficult for even the most powerful leaders to fulfil *all* of these criteria simultaneously in any consistent way.

Far more commonly perhaps a leader can claim to pass only a single test of influence. For example, she might make a welfare gain, but could just have been ‘lucky’ because other forces were strongly fostering this outcome anyway (area 4 in Figure 1). Or a leader might be able to demonstrate that because their involvement was late-on in the decision process she exercised a ‘last word’ influence in specifying the precise place where policy formation processes come to a halt, but without being able to say anything plausible on the other tests (area 7 in Figure 1). Alternatively a

Figure 1: Three criteria of leadership influence on public policy-making, and their intersections



leader might be able to claim that her interventions made a difference to outcomes, without being able to show either that her welfare was thereby improved, or that she personally fixed the final location of policy outcomes (area 3 in Figure 1). The observational evidence that leaders can often do great harm (rather than achieve the positive effects stressed in transformational models) is suggestive here.

Far more rarely leadership might be demonstrated on two of the three criteria at the same time, but not the third. Thus some leaders might be able to point to a welfare gain achieved from their point of view, and a counter-factual effect also, but without controlling the eventual outcome achieved in any detail – area 2 in Figure 1. Perhaps President Obama’s ‘saving’ of General Motors qualifies here, for his ‘car czar’ Steve Rattner (2010) explains how GM was saved (with \$39 billion of federal taxpayers’ aid) in a policy process where Obama apparently featured as an active participant only five times, and even then mostly in bit-part or formalistic roles. A leader who can show both that the outcome moved towards their bliss point, and that they personally fixed the outcome point in some detail (area 5 in Figure 1) can still be a long way off proving that they achieved any counter-factual impact. Perhaps they might achieve Nye’s ‘eventful’ leader status

from their ability to credit-claim (as Ronald Reagan none the less still did at the end of the Cold War in the late 1980s, despite the onset of alzheimers) but no more than that. And finally a leader might be able to show that she achieved some counter-factual effect, and thus fixed where outcomes came out with some precision, but yet not be able to show a welfare gain resulting (area 6 in Figure 1).

Now if leaderist claims are in fact so hard to stand up in any plausible way, it is unsurprising that the ‘mini-leaderist’ view of academic or research influence on policy-making is even less sustainable or realistic. For excellent reasons the Dunleavy and Steunenberg criteria are strongly inapplicable:

- *In welfare or proximity terms* we might rate research as being successful if policy outcomes move towards a configuration that is closer to the viewpoint that a given researcher or team expressed, and which we can also show was taken account of as part of the decision-making process. Here the model is that the academic ‘benefits’ if government in some way shifts towards recognizing or accepting their arguments. But what is the nature of this benefit – is it just plausible credit-claiming? And how (if at all) could that bear upon the achievement of social welfare gains? Academics, scientists and researchers are supposed to be disinterested – like other professional occupations standing aside from the main lines of class conflicts, and disinterestedly developing knowledge. Yet applying a welfare interpretation of ‘mini-leaderist’ ideas just seems to make them a partisan player like any other – just another member of an eclectic and temporary ‘advocacy coalition’ (Sabatier, 1987)?
- *In counter-factual terms*, academics or researchers as mini-leaders makes no sense. If indeed any counter-factual impact could be demonstrated, it is is inherently unlikely to be anything more than an eddy that in the political process, one that manifests and dies out almost as soon as it becomes visible as a thousand other influences are brought to bear. Mini-leadership here can only be the most evanescent swirl of Kingdon’s (2003) ‘policy stream’.
- Finally *in terms of fixing outcomes* academics are rarely involved in the detailed implementation of policies for very good reasons. The involvement of researchers typically comes early on, giving way to the roles of bureaucrats, expert officials, government lawyers, special assistants, political leaders and consultants in specifying what exactly is to be implemented and how. We show in section 3 below that most implementation-level research is dominated by consultancies and intermediaries, for reasons well explored in Bastow et al (2014). It would be credulous to think that in

advanced industrial economy, where the state is a complex multi-system needing dedicated management (Dunleavy, 2014) it could widely be otherwise.

So even for the most experienced and engaged academics, knowing whether or not their research has impact on policy can be an elusive thing. There is no guarantee that their efforts will result in actual discernible change in policy outcomes. This is not in any way a reflection of any aspect of their skills or research, but rather a consequence of the machinations and vagaries of the policy process itself. Of course there may be ‘windows of opportunity’ or ‘waves that can be ridden’, and if they get the timing right researchers may be able lead debates and discussions to a certain extent. Kingdon (2003), for example, argues that we may trace how political agendas are shaped by convergence of three streams, ‘political, policy, and problem’, and that researchers are likely to play their main parts in the ‘policy’ stream. As he puts it, ‘proposals, alternatives, and solutions float about, being discussed, revised, and discussed again’ (p. 272), and we can envisage the opportunities for academics to direct the debate with research and well-targeted argument (Bastoe et al, 2014, Ch. 6). But this may only offer a first step in the policy impact process, and there is no guarantee that any second or third steps will follow.

Waves and windows are usually fleeting, the ability to take advantage of them is often down to serendipity or luck, and it may be difficult for researchers to predict where or when these windows of opportunity will open up. Similarly, there is further difficulty for researchers in being able to predict which of their interventions or actions will be the one that makes a critical difference. Even if they are able to anticipate an opportunity, there is no guarantee that a particular occasion of influence will lead to a discernible outcome (let alone a desired one). Of course, the opposite is also true whereby considerable amounts of research that academics may think will lead to opportunities for impact end up ‘disappearing’ into the complex and continually churning ‘mill’ of information-processing and decision-making in government bureaucracies. As one academic told us:

The last thing I did for them was a review of family policies. We gave a presentation to some sort of working group to do with population. The work fed into a group. Any impact was going to be indirect. It disappeared into this group. The commission published a report on its website, then you don’t see or hear anything. They would have been commissioning a number of different studies, and you are one of many that might filter up to somewhere in the commission.

The burden of proving a counterfactual is especially intense for research involved in public policy, where often policy makers tend to decide on a policy direction, and then search for research and evidence that can help to substantiate that line. As one well-connected policy maker told us in interview about how major parties develop their election manifestos:

The first thing to say is that academic research primarily doesn't have an impact on that sort of thing. Typically you start at totally the other end of the telescope. You start by thinking 'what is it that British people want and need?' And then look around for some evidence to help you.

Of course policy can be evidence-based in this circumstance but only in a 'retro-fitted' way – and that is not necessarily the fault of academics. For example, an academic reflected on the impact of her work in the context of government determination to follow competition-based policies in the health service:

Whether that actually influenced decision making at all, I don't know. More likely they wanted to go for more competition anyway. That was their ideology. And so this piece of evidence comes up that suits them, and so they cite it. I'm a little bit sceptical about that. But then if they'd found the opposite, then maybe the Conservatives would have been a bit less gung-ho about the reforms. It is so intangible...it is impossible to pin down.

Other theoretical approaches in the impact literature remind us too that even when windows of opportunity for research materialize they are often unpredictable and follow non-linear dynamics patterns. Concepts such as 'punctuated equilibrium' (Baumgartner and Jones, 2004) and 'self-organizing criticality' (Bak, 1996; Brunk, 2001; Van Buuren & Gerrits, 2008) emphasize unpredictability, which again casts doubts on even claims of influence backed up by apparently plausible associations of research influence and policy change.

