

Research metrics, labor markets, and epistemic change: evidence from Britain 1970-2018

Juan Pablo Pardo-Guerra

Department of Sociology

University of California, San Diego

9500 Gilman Dr, La Jolla, CA 92093

jpardoguerra@ucsd.edu

Abstract

In this paper, I study the effects quantifying knowledge on the organization of disciplinary fields. Using an original dataset of about 44,000 British social scientists and bibliographic information for their published, peer-reviewed articles, I show that the introduction of standardized research evaluations disturbed local academic labor markets in British higher education, leading to patterns of interinstitutional mobility that altered the epistemic diversity of social science disciplines and the organization of their academic fields. Much like a market-based intervention, research evaluations lead to a form of epistemic matching that has distinct consequences on how knowledge is generated. In particular, when evaluations affect organizational units (such as academic departments) and stress disciplinary norms, they foster forms of isomorphism that lead to reductions in a discipline's thematic diversity and a more homogeneous structure for the field.

Keywords: Quantification, knowledge, epistemic change, institutions, labor markets

Research metrics, labor markets, and epistemic change: evidence from Britain 1970-2018

What happens when we measure the worth of scientific knowledge? For economists, ascertaining the quality of scientific claims is arguably a precondition for an efficient “market for ideas”. By revealing the “degree of success” of individual publications, researchers, and organizational units, evaluations provide greater efficiency to the complex and uncertain allocation of resources to innovation (Arrow 1969). The expanding, instituted efforts to quantify scientific outputs through a panoply of measuring devices are a testament to this logic (Gingras 2016; Godin 2004), seeking to provide instruments that both identify the value and contributions of knowledge to particular disciplinary fields while allowing for their management in organizations devoted to research and higher education. To be reliable, nevertheless, it is acknowledged that evaluations must refrain from distorting science’s practices of production (David and Foray 2015; Hicks et al. 2015; Stephan, Veugelers, and Wang 2017). In their most basic form, evaluations should act as simple sieves that distinguish the quality of inputs without modifying the operation of the freely competitive and self-regulated institutions of research (Merton 1973). Evaluating knowledge certainly shifts the balance of resources, but to be effective, it must remain external to the establishments that generate knowledge itself, providing measures that reveal what is valuable and available without transforming the makeup of epistemic fields. To wit, impact factors may well shape where authors decide to publish, but not the lines of inquiry that they pursue throughout their careers or the structure of the disciplines they inhabit.

Sociologists of science demonstrate, however, that the context of knowledge production can have distinct effects on both who knows and what is ultimately known (Barnes, Bloor, and Henry 1996; Bloor 1991; Collins 2010; Frickel and Gross 2005; Schofer 2003). As evaluations become endemic to contemporary science and its management, they do not only disclose putative values but fundamentally transform how authors and institutions conceptualize worthy projects and invest in the production of knowledge about

the world. In recent years, much work has focused on understanding how evaluations shape the experiences of scholars in making sense of their research and work-identities. From studies of reactivity (Espeland and Sauder 2007) to questions about the impact, design, and democratization of indicators (Hicks et al. 2015; Ràfols 2018; Rijcke et al. 2016), there is now an growing body of literature documenting the effects of measuring knowledge on the practices and imaginaries of contemporary academics (Burrows 2012; Clarke and Knights 2015; Gill 2014; Hammarfelt and De Rijcke 2015; Vostal 2015).

Yet beyond the individual experience of the quantified scholar and relatively sparse anecdotal claims, there is little evidence that research evaluations change the *organization* of epistemic fields. Does counting, measuring, ranking, or valuing disciplines change their composition? Answering this question is the task I take on in the following pages. In this paper, I address the question of how measuring the worth of knowledge changes the intellectual organization of disciplinary fields. For this, I provide novel evidence of the effects of a long-standing, paradigmatic research evaluation conducted by the British state since 1986. I focus on the social sciences as particularly reactive and less infrastructurally-bound disciplinary formations in academia. As I show, research evaluations produced a competitive field that perturbed the patterns of interinstitutional mobility of social scientists (Allison and Long 1987; Fuller 2008), leading to changes in the epistemic makeup of disciplines. Specifically, these changes consist of a decrease in the thematic and organizational diversity of disciplines, resulting in more homogeneous fields of knowledge. Using an original dataset of career transitions and academic outputs of over 44,000 British academics, I show compelling evidence that this process of homogenization (which I refer to as ‘epistemic matching’) was catalyzed by the way knowledge was evaluated by the apparatus of the British state. Quantification, in sum, changed the British social sciences at their very core.

Changing knowledge through research evaluations

Organized evaluations of science are a historically recent phenomena tied in large measure to the problem of managing the expanding research, development, and higher education

efforts of the second half of the twentieth century. The urgency to direct the growing crews of scientists and institutions across an increasingly complex postwar research and higher education ecology had notable consequences. The most prominent, perhaps, is the emergence of fields devoted to studying science through a scientific lens. While quantitative studies of scientific productivity exist since at least the 1920s (Lotka 1926), much of the contemporary literature on research evaluation traces its origins to policy-oriented efforts in the late 1960s to measure scientific outputs to better manage public investments in innovation under increasing fiscal constraints (Gingras 2016; Godin 2006; Sugimoto and Larivière 2018). Examples of this include the development of scientometric and bibliometric studies that, by quantifying and modelling patterns of knowledge production, dissemination and use, seek to partly determine where resources are best spent. These developments were widely successful, particularly as they were coupled to the creation of metrics that travelled intra and transnationally between funding agencies as heuristics that facilitated comparing heterogeneous sets of scholars, institutions, and disciplines. Today, the use of metrics as means for evaluating scholarly quality is a standard approach throughout the world (Abramo, D'Angelo, and Caprasecca 2009; Geuna and Martin 2003; Hicks 2012).

Despite their pervasiveness, this wealth of evaluations is oddly ignored by the sociologies of knowledge and science which, while attendant to situated practices of knowledge production, have yet to examine how they shape and affect the evolution of scholarly fields over time. Work on the related field of research evaluation (Rijcke et al. 2016; de Rijcke and Rushforth 2015) has nevertheless produced numerous discussions about the measurement of science, including important studies on how the use of metrics and indicators shape the practices of scientists and their host institutions. Reflecting some of the key concerns and debates from the sociology of quantification (Berman and Hirschman 2018; Espeland and Stevens 2008; Mennicken and Nelson Espeland 2019), contemporary studies of research evaluations have located these within a “broader surge of accountability measures” (Rijcke et al. 2016:162) across public institutions, where audit logics transform the structural affordances and strategic behaviors available to actors. By quantifying, making visible, and incentivizing certain publications, outputs, and research

activities over others, evaluations and their associated logics are seen to affect the production of knowledge in three distinct ways.

First, because they are public and visible exercises, research evaluations are deemed to promote *strategic behavior* among scientists who, knowing the rules of assessment, adopt behaviors that seek to maximize their potential scores. Research by Butler (2003, 2005) shows, for example, that the introduction of an evaluation system in Australia linking research funding to peer-reviewed journals led to a noticeable rise of ISI-indexed publications, with a heavy concentration in the less selective quartiles. Similar changes in publication patterns have been documented elsewhere and in connection to the implementation of research evaluations. Bibliometric studies in economics show, for instance, that the use of scientometrics by members of the disciplinary community lead to both the reinforcement of patterns of knowledge production and prestige and the stabilization of dominant scientific paradigms (Aistleitner, Kapeller, and Steinerberger 2018).

While evaluations may lead to strategic planning of publications, this needs not affect the contents of knowledge production. A more direct mechanism of epistemic change, however, is provided by a second effect, namely, a shift in the focus of individual research decisions to privilege those most congruent with evaluations (what (Rijcke et al. 2016) identify as '*task reduction*'). Gläser and Laudel (2016) convincingly show that tensions between formalized evaluation criteria and a scholar's own conceptualizations of worthy research lead to the abandonment of some projects over others. A simple example of this is provided by evaluation systems that privilege articles over books as ideal research outputs, resulting in fewer published monographs and edited collections (February 13th et al. 2018). Along with these shifts in investments by individual scholars, evaluations might transform disciplines through a third mechanism, namely, by creating disincentives to interdisciplinarity. Added to changes in what type of research scholars pursue, the institutionalization of evaluations as formal and informal workplace regulations can create barriers to conducting *interdisciplinary work* that is not commensurable with the parameters of the evaluation (Rijcke et al. 2016). In addition to more being difficult to

publish, interdisciplinary work might suffer from the added cognitive penalties of spanning established categories of knowledge (Zuckerman 1999). By strengthening the shared norms of quality within the field and creating symbolic incentives to patrol these, research evaluations can lead to disciplinary entrenchment at the expense of epistemic diversity.

As these studies suggest, evaluating knowledge has potential performative consequences: by transforming individual and institutional incentives to the production of scientific claims, they change what can be known in distinctly patterned ways. For example, the mechanisms explored by research evaluation scholars provide an important account of how scientists respond to evaluations by either shifting publication and research strategies or avoiding research programs that do not fit with the dominant logic of assessment. Yet despite its richness, the literature is also limited by its analytical scope: most current discussions focus on individual researchers, addressing the possible effects of evaluation regimes on short-term decisions about what to publish and where. The collective problem of epistemic change over longer periods is thus reduced to the myriad individual actions by researchers who chose to “play the game”. If knowledge changes in response to how it’s worth is measured, we learn, it is mostly because of how scholars react, embrace, or begrudgingly accept how their work is counted and compared.

The state of the art on research evaluation in particular and the sociology of science more generally thus lacks an institutional mechanism explicitly connecting the actions of quantified scientists with systemic, field-level transformations. While it is true that generic ‘reactivity’ to rankings and quantification is well documented, less is known about how it and similar processes shape the organization of knowledge over time. We know that scientific practices likely shift as a result of research evaluations (Gläser and Laudel 2016; Hammarfelt and De Rijcke 2015), but it is not entirely clear if these changes have effects on the direction of knowledge production with longer-term consequences on epistemic organization.

