The extended computational case method: a framework for research design

Juan Pablo Pardo-Guerra*

UC San Diego

Department of Sociology

jpardoguerra@ucsd.edu

Prithviraj Pahwa

UC San Diego

Computer Science and Engineering

pspahwa@ucsd.edu

This paper considers the adoption of computational techniques within research designs modeled after the extended case method. Echoing calls to augment the power of contemporary researchers through the adoption of computational text analysis methods, we offer a framework for thinking about how such techniques can be integrated into quasi-ethnographic workflows to address broad, structural sociological claims. We focus, in particular, on how this adoption of novel forms of evidence impacts corpus design and interpretation (which we tie to matters of casing), theoretical elaboration (which we associate to moving empirical claims across scales and empirical domains), and verification (which we see as a process of reflexive scaffolding of theoretical claims). We provide an example of the use of this framework through a study of the marketization of social scientific knowledge in the United Kingdom.

1. Introduction

Among the fast-growing contributions to computational social science, research designs that use machine learning techniques within strategies of interpretative analysis have become increasingly important instruments in the repertoires of qualitative sociologists. Rather than reinforcing longstanding methodological divides, this emerging scholarship uses computational techniques—particularly those applied to textual, machine-readable data—as aids to interpretative analysis. Harnessing techniques for data reduction and pattern recognition, this scholarship leverages computational models as means for generating “insight[s] about what to look for in the data and how to theorize what is being observed”. (Baden et al. 2021; Grigoropoulou and Small 2022). Arguably most prominent among these, the program of computational grounded theory (CGT) developed by Laura K Nelson (2020a) and the framework of computational ethnography proposed by Corey Abramson et. al. (2018) demonstrate the exceptional payoffs of this strategy, tying the rigorous study of digital textual data to fine-grained qualitative explorations of patterns, mechanisms, and causality (Edelmann et al. 2020; Karell and Freedman 2019; Nelson 2020b, 2021a, 2021b).

Building on these contributions, in this paper we address a different instantiation of interpretative computational social science. Our focus is exploring what a form of computational social science looks like when shaped by the logics of the extended case method (ECM) (Burawoy 1998a, 2009; Van Velsen 1979). The issue that motivates this exploration is the same question of the extension of claims (Barnes 2013; Hesse 2020) that animates other debates in research design. By this problem of extension, we mean the following: what are the conditions under which we can make assertions about general social processes based on a constrained set of empirical instances? This challenge involves the problem of generalizing empirical findings, central to a variety of research literatures, from techniques of statistical inference to debates informing decision points in qualitative analysis. As we argue in this paper, this issue also underpins the contributions of computational social science when inflected towards ethnographic, interpretative, and reflexive sensibilities.

The contribution of this article stems from presenting a way of integrating computational workflows into the research practices of qualitative social scientists along the lines of the extended case. The point here isn't about using computational methods *in* ethnographic/interpretative research or importing ethnographic logics *onto* computational workflows. Rather, our proposal for an *extended computational case method* (ECCM) combines different data and analytical approaches in a strategy that builds theoretical claims by comparing findings across empirical and methodological domains—it tackles the problem of extension through the sequential articulation of computational and non-computational evidence. As such, this article does not present a specific technique of computational data reduction or benchmark a new mechanism for analyzing text. The contribution resides elsewhere, in presenting a framework where computational text analysis and qualitative research are linked within a single comparative workflow that addresses and elaborates general theoretical claims. This involves, for example, using results from computational workflows as prompts shared with informants in ethnographic research, or designing computational analyses that respond to questions raised by informants in the ethnographic field.

Our contribution necessarily builds on computational grounded theory (Nelson 2020b) and use of big data in ethnographic analysis (Abramson et al. 2018; Bjerre-Nielsen and Glavind 2022), although it is distinct from these two approaches. In these, computational techniques are primarily used to scale tasks of pattern recognition and data reduction (Baden et al. 2021; Deterding and Waters 2021). Computational text analysis methods, in particular, are means for empowering in-depth grounded coding (e.g. Nelson 2021a) and hybrid ethnographic coding (Li, Dohan, and Abramson 2021) while preserving the interpretative role of the expert analysts (Deterding and Waters 2021; Nelson 2020a).

By integrating computational techniques and the analytical strategies of the extended case method, however, our approach uses an abductive strategy based on the sequential iteration of findings across bodies of evidence with the aim of refining theoretical claims. The metaphor driving our proposal is not augmentation, but the articulation of evidence across empirical and methodological domains. This changes how

we approach the issue of the extension of claims with respect to these other approaches. Computational ethnography, for example, uses techniques to render manageable the ethnographic field, centering the problem of extension on scaling the analysis to “include more sites, more researchers, more subjects, and greater amount of data”, hence mirroring the systematicity of traditional statistical studies (Abramson et al. 2018:260). Similarly, computational grounded theory tackles the construction and generalization of claims by “building them up” from a large-N, empirically rich ground. A feature of this approach is the ostensive lack of commitment to theory accompanied by transparent protocols for collecting, coding, and analyzing data. These allow tracing how abstract claims are made based on the analysis of specific empirical instances. The extended case method, by contrast, begins with an explicit logic of elaboration: rather than building up from empirical data and a position of theoretical agnosticism, or seeking systematically sampled pools to address questions of generalizability, it uses observations from different domains to shift theoretical claims as part of an ongoing, reflexive rewriting of theories of general mechanisms (Burawoy 2009). As Burawoy famously noted, the extended case method starts with theory to refute and deepen its content; it is a means for putting theoretical claims to the test that, unlike falsification or hypothesis testing, involves a reflexive reformulation of theory across empirical settings and cases (Burawoy 1998a).