Rejecting 'mini-leaderism' does not in any way mean abandoning a quest for a closer woven and more plausible account of policy-making and of evidence-based and cognitive processes within that. Instead it means rejecting a completely disabling and counterproductive set of expectations and criteria, that cannot even be stood up unambiguously for the most powerful leaders exercising the most direct and final of influence upon policy-making. Rejecting mini-leaderism as a relevant template mean turning our backs only on the most 'impossibilist', ineffective and irrelevant of criteria for assessing research influence – and doing so for an approach that is far better, more relevant and systematically applicable.

Nor does rejecting the 'mini-leaderist' perspective mean dismissing the value of some case studies out of hand, far from it, but simply recognizing that a blanket application of case study or narrative approaches (especially undertaken at the behest of a quasi-academico/bureaucratic apparatus like the REF system), is highly unlikely to be systematically enlightening. Amassing 7,200 refrains on a single, bureaucratically-dictated tune will not allow us to compile a relativized or comparative picture of impacts across government as a whole. In acute contrast, the field of bibliometrics over the decades has allowed academics to measure the impacts of their work amongst their colleagues, and in recent years, 'alt(ernative)-metrics' have begun to explore the possibility for

using similar techniques in external non-academic realms of society (Hammarfelt, 2014). There are already many signs that this altmetric movement can offer a potential (less labour-intensive) alternative to peer-reviewed case studies for assessing research funding for UK universities (Jump, 2014; Dunleavy, 20xx).

However, it is still early days however in this field of endeavour, and so far academia has done relatively little to shed light on how research is accessed by organizations in society. Indeed, even after numerous decades of research evidence and impact in policy making, we still know relatively little about the scale and patterns of research usage and research impact across government, either from a demand or supply-side perspective. In the next two sections, we set out to show how it is possible to begin to fill this gap in our empirical knowledge about policy impacts.

2. How social science academics actually exert influence in government

The next step is to evaluate the extent to which social science research done by UK-based academics impacts on UK central government. Perhaps not surprisingly given the discussion above, our definition of impact is a modest one: ‘a recorded or otherwise auditable occasion of influence from university research upon another actor or organization’ (Bastow et al., 2014). The emphasis here is on counting the potential ‘occasions of influence’ rather than trying to chart the unmeasurable (perhaps unknowable) outcomes of this influence. Ours is a more encompassing conception of impact that accepts that researchers may generate a great many occasions of influence that, for one reason or another, do not transform into outcomes of change (see Bastow et al, 2014, *passim*). This should not make them any less important in terms of our assessment of academic efforts on impact.

As a proxy for these occasions of influence we look at the visibility of academics and their research. We rely here on unobtrusive and web-based search methods to build a picture of academic visibility across government. For further explanation of these methods, see our online methodology (Bastow et al., 2014, Ch. 3). We randomly sampled 370 academics based at UK universities, 270 of which were from social science disciplines and a further 100 chosen for comparative purposes from STEM disciplines (see Appendix 1 for a list of the disciplines and the number of academics in each). For each group of academics, we carried out two fine-grained searches using Google.

[1] We typed the full name of each academic into Google, and worked through the search results to record each time the academic was referenced on a website belonging to organizations that could not be classed as academic or related to academic publishing (i.e.

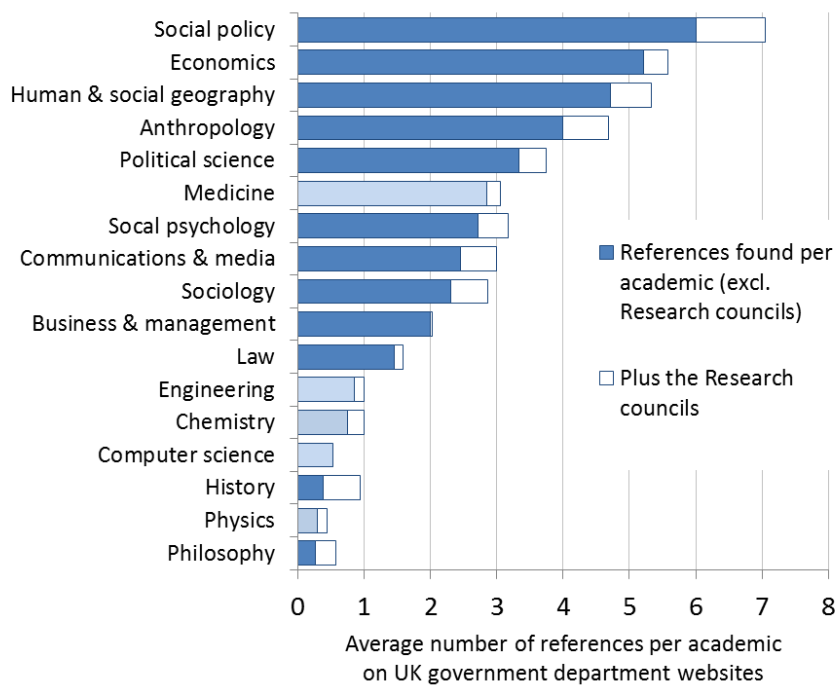
libraries, publishers, journals, universities). For each academic we were able to compile a composite picture of every mention found in any non-academic organization. We predominantly found English language references, but also coded up any foreign language references.

[2] As a second part to the research, we used Google advanced search to limit the domain to gov.uk and search specifically for each of our 370 academics. This allowed a much deeper search of the government domain than the previous more wide-ranging Google search, and served to supplement the findings. Our combined central government search allowed us to develop a comprehensive picture of all references made to our 370 academics or their research.

Figure 1 shows the average rate at which our academics were referenced by central government, distinguishing between different disciplines, and also between references made on departmental websites and those made to academics on the websites of research councils. It is immediately obvious that social science academics are much more visible on central government websites than their STEM counterparts. Only academics from medicine are able to demonstrate comparable visibility with the main group of social sciences, but as we move up the graph the average rate of referencing per academic almost double beyond the rate of medicine. A social policy academic, for example, is twice as likely to be referenced on UK central government websites as one working in the field of medicine. Looking at the four other STEM disciplines, it is clear that their individual researchers are comparatively much less visible than practically all other academics in our social science disciplines.