The institutional link explored in this paper comes from considering research evaluations as interventions not only upon the bench and desk practices of scientists but,

as importantly, on the larger labor markets that structure academic fields. The existing research on the quantification of knowledge largely sees scholars as institutionally fixed, despite the vast evidence that shows the importance of research mobility on epistemic-level field change (Czaika and Orazbayev 2018; Frickel and Gross 2005; Latour 1987). Furthermore, focusing on academic fields as spaces structured through capital and mobility matters because, as Whitley notes, they are what gives rise and logic to disciplines, the “units of labor market control” that define the training of new scholars and monopolize the meaning of intellectual contributions (Whitley 2000). This analytical shift from the effects of measurement on scholars in isolation to its propagation across employment structures fields stems from the twofold observation that research evaluations are intimately connected to the very generation of fields of competition between scholars and institutions (Whitley, Gläser, and Engwall 2010) that tie both the labor markets of specific disciplines and the career trajectories and research decisions of individual scholars.

The idea that research evaluations generate novel fields of competition or substantially alter existing ones that are intimately tied to the logics of disciplines is partly a consequence of the administrative origins of these interventions. Both historically and presently, research evaluations operate as devices of governance created and sustained under the auspices of the state to deal with the constrained allocation of financial support to scientific work in public institutions. Through their official sanctioning, the metrics, rankings, and other forms of commensuration entailed by most research evaluations create signals that are legible to, and valuable within, a context of competition both for funding and talent, making tangible “the interests directly associated with [an agent’s] position in the academic field” (Bourdieu 1988). Within otherwise uncertain and potentially multipolar academic fields (Whitley 2000), evaluations increase the “stability and predictability of reputational norms”, establishing field-level as opposed to local exigencies on research strategies: expectations of quality of local institutions (such as colleagues, heads of department, and university managers) will be better aligned to those of the evaluation as a whole. These exigencies will invariably encode particular rules of engagement that reflect dominant ideas of what is valuable within disciplinary traditions but that simultaneously allow for comparison and competition among scholars and institutions as well as the

latter's individual task realignment. For example, an evaluative framework that links worth to a ranking of journals will change the field by making specific titles more attractive than others, all other things considered. If bootstrapped to the institutionalized ways meaning is collectively produced (Barnes 1983), both the individual intellectual worth of scholars and the better market outcomes will be tied to publishing in top ranked journals. The logics of individual valuation and the career structures will be different if, instead of an explicit ranking, a loose but shared hierarchy of institutional status positions is what determines a scholar's worth (Burris 2004). Both situations will also produce different configurations for the epistemic field: in the first, elite journals and their gatekeepers are central to value determination leading, perhaps, to stronger field-wide notions of quality; in the second, elite departments retain some degree of flexibility in establishing and reproducing their own traditions.

The above examples already suggest how, by establishing common metrics that are taken as given within disciplinary communities, research evaluations create novel resources that allow some actors to attain better positions in the field. A scholar whose work is highly cited, for example, will have even more visibility in her field if research evaluations place greater worth on citation metrics, and so might be able to transform this 'evaluative capital' into concrete economic gains through either intraorganizational movements (e.g. promotion) or via labor markets. Indeed, we can think of research evaluations as constituting, rather than simply describing from afar, the marketplace for ideas and its connected labor market outcomes: they create, or at the very least alter a metrics-sensitive academic field, and introduce a new form of capital centered around evaluation attainment to its competitive struggles. A recent finding from the United Kingdom provides evidence for both the existence of this capital and its transfer into other forms of resources. In their work, De Fraja and his colleagues (De Fraja, Facchini, and Gathergood 2016) show that, over the last cycle of research evaluations in Britain, professorial salaries increased at higher rates than salaries at lower ranks at the same time as wage inequalities between elite and non-elite institutions also grew.

When performative, research evaluations can therefore shape the career patterns of scientists, leading to distinct disciplinary shifts catalyzed as much by choices about what to investigate, as highlighted by the research evaluation literature, as by the transformation of patterns of interinstitutional mobility. Evaluations may not structure the labor market directly, but they provide further incentives—both for departmental managers, units, and individual scholars—to sort the field in ways that are proximally commensurable to the assessments. What occurs through evaluation, then, is a form of isomorphism prompted as much by institutions imitating organizational designs as by the field-driven movement of scholars through labor markets. To explore these, I use an original dataset of British academics in the social sciences, briefly described below.

The case

Of all the possible examples of research evaluation influencing epistemic field change, the regulation of science funding in Britain arguably comes closest to a natural experiment. When thinking about the links between quantification and changes in disciplinary makeup, an important case is given by assessments where the focus and administrative consequences of evaluation are not on individual scholars—as tends to be the norm in most science evaluation frameworks—but rather sub-units that serve as avatars of disciplines within the academic field (Abbott 2005, 2010). This bears the highest resemblance to the forms of reactivity studied by Espeland and Sauder (Espeland and Sauder 2007; Sauder and Espeland 2009), where metrics are relevant to organizational units rather than isolated scholars but nevertheless have influence on the latter's workplace experiences. In Espeland and Sauder's canonical work, rankings (i.e. measurement) changed the way law schools deployed resources strategically. Yet they did not substantively change how schools approached the law as an academic field—they might have played around with student/teacher ratios or other components of the metric, but the practice of legal scholarship remained largely unaltered. This is why the British case offers such a paradigmatic example: because of their emphasis on research quality and breadth (they involved not one but all disciplines across every institution of higher education over more than three decades) these evaluations are centrally about how

academic knowledge at large is re-disciplined and transformed (Campbell and Carayannis 2012; Gläser and Laudel 2016; Lewis 2015).

Research evaluations were introduced to the British higher education landscape in 1986 when public universities—representing the vast majority of higher education establishments in the country—were experiencing intense financial pressures stemming from the rapid growth of the sector over the past quarter century and the country’s ongoing fiscal crisis (Brown and Carasso 2013). Up to this point in time, the government body that oversaw higher education in Britain, the University Grants Committee (UGC), had allocated block funding ostensibly earmarked to cover teaching and research expenses by accounting for student enrollment across institution. This relatively egalitarian approach, however, faced mounting criticisms. With dwindling budgets, a growing sector, and a funding mechanism that didn’t discriminate research quality, key groups of policy makers and scientists pushed for a ‘selective’ allocation to provide centers of excellence with the ability to maintain their research efforts and infrastructures despite the endemic and seemingly irresolvable fiscal problems of the nation (Whitley et al. 2010).

The government’s commitment to quantification as an instrument of prudential public management meant a strong reliance on performance indicators and sectorial evaluations as leverages for resource distribution. This was particularly so for higher education, where the government clearly saw indicators as devices that, more than representing, could loop values, behaviors and expectations back into the sector (Hacking and Hacking 1983). As a 1985 whitepaper noted, the government believed “that there would be an advantage in the regular publication of a range of unit costs and other performance indicators by institution and by department [that will be] important for the internal management of institutions” (Education and Science 1985). Measurement could not only reveal worthy knowledge, products, and institutions, but also foster incentives for internal reorganization. Inspired in part by the proliferation of other informal rankings and metrics of institutional quality, the UGC thus trialed a system aimed at evaluating the research conducted across Britain’s universities to guarantee some degree of ‘value for money’ legible to external publics and the state (Power 1997; Shore et al. 2015; Strathern

2003). The result was the first of a series of nation-wide assessments of institutional quality that, albeit much changed, remains in place.

Like any distributional intervention, this first assessment was not without controversy, criticized for relying on a method “shrouded in secrecy” (Phillimore 1989). Operationally, the exercise required all universities in the United Kingdom to submit information for each of their ‘cost centers’ to subcommittees of experts in each discipline (cost centers were accounting constructs created and defined by the government that broadly mapped onto traditional subjects and departments; e.g. physics, medicine, sociology, biological sciences). Each file contained a two-page description of the achievements of the cost-center, five examples of the best publications produced in the previous 5-year period, data about new hires and success in obtaining competitive research funding, along with peer reviews of research performance (Anderson 1986). The outcome of the deliberations was a ranking of institutions across four possible points (‘outstanding’, ‘above average’, ‘average’, and ‘below average’). These were then used to weigh the number of students enrolled in each institution to determine funding outcomes. Although the categories informing the first exercise were rapidly institutionalized as criteria in the allocation of future UGC grants, their precise weights, benchmarks, standards, and mechanics of consideration were not initially clear. Indeed, as John Phillimore wrote, the impact of the assessment was “more divisive and pronounced than the more general funding cuts imposed since 1981 [...] Misery will no longer be shared equally” (Phillimore 1989:258–59).

Opposition notwithstanding, policy experts and government officials pressed for a refinement, expansion, and re-implementation of evaluations. This first pilot assessment was followed by the more exacting Research Selectivity Exercise of 1989. Unlike the previous ad hoc exercise, the 1989 iteration made a larger portion of the government’s funding dependent on the outcome of the evaluation (Lee and Harley 1998). It also expanded the breadth of the assessment. Instead of assessing only the five most prominent publications for each cost-center, the nearly three hundred members of the advisory groups and assessment panels of the Research Selectivity Exercise were expected to

produce a consensus about the quality of each unit of assessment on the basis of informed peer review, data on the number of publications of each department, information on income from extramural research grants, data on the number of research studentships, and critically, detailed bibliographic information of up to two publications from each full-time academic (Johnes and Taylor 1992:70).

Since this early exercise, research evaluations grew in size, complexity, and intensity. The Research Selectivity Exercise of 1989 was followed by four additional evaluations in 1992, 1996, 2001 and 2008 under the name of Research Assessment Exercise, itself substituted by the Research Excellence Framework of 2014 and 2020. All in all, British academics who started their careers in the early 1980s have experienced seven of these exercises with rules and standards that change under each iteration. [] The fact remains, however, that since the introduction of the Research Assessment Exercise of 1992, British scholars have seen their work evaluated by panels of peers, with the outcome determining how the government's unencumbered research funds flow back into their institutions.