Despite sharing the expansive logic of the traditional extended case method (it is oriented towards general, ‘macro’ processes through a theory-driven exploration), the extended computational case method is consistent with the interpretative emphasis of grounded theory and computational ethnography (which inductively create theoretical accounts out of specific empirical settings). Computational grounded theory has demonstrated how computational methods of text analysis and natural language processing complement the coding practices associated to grounded theory approaches. The question is, however: can these methods be transported to query larger theoretical systems (theories of capitalism, of power, of systemic inequalities) without losing interpretative richness and without the agnosticism of grounded theory? The challenge isn’t just expanding computational methods through the well-known hypothesis-testing

frameworks or tacking them on as confirmatory evidence to qualitative research but, rather, blending them directly with the logics of the extended case.

In addressing this question, we propose three specific opportunities for the synthesis of computational techniques and research designs consistent with the traditions of the extended case method. These speak to the practical and methodological differences between the ethnographic foundations of the ECM and the qualitative, inductive theory-development of grounded theory and other mixed, augmented approaches: specifically, we focus on issues of *casing*, *scaling*, and *scaffolding* in relation to the production of sociological claims. After describing these three dimensions and how they diverge between computational grounded theory and the extended computational case method, we then explore a recent contribution that outlines the former in the domain of the ‘science of science’.

2. The extended computational case method in contrast

To establish the payoffs of the extended computational case method, let us start by building it in comparison to qualitative designs that focus on augmenting the researcher’s ability to recognize, classify, and query patterns in empirical data, as represented by computational grounded theory (CGT) and computational ethnography (CE).

The rise of computational social science over the past decade was tied to a rapid growth of studies that investigate human behaviors, mostly using novel digital datasets (Burrows and Savage 2014; Lazer et al. 2009; McFarland, Lewis, and Goldberg 2016; Molina and Garip 2019; Mützel 2015). From historical research on modernization (Knigge et al. 2014) to studies that benefit from new types of social network and behavioral trace data (Lewis and Kaufman 2018), most of the ‘large number’ studies associated to computational social science follow relatively traditional research designs: theoretical claims from the literature (about, for example, the predictors of modernization in Europe), are rendered into testable hypothesis that are then explored through some statistical model (for example, of occupational similarity of siblings living in different contexts (Knigge et al. 2014)). This mode of research has led to important contributions yet remains

an extension of established forms of (mostly quantitative) research where theory is verified (or falsified) according to the findings of a specific statistical (often regression) model. In significant ways, computational techniques that deal with textual data fall askew of this model, as exemplified by computational grounded theory.

Grounded theory and its computational extensions are reactions to the ‘traditional’ analytical framework of statistical verification, seeking “patterns of action as well as changes in these action patterns over time” (Decoteau 2017:60). Unlike verificationist approaches, grounded theory generates hypotheses and concepts from the data itself. Importantly, the process of constructing claims in grounded theory avoids commitments to existing theory. As Glaser and Strauss noted, “if the analyst starts with raw data, [they] will end up initially with a substantive theory” while those who start with theoretically bound findings “will end up with formal theory pertaining to a conceptual area” (Glaser and Strauss 2017). Through constant comparative analysis, grounded theory researchers are thus able to generate theory that speaks to the data and works beyond its local constraints.

The commitment to “raw data” as the starting point for the inductive construction of theoretical claims explains, in part, the affinities between computational text analysis methods and grounded theory. As Nelson writes, computational grounded theory is an update where automated “techniques [of text analysis provide] the ability to incorporate massive amounts of data into theory-generating research in a rigorous and reliable fashion, mitigating the shortcomings of purely qualitative research” (Nelson 2020a). CGT thus mirrors other forms of computational social science insofar as it stands as a transposition of previous research strategies onto larger datasets. The fact that techniques of computational text analysis and classification afford researchers the capacity to identify patterns in vast corpora is directly tied, within CGT, to making “more substantive” claims: “combining the interpretive and computational approaches, [CGT delivers] both *quantity* and quality, breadth and depth, and allows researchers to leverage both close and distant reading to better measure meaning” (emphasis added) (Nelson 2020a). This mirrors other modes of computational augmentation—including computational ethnography where new

tools “offer resources for moving beyond ‘small n’ exploratory work, allowing larger samples through scaling” (Abramson et al. 2018:261).