As ever, looking just at averages here could be potentially misleading. Looking at our social policy academics, for example, it is not the case that all of these are scoring 8 or 9 references each. In reality the distribution is far more scattered. In fact we can show that the distribution of all 370 academics closely resembles a 'power law', in that there are comparatively few academics that account for a large proportion of the references, and a long-tail of academics that have little or no visible presence in central government. Figure 2 illustrates this power law, plotting the

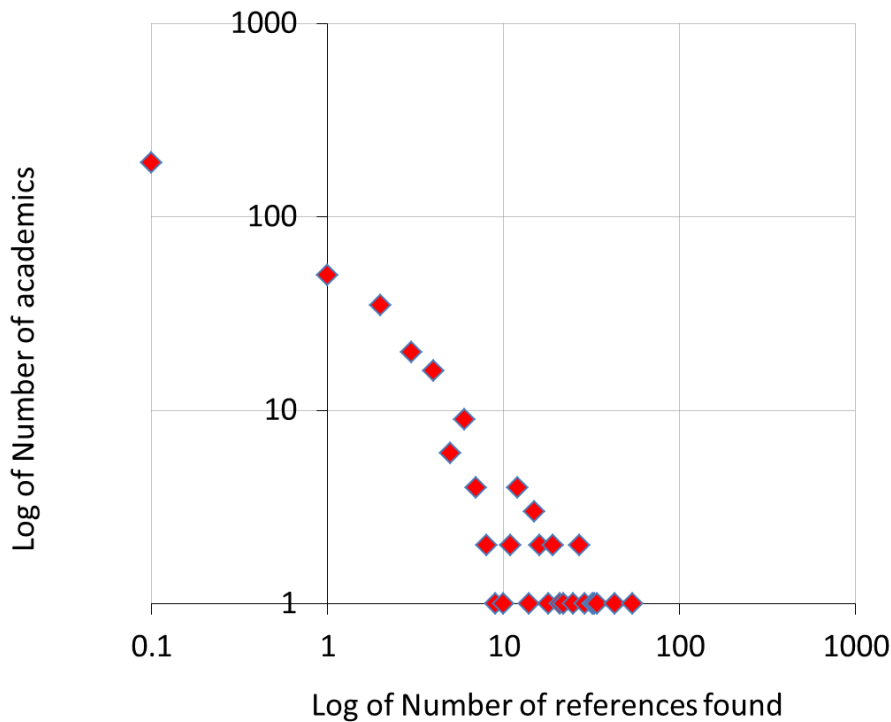
Figure 1: Average number of references per academic found on UK central government department websites, by discipline group



Source: LSE PPG analysis of dataset of 370 academics, 2012-13.

horizontal axis the (log of) number of the references per academic, and on the vertical axis, the (log of) number of academics with that amount of references. Exactly 191 academics have zero references in the central government web domain. We find 50 with one reference, 35 with two, 20 with 3, and so on. The most highly referenced academic is in social policy and has a total of 54 references. We observe variations of this power law distribution across all 17 disciplines covered. This result is in fact predictable from other work in citations research. We would not expect every academic in our sample to produce research that is useful or relevant for government, even more so given that this is a relatively narrow study of visibility in UK central government only. Many researchers might be highly impactful in other areas, in sub-national governments, grass roots public service delivery (like the NHS), the third sector or private sector. It is worth pursuing this

Figure 2: Average number of references per academic found on UK central government department websites, by discipline group



Source: LSE PPG analysis

Note: For 191 academics in our sample of 370 we found zero references to them or their research on UK central government department websites. This is illustrated in the Figure above by 0.1 references.

Inter-relationship here because it can help us understand more about the relationship between broader visibility and impact in society, and more specifically, impact in the narrower and more specific field of UK central government and policy. Our dataset allows us to compare for each academic their research visibility in UK central government with their research visibility across all other non-academic ‘external society’ domains. In this category we include all references to each academic found in all other areas of government (local and international), all third sector organizations, all private sector firms, all media and press, and all professional and representative associations (but not learned societies).

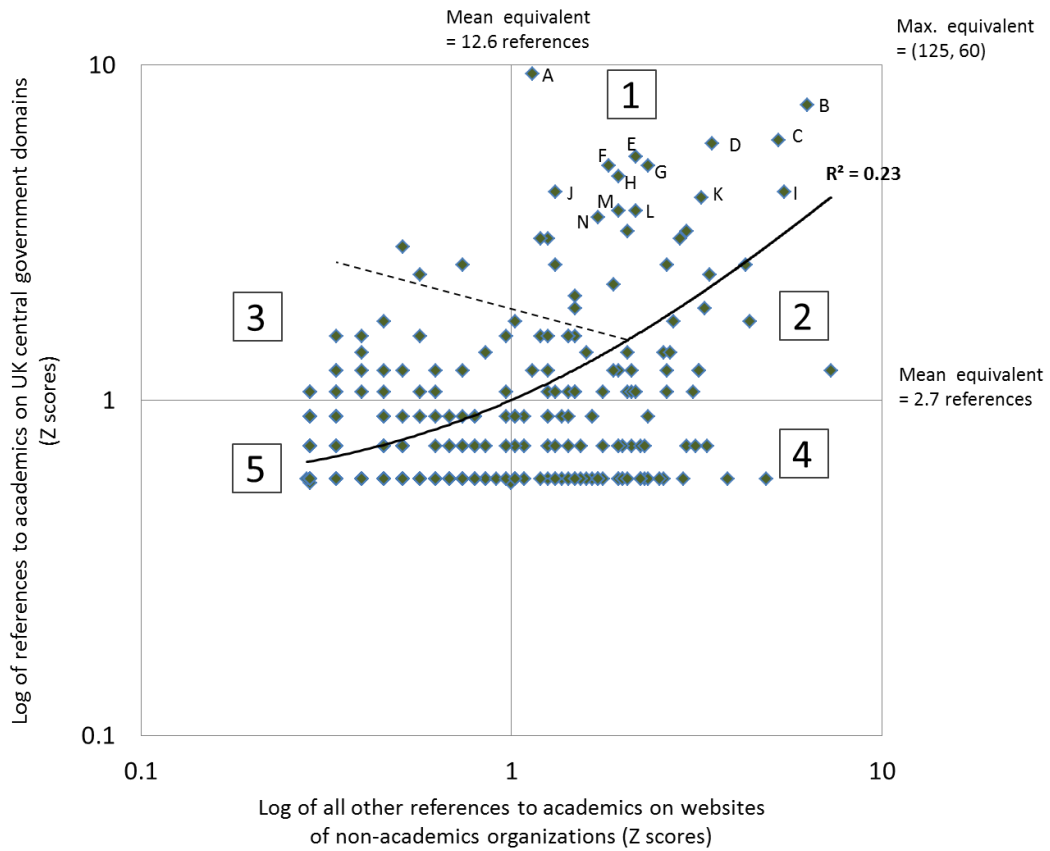
What might we expect to see in this relationship? First, it is likely that researchers academics who have strong visibility in wider societal domains would see some of this visibility ‘seep through’ into central government. Policy makers must start somewhere when finding research and expertise, and if academics are profiled visibly in these external domains, then it is not surprising if they are courted or consulted by policy makers. In this case, we would expect a strong(ish) relationship between both types of visibility. Second, we might find that certain

academics have developed niche areas of interaction with government departments, and that this leads to high visibility in government that does not necessarily translate into high visibility in society at large or with other sectors. We cannot expect the same seepage effects to occur from government outwards, especially if the areas of expertise involved are technical or highly specific. In this case, we might expect a group of academics to have high but concentrated visibility in government but not outside. Third, we might expect to find a group of academics and researchers who are highly visible in other sectors, but whose work does not translate or have relevance in the policy making world. Finally, we will inevitably find a group of academics and researchers who have little visibility in government of any kind and little visibility in society at large = for instance, because they have just started on the research ladder, or because their work has primarily intra-academic significance, or perhaps is not very memorable or worthy of notice.