While not directly or overtly, these evaluative frameworks certainly put a price on research by measuring its quality and linking it to distinct allocative consequences. It is thus not odd to see why scholars have often identified RAE/REF as part of the increased marketization of higher education in the United Kingdom (Brown and Carasso 2013; Hamann 2016; Sayer 2014). While a 'market of ideas' (Arrow 1969) certainly existed prior to the implementation of these systems, they surely perfected exchange by providing rankings, comparisons, and performance indicators that putatively signaled consensus about research quality within disciplinary fields, across institutions, and to managerial cadres. In effect, by tying some portion of the state's funding of institutions to their employee's work, these exercises intervened directly in the labor markets of disciplines by making some scholars more valuable (both symbolically and financially) than others. There are several concrete demonstrations of the linkage between evaluations and academic labor markets. In my interviews with British academics, I often heard from early career scholars about how they were informally made aware of the importance of being 'REFable'

before searching for jobs in the market (Interview A8, Interview B7, (Leathwood 2017; Locke 2014; Sayer 2014)). Mid-career and senior academics echoed this view, noting that search committees would often include a scholar charged with looking at the applications “from a REF perspective” (Interview B6). More directly, there is some evidence that REF both instigated so-called transfer markets for ‘star researchers’ (Elton 2000; Scott 2003; Talib and Steele 2000) and had other temporal effects on the retention of teaching and research staff across evaluations (Metcalf, Great Britain, and Department for Education and Skills 2005). Whichever the case, evaluations likely distorted labor markets and, potentially through these, the organization of disciplinary fields.

How did research evaluations transform British science, if at all? To answer this question, I focus on the evolution and labor market dynamics of four social sciences: social and cultural anthropology, economics, politics and international relations, and sociology. The choice of these over other disciplines is threefold. First, unlike laboratory-based or other infrastructure-intensive fields, the social sciences have relatively low fixed costs that ostensibly translate into both higher mobility of researchers between institutions and less pressure towards centralization. Second, the production of knowledge in the social sciences is more closely aligned to individual creators rather than groups of collaborators, while preserving some degree of coupling to organizational units. In comparison to other fields, the social sciences have lower numbers of authors per published article and lower levels of recurring collaborations (Parish, Boyack, and Ioannidis 2018). Simultaneously, the social sciences are more accommodating of disciplinary *styles* associated to identifiable units, departments, or institutions. We speak of the Chicago School as a reference to a particular variant of economics in the same way as we think of the Manchester School as a tradition in anthropology, Birmingham’s Center for Contemporary Cultural Studies as a distinct intellectual environment, or the Edinburgh School as shorthand for a strand of literature on the sociology of science and knowledge. This tension between mostly individualized intellectual production and disciplinary traditions anchored on departmental avatars is partly what motivates the emphasis on the matching and sorting effects of labor markets (cf. (Abbott 2010)). If not aligned with the evaluation’s metrics, otherwise resilient departmental styles might change through both the hiring of new scholars trained under

different traditions or the movement of scholars to other institutions. An empirical manifestation of this process would involve changes in some measure of epistemic diversity, moderated by labor market movement, and associated to the adoption of research evaluations.

Third, and despite their apparent similarities, the dynamics of these disciplinary fields vary greatly in both epistemic and institutional terms. Over the period covered by the research evaluations, for example, economics was fundamentally reshaped by processes of internationalization and diffusion that pegged value to a global, transnational professional standard (Backhouse and Backhouse 2002; Backhouse and Cherrier 2017; Fourcade 2006, 2009). While outwardly oriented, anthropology and sociology were not exposed to similar globalizing processes, thus preserving local labor markets, institutional styles, and longstanding departmental hierarchies as features of their disciplinary organization (Halsey 2004; Kuklick 1991, 2009; Urry 2012). Politics and international relations, on the other hand, grew to combine national traditions of political theory with empirical approaches from both Continental Europe and North America (Easton, Graziano, and Gunnell 2002). This variability matters, for while these four disciplines have relatively comparable modes of knowledge production and require similar institutional investments, they embody different ways of organizing the academe.

Labor markets, mobility, and epistemic matching

To identify the effects of research evaluations on the epistemic organization of scientific fields, I will focus on mobility of social scientists between academic institutions as a process of market-mediated matching that slowly changes departmental composition and, through this, the structure of each discipline. Most of the literature on academic mobility focuses on the predictors and experiences of international scholars. Research on internal interinstitutional mobility is rather sparser, although existing studies point to some key factors that influence shifts in scholars' careers. Research has shown, for example, a marginal though positive effects of scholars' productivity, citation, and degree of

specialization on mobility (Crane 1970), their future productivity (Allison and Long 1987; Long 1978), and overall career progression (Grimes and Register 1997; Leahey 2007).

These studies suggest two clear sets of predictors of interinstitutional mobility. The first concerns variables that proxy visibility and productivity. Briefly, previous studies suggest that academic follow similar patterns to general job changers: scholars occupying higher levels move more slowly than those in lower-level positions (Rosenfeld and Jones 1986). Research also shows that scholars with more citations and higher levels of productivity (measured, for example, through numbers and frequency of publications) are promoted more rapidly, making of these forms of prestige and visibility anchors that reduce mobility by moving individual into higher level positions (Allison and Long 1987).

A second predictor of mobility is gender. Studies indicate clear gendered differences in both patterns of rank promotion (Long 1978; Long, Allison, and McGinnis 1993). Overall, women both suffer wage penalties in the academic workplace (Barbezat and Hughes 2001; Leahey 2007; Leathwood 2017) and experience higher rates of mobility than men. Authors have presented several reasons why female academics might experience higher levels of job market mobility than their male counterparts. One argument is that female scholars are more likely to change institutions because of spousal relocation. In their study of two cohorts of economists, for example, Hilmer and Hilmer (2010) suggest that women move for non-pecuniary reasons while men use institutional transitions as a mechanism for upward rank and wage mobility. Other arguments involve career interruptions as explaining differing patterns of promotion, earnings, and job market stability. In Long, Allison and McGinnis' (1993) now classical work, for example, much of the variation between men and women derives from how interruptions affect rank advancement, particularly among early career scholars, which may translate into higher rates of mobility.

There is an important caveat with the literature on interinstitutional mobility: because it is disproportionately focused on American academics, it may not describe mobility rates for academics working in the United Kingdom. The defining feature of the American academic system (tenure as a milestone career event that guarantees de facto job stability) is simply

lacking in Britain. Since the structural, system-wide reforms to British higher education in the mid 1980s, academics have not had the same type of contractual stability that scholars employed in America enjoy. Indeed, if anything, British academics should be closer to their non-academic peers in terms of both labor market outcomes and dynamics, shifting the role that some of the above predictors play in explaining interinstitutional transitions. For example, status and productivity signals might be even more important in the United Kingdom than in the United States, given that they facilitate commensuration in a more fluid, active market. Similarly, recent productivity records may have a more important role in allowing for mobility (and, in particular, upward mobility) given their ties to the incentives of research evaluations. Given that they ground existing discussions about academic careers, however, I retain variables connected to these two sets of predictors.

Data and Methods

I draw on an original dataset containing information on 44,956 social scientists active across the 116 public universities in the United Kingdom between 1970 and 2018. The dataset was compiled using the bibliographic records of Web of Science's Social Science Citation Index for this period. The records obtained from SSCI included titles for 149,796 research articles along with journal titles, dates of publication, corresponding addresses, funding data, and complete abstracts for 76,898 articles. By querying Web of Science for all research articles published in anthropology, economics, political science, and sociology¹ containing at least one author with a British affiliation, we were able to generate a longitudinal data set that proxies career mobility for individuals across each evaluation period and for the entire sector.

The data obtained through these queries was processed to standardize both researchers' names and those of their institutional affiliations. We first eliminated duplicates in the data. Then, we mapped names and institutional affiliations. Because of

¹ An example of the WoS search query for sociology is: (AD=(england OR scotland OR wales or UK) AND SU=(sociology) AND WC=(sociology)) AND DOCUMENT TYPES: (Article OR Book OR Book Chapter) Timespan: 1970-2018. Indexes: SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC.

changing archiving standards in the Social Science Citation Index, the association between corresponding addresses and listed authors in the bibliographic data was non-trivial: the reported affiliations did not always match the number and order of authors, requiring us to constrain the data to researchers associated at least twice across different publications with the same institution over an evaluation period. Because of changes in the names of various institutions of higher education, we also standardized author affiliations using current nomenclature. For example, Thames Polytechnic, one of the oldest polytechnic institutions in the United Kingdom, was matched onto University of Greenwich, its name since 1992. Finally, we dropped authors for which there was no guarantee of uniqueness. Prior to the mid 1990s, the SSCI only recorded authors using their first name initials followed by their last name. This meant that an imagined Amanda Smith is indistinguishable from a similarly imagined author, Adam Smith. Although it might have been possible to use institutional information to disambiguate data, this could have potentially introduced further mistakes in the model by confounding moves across institutions for the same researcher. The data therefore only contains information on authors for whom there is little ambiguity about both identity and interinstitutional mobility. The result is a longitudinal dataset of 16,531 individual academics over the four fields (2,208 anthropologists, 6,384 economists, 4,271 political scientists, and 3,668 sociologists). Having the full first names for around 80% of the authors, we were also able to infer their gender using information from the 1991 UK Census. The resulting sample contained 61.63% male and 38.37% female academics.