This strategy of augmentation is visible in the way computational grounded theory is operationalized in practice (Nelson 2020b). Given a textual corpus, computational techniques are used to, first, detect patterns—by, for instance, identifying clusters of terms in the corpus, or structural topics that classify texts in some manner—to then refine them through either a revision of the findings or of the corpus itself. With this, a potentially large collection of documents can be reduced into “interpretable groups of words, helping researchers cut through the noise inherent to text-based data”. These textual features lack meaning, requiring the intervention of an expert analyst and so the second step involves “a deep reading of the text and incorporates holistic interpretation”. In this process, researchers revise their categories and classifications, which effectively constitute the coding schemas of their grounded theory, reconsidering both the parameters of models and the contents of the corpus. The result is an iterative process where statements are built through a recursive computational-interpretative loop. The third and final step involves testing the validity of the patterns identified in the text. This step allows for the extension of claims from the corpus to broader empirical phenomena.

The extended case method relies on a different strategy—as does the computational variant we present. Here, the focus is refining or elaborating structural theories by selecting and analyzing cases that reveal the “social forces that impress themselves on the ethnographic locale” (Decoteau 2017). This places emphasis on the theoretical repertoires that influence case selection and moves us away from questions of systematic sampling as a condition for the generalization of claims. Given their articulations to a general, structural theory, cases allow for an extension of claims by reflexively moving away from individual empirical instances in the field and onto a wider analysis of the regimes of power that are at play in observed processes. Theories are thus elaborated by stringing reports across domains and linking observations and refutations of these with a broader theoretical claim. Thus, while GT is in principle agnostic to theory, ECM is committed to specific theoretical concerns. Similarly, while GT critically depends on the boundedness of its empirical

foundations from which it derives legitimate coding and analytical claims, ECM allows for a certain flexibility of locales, times, and data insofar as these speak to a form of casing driven by theoretical concerns. In this, the ECM echoes ethnographic ‘structural stories’ that “extend into but also beyond the studied scene” (Van Maanen 2011:168).

How does this translate into a computationally informed workflow? In the same way that CGT is skeptical towards theory yet committed to the empirical primacy of textual data, we see the extended computational case method as agnostic to specific computational methods yet committed to an empirical delineation of what constitutes the field of evidence as well as a reflexive strategy of theory-building. Specifically, just like natural language processing and classification augment the ability of grounded theorists to process data, a variety of computational techniques—from statistical models to complex visualizations—can become additional mechanisms for “framing, focusing, and generalizing” the ethnographic account (Van Maanen 2011). This partly echoes the logic of computational ethnographies, where visualization techniques like modified heatmaps allow identifying and visualizing patterns in the data (see Abramson et al. 2018:262–70 for a discussion of these “ethnoarrays”); yet unlike these approaches, a computational form of the extended case method uses these not as forms of data management but, rather, as instruments for querying ethnographic informants (or *other* empirical locales) and articulating evidence across domains. Computational techniques, in other words, are not just forms of organizing data and identifying patterns, but also means for eliciting data in the ethnographic field. In practical terms, the differences fall across three dimensions (*casing*, *scaling*, and *scaffolding*), which we will now proceed to describe (Figure 1).

Computational Grounded Theory

FIGURE 1. Schematic differences between Computational Grounded Theory (above) and the Extended Computational Case Method (below)

2.1 Casing

In their analysis of ethnographic methods, Timmermans and Tavory identify ‘casing’ as a useful way of framing the differences between grounded theory and the extended case method (Timmermans and Tavory 2012). By ‘casing’, they refer to the way cases are defined as analytical units, of setting “parameters for later theoretical work without predetermining” outcomes (Lichterman and Reed 2015:593). These two approaches diverge in how they constitute and conceptualize cases in relation to ethnographic narratives and theoretical claims. Grounded theory, they argue, assumes boundaries that reflect the way “institutional and inter-subjective mechanisms” order social life and narratives. These ordered narratives encode patterns of action and supply “coherence, predictability and structure” that can be analyzed by the researcher. There is an assumption that boundaries that are readily visible in social life give coherence to narratives and the ethnographic field: if one studies practices of knowledge production in science, for example, the organizational structures of the field (given by laboratories,

departments, units, and institutions) provide the empirical limits of the case. Indeed, the actual casing (how observations ultimately relate to theory) only emerges from this empirical bounding after exploring the field (Tavory and Timmermans 2014) (This point is also highlighted by (Deterding and Waters 2021) who, in their review of forms of computer-assisted flexible coding consistent with grounded theory, note that qualitative researchers “are sensitized by the previous research early on”, informing how they select, approach, and delineate the empirical field).

The extended case method, on the other hand, upholds a less bounded conceptualization of narratives—and thus, of the structure of cases. As Timmermans and Tavory suggest, the ECM works with “the implicit idea that the social world is not neatly or naturally divided into parts and cases” (Timmermans and Tavory 2012). Closure of the field is only possible through theoretical arguments, which indicate what is within and without the space of analysis. Such closure also relies on an iterative process of re-casing, where practices of encasing are informed by a broader “community of inquiry” that gives meaning to the value of particular theoretical claims and their salience for bounding the empirical field (Lichterman and Reed 2015). In a sense, the coherence of narratives as elements of the case comes from how they are rendered relevant or not through a both specific theoretical lenses and how these are made sense of by the ethnographic audience.