Figure 3 plots the distribution of the logs of these two variables for all 370 academics in our sample. In order to make these two variables more directly comparable we have converted the total scores for each academic into 'z-scores', a process that normalizes both series around a common mean and standard deviation of 1. The intersection of the means of both series is located at the very centre of the graph. This point is equivalent to 2.7 references in the government domain, and 12.6 references in the wider societal domains. The trend-line intersects this mean point, and shows only a relatively weak (as expected) correlation between these two variables ($r^2 = 0.23$). As it is logged, the scales are not linear and hence the higher scoring points are contracted so that all observable data points can be shown on one diagram. Hence the top right corner of the graph is equivalent to a government score of 60 and a wider societal score of 125.

We can examine the distribution of the sample in more detail by counting up the number of academics in different sections of the graph (see Figure 4). Starting from the bottom left quadrant (labelled 5), around 42 per cent have zero visibility in UK central government as well as 'below average' visibility in wider societal domains. This figure increases to 57 per cent if we count all those academics that are below average on both counts – suggesting a relative lack of visibility in society and in central government too. In the bottom-right quadrant, we move into the realm of individuals that are above average in terms of their wider societal visibility, but have not, for one reason or another, translated this into high visibility in central government. This accounts for a further 19 per cent. More than three in every four of academics are in either areas 4 or 5 of Figure 3.

Figure 3: Average number of references per academic found on UK central government department websites, by discipline group



Source: LSE PPG analysis.

Figure 4: Average number of references per academic in our dataset found on UK central government department websites, by discipline group

Area	In UK central government (y-axis)	More widely amongst all other non-academic bodies (x-axis)	Total %	Cum. Total %
1	Above trend-line	Right of the dotted line	8	15
2	Below trend-line	Above average	7	
3	Above average	Left of the dotted & trend line	9	19
4	Below average	Above average	8	
4	No reference	Above average	11	57
5	Below average	Below average	15	
5	No reference	Below average	27	
5	No reference	No reference	15	

The individuals in area 3, demarcated by the dotted line and the trend-line above the horizontal axis, appear to have a relatively stronger visibility in UK central government than in other societal domains. Around 9 per cent of the total sample show signs, to varying degrees, that they have had their work picked up more readily by central government, but this has not translated into wider appeal. Meanwhile in area 2 the opposite applies, in that academics have built up a strong visibility in society, but have not quite managed to convert that into an equivalently strong visibility in central government. Some of these academics are highly experienced and impactful in these wider domains, but have not had their work picked up to same extent by government departments. Finally, in section 1 (everything above the trend-line and the dotted line) we find academics that that achieved high visibility in government, with varying degrees of success more widely. It is noticeable here how the majority of those academics that have achieved very high visibility in government have also tended to have comparatively high visibility more widely – and arguably vice versa. This amounts to around 8 per cent of all academics.

For the most visible academics in UK central government, it is interesting to look more closely at the possible trade-off between concentrated or diffuse impacts. We have labelled the fourteen most visible academics in central government (see Figure 5). These all have, to varying degrees, above-average visibility in wider societal domains, and it is likely that this wider visibility has done much to boost their specific visibility in government domains. But what are the signs that their visibility is concentrated or diffuse across central government? Around half of these academics can be shown to have almost all their references in one department or one specific area of policy. For example, Prof A in social policy has the highest government score, and practically all of his references are found in the area of Work and Pensions. On further analysis, it is clear that this academic has built up a strong series of consultancy reports and evaluation papers for DWP programmes, signs of a highly concentrated impact and a fairly lasting relationship with DWP policy makers over the years. Similarly, lecturer D in social anthropology shows concentrated impact in the Department for International Development, with very few references anywhere else in government. At the more diffuse end of the spectrum, lecturer B in communications and media, for example, has her work picked up across no less than seven departments, an impressive count for someone in relatively early stages of her career. As her position on the graph suggests, this is underpinned by a very strong visibility more widely, and it is likely here that policy makers have responded to this broader visibility and picked up her work where applicable.

Figure 5: Distribution of references to the fourteen academics most visible to / influential with central government departments in our data set

	Title	Discipline	Age group	Distribution of references across UK central government departments
A	Prof	Social policy	Late 50	High concentration in Work & Pensions
B	Lec	Comms & media	Early 30	Across 7 depts – incl. Communities & Local Gov, Education, Work & Pensions
C	SLec	Social policy	Early 60	Across 4 depts – majority in Health
D	Lec	Anthropology	Late 40	High concentration in International Development
E	Prof	Economics	Late 50	High concentration in Business & Innovation
F	Prof	Economics	Late 40	Across 7 depts – strong in DEFRA
G	Prof	Economics	Early 50	Across 8 depts – strong in Business & Innovation
H	Prof	Social geography	Early 50	High concentration in DEFRA
I	Prof	Social psychology	Early 50	Strong in Education & Devolved governments
J	Prof	Political science	Early 40	Strong in Devolved governments
K	SLec	Social psychology	Early 60	Across 6 depts – strong in Home Office
L	Prof	Medicine	Early 50	High concentration in Health
M	Prof	Social geography	Late 30	Across 7 depts – strong in DEFRA
N	Prof	Law	Late 50	Strong in Justice

3. The demand for research and knowledge in UK central government

Getting to an accurate picture of demand across government for university research has been hampered over the years by a lack of available comparable data. This is particularly true for social sciences and humanities research. For many years government accounts have published R&D investment statistics, but these have been biased towards STEM-based research, and have largely overlooked money spent by government on consultancy and other knowledge transfer. In 2006 the UK National Audit Office published a report on government’s use of private sector consultants, but again, this said little about the comparative sourcing of research from universities, research institutes, and think tanks (NAO, 2006). Similarly, the Higher Education Statistics Authority (HESA) collects data from universities directly about the amount of funds they receive from UK central government, but this collection process is fraught with inconsistencies across universities and problems of incomplete returns.