The dataset that I use is thus admittedly imperfect. Firstly, it excludes books and other forms knowledge dissemination. Second: because it relies on the institutional affiliation as reported on articles, it is affected by both a temporal lag for career transitions and self-reporting. Scholars on sabbatical leave or on visiting appointments, for example, may have published under their temporal, rather than permanent affiliation, creating the appearance of mobility without an underlying change in employment. Similarly, owing to uneven publication lags, the data does not allow studying the 'transfer market' directly: we cannot ascertain with clarity when transitions occurred but only if a scholar changed institutions between evaluation periods. The outcome variable is, indeed, quite course

taking a value of 1 if a scholar identifiably changed institutions between periods and 0 if she did not. There are also some assumptions about scholars and disciplines that are relatively unavoidable: because the search query that generated the source data hinges on the classification of journals used by Web of Science, it doesn't map neatly and perfectly to departmental structures and affiliations. (A random sample of 200 academics from the dataset suggested an 89% accuracy of departmental affiliation, a 96% accuracy of mobility, and a 94% accuracy in assigned gender.) Lastly, since we constrained the analysis to 116 institutions of higher education, our dataset doesn't cover transitions from academia into non-academic institutions and vice-versa. Institutions not recognized as higher education providers were labeled 'Other' and do not feature in the analysis.

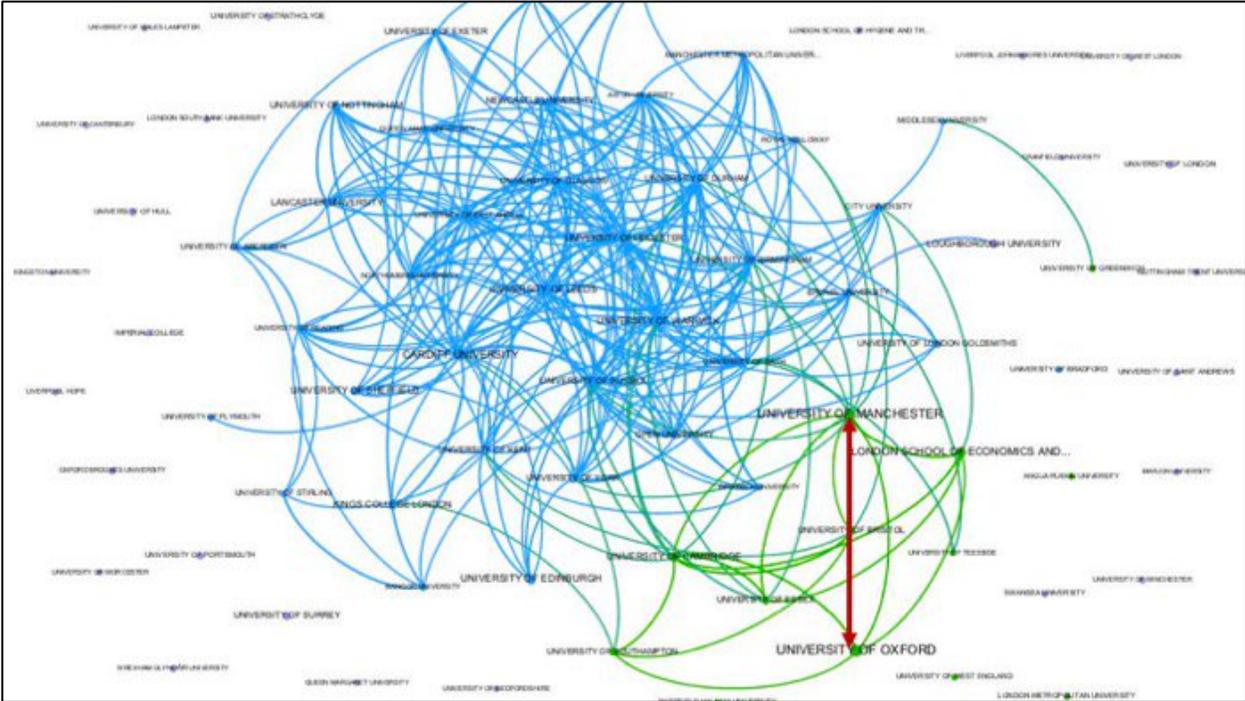
The data is nonetheless revealing for other reasons. In addition to providing an exceptionally detailed, if imperfect, view of academic labor markets in the British social sciences, relying on bibliographic data allows studying how individual scholars' intellectual production changed over time and was matched by their departments and fields. For this, this paper relied on two measures of textual similarity that allowed comparing scholars to their home institutions, departments within disciplines, and evaluating the evolution of academic fields over time. These measures are all based on a Latent Dirichlet Allocation topic model of the corpus of paper abstracts for each discipline (Blei, Ng, and Jordan 2003; DiMaggio, Nag, and Blei 2013; Mohr and Bogdanov 2013). Taking the collection of all abstracts for each of the 4 social science, we determined four 40-topic solutions that neatly described most of the key themes covered by articles (and, ostensibly researchers' intellectual agendas) over time. The 40-topic solution was arrived at iteratively, by examining the heuristic coverage of models ranging from 15 to 70 topics (Karell and Freedman 2019). These four 40 topic solutions, corresponding to the different disciplines, were then used to describe first, the aggregate periodic intellectual production of each subject/institution as a 40-dimension vector, and secondly, the production of individual scholars with respect to their institutions of employment. An example of topics for the four disciplines is provided in figure 1.

Figure 1. Selection of topics for the 4 disciplines, 1970-2018	
<i>Anthropology</i>	
Topic 29	archaeologi; data; develop; record; past; review; interpret; issu; discuss
Topic 30	network; structure; social_network; commun; reserve; ti; central; social; measur; time
Topic 31	anim; dog; owner; behavior; attitud; human; emot; pet; hors; cat
<i>Economics</i>	
Topic 29	optim; process; option; time; function; price; stochast; method; deriv; dynam
Topic 30	inform; contract; agent; incent; effici; optim; auction; mechan; effort; cost
Topic 31	polici; reform; market; govern; regul; privat; service; competit; sector; uk
<i>Politics and International Relations</i>	
Topic 29	parti; issu; left; elector; posit; polit; polit_parti; support; polici; ideolog
Topic 30	public; tax; cost; spend; budget; govern; effect; polici; expenditure; fund
Topic 31	labour; conserve; polit; british; parti; leader; leadership; govern; british_polit; britain
<i>Sociology</i>	
Topic 29	class; culture; music; bourdieu; field; habitu; middl_class; social; capit; distinct
Topic 30	knowledge; scienc; technologi; genet; practice; scientif; human; power; form; natur
Topic 31	manag; worker; labour; profession; employ; service; industry; organis; organ; employe

The first measure allowed us to assess the degree of similarity between departments: calculated as a Euclidian distance, two departments having exactly the same research profile (that is, the same distribution across the 40 discipline-specific topics) will have a minimal distance of zero. Similarly, two departments focusing on entirely different topics will have a maximal distance of 1. The collection of normalized distances between all pairs of departments can be used to construct a graph that captures information on the structure of each disciplinary field (figure 2). This graph then allows us to identify departments that are more structurally central to the discipline. The measure of betweenness centrality (Leydesdorff 2007; White and Borgatti 1994), in particular, provide a proxy for *departmental typicality*: departments or institutions more central to the graph will share greater thematic features and will be more similar to each other, particularly when compared to institutions with a lower degree of betweenness centrality. Importantly,

the distribution of distances also allows visualizing changes to the typicality within the field: a field with very lower average distances will be populated by more similar departments than one with higher average distances.

Figure 2. Graph representing the organization of British sociology, 1998-2004



The second measure is based on a similar logic. In this case, we estimated the difference in topic models for each author-department couple by first calculating the model for the scholar’s department without her contributions, to then calculate the topical distribution for her work over the period. We then determined the cosine similarity between the vector representing each scholar’s publications and the vector representing the work of the remainder of her department. This produced a metric of the *relative distinctiveness* for the work of each social scientist. A scholar whose work is identical to that of the rest of the department will have a relative distinctiveness of 1 whereas a scholar whose work is completely orthogonal to that of the department would have a minimal relative distinctiveness of 0.

The binary outcome variable in this study is institutional mobility, measured as an and identifiable change in a scholar's affiliation across periods and taking values. Formally, the overarching model can be expressed as

$$\Pr(\text{MOBILITY}_{it}) = \beta_0 + \beta_1 \text{AGE}_{it} + \beta_2 \text{GENDER}_{it} + \mathbf{B}_1(\text{PRESTIGE}_{it}) + \mathbf{B}_2(\text{PRODUC}_{it}) + \beta_3 \text{SCORE}_{it} + \beta_4 \text{DEP_TYP}_{it} + \beta_5 \text{DIST}_{it} + \text{DISC} + \text{PERIOD} + U_{it} + E_{it}$$

The identifying variable in the regression was the individual scholar with each period serving as the time variable. To control for idiosyncratic variations, all models included fixed effects for periods and disciplines but and worked with random effects at the level of individual scholars to account for variance in the panel data.

The probability that a scholar moves institutions between periods was analyzed with respect to six distinct models. Models 1 and 2 provide a baseline that reflects variables identified in the existing literature on academic labor markets as predictors of job change, namely, individual productivity (measured through the number of articles published cumulatively and over the most recent period by each scholar), prestige (measured through both cumulative and with recent citations, success at obtaining extramural funding, as well as average impact factor of the journals where each scholar published), and demographic factors (age, approximated as the time that the scholar has been active in the field, and attributed gender).

Models 3 to 5 include variables directly related to the existence of research evaluations. Model 3 adds to the baseline analysis by introducing the institution's most recent outcome in the RAE/REF as a way of testing for the contributions of these exercises as a form of status that affects the mobility of scholars. If evaluation outcomes constitute a form of capital that can be transformed into mobility, then a higher standardized grade ought to translate into higher odds of transitioning between institutions. Model 4 tests the relation between mobility and departmental typicality. For research evaluations to have isomorphic effects on the disciplinary field, more typical departments would have to see

less losses to their ranks than more atypical departments that, over time, would increasingly conform to a disciplinary standard. Model 5 tests an individual scholar’s distinctiveness with respect to her institution as a predictor of mobility. Finally, model 6 includes both departmental typicality and scholar’s distinctiveness but adds an additional variable that speaks directly to ideas about evaluations as instruments of epistemic disciplining: it considers the effects of being ‘structurally coherent’ within the logic of the evaluation (Zuckerman 1999). A feature of British research evaluations is that they operate on ‘units of assessment’ rather than individual departments. Thus, institutions decide how scholars are submitted: an anthropologist working in a medical school, for example, can be part of the institution’s submission for the anthropology assessment panel, reflecting the quality of the discipline in the university as a whole rather than in one of its departments. When departments are small or expertise are distributed across units, this may lead to scholars being submitted to assessment panels that are just proximally related to their main area of research. For example, an institution lacking a department of anthropology may submit its anthropologists as sociologists. If disciplining devices, research evaluations would make this structural incoherence into a motive for changing jobs or research avenues. Indeed, we would expect scholars who are not structurally coherent to show higher rates of mobility as evidence of disciplinary matching.