When taken to computational research, these divergences have practical implications for the way empirical objects (the digital corpora) are defined and constructed: cases are not assumed as ‘just existing’ but, rather, as conventions shaped and constrained by the practices and concerns of the social sciences (cf. (Ragin and Becker 1992)). Claims in computational social science are anchored to the corpus, in the same way that claims in traditional statistical inferential work remit to the random sample as a representative slice of empirical reality. How a corpus is defined and constituted matters for the type of claims that can be made with its contents as well as for their extension and generalization. Here we find two possibilities of divergence between computational grounded theory and the extended computational case method: echoing the construction of cases, computational grounded theory sees the corpus as the field itself. The decisions,

organizational dynamics, and social forces that led to the creation of each corpus—be they archived documents curated under particular logics, social media content constrained by the algorithmic affordances of digital platforms, or the totality of documents contained within a particular organizational setting—are assumed in the narratives and patterns of its embedded texts. The corpus here *is* the case, establishing a clear empirical boundary as to what is left within and without. Textual models of language are then used to parse, classify, and identify patterns in the digital corpus that serve as the basis of interpretative close readings which add context to the analysis—effectively grounding automated classifications and coding schemes through expert analysis. The claims are then extended, by testing their accuracy with respect to the corpus and abductively building them into theoretical statements.

The theory of data underlying the extended computational case method (ECCM) is different: the corpus is *not* the field but an element in a longer string of interconnected empirical/ethnographic instances; it is a collection of narratives that only hold value when placed in conversation with narratives *outside* the corpus, where they can be assessed by informants and communities of inquiry.

What does this mean in practice? Consider studies that focus on organizations. A grounded strategy might entail using institutional archives of some sort as a proxy of organizational cultures. Such corpus might include speeches, correspondence between members, or other records of interactions that contain meaningful texts as constructed in that organizational setting, which serve as the inputs for computational models. Because these archives are avatars of a richer organizational life, they must be as comprehensive as possible—a shallow archive would not only create computational difficulties (too few texts for models to perform adequately) but would also be a poor representation of existing cultural forms. If cleaning is performed on this textual corpus, it is mostly to clarify its contents (for example, to disambiguate terms) or to remove outliers that somehow break with the proxy-ness of the archive (for example, excluding advertisements from personal communications). A computational grounded theory approach would use this as the basis

of analysis and interpretations, seeking to maintain as much as possible the integrity of the archive.

In the ECCM, the integrity of the corpus as a proxy is less important than the question of how it highlights a theoretical concern. In this case, what data is included and what is excluded from the analysis is dictated not by whether it reflects an underlying empirical referent so much as whether it foregrounds concrete themes of interest. In the archive of an organization, relabeling data in what might seem arbitrary groups or parametrizing the analysis in certain ways may not reflect the everyday cultural logics of meaning-making of the members, but it may provide a way for exploring, for example, questions of power that can only be derived from applying a theoretical lens onto the data. These changes and segmentations of the corpus break with the assumptions of “raw data” of a grounded approach, introducing theoretical considerations into the analysis, before computational or statistical models are even performed.

In this conceptualization of evidence, ECCM understands the corpus as the product of specific forms of work and structures of power that need to be evaluated through theory. This places an additional onus on the researcher for understanding curatorial logics in the constitution of the corpora: knowledge of how a collection of texts was assembled is necessary for evaluating their theoretical significance—it is, in a way, akin to understanding the expectancies of practice (Garfinkel 1963; Wolfinger 2002) that are coded into the data structures of a corpus. This knowledge matters both for casing (that is, for defining what is considered and what is not) and for interpretation: patterns observed within a collection of texts acquire meaning not only through an internal analysis and coding of said textual elements but rather through their contraposition to other narratives available to the researcher that speak to a ‘well identified’ theoretical concern (Burawoy 2009; Lichterman and Reed 2015). This could include, for example, iterating the findings of a computational analysis of a specific corpus with interviews, ethnographic observations, or other participative processes, where informants are able to evaluate patterns and claims about their social worlds derived from computational techniques of pattern identification and information reduction.

Note that the above also constitutes a point of divergence from other mixed, computationally augmented methods like computational ethnography. The key issue is not simply scaling the capacity to analyze empirical instances through computational tools, but rather integrating, through a contrastive logic, automated pattern recognition with ethnographic deliberation. This implies an encasing that is “responsive to actors’ own meanings” yet simultaneously informed by the theoretical concerns of the community of inquiry. And in this, it is distinct. While proposals to integrate computational tools with ethnographic analysis show many payoffs of this strategy, they tend to separate computational analysis from in-depth interviews and ethnographic data collection. Computational processing works as either a way for the analyst to observe *ex post* patterns in data obtained through fieldwork (Abramson et al. 2018; Li et al. 2021) or to infer particular outcomes from ethnographic observations (Bjerre-Nielsen and Glavind 2022).