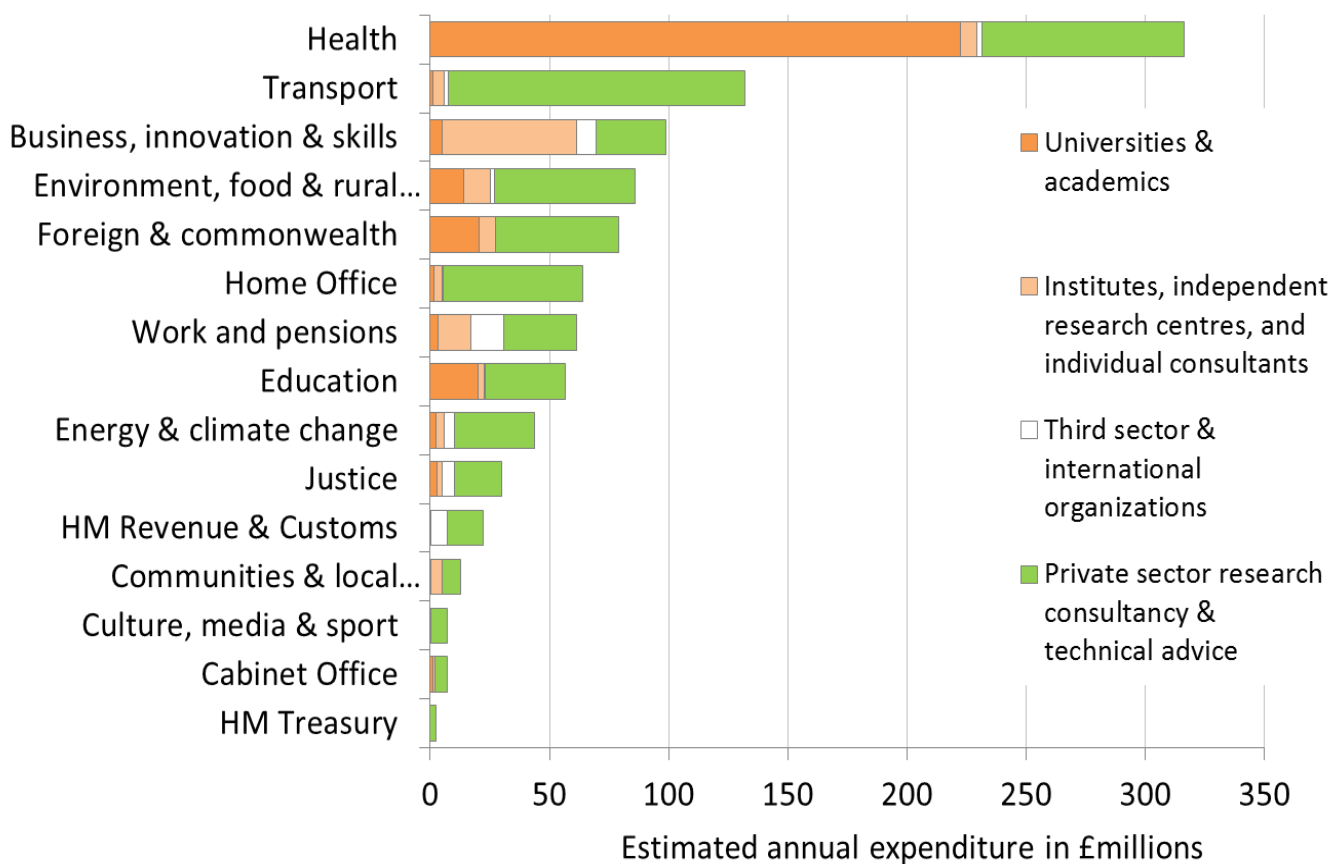
In recent years however the government ‘open data’ initiative, in particular the publication of government spending by line item and by department has made it possible to look in more systematically comparable ways at the demand in government for commissioned research (Bastow,

2014). For the first time we can estimate central government department expenditure on research and technical consultancy, and to be able to compare its provision across different types of providers. We downloaded one year of data from the data.gov.uk website for each major civilian government department, and through a combination of automated look-up and manual inspection and coding, we estimated for each department the amount spent on research. Of course, we are relying on the comprehensiveness of government data, and this is fraught with risks, not least through patchy reporting and inconsistencies in methods of categorisation. Nevertheless the fact that these figures have been through central accounting processes in each department suggest that their level of completeness will be consistently better than data provided by HESA.

Figure 6 summarizes at least five months of automated and manual coding processes by our team. It is a best-case estimate of the amount of money flowing to universities and academics for commissioned research, compared to private consultants, research institutes, and third sector organizations. It is important at this stage to note a couple of points. First, this data covers only research and knowledge commissioned directly by departments, and does not include ‘quality-related’ (known as QR)ⁱ research funding flowing to universities through the Higher Education Funding Council of England (HEFCE), or grants and funding flowing to universities and academics from government research councilsⁱⁱ. It is also important to note that these data only cover direct spending by central departments themselves, and do not cover spending by regionalised or local bodies to which the departments make transfer payments or grants. For example, figures for the Department of Health do not include very considerable research spending by regional NHS bodies, just as the figures for Department of Communities and Local Government (DCLG) do not include research expenditure for local government authorities.

It is clear from Figure 6 that universities and academics collectively are far from dominant suppliers of commissioned research to government departments. In only one policy area, health, universities and academics account for a majority of expenditure, and this can be explained by large-scale funding of clinical research in leading medical universities and NHS teaching hospitals. At least nine tenths of this health research expenditure flows to STEM disciplines, with only a very small minority absorbed by social sciences and humanities. In no other policy area does the university sector account for than half of the expenditure, and as the figure shows, in most the proportion of research expenditure flowing to universities is relatively small in comparison to the amount going to private sector consultants and legal services firms. We might argue here that many independent research institutes and think tanks would be important mediators or translators of academic research. Yet even if we include this group of actors, still in

Figure 6: Estimated expenditure by UK central government departments (only) on commissioned research and technical consultancy, 2012 to 2013



Source: PPG analysis of government spending data at data.gov.uk, 2012 to 2013.

Note: We have intentionally left the Department for International Development (DfID) out of this graph as its figures are considerably larger than all other departments, and it skews presentation. For the period July 2012 to June 2013, DfID spent an estimated £1.2bn for research and consultancy, broken down in each of these four sectors: Universities and academics £61m; Research institutes and research centres £103m; Third sector and international organizations £694m; and Private consultancy £354m.

only one area, business, innovation, and skills (BIS), does this combined percentage exceeds 50 per cent. The domination of the private sector in supplying research, consultancy, and technical expertise to government is clear. It is confirmed by a great many civil servants that we have interviewed across different departments.

These data provide a more complete picture of research spending across central government, but they will only tell so much about how departments prioritise or value research. How can we develop an equally relativized proxy for how government departments value research? One approach is to use our dataset of 370 academics, to track which departments are referencing the academics and their research on their websites, and to extrapolate from that an indicative metric that allows us to estimate research intensity for each department (see Figure 7). For each of the