Table 1 – Summary Descriptive Statistics

	count	mean	sd	min	max
Change of affiliation	79284	.1312749	.3377029	0	1
Age (log)	79284	1.092411	1.163642	0	3.7612
Gender (1=female)	47471	.3508036	.4772266	0	1
Number of publications (log)	79284	1.042002	.5788998	0	4.89784
Cumulative publications (log)	79284	1.289131	.7520639	0	5.209486
Number of citations (log)	79284	.7700703	1.150336	0	6.993933
Cumulative citations (log)	79284	.8885759	1.274042	0	7.731931

Average journal impact factor	79284	.7025544	.8815296	0	7.863
Cumulative journal impact factor	79284	.9328543	1.152946	0	13.0829
UK Funding	79284	.1367111	.3435442	0	1
European Union Funding	79284	.0212527	.1442265	0	1
Departmental typicality	60427	20.26517	5.389539	2.602462	29.48124
Scholar's relative distinctiveness	44048	.1320621	.0491929	0	.8313711
Structural incoherence	79284	.4694516	.4990691	0	1

Results

The results across models suggest one intuitive finding: age has a negative association to interinstitutional mobility. Similarly, and consistent with the literature (Long 1978), models 2-6 also show a positive effect of gender on mobility: the models suggest that women are up to 1.34 times (model 6) more likely to move institutions between evaluation periods when compared to men. Variables that proxy prestige and field visibility (numbers of publications and citations) suggest a slightly more complicated interaction between recent and cumulative productivity. Cumulative measures of both productivity and citations are similar to age, in the sense that they increase temporally by a sheer process of accumulation. These were included in the model along with the same measures for the latest period to assess the interplay between accumulated prestige and recent visibility in the field (Leahey 2007). The implications of the models are surprising, insofar as they suggest very weak associations between interinstitutional mobility and the publications of individual scholars. While the number of publications contributes positively to the odds of mobility, it is offset by the scholar's cumulative publication record. Since there will always be as many, if not more, cumulative publications than recent publications, the effects of these in models 2 to 4 will vary from positive to negative as accumulated publications and citations weigh more than those generated in the most recent evaluation period.

Table 2 – Panel logistic models with individual scholar random effects						
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	Interinstitutional mobility					
Age (log)	-0.158*** (-5.93)	-0.135*** (-4.08)	-0.120** (-2.64)	-0.0802* (-2.22)	-0.181*** (-3.94)	-0.186*** (-4.05)
Number of publications (log)	0.0336 (0.48)	0.264** (2.75)	0.448*** (3.38)	0.438*** (4.12)	0.221 (1.66)	0.236 (1.77)
Cumulative publications (log)	0.0602 (0.75)	-0.221* (-2.05)	-0.343* (-2.32)	-0.374** (-3.14)	-0.134 (-0.89)	-0.120 (-0.80)
Number of citations (log)	0.117 (1.53)	0.170 (1.66)	0.355* (2.42)	0.240* (2.13)	0.356* (2.53)	0.360* (2.55)
Cumulative citations (log)	-0.141 (-1.86)	-0.174 (-1.71)	-0.404** (-2.77)	-0.248* (-2.21)	-0.355* (-2.53)	-0.359* (-2.55)
Average journal impact factor	0.00602 (0.09)	-0.0965 (-1.24)	-0.0350 (-0.36)	-0.0804 (-1.01)	-0.132 (-1.45)	-0.127 (-1.39)
Cumulative journal impact factor	-0.0795 (-1.37)	-0.0146 (-0.22)	0.0243 (0.30)	-0.0240 (-0.36)	-0.0149 (-0.20)	-0.0102 (-0.13)
UK funding	-0.0482 (-0.74)	-0.0464 (-0.63)	-0.0417 (-0.45)	-0.0617 (-0.82)	-0.0605 (-0.75)	-0.0460 (-0.57)
European Union Funding	0.719* (2.24)	0.934* (2.56)	0.979* (1.99)	0.915* (2.46)	0.696 (1.77)	0.654 (1.67)
Anthropology	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)
Economics	-0.0619 (-0.99)	-0.0780 (-0.98)	-0.0333 (-0.29)	0.180 (1.53)	-0.125 (-1.13)	0.0966 (0.65)
Politics and IR	0.0298 (0.45)	0.0167 (0.20)	0.0357 (0.32)	0.190 (1.76)	0.0239 (0.21)	0.247 (1.81)
Sociology	0.262*** (3.94)	0.177* (2.16)	0.131 (1.13)	0.151 (1.68)	0.176 (1.55)	0.148 (1.30)
1986	0 (.)	0 (.)				
1989	0.133 (1.54)	0.185 (1.38)				
1992	0.128 (1.50)	0.0478 (0.36)		0 (.)	0 (.)	0 (.)
1996	0.198* (2.54)	0.144 (1.22)	0 (.)	0.416* (2.15)	0.0607 (0.12)	0.250 (0.50)
2001	0.250*** (3.33)	0.211 (1.92)	0.113 (0.92)	0.583** (2.92)	0.199 (0.41)	0.436 (0.86)

2008	-0.303*** (-4.41)	-0.517*** (-5.19)	-0.673*** (-5.83)	-0.0796 (-0.39)	-0.543 (-1.11)	-0.287 (-0.56)
2014	-0.438*** (-6.10)	-0.562*** (-5.45)	-0.783*** (-6.31)	-0.146 (-0.71)	-0.607 (-1.24)	-0.380 (-0.74)
Male		0 (.)	0 (.)	0 (.)	0 (.)	0 (.)
Female		0.272*** (5.23)	0.188** (2.75)	0.241*** (4.31)	0.283*** (4.39)	0.296*** (4.59)
Standardized evaluation rating			-0.0684* (-2.11)			
Departmental typicality				-0.0304*** (-3.50)		-0.0233* (-2.15)
Scholar's relative distinctiveness					1.385* (2.41)	1.146* (1.98)
Structural incoherence						0.333*** (5.03)
Constant	-0.640*** (-8.61)	-0.459*** (-4.22)	-0.506*** (-3.71)	-0.476** (-2.66)	-0.635 (-1.27)	-0.656 (-1.31)
lnsig2u						
_cons	0.339*** (4.14)	0.196 (1.71)	0.416* (2.51)	0.330** (2.59)	0.144 (0.78)	0.131 (0.69)
N	24745	15362	9393	13502	9456	9456

t statistics in parentheses

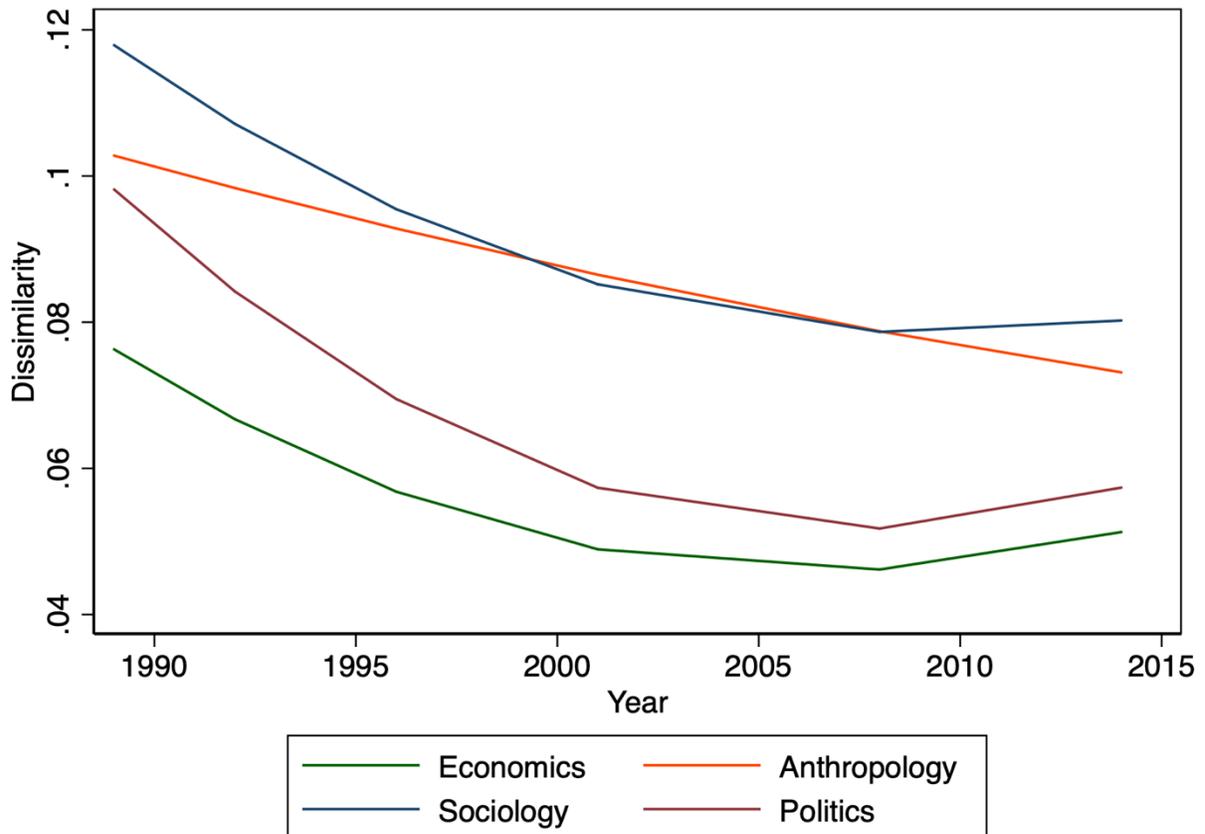
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Perhaps the more interesting findings come from models 3 to 6, which include measures that originate in the evaluation exercises and speak of the structure of the academic field. Model 3, for example, suggest that being in a unit of assessment that was highly rated in the previous evaluation will have a slight negative on interinstitutional mobility. This finding is consistent with the broader literature on mobility, which suggests lower transition rates for workers in high status positions when compared to cohort peers (Fuller 2008; Rosenfeld and Jones 1986). Supplemental analysis that looked at transition probabilities across different tiers of universities did not provide evidence of status-based matching (the odds of moving from, say, a former polytechnic to an elite Oxbridge institution and vice-versa are similar). There is, however, a slight negative association between the standardized score of an institution in the most recent evaluation and academic mobility. Scholars at institutions with the highest possible standardized scores ($z=1$) are just 7% less likely to change affiliations than scholars at modal institutions ($z=0$).