The above results in two characteristics of ECCM. First, it is more flexible about the types of computational tools and data reduction approaches that may be incorporated into research strategies when compared to CGT and other approaches, which emphasize techniques of pattern-recognition as augmentations to manual, interpretative coding. In the ECCM, the purpose is not only identifying trends and salient features in a corpus but, more broadly, to generate ‘mappings’ that are legible to others in the ethnographic field (see Scaling below). These may include textual patterns discerned through such techniques as topic models and word embeddings yet could also involve visualizations that result from statistical models, novel quantitative metrics, or general claims about behavior that derive from some large-scale quantitative analysis. These instruments matter not uniquely because of how they reveal patterns in isolation but rather in terms of how these same patterns are made sense of by ethnographic informants. Second, this iterative logic of moving back and forth between the corpus and other empirical domains in the field provides opportunities to identify negative cases, that is, narratives that seem to contradict existing accounts and that allow for the refinement of theoretical frameworks (Pacewicz 2020). Taking surprising or peculiar findings of computational results to the ethnographic field becomes a means for exploring when such patterns hold and when they break,

prompting revisions and expansions of theory that are impossible by constraining the analysis to a single digital corpus.

2.2 Scaling

Iteration is not unique to the ECCM to the point that it figures prominently in Nelson's computational grounded theory: as a three-stage process, Nelson reminds researchers of the importance of revising the patterns that they identified through automated, unsupervised methods and of using findings from these to guide the close, interpretative reading of empirical exemplars. Through this abductive process (Timmermans and Tavory 2012) of going back and forth between data and patterns, argues Nelson, researchers can "check their interpretations" to "potentially modify the identified patterns to better fit a human, and holistic, reading of the data" (Nelson 2020b:25).

The difference between approaches resides in how these forms of iterative analytical work occur across scales: because the corpus is effectively the field in CGT, iteration is necessarily *internal*: it is about the robustness of the coding process and the statistical validity of the patterns identified through computational methods which are tested against slices of the corpus itself. By contrast, in ECCM, iteration is *external*, taking observations from computational workflows into ethnographic conversations. In ECCM, iterative and abductive practices are tied to a process of extension across domains rather than a closed bracketing of the empirical field upon which knowledge is generated (Burawoy 1998b; see also Timmermans and Tavory 2012:173). Results in ECCM are then the product of the articulating computational findings (patterns, statistics, claims) across places and time, putting them to the test without an artificial "closure 'grounded' in social life". Through this articulation, the extended computational case method strives "for saturation rather than representation as the stated aims of research" (Small 2009).

What does this look like in practice? Following from the example above, we can imagine using a corporate archive to try to identify some theoretically interesting pattern between, for instance, workers and managers. This may involve a modified corpus which, by design, focuses on specific types of conflicts or dynamics between these two groups.

Because the corpus is an artefact of our initial theoretical concern, it is a poor basis for making generalizable claims. It is, nonetheless, a solid basis for finding patterns that may be relevant for members of the organization. A computational analysis may show, for example, that descriptions of job satisfaction have textual patterns that follow from differentials in organizational power, perhaps in counterintuitive ways. Rather than constraining the analysis to the corpus, scaling would involve taking these findings to the members of the organization (or some other ethnographic locale) as prompts for conversation. (Note that scaling does not have to be from computational to ethnographic but can involve taking findings from one site and using these to design further studies in a different, but comparable, locale. The logic of scaling is about taking findings from one field to another, in order to query them—this can take various forms).

The internal and external orientations of CGT and ECCM respectively, then, are tied to how claims are extended in each of these research strategies. While CGT keeps analysis linked to data—leading to findings that are often associated with mid-range theories—iterative practices in ECCM seek to connect findings across scales to tackle macro-range, structural accounts. Because textual and computational corpora are often the result of arbitrary curatorial decisions and lack the features of random samples, ostensibly constructed to allow for the generalization of claims into populations, structural claims require ‘scaling up and sideways’ the analysis. In this process of scaling, then, the issue is not the primacy of measurements and quantification but using these to *amplify* interpretation and elaborate theory—echoing, for example, Mische’s work (Mische 2009), which moved from ethnography to measurement, “only to return to the actors on the ground” (Mohr et al. 2020:145). This scaling of claims across domains appeals to what Seim calls the “awkward union of participation and observation” (Seim 2021) characteristic of ethnographic work, where ‘participation’ involves presenting informants claims about their social worlds and structural conditions to elicit observations about their assumptions of social order (Garfinkel 1967). Such strategy of moving across scales is also consistent with the forms of ethnographic work proposed by Schritt, where the account of a public event is necessarily tied to different sites and times of observation—rather than a singular ethnographic moment, captured by the embedded observer. As he notes, an application of

the extended case method to these settings involves ascribing to ‘methodological situationalism’, where observations, events, and narratives are placed in a historicized context that allows for their evaluation (Schritt 2019).

2.3 Scaffolding

There is a certain temptation to describe the ECCM through established spatial metaphors. Is this not, for example, just a specific instance of ‘triangulation’ (Flick 2004; Jick 1979), long held as a paradigm for social scientists seeking to establish the credibility of their claims? Triangulation is meant to provide researchers with a ‘comprehensive understanding’ of the phenomena, exploring processes and events through different methods that, while perhaps overlapping in data sources and logic, offer evidence of conceptual convergence in their findings and accounts. We may well prefer cartographic renderings (Lee and Martin 2015), presenting each technique of analysis in the ECCM as a specific projection of the world onto a map that we then use to navigate and make sense of social phenomena. What is, in other words, the uniqueness of the ECCM?