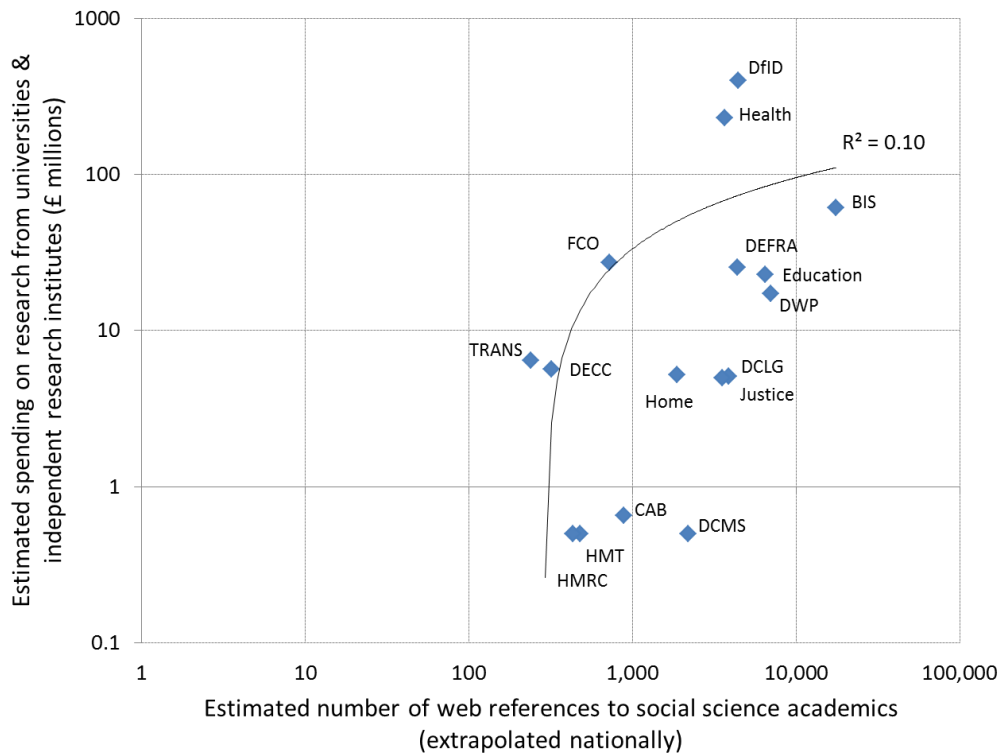
academics sampled, we have coded the number of references made to them in the UK central government web domain. For each of the 17 disciplines therefore, we are able to calculate a total number of references made to academics in each discipline by each government department. In other words, we are able to count how many times each department makes a references to each discipline, and because the dataset is a randomized sample of the population of all UK-based academics, we can present this as a proxy representation of the whole. We can extrapolate from this mini-sample to the UK population simply by multiplying the number of references by the total number of research academics in that discipline, and then weighting each discipline according to how many research academics a particular discipline has compared to the other 17 disciplines. We can use these scores as a general proxy for how much departments value the research outputs of our social science academics. The Department for Business, Innovation and Skills (BIS) is shown to have the highest research intensity score, and perhaps not surprisingly, nearly three quarters of these are made up of references to economics, business, and management academic research. Nevertheless, references were also found to a further eight of our academic disciplines, denoting a high spread of interest across different types of research knowledge. We might compare this to the Department of Work and Pensions (DWP), for example, where we find four fifths of research references in areas of social policy, media and communications, with a narrower spread of other references across four other disciplines. In the Department of Health, for example, the interest is very much on social policy and sociology, with 99 per cent of all references accounted for by these areas. This includes only social science disciplines, and clearly a vast proportion of actual health research references will be to clinical health disciplines (not included in this analysis).

We can draw together this analysis by plotting department spending and research referencing intensity on a spatial dimension. Do we find, for example, that the departments that spend most on academic research are more inclined to have high research intensity scores? Any clear relationship along these lines is not immediately obvious. Figure 8 above shows the logs of the two variables discussed in this section of the paper, and computes only a very weak r-squared score (0.1). Not surprisingly, it does not necessarily hold that big spenders on academic research should necessarily be the most prolific ‘referencers’ of academic research. In reality, the organizational dynamics are likely to be more complicated. We can however make some worthwhile observations about departmental variations, particularly by looking at clusters of departments on the graph. For example:

Figure 7: How visible is social science academic research on government department websites, by departmental grouping

	Total research refs	Subject 1	Subject 2	Accounts for % of refs	Disciplines referenced (out of 12)
Business, Innovation & Skills (BIS)	17,580	Economics	Business & management	73	10
Devolved governments	12,670	Political science	Business & management	60	10
Work & pensions (DWP)	7,030	Social policy	Communications & media	80	6
Education	6,510	Business & management	Economics	49	10
International development (DfID)	4,430	Anthropology	Economics	74	6
Environment, Food & rural affairs (DEFRA)	4,360	Economics	Human geography	72	5
Communities & local government (DCLG)	3,860	Sociology	Human geography	53	7
Health	3,660	Social policy	Sociology	99	3
Justice	3,560	Law	Social policy	40	11
Culture, Media & Sport (DCMS)	2,200	Sociology	Business & management	71	7
Home Office	1,880	Sociology	Social psychology	81	4
Cabinet (CAB)	890	Economics	Social policy	51	6
Foreign and Commonwealth (FCO)	720	Human & social geography	Business & management	77	4
HM Revenue and Customs (HMRC)	480	Business & management		100	1
HM Treasury (HMT)	430	Economics	Social policy	82	3
Energy & climate change (DECC)	320	Social policy	Economics	100	2
Transport	240	Business & management		100	1

Figure 8: Relationship between research funding by government departments, and visibility of academic research on departmental websites



- Treasury, HM Revenue and Customs, and the Cabinet Office appear to be low commissioners of academic research and be some of the least frequent referencers of research as well;
- Transport, DECC, and FCO all appear to commission more academic research but do not appear to reference it readily. In the case of FCO this may be explained by the secret or confidential nature of the research or advice. In the case of DECC and Transport, it may be related to the highly technical or specific nature;
- Two related clusters of departments on the mid-right of the graph, from Justice up to DEFRA, suggest modest levels of commissioning but comparatively higher levels of referencing and, perhaps, usage. This is particularly the case for DWP and Education;
- DfID and Health both appear to spend considerably more than other departments on research, even if they appear to reference it less frequently than other departments perhaps more geared towards social policy. Although DfID is an intensive commissioner and user of research, it is operating in a relatively narrow or specific field, and hence will tend to have relationships with a relatively smaller group of academics. For Health, these figures only

include social scientists, and adding in the vast number of STEM researchers whose work the Department would reference would have the effect of pushing health much further to the right on this graph.

4. Relativized, robust (perhaps replicable) and realistic – the ‘three Rs’ of impact mapping

The complexity and messiness of policy-making makes it very difficult to identify with any surety where causal links exist and suggest that academics can rarely if ever be ‘mini-leaders’ in the way that contemporary ‘leaderist’-contaminated official discourses often contrive to suggest. Our response here has been to put together a more modest, but no less sophisticated picture of impact across UK central government using unobtrusive web-based techniques. We do not claim that this should be the last word on impact, far from it. It is merely a first step on what will be a much longer process of developing better knowledge and information about how academics impact the thinking and actions of policy makers. But it seems clear that the way forward lies with approaches like this that are— *relativized, robust (perhaps even replicable) and show realism.*

1. Building a *relativized picture* of how different parts of government use academic research is not without some difficulties, but is relatively easy to make progress on with a degree of commitment and ingenuity. We have presented in this paper two original perspectives based on unobtrusive collection of data, both in terms of how visible academics are in central government and how much money central government spends on commissioning research. From this analysis we are able to get a sense about which departments, for example, are most research intensive, perhaps also which departments access research outputs most frequently and therefore value it more, and which academics and disciplines have been most integrated and visible in the government domain. Clearly, these do not tell us all we need to know about supply of and demand for research, but for the first time they give us at least an initial picture of visibility and value of research in a relativized way across UK government. Metrics and indicators such as these also provide little meaning in terms of their absolute values, but in their relationship to each other, *i.e.* their relativized values, we are able to infer considerable meaning. By presenting social policy as twice as visible in UK central government as, say, medicine, we are saying nothing about their absolute visibility per se but drawing conclusions about how in general academics from each broad discipline compare. This is something that case studies alone, however systematic, will never be able to do.

Another important aspect of relativization concerns the combination of methods used and the technique of triangulation between different methods. In this paper we have combined web-based unobtrusive analysis of our sample of 370 academics with an extensive trawl of government spending data, in order to match together the demand and supply sides. As the analysis of academics developed in section 2 and 3, the focus narrowed down to more qualitative aspects of impact. For example, we focused down on the 14 most visible academics in our sample, their disciplines, their visibility profiles, and so on, and it is easy to see how this can push towards a more mixed qualitative comparative analysis (QCA) approach. One could envisage looking in much more detail at these individual academics and learning more about their particular strategies for concentrated or diffuse impacts in government.