Measures of field structure tested in models 4 to 6 provide further insights into how evaluations sort scholars across departments and, in doing so, transform the organization of the disciplinary field. In model 4, departmental typicality is shown to have a negative association to mobility. Scholars at very typical departments are less likely to change institutions across periods. For the most atypical department (departmental typicality = 2.6), the probability of moving jobs is reduced by 7.6% with respect to a baseline typicality of zero; scholars working in the most typical department in the dataset (departmental typicality = 29.48), by comparison, will be 60% less likely to change jobs with respect to the same baseline. Scholarly distinctiveness also plays into potential epistemic matching. As model 5 shows, scholars whose work is very close to that of their departmental colleagues are disproportionately more likely to change jobs (maximally, 3.16 times) than those whose work is completely unlike that of their departmental peers. Model 6 explores cases of scholars not submitted to the research evaluations in their fields of publication. This rather artificial measure gives a sense of the structural incoherence with respect to the disciplinary logics of the field and their operationalization in the research assessments. Surprisingly, the model suggests that scholars in positions of structural incoherence are almost 40% more likely to change institutions across evaluation periods than their structurally coherent peers. (Supplemental work using structural incoherence as part of a difference-in-differences analysis suggests a statistically significant causal effect of not being evaluated in the coherent discipline and interinstitutional mobility.)

In aggregate, the models tested using the career data for the British academics suggest a form of sorting between and across institutions linked to how the disciplinary field is organized. The question remains, however, as to whether the organization of the field varied over time. For this, evidence is provided by graphing the average normalized distance between departments over time to gauge a potential shift in the organization of the disciplinary fields. As figure 1 shows, the fractional polynomial fits of the topical distances between departments from the late 1980s to 2018 show a reduction in the average dissimilarity between institutions across all disciplines.

Figure 3 - Fractional polynomial fit of topical distances between departments, across all disciplines



Discussion and conclusions

Coupled to metrics that describe the publications of British social scientists over the last half century, the approximate though detailed career data used in this paper suggests the presence of connections between interinstitutional mobility and changes in the structure of disciplinary fields that are directly tied to the implementation of research assessments. Since their introduction in 1986 and expansion throughout the institutional landscape of British higher education, research evaluations have transformed both academic labor markets and the structure of the disciplinary field. Figure 1 provides telling evidence of what this looks like: a shift in the thematic composition of departments across all disciplines. As a point of comparison, the topical organization of British anthropology today is similar, in this structural sense, to that of British economics in the late 1980s.

These findings lend support to other studies that have documented a profound transformation of British social sciences in the last half century. Economics is, perhaps, the most paradigmatic example of this process. From having had a higher concentration of specialist departments prior to the 1980s, the discipline has become more isomorphic and transnational in the past thirty years (Fourcade 2009). An indicator of this transformation may well be how heterodox traditions of economic thought fared in this changing field: as Frederick Lee and colleagues argue (Lee, Pham, and Gu 2013), the steady decline in the hiring of heterodox economists by economics departments and an “increasing number of departments with only mainstream economists on staff combined with the progressive decline in heterodox submissions to the economics RAE panels has created the all-too-accurate impression that economics is a discipline with only a single paradigm of mainstream economics”. These changes are not unique to economics. British sociology has, too, experienced an important shift in its disciplinary structure. Although still intellectually diverse, a 2010 international benchmarking exercise noted distinctive strengths in British sociology (social and policy, health and medical research, science and technology studies, economic sociology, cultural sociology, gender and women’s studies, and social theory) that are topics commonly found across organizational units. This contrasts with sociology’s earlier disciplinary organization, where specialization was more often than not the defining feature of departments (Savage 2016; Soffer 1982).

The slight move away from thematically differentiated departments to increasingly similar units that echo dominant disciplinary logics is more convincing when we take into account the model of interinstitutional mobility. Disciplinary transformations were modulated not only by processes of internationalization and endogenous disciplinary change, but also by how external research assessments shaped the labor markets for UK social scientists. Here, the most notable indicator that research evaluations impinged on patterns of interinstitutional mobility is provided by the variable that proxies the structural incoherence of a scholar under the assessment. If a scholar is employed at an institution where her work is not assessed in the relevant discipline (for example, an anthropologist by publications working at an institution that isn’t assessed in anthropology), she will have

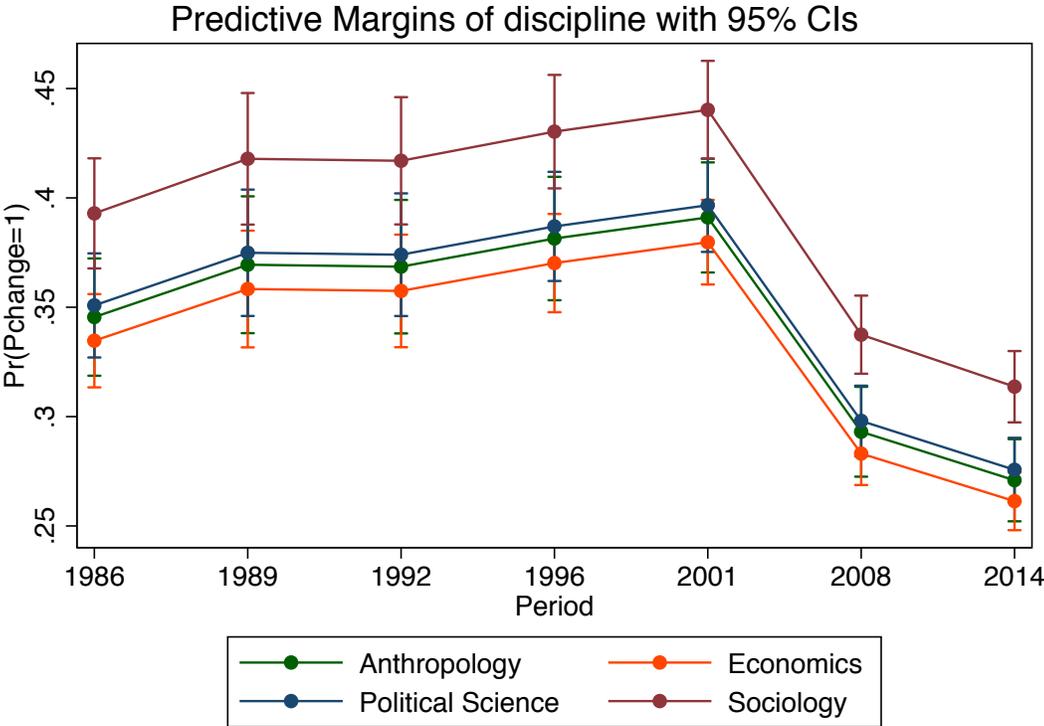
a predicted rate of mobility 40% higher than her structurally coherent peers. If interinstitutional mobility were not affected by research assessments, this completely artificial, organizationally bound variable would have no effects on transition probabilities.

The two other variables that encode positions in the field are similarly revealing. Because departmental typicality and scholarly distinctiveness are both associated to interinstitutional mobility, the model suggests that individuals are sorted in particularly patterned ways within their disciplinary fields. Indeed, the evidence from this study suggests a pattern not too dissimilar from market-based matching (Roth 1986; Roth and Peranson 1999; Roth and Sotomayor 1992), whereby the movement of scholars across institutions reflects specific disciplinary complementarities: institutions are matched onto institutions that make them less atypical while retaining those that preserve their centrality. Research evaluations are effectively market devices (Callon, Millo, and Muniesa 2007; Muniesa, Millo, and Callon 2007) framing and acting not on knowledge itself (this is not a case of the marketization of claims or the commoditization of science) but on the often-unattended labor markets that slowly transform disciplinary formation. From this perspective, the reduction in thematic diversity across the different fields presented in figure 1 can be understood as resulting from a market-mediated form of mimetic isomorphism (DiMaggio and Powell 1983) that emerges not from the actions of individuals reorienting their research (that is, shifting from one topic to another, as emphasized in the research evaluation literature (Rijke et al. 2016)) but through their institutional mediation in labor.

What is similarly telling from the models is that research evaluations didn't only perturb labor markets, acting as exogenous forces shaping an otherwise self-contained but well bounded and defined competitive field. As noted above, research evaluations are best thought of as constitutive of fields of competition and commensuration that would have otherwise not existed. This is partly visible in the predicted probabilities of mobility, which suggest variations in predicted interinstitutional mobility throughout evaluations, rising in the first iterations of the Research Assessment Exercise from the observed 1970s baselines to then fall with later iterations of the RAE and the implementation of the Research

Excellence Framework (Figure 3). While it is impossible to infer anything from these predicted probabilities, supplemental analyses using a random sample of 200 scholars who changed institutions at least once (100 men, 100 women) suggest that men, particularly those in higher status institutions, were about 1.3 times more likely to transition into higher ranks (using mobility as a means for promotion) than women who often moved horizontally into similar ranks. In the context of the broader literature on labor market dynamics, this hints at the existence of some form of capital that, tied to status the positions of organizations, grants greater degrees of mobility to some scholars over others. The critical piece of data missing here, of course, pertains to salaries. Unfortunately, the type of analyses that would reveal the presence of forms of capital that, intersectionally, modulate rates of mobility, earnings, visibility, and ‘professional’ success in the field would require detailed wage data for British academics which is surprisingly sparse and difficult to obtain.