Although useful, these metaphors focus on methods as projections, perspectives, frames, and instruments, disregarding the forms of contingency and indexicality that characterize empirical data and make difficult extending claims into the realm of structural theories. Elaborating such structural theories through a movement across scales is not merely an act of triangulation but is closer, in a sense, to the types of translational process through which researchers in other fields build general claims of the world: they do so not by amassing independent pieces of evidence but by articulating theories across evidentiary sets—for example, in the domain of biomedical sciences, by *translating* claims about the effects of particular pharmacological compounds on mice models to their likely effects on complex human conditions (Nelson 2013). As Nicole Nelson writes, establishing “plausible relationships between mouse experiments and human disorders” in these settings requires “building up a structure of arguments and evidence to substantiate the use of the model to do knowledge production work”. This process of ‘stacking claims’ both about process and models, from animal to human, requires various inferential jumps that, when poorly

supported, do not allow translating findings from one empirical domain (ethological mouse models and physiological data on the effects of anxiolytics, for example) to another (complex, psychiatric human conditions). Building these epistemic scaffolds is not a process of ‘triangulation’ but is rather more complex: it does not occur in the same ‘space’, requiring various forms of non-overlapping expertise, methods, and arguments that, by ‘building up and sideways’ from various forms of evidence, allow comparing and reframing findings across domains.

The extended computational case method thus involves similar types of translational work, with abduction ‘building’ and domain-spanning iterations ‘shaking’ the scaffold of claims to see how well (or not) it bears the weight of new observations. This occurs from the definition of the corpora and how it is delimited by theoretical concerns to the strategies for extending work into the ethnographic field and back again into the digital dataset. In each of these steps, shifts of scale and locale involve acts of translation that test both the nature of empirical findings and the usefulness of theory. For example, presenting findings of a computational study to informants may elicit not only reactions that expose expectancies about social order but, equally, alternative strategies for neutralizing or making sense of such findings and that provide access to other structures of empirical knowledge. This contrastive logic leads to scaffolds that better link knowledge claims and, hence, lead to general theoretical elaboration. This connection is only possible because claims are taken to the field for their translation and evaluation—it is there where attributions of causality, agency, and structure can be had, for example, which the researcher can use to query the appropriateness of existing theory.

In practice, and following from previous examples, this may involve performing further corroborative analyses suggested by the scaling of claims across domains. In the example from the previous section, taking the computational analyses to members of the organization as points of discussion may elicit new questions that serve as corroborations of hypotheses. These may call for an additional leg of research—for example, structured interviews or surveys of members, or further computational analyses—that, like additional scaffolds, provide confidence in the claims derived from the workflow.

Importantly, and mirroring the affordances of both computational ethnography and computational grounded theory, the ability to reproduce some analyses—even if partially—contributes to the quality of the scaffolds, if not by making them entirely reproducible (ethnographic extensions can, at best, be revisits to the field), then by making workflows transparent and accessible to the community of inquiry where the analysis is relevant. Here, ECCM finds agreement with the calls for transparency in data, code, and workflow of other computational approaches, though it is more constrained—scaling results across computational and ethnographic fields is not altogether reproducible, nor are all the considerations behind corpus construction and casing. This does not preclude, however, sharing both code and ethnographic fieldnotes, but simply calls on researchers to describe, in more detail, various decision points and interpretations as supplements to their analysis.

3. Extended Computational Case Method: an application

A domain where the ECCM might prove fruitful concerns mixed-methods studies of science. In recent years, studies of science and of scientific work have flourished, partly owing to the increased availability of both machine-readable data related to scientific publications, policies, and institutions, and new computational techniques for data processing and text classification. An area that was once mostly the province of historical and ethnographic work (classically, (Ben-David and Collins 1966; Ben-David and Sullivan 1975)), studies of science are now increasingly informed by large datasets and quantitative analyses that tackle central questions about meaning and the institutionalization of knowledge (Azoulay et al. 2018).

Captured under the umbrella of the “Science of Science”, this new genre of research has advanced some important sociological claims about the production of knowledge through a growing number of large, quantitative, computational studies. For example, using nearly 2 million abstracts from Clarivate’s Web of Science, McMahan and Evans (2018) show that, instead of hindering discovery, conceptual ambiguity offers spaces for

uncertainty, creativity, and exploration in scientific practices. Similarly, using the bibliographic records of Medline, Foster, Rzhetsky, and Evans (2015) explore the strategies of biomedical researchers who have to navigate between forms of more paradigmatic, tradition-bound research and riskier innovations that somewhat break with existing lines of compound development. Comparable research by Teplitskiy and colleagues (2020) used large scale survey data of citation practices to document the role of journal prestige in how knowledge is referenced by authors. The list of exceptional studies is long and growing.