2. Using impact-tracing methods that are robust (and ideally replicable) is a second key component. In order to account for change in impacts over time, it is necessary to be able to revisit the methods used and replay them in a systematic way to generate comparable results. The more we move into the qualitative domain of case studies, the more difficult it becomes to find enough solid ground to replicate assumptions, techniques or researcher judgements. We are left only with contesting narratives and perspectives, highly sensitive to the role of initial starting points and hidden assumptions in shaping interpretations. We might well sustain some evidentially-based discourse here that is worthwhile and has some potential to cumulate over time.

But the methods we have used here have an advantage in this sense. They rely on standardised coding frameworks that are applied to government websites in exactly the same way. Theoretically it would be possible to replay this analysis in five or ten years, and have at least a basic platform for replicating these visibility tests. We would of course expect new web-based landscapes to have grown in that time, in particular social media profiles of departments and agencies, but fundamentally, exactly the same systematic techniques would be applicable. To some degree we would be able to assess to what extent academics had become more or less visible over time. The fact we are limiting the analysis to a defined web domain also means that we are able to control fluctuations in the size of the domain and hence the level of research intensity.

There are however legitimate voices of scepticism. To what extent can we rely on government websites to be an accurate reflection of actual impact activity? What are the risks that such unobtrusive analysis might under-report or over-report actual activity? We might reasonably assume that academics spend much of their time advising governments in confidence or behind the scenes – in ways that would not get written up on department websites. For example, our government spending data above suggests that HM Treasury spends comparatively little on

commissioned research. But it would be unrealistic to suggest that many Treasury officials do not have close relationships with senior academics in all sorts of ways. We must be careful therefore of reading too much into un-triangulated sources of evidence. Clearly, HM Treasury do rely on expert advice through committees, working groups, and informal briefings, and this may not necessarily show up in web-based domains. Nevertheless, our coding covers a wide range of ‘potential’ occasions of influence and it becomes very difficult over time for organizations to keep these sorts of expert council links entirely secret. Membership lists of committees are often published, or these academics concerned are invited to speak at departmental events or seminars, and their names appear on related documentation. Over time, if academics are engaging in sustained and significant interactions with government departments, these traces inevitably become visible in web-based format, even though traces of private briefings may never see the light of day. It is unlikely however that academics will exclusively brief in private – the social relations beneath suggest that they would be asked to speak at a department event or perhaps write commentary piece. In this sense, any traces may be treated as ‘gold dust’ – one or two may well be all it requires to indicate a more interaction and social relations behind the scenes.

Similarly, we might legitimately point out that government departments do not tend to cite systematically in the policy documents any academic research that they produce – at least with the same level of diligence that academics are expected to cite other academics. Often policy may emanate from scientific research, but in the churning and development of policy discussions, mini-decisions, further deliberations, and a great many potential advocates and stakeholders involved, the link is often broken and these actual citations are neither acknowledged nor visible. It is worth pointing out that a great many of our references recorded in this analysis have been in government policy documents or working papers, so to say flatly that governments do not cite is perhaps exaggerating the problem. Nevertheless, if we follow this sceptical line and accept a certain level of under-citation in government, we can also make the point that each trace or grain of referencing again becomes a kind ‘gold dust’ sign of likely impact. It is clear from the data above that the average reference score for academics is 2.8; in other words, the academic does not have to register many references in order to get an above-average score. Of course, there may be a risk here that academics might hit a jackpot by scoring a great many references in one hit, perhaps in a piece of research that is continually cited, and that this could easily skew them well into the higher realms of impact (while their actual impact is short-lived and rather artificial).

3. Both the Platonic and the quasi-ideological ‘leaderist’-influenced, or rather mini-leaderist, approaches to understanding impact essentially posit criteria for assessing the impact of research on

policy making that are disabling and impossibilitst from the outset. Hence it is all the more important that the way forward is predicated on a realistic appreciation of the diversity, complexities and subtleties of policy making in advanced industrial democracies, and that embody from the outset an intelligent and critical awareness of the inherent limitations on academics as constrained mini-leaders (at best) of change. It is highly unlikely that any one piece of research, or even a whole connected stream of research, will be strong and influential enough to reshape an entire policy and decision-making system from the outside. But in the social sciences as a discipline group, cumulative effects often operate - this piece of research will lead to another, and the massing of evidence across many studies and research team (and perhaps across many contexts or countries also) will sometimes have the effect of generating an opportunity to ‘open a door’ or contrive a ‘serendipitous’ meeting, and so on.

An important and inherent problem may be that it is difficult for academics to know which ‘grains’ of research are the ones that will have most impact. As things are rarely linear (although they may seem so in hindsight), perhaps the most rational response of external academics to this uncertainty is to broadly invest care and commitment in *most* opportunities or occasions for influence. In sandpile that is a complex system it is not possible to know exactly what kind of disruption each grain of sand (or action) will have on the overall system, and the last grain of sand that triggers a reconfiguraing avalanche of sand is inherently no more critical or perhaps significant than the hundreds that preceded it. Many policy-world processes essentially rely on boosting the visibility of research and increasing the probability of making the ‘small world’ connections that often get explained away as serendipity (Watts, 1999). As researchers of impact, it is no less difficult to identify these ‘decisive’ grains of impact, and hence we designed a method to track all possible traces in whatever size or shape they appear. If we are systematically searching a government website for signs of researcher impact, we cannot know which trace or grain of impact is critical in this respect. We can only record each one in a fine-grained way, and look at the overall cumulative picture. By building up a cumulative picture, grain by grain, we can get to a more encompassing story than by reifying any one grain (or series of grains) in particular.

Conclusions

It is probably not an exaggeration to say up to now that knowledge about the societal impacts of the social sciences and humanities over the decades has relied predominantly on case studies or ‘fairy-tales of influence’. Case studies of course have their part to play, but by nature they present impact in narrative form, something akin to Walter Benjamin’s ‘angel of history’ that is blown forward by

a great wind into the future causing great tumult and chaos, while all it sees in its wake is the illusion of a logical series of events. Tales of mini-leaderism in academia can often run the risk of reproducing this kind of fallacy, hunting for an apparently series of causal processes whereby academic research is done at point *A* and by dint of its persuasiveness and timing has visible impact at point *B*. The problem is that this assumed causality is desperately hard to prove, not to mention isolate in separation to many other factors or causal influences that come to bear on policy decisions.

Yet out of the sheer massing of REF impact case studies may yet emerge a unique ‘big data’ opportunity to map comprehensively the ways that higher education research has impact on policy making, an analysis that could potentially be very fruitful. Until that time we should move on from case studies alone to make the maximum feasible use of more comparative and systematic methods and metrics for assessing the relative influence of academic research across discipline groups, universities and by comparison with other sectors and components of advocacy coalitions around public policy. This is a journey that has just begun, but one where rapid progress is feasible if we start out with appropriate expectations and methods.