Figure 4. Predicted probabilities of mobility for scholars across the 4 disciplines



This study reveals, however, some of the boundary conditions that make the quantification of knowledge a factor for epistemic change. The British case highlights at least three features: first, a mobile market where individuals have higher possibilities of changing institutions. The lack of tenure protection in the higher education sector makes interinstitutional mobility more likely than in settings where academics have higher levels of contractual stability. Second, the evaluation's focus on units rather than individuals creates a looser coupling between notions of disciplinary quality and a scholar's research decisions, leading to a more prominent role of the labor market. If evaluations placed incentives directly on individuals, we should expect matching to be less relevant since it would be in the interest of academics to conform to the dominant conceptions of value of the assessment regardless of their affiliation. Because of the relative ambiguity and decentering of the British assessment system, there remains some flexibility in scholars' preferences that allows for a stronger commitment to their research programs. Change in the field comes not from task realignments or a direct punishing of interdisciplinarity, but rather by how the departmental system rearranges around disciplinary norms and conceptions. This highlights a third feature: centralization. Scholars at least in the period covered by this paper were not concerned with multiple evaluative ranking. When these existed (as in economics, where an internal 'ranking' of journals matters for decisions of contractual renewal and promotion), they were entirely aligned with the evaluative frameworks of the assessment (Interviewer B.7). Centralization made these assessments powerful not by force of coercion but by the virtue of uniqueness.

This paper also signals an important constraint to how fast and how intensely research evaluations can transform epistemic fields. In their classical work on interinstitutional mobility, Allison and Long (1987) remind us that the labor market only goes so far in explaining career outcomes for academics. "Moves are voluntary", they note, and scientists may choose their jobs on the basis of complex, individual criteria. If a matching process, then, the type of distortions introduced to labor markets by quantifying knowledge are imperfect in the sense that decisions to move, retain, dismiss, promote, and remain in positions are not only driven by the pecuniary incentives of research funding allocation and status competition in the field but, as importantly, by the uneven and

idiosyncratic patterning of organizational dispositions in the field. Distinct organizational cultures will mediate, with different intensities, foci of attention, and logics of practice, how evaluation is interpreted and made sensible. For some institutions, it may create environments where the cognitive and economic costs of searching and pursuing other positions is acceptable; for others, evaluations may simply not matter because of how administrators and scholars are buffered from the logics of precarity and austerity of the exercise. As in other labor markets, the results of this study are suggestive of how, on average, research evaluations change fields. How this occurs in practice will depend on the organizational ecology of the field and the variability of circumstances, resources and cultures across institutions, features that are simply not captured by the data at hand. In the end, however, the conclusions hold: quantifying knowledge changes our world by shifting, even is slightly, how we build the institutions through which scholarship is produced.

References

- Abbott, Andrew. 2005. "Linked Ecologies: States and Universities as Environments for Professions." *Sociological Theory* 23(3):245–274.
- Abbott, Andrew. 2010. *Chaos of Disciplines*. University of Chicago Press.
- Abramo, Giovanni, Ciriaco Andrea D'Angelo, and Alessandro Caprasecca. 2009. "Allocative Efficiency in Public Research Funding: Can Bibliometrics Help?" *Research Policy* 38(1):206–215.
- Aistleitner, Matthias, Jakob Kapeller, and Stefan Steinerberger. 2018. "The Power of Scientometrics and the Development of Economics." *Journal of Economic Issues* 52(3):816–834.
- Allison, Paul D. and J. Scott Long. 1987. "Interuniversity Mobility of Academic Scientists." *American Sociological Review* 643–652.
- Anderson, Alun. 1986. "Research Gradings Stir Emotions." *Nature* 322(6077):299–299.
- Arrow, Kenneth J. 1969. "The Organization of Economic Activity: Issues Pertinent to the Choice of Market versus Nonmarket Allocation." *The Analysis and Evaluation of Public Expenditure: The PPB System* 1:59–73.
- Backhouse, Roger E. and Roger Backhouse. 2002. *The Penguin History of Economics*. Penguin UK.
- Backhouse, Roger E. and Béatrice Cherrier. 2017. "The Age of the Applied Economist: The Transformation of Economics since the 1970s." *History of Political Economy* 49(Supplement):1–33.
- Barbezat, Debra A. and James W. Hughes. 2001. "The Effect of Job Mobility on Academic Salaries." *Contemporary Economic Policy* 19(4):409–423.
- Barnes, Barry. 1983. "Social Life as Bootstrapped Induction." *Sociology* 17(4):524–545.
- Barnes, Barry, David Bloor, and John Henry. 1996. *Scientific Knowledge: A Sociological Analysis*. University of Chicago Press.
- Berman, Elizabeth Popp and Daniel Hirschman. 2018. *The Sociology of Quantification: Where Are We Now?* SAGE Publications Sage CA: Los Angeles, CA.
- Blei, David M., Andrew Y. Ng, and Michael I. Jordan. 2003. "Latent Dirichlet Allocation." *Journal of Machine Learning Research* 3(Jan):993–1022.
- Bloor, David. 1991. *Knowledge and Social Imagery*. University of Chicago Press.

- Bourdieu, Pierre. 1988. *Homo Academicus*. Stanford University Press.
- Brown, Roger and Helen Carasso. 2013. *Everything for Sale? The Marketisation of UK Higher Education*. Routledge.
- Burris, Val. 2004. "The Academic Caste System: Prestige Hierarchies in PhD Exchange Networks." *American Sociological Review* 69(2):239–264.
- Burrows, Roger. 2012. "Living with the H-Index? Metric Assemblages in the Contemporary Academy." *The Sociological Review* 60(2):355–372.
- Callon, Michel, Yuval Millo, and Fabian Muniesa. 2007. *Market Devices*.
- Campbell, David FJ and Elias G. Carayannis. 2012. *Epistemic Governance in Higher Education: Quality Enhancement of Universities for Development*. Springer Science & Business Media.
- Clarke, Caroline A. and David Knights. 2015. "Careering through Academia: Securing Identities or Engaging Ethical Subjectivities?" *Human Relations* 68(12):1865–1888.
- Collins, Harry. 2010. *Gravity's Shadow: The Search for Gravitational Waves*. University of Chicago Press.
- Crane, Diana. 1970. "The Academic Marketplace Revisited: A Study of Faculty Mobility Using the Cartter Ratings." *American Journal of Sociology* 75(6):953–964.
- Czaika, Mathias and Sultan Orazbayev. 2018. "The Globalisation of Scientific Mobility, 1970–2014." *Applied Geography* 96:1–10.
- David, Daniel and Petre Frangopol. 2015. "The Lost Paradise, the Original Sin, and the Dodo Bird: A Scientometrics Sapere Aude Manifesto as a Reply to the Leiden Manifesto on Scientometrics." *Scientometrics* 105(3):2255–2257.
- De Fraja, Gianni, Giovanni Facchini, and John Gathergood. 2016. "Professorial Salaries and Research Performance in the 2014 Research Excellence Framework." *Voxeu. Org.*([Http://Voxeu.Org/Article/Professorial-Salaries-and-Researchperformance](http://Voxeu.Org/Article/Professorial-Salaries-and-Researchperformance)).
- DiMaggio, Paul J. and Walter W. Powell. 1983. "The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields." *American Sociological Review* 147–160.
- DiMaggio, Paul, Manish Nag, and David Blei. 2013. "Exploiting Affinities between Topic Modeling and the Sociological Perspective on Culture: Application to Newspaper Coverage of US Government Arts Funding." *Poetics* 41(6):570–606.
- Easton, David, Luigi Graziano, and John Gunnell. 2002. *The Development of Political Science: A Comparative Survey*. Routledge.

- Education, Department of and Science. 1985. *The Development of Higher Education into the 1990s*. HMSO London.
- Elton, Lewis. 2000. "The UK Research Assessment Exercise: Unintended Consequences." *Higher Education Quarterly* 54(3):274–83.
- Espeland, Wendy Nelson and Michael Sauder. 2007. "Rankings and Reactivity: How Public Measures Recreate Social Worlds." *American Journal of Sociology* 113(1):1–40.
- Espeland, Wendy Nelson and Mitchell L. Stevens. 2008. "A Sociology of Quantification." *European Journal of Sociology/Archives Européennes de Sociologie* 49(3):401–436.
- February 13th, 2018|Higher education, Impact, REF2014, REF2021, Research funding, and Research policy|0 Comments. 2018. "The RAE/REF Have Engendered Evaluation Selectivity and Strategic Behaviour, Reinforced Scientific Norms, and Further Stratified UK Higher Education." *Impact of Social Sciences*. Retrieved January 21, 2020 (<https://blogs.lse.ac.uk/impactofsocialsciences/2018/02/13/the-raeref-have-engendered-evaluation-selectivity-and-strategic-behaviour-reinforced-scientific-norms-and-further-stratified-uk-higher-education/>).
- Fourcade, Marion. 2006. "The Construction of a Global Profession: The Transnationalization of Economics." *American Journal of Sociology* 112(1):145–194.
- Fourcade, Marion. 2009. *Economists and Societies: Discipline and Profession in the United States, Britain, and France, 1890s to 1990s*. Vol. 38. Princeton University Press.
- Frickel, Scott and Neil Gross. 2005. "A General Theory of Scientific/Intellectual Movements." *American Sociological Review* 70(2):204–232.
- Fuller, Sylvia. 2008. "Job Mobility and Wage Trajectories for Men and Women in the United States." *American Sociological Review* 73(1):158–183.
- Geuna, Aldo and Ben R. Martin. 2003. "University Research Evaluation and Funding: An International Comparison." *Minerva* 41(4):277–304.
- Gill, Rosalind. 2014. "Academics, Cultural Workers and Critical Labour Studies." *Journal of Cultural Economy* 7(1):12–30.
- Gingras, Yves. 2016. *Bibliometrics and Research Evaluation: Uses and Abuses*. MIT Press.
- Gläser, Jochen and Grit Laudel. 2016. "Governing Science: How Science Policy Shapes Research Content." *European Journal of Sociology/Archives Européennes de Sociologie* 57(1):117–168.
- Godin, Benoît. 2004. *Measurement and Statistics on Science and Technology: 1920 to the Present*. Routledge.
- Godin, Benoît. 2006. "On the Origins of Bibliometrics." *Scientometrics* 68(1):109–133.