The contributions of the science of science are particularly salient in the context of the larger literature in the sociology of science because of their anchoring on large datasets that ostensibly lead to generalizable claims. This stands in contrast to the type of qualitative, interpretative studies that characterized an earlier generation of research on science and scientific knowledge. Indeed, at least two of the papers mentioned above have descriptors of the magnitude of their data (“large scale publication data”, “thousands of scientists”, and “millions of abstracts”) in how they frame their contributions. Echoing much computational social science, an important component of the strength of the claims produced by the science of science today rests precisely on this argument of scale and augmentation: these studies cover enough cases to describe not just one science in isolation (astronomy after the renaissance (Kuhn 1957), microbial theory in the nineteenth century (Latour 1993), or collaborative projects in high energy physics (Cetina 2009)) but science *in its entirety*. Claims of scale (which evoke the overall generalizability of positive science) seem to clash with the qualitative/historical approach that characterizes existing research paradigms in specialized sociologies and anthropologies.

This is the substantive crux of the problem: computational studies of science are useful and methodologically interesting contributions to ongoing academic and policy discussions, yet their overall logic seems to clash with previous research paradigms about how best to study change in knowledge and professions. The science of science observes knowledge practices “through the outside, interrogating them through intermediaries” in such a way that they “suspend [their] participation in the world” (Burawoy 1998b). Quantitative science of science “extends out” singularly from the dataset, forgoing

participation in the world and the data-generation process. It is true that previous studies also extended from their qualitative, ethnographic, and historical data, but they often did so with an insistence on reflexivity and co-participation. This contrast seems particularly important today, given the potential prominence of computational studies of science in the public sphere. As Brad Wible wrote in *Science* magazine, this new family of large-scale studies of knowledge and scientific production can perhaps move us away from “intuition-based policy” (Azoulay et al. 2018)—as if previous work were devoid of structured, scientific insights or concrete empirical foundations.

How to reconcile these approaches, harnessing both the empirical and methodological wealth of computational studies of science and the reflexive sensibilities of alternative traditions of scholarship? The extended computational case method offers one possible avenue, combining computational techniques of analysis with the type of ethnographic sensibilities that link scales and domains, allowing for robust generalizable claims.

4. Science and the politics of quantified work

We exemplify these possible contributions through the application of an extended computational case method approach to the question of how modes of quantification shape knowledge production practices. While discussed in the literature on science and research policy, existing accounts of the effects of quantification are restricted in their generalizability and do not address larger theoretical questions about the articulations between science and structural socioeconomic forces. The larger systemic question about how modes of evaluation of scientific productivity are tied to logics of late capitalism and its interventions on labor practices is largely unaddressed.

This structural question (“How does marketization transform scientific knowledge?”) could have been tackled with either a purely computational study or a more traditional historical/ethnographic design. We considered the first route as only partially informative: quantitative studies within the science of science too often treat knowledge

production as an epistemic, rather than labor, puzzle. A historical or ethnographic approach might be able to frame the labor question more prominently, but it would not address larger systemic claims about scientific work owing to the specificity of field sites. A research design that combines computational studies as part of the ethnographic data generation process might, however, resolve these issues.

Organized evaluations of science are 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. Today, the use of metrics as means for evaluating scholarly quality and managing professional careers is a standard approach throughout the world (Abramo, D'Angelo, and Caprasecca 2009; Geuna and Martin 2003; Hicks 2012).

Despite their pervasiveness, the 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 (the “macro” theoretical question). Work in the related field of research evaluation (Rijcke et al. 2016; de Rijcke and Rushforth 2015) has 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.

Most of these studies, however, center around local cases, often forgoing larger theoretical articulations. Gläser and Laudel (2016) convincingly show, for example, 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). The state of the art on research evaluations within sociology thus lacks a generalized 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 are well documented (Espeland and Sauder 2007), less is known about how they and similar processes shape the organization of knowledge and epistemic labor 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 first step in this project thus consisted in using computational techniques to identify large-scale patterns in scientific labor markets that would become *points of conversation* in the field about epistemic changes. The emphasis wasn’t in finding generalizable patterns through a statistical model by, rather, identifying empirical regularities through computational means that would then permit ethnographic exploration.

The focus of our computational analysis was constrained to the social science labor markets in the United Kingdom. This case is salient for two reasons: first, it offers what was, perhaps, the closest thing to a controlled field site: in addition to having a mature higher education sector, the UK’s higher education sector is organizationally less complex than those elsewhere, with most research occurring in public universities that respond to the same central policies of quality and evaluation; second, since 1986, British academics have been the subjects of intense, periodical research evaluations that seek to assess the “excellence” of scholarship across all disciplines, tying these to state funding outcomes: only the highest ranking institutions in each field receive centralized research funding over the next evaluation cycle. These evaluations operate as ‘market devices’ (Callon, Millo, Muniesa 2007), creating both competition and publicly observable scores that are financially consequential for departments and their universities. A useful metaphor for

these evaluations is that of a closed-envelop auction where institutions submit blind bids and are rewarded according to a central selection process. The participants of the auction (academics) have little power over the processes, with strategy about submission coordinated by university managers.