Bibliography

- Bak, P. (1996). *How Nature Works: The Science of Self-Organized Criticality*. New York: Copernicus.
- Bastow, S. (2014). ‘Open data sheds light on how universities are minority providers of commissioned research to government’. Blogpost on LSE Impact blog, 20th March 2014.
- Bass, Bernard M. (1990) From transactional to transformational leadership: Learning to share the vision’, *Organizational Dynamics*, vol 18, no 3, pp. 19–31
- Bass, Bernard M. and Riggio, Ronald E. (2005) *Transformational Leadership* (xx; Psychology Press).
- Bastow, S., Dunleavy, P. and Tinkler, J. (2014). *The Impact of the Social Sciences: How academics and their research make a difference*. London: Sage Publications.
- Baumgartner, F., and Jones, B. (2004). ‘A Model of Choice in Public Policy’. *Journal of Public Administration Research and Theory*, 15(3), 325-351.

- Brunk, G. (2001). 'Self-Organized Criticality: A New Theory of Political Behaviour and Some of Its Implications'. *British Journal of Political Science*, 31(2), 427-445.
- Cairney, P. (2014). 'Evidence-based Policy Making: If you want to inject more science into policymaking you need to know the science of policymaking'. Paper presented at the Political Studies Association Annual Conference 2014, Manchester.
- Campbell, S., Benita, S., Coates, E., Davies, P., & Penn, G. (2007). *Analysis for Policy: Evidence-based policy in practice*. London: HM Treasury.
- Davies, H., Nutley, S., and Smith, P. (Eds.). (2000). *What Works? Evidence-based policy and practice in public services*. Bristol: The Policy Press.
- Dunleavy, Patrick (2010) 'New worlds of political science', *Political Studies*, 58(1): 239–65.
- Dunleavy, Patrick (2014) 'The State is a Multi-system: Understanding the Oneness and Diversity of Government', Paper to the UK Political Studies Association Conference, 15 April. Available at: <http://eprints.lse.ac.uk/56492/>
- Dunleavy, P. and Steunenberg, B. (2006). *The Elusive Quality of Leadership: Leaders and Delegates*. Government Department. London School of Economics. London.
- Flyvbjerg, B., Landman, T. and Schram, S. (Eds.). (2012). *Real Social Science: Applied Phronesis*. Cambridge: Cambridge University Press.
- Hammarfelt, B. (2014). Using altmetrics for assessing research impact in the humanities. *Scientometrics*, <http://dx.doi.org/10.1007/s11192-014-1261-3>.
- Hood, C. (1998). *The Art of the State: Culture, Rhetoric and Public Management*. Oxford: Oxford University Press.
- Jump, P. (2014). Overhaul of metrics usage could cut frequency of REF, *Times Higher Education Online*, 3 April.
- Kingdon, J. (2003). *Agendas, Alternatives and Public Policies* (2nd ed.). Boston: Little Brown.
- National Audit Office (NAO). (2006). *Central government's use of consultants*. A report by the Comptroller and Auditor General, HC128 Session 2006-7. London: HMSO.
- National Audit Office (NAO). (2007). *Government on the Internet: Progress in delivering information and services online*. Report by the Comptroller and Auditor General HC529 Session 2006-2007. London: UK National Audit Office.
- Nutley, S., Davies, H., & Walter, I. (2002). *Evidence Based Policy and Practice: Cross Sector Lessons from the UK*. ESRC UK Centre for Evidence Based Policy and Practice.
- Nye, Joseph. (2008) *The Powers to Lead* (New York: Oxford University Press).
- O'Reilly, Dermot and Reed, Mike. (2008) 'Leaderism' and UK Public Service Reform Paper for EGOS 2008 Colloquium VU University Amsterdam. Available at:

<http://www.esrc.ac.uk/my-esrc/grants/RES-000-23-1136-A/outputs/Download/59fb7df7-dd88-4659-af3a-c817f4f40b84>

- O'Reilly, Dermot and Reed, Mike. (2010) "'Leaderism": An Evolution of Managerialism in UK Public Service Reform', *Public Administration* Vol. 88, No. 4, 2010 (960–978)
- O'Reilly, Dermot and Reed, Mike. (2011) 'The Grit in the Oyster: Professionalism, Managerialism and Leaderism as Discourses of UK Public Services Modernization, *Organization Studies*, 32(8), pp. 1079 –1101
- O'Reilly, Dermot and Reed, Mike. (2012) "'Leaderism" and the discourse of leadership in the reform of UK public services', In Christine Teelken, Ewan Ferlie, Mike Dent (eds) *On Leadership in the Public Sector: Promises and Pitfalls* (London: Routledge, 2012).
- Parsons, W. (2002). 'From Muddling Through to Muddling Up: Evidence-Based Policy Making and Modernization of British Government'. Queen Mary, University of London. London.
- Pawson, R. (2006). *Evidence Based Policy Making: A Realist Perspective*. London: Sage Publications.
- Rattner, Steven. (2010) *Overhaul: An Insider's Account of the Obama Administration's Emergency Rescue of the Auto Industry*. (New York: Houghton Mifflin Harcourt).
- Rutter, J., Marshall, E., and Sims, S. (2012). *The 'S' Factors: Lessons from IFG's policy success reunions. A report by the Institute for Government*. London: Institute for Government.
- Sanderson, I. (2009). Intelligent Policy Making for a Complex World: Pragmatism, Evidence and Learning. *Political Studies*, 57, 699-719.
- Sabatier, Paul (1987) 'Knowledge, policy-oriented learning, and policy change: an advocacy coalition framework', *Knowledge: Creation, Diffusion, Innovation*. 8 (June): 649–92.
- Sechrest, L., Schwartz, R. D., Webb, E. J., and Campbell, D. T. (1999). *Unobtrusive Measures: SAGE Publications*.
- VanBuuren, A., and Gerrits, L. (2008). Decisions as Dynamic Equilibriums in Erratic Policy Processes. *Public Management Review*, 10(3), 381-399.
- Watts, D. J. (1999). *Small Worlds: The dynamics of networks between order and randomness*. Princeton and Oxford: Princeton University Press.
- Weiss, C. (1979). 'The Many Meanings of Research Utilization'. *Administration Review*, 39(5 (Sept - Oct 1979)), 426-431.
- Weiss, C. (1995). *The Haphazard Connection: Social Science and Public Policy*. Ithaca, NY: Cornell University Press.

ⁱ Quality-related or QR funding consists of transfer payments directly to universities in recognition of their research excellence. This is not directly feeding into specific research outputs in the same way as commissioned research. It

contributes to the underlying research environment, helps to employ excellent researchers who publish and have impact. In 2012-13, we estimate that QR funding to all UK universities totalled around £1.5 billion, around £312m of which flowed to social science departments (Bastow et al., 2014).

ⁱⁱ Research councils include Arts and Humanities Research Council (AHRC), Biotechnology and Biological Sciences Research Council (BBSRC), Engineering and Physical Sciences Research Council (EPSRC), Economic and Social Research Council (ESRC), Medical Research Council (MRC), Natural Environment Research Council (NERC), Science and Technology Facilities Council (STFC), Royal Society, British Academy, and the Royal Society of Edinburgh. We estimate that research councils transferred around £1.6 billion to UK universities in 2010-11.