- Grimes, Paul W. and Charles A. Register. 1997. "Career Publications and Academic Job Rank: Evidence from the Class of 1968." *The Journal of Economic Education* 28(1):82-92.
- Hacking, Ian and Jan Hacking. 1983. *Representing and Intervening: Introductory Topics in the Philosophy of Natural Science*. Cambridge university press.
- Halsey, Albert Henry. 2004. *A History of Sociology in Britain: Science, Literature, and Society*. Oxford University Press.
- Hamann, Julian. 2016. "The Visible Hand of Research Performance Assessment." *Higher Education* 72(6):761-779.
- Hammarfelt, Björn and Sarah De Rijcke. 2015. "Accountability in Context: Effects of Research Evaluation Systems on Publication Practices, Disciplinary Norms, and Individual Working Routines in the Faculty of Arts at Uppsala University." *Research Evaluation* 24(1):63-77.
- Hicks, Diana. 2012. "Performance-Based University Research Funding Systems." *Research Policy* 41(2):251-261.
- Hicks, Diana, Paul Wouters, Ludo Waltman, Sarah De Rijcke, and Ismael Rafols. 2015. "Bibliometrics: The Leiden Manifesto for Research Metrics." *Nature News* 520(7548):429.
- Hilmer, Christiana and Michael Hilmer. 2010. "Are There Gender Differences in the Job Mobility Patterns of Academic Economists?" *American Economic Review* 100(2):353-57.
- Johnes, Jill and Jim Taylor. 1992. "The 1989 Research Selectivity Exercise: A Statistical Analysis of Differences in Research Rating between Universities at the Cost Centre Level." *Higher Education Quarterly* 46(1):67-87.
- Karell, Daniel and Michael Freedman. 2019. "Rhetorics of Radicalism." *American Sociological Review* 84(4):726-53.
- Kuklick, Henrika. 1991. *The Savage within: The Social History of British Anthropology, 1885-1945*. Cambridge University Press.
- Kuklick, Henrika. 2009. *New History of Anthropology*. John Wiley & Sons.
- Latour, Bruno. 1987. *Science in Action: How to Follow Scientists and Engineers through Society*. Harvard university press.
- Leahey, Erin. 2007. "Not by Productivity Alone: How Visibility and Specialization Contribute to Academic Earnings." *American Sociological Review* 72(4):533-561.

- Leathwood, Carole. 2017. "Women Academic Researchers: Still Interlopers in the UK Academy?" Pp. 227–242 in *The changing role of women in higher education*. Springer.
- Lee, Frederic S. and Sandra Harley. 1998. "Peer Review, the Research Assessment Exercise and the Demise of Non-Mainstream Economics." *Capital & Class* 22(3):23–51.
- Lee, Frederic S., Xuan Pham, and Gyun Gu. 2013. "The UK Research Assessment Exercise and the Narrowing of UK Economics." *Cambridge Journal of Economics* 37(4):693–717.
- Lewis, Jenny M. 2015. "Research Policy as 'Carrots and Sticks': Governance Strategies in Australia, the United Kingdom and New Zealand." Pp. 131–150 in *Varieties of Governance*. Springer.
- Leydesdorff, Loet. 2007. "Betweenness Centrality as an Indicator of the Interdisciplinarity of Scientific Journals." *Journal of the American Society for Information Science and Technology* 58(9):1303–1319.
- Locke, William. 2014. "Shifting Academic Careers: Implications for Enhancing Professionalism in Teaching and Supporting Learning."
- Long, J. Scott. 1978. "Productivity and Academic Position in the Scientific Career." *American Sociological Review* 889–908.
- Long, J. Scott, Paul D. Allison, and Robert McGinnis. 1993. "Rank Advancement in Academic Careers: Sex Differences and the Effects of Productivity." *American Sociological Review* 703–722.
- Lotka, Alfred J. 1926. "The Frequency Distribution of Scientific Productivity." *Journal of the Washington Academy of Sciences* 16(12):317–323.
- Mennicken, Andrea and Wendy Nelson Espeland. 2019. "What's New with Numbers? Sociological Approaches to the Study of Quantification." *Annual Review of Sociology* 45.
- Merton, Robert K. 1973. *The Sociology of Science: Theoretical and Empirical Investigations*. University of Chicago press.
- Metcalf, Hilary, Great Britain, and Department for Education and Skills. 2005. *Recruitment and Retention of Academic Staff in Higher Education*. London: Department for Education and Skills.
- Mohr, John W. and Petko Bogdanov. 2013. *Introduction—Topic Models: What They Are and Why They Matter*. Elsevier.

- Muniesa, Fabian, Yuval Millo, and Michel Callon. 2007. "An Introduction to Market Devices." *The Sociological Review* 55(2_suppl):1–12.
- Parish, Austin J., Kevin W. Boyack, and John PA Ioannidis. 2018. "Dynamics of Co-Authorship and Productivity across Different Fields of Scientific Research." *PloS One* 13(1):e0189742.
- Phillimore, A. J. 1989. "University Research Performance Indicators in Practice: The University Grants Committee's Evaluation of British Universities, 1985–86." *Research Policy* 18(5):255–271.
- Power, Michael. 1997. *The Audit Society: Rituals of Verification*. OUP Oxford.
- Ràfols, Ismael. 2018. "S&T Indicators in the Wild: Contextualization and Participation for Responsible Metrics." *Research Evaluation* 28(1):7–22.
- de Rijcke, Sarah and Alexander Rushforth. 2015. "To Intervene or Not to Intervene; Is That the Question? On the Role of Scientometrics in Research Evaluation." *Journal of the Association for Information Science and Technology* 66(9):1954–1958.
- Rijcke, Sarah de, Paul F. Wouters, Alex D. Rushforth, Thomas P. Franssen, and Björn Hammarfelt. 2016. "Evaluation Practices and Effects of Indicator Use—a Literature Review." *Research Evaluation* 25(2):161–169.
- Rosenfeld, Rachel A. and Jo Ann Jones. 1986. "Institutional Mobility among Academics: The Case of Psychologists." *Sociology of Education* 212–226.
- Roth, Alvin E. 1986. "On the Allocation of Residents to Rural Hospitals: A General Property of Two-Sided Matching Markets." *Econometrica: Journal of the Econometric Society* 425–427.
- Roth, Alvin E. and Elliott Peranson. 1999. "The Redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic Design." *American Economic Review* 89(4):748–780.
- Roth, Alvin E. and Marilda Sotomayor. 1992. "Two-Sided Matching." *Handbook of Game Theory with Economic Applications* 1:485–541.
- Sauder, Michael and Wendy Nelson Espeland. 2009. "The Discipline of Rankings: Tight Coupling and Organizational Change." *American Sociological Review* 74(1):63–82.
- Savage, Mike. 2016. "The Fall and Rise of Class Analysis in British Sociology, 1950-2016." *Tempo Social* 28(2):57–72.
- Sayer, Derek. 2014. *Rank Hypocrisies: The Insult of the REF*. Sage.
- Schofer, Evan. 2003. "The Global Institutionalization of Geological Science, 1800 to 1990." *American Sociological Review* 730–759.

- Scott, Peter. 2003. "The Ethical Implications of the New Research Paradigm." *Science and Engineering Ethics* 9(1):73–84.
- Shore, Cris, Susan Wright, Vered Amit, Judy Brown, Casper Bruun Jensen, Mark Maguire, Sally Engle Merry, Michael Sauder, Wendy Espeland, and Gavin Smith. 2015. "Audit Culture Revisited: Rankings, Ratings, and the Reassembling of Society." *Current Anthropology* 56(3):431–432.
- Soffer, Reba N. 1982. "Why Do Disciplines Fail? The Strange Case of British Sociology." *The English Historical Review* 97(385):767–802.
- Stephan, Paula, Reinhilde Veugelers, and Jian Wang. 2017. "Reviewers Are Blinkered by Bibliometrics." *Nature News* 544(7651):411.
- Strathern, Marilyn. 2003. *Audit Cultures: Anthropological Studies in Accountability, Ethics and the Academy*. Routledge.
- Sugimoto, Cassidy R. and Vincent Larivière. 2018. *Measuring Research: What Everyone Needs to Know*. Oxford University Press.
- Talib, Ameen and Anthony Steele. 2000. "The Research Assessment Exercise: Strategies and Trade-Offs." *Higher Education Quarterly* 54(1):68–87.
- Urry, James. 2012. *Before Social Anthropology: Essays on the History of British Anthropology*. Routledge.
- Vostal, Filip. 2015. "Academic Life in the Fast Lane: The Experience of Time and Speed in British Academia." *Time & Society* 24(1):71–95.
- White, Douglas R. and Stephen P. Borgatti. 1994. "Betweenness Centrality Measures for Directed Graphs." *Social Networks* 16(4):335–346.
- Whitley, Richard. 2000. *The Intellectual and Social Organization of the Sciences*. Oxford University Press on Demand.
- Whitley, Richard, Jochen Gläser, and Lars Engwall. 2010. *Reconfiguring Knowledge Production: Changing Authority Relationships in the Sciences and Their Consequences for Intellectual Innovation*. Oxford University Press.
- Zuckerman, Ezra W. 1999. "The Categorical Imperative: Securities Analysts and the Illegitimacy Discount." *American Journal of Sociology* 104(5):1398–1438.