A population-level account of career structures in the British social sciences would potentially offer a mapping of the effects of research evaluations on the system. Computational techniques were hence used for producing a mapping of the field, namely, a large-scale overview of the forces shaping academic careers in the UK over the period during which the evaluations were active. Anchored on the existing literature on academic careers (Leahey 2007; Long 1978; Long, Allison, and McGinnis 1993; Rosenfeld 1992), the first computational iteration of the project identified and visualized how research evaluations might have reinforced systems of incentives around increasing returns to productivity, a common trope associated to ‘academic capitalism’ (Petersen et al. 2014). For this, we focused on career mobility as our outcome variable: while mobility may be downward or lateral, moves between institutions were conceptualized as perturbations of the labor market that exposed the fragility of certain career structures and the stability of others. Inspired by recent work in computational social science (Vaccario et al 2020), our initial intention was to infer a population-level census accounting for a vast majority of scholars active in the British social sciences from 1970 to the present.

4.1 Casing

The first decision point in this project was defining the corpus that matched our theoretical concerns. Lacking a centralized database of British academics, we turned to the Web of Science’s Social Science Citation Index (SSCI) as a means for producing a dataset that represented the career trajectories of UK social scientists. As of 2020, Web of Science’s Social Science Citation Index contained more than 9.37 million records published across more than 3,400 social science journals tracing back over a century. The SSCI was introduced in 1973 after its equivalent in the ‘hard’ sciences, the Science Citation Index (SCI), was launched in 1964. The bibliographic records of SSCI include data about author

affiliations which, as per Vaccario et al (2020), can be mined to create a longitudinal record of institutional affiliations and hence a proxy of career mobility. The bibliographic records we obtained from Web of Science's SSCI included all contents with at least one UK author published between 1970 and 2018 in anthropology, economics, political science, and sociology. This was the "raw data" from which we built our theoretically constrained corpus.

The corpus had to perform two jobs: at one level, it had to provide an account of career structures in British academia that also represented, in the most exact way possible, the disciplinary affiliations of scholars. At another level, the corpus had to provide an account of research contributions over time on which we could observe the effects of evaluations on knowledge. Our approach to corpus construction understood the "raw data" as a collection of objects that required manipulation and contextualization before they were relevant for analysis (Ragin and Becker 1992). Specifically, we used this raw data in two ways. First, we used all entries, which included bibliographic records of research articles, book reviews, commentaries, and other items, to extract and identify pairs of authors and institutional affiliations. Considering that each record had an attributed disciplinary affiliation, we then used this information to disambiguate the pairs of author-institution: this allowed distinguishing an A Smith working at the University of Edinburgh in economics from an A Smith working in the same institution at a different period but in anthropology. This allowed identifying distinct authors in the data. With this information, and observing the reported affiliations over time, we could determine if they experienced mobility across evaluation periods by comparing their affiliations before and after a research assessment.

This first approach used all the bibliographic records, which is inadequate for capturing changes in knowledge since they include a variety of non-peer-reviewed, non-research items. Having identified and disambiguated authors, however, a second manipulation of the raw data consisted in restricting the corpus to research articles and using information on authorship from the previous step to identify productivity and citations across periods (see figure 2). Our resulting corpus was thus not simply the

product of data cleaning and disambiguation, but was an artefact that was “indexed” (Deterding and Waters 2021) with additional contents and segmented in specific ways that reflected our theoretical concerns (Boellstorff 2013).

FIGURE 2. Corpus construction. The corpus for this study was not a collection of raw data obtained from a digital repository but, rather, an altered dataset designed to represent our theoretical concerns.

This corpus construction process had thus both an epistemic and methodological dimension, as Campagnolo (2020) suggests: the actual objects we were studying were not scientific careers but their approximations through the records of SSCI, making data features and constraints directly relevant to the types of claims we could answer and take into the field. Careers themselves were inaccessible; the only data we had access to were affiliations as reported in publications—themselves imperfect reflections of “scientific work”. Dealing with data involved, in this respect, constant sense-making in our research team, triggering discussions about how records might have been constituted, what they signified, how they could be best processed, and how these decision points were tied to theoretical considerations and hence the boundaries of the case.

4.2 Scaling

The corpus we obtained resulted in a dataset covering 16,531 individual academics over four disciplinary fields (2,208 anthropologists, 6,384 economists, 4,271 political scientists, and 3,668 sociologists). Our disambiguation workflow allowed identifying the full first names for around 80% of the academics in our dataset, permitting us to infer their gender using name frequencies from the 1991 UK Census. We tested the accuracy of our gender attribution strategy (92%) by manually checking a random sample of 200 academics.

The dataset we obtained allowed cross-referencing academics with the standardized scores that their departments obtained in the multiple research assessments as well as data on institutional prestige. With this, we were able to study some predictors of mobility within our dataset, testing for the role of productivity (measured as citation counts and average impact factors), gender, institutional status, performance in the evaluations, and self-reported external funding (as reported in the Web of Science data). Predictions from models using our dataset came close to values reported in the literature. For example, our data tracked comparatively well the gendered productivity gap in academia, as well as disciplinary variations in citations, publications per period, and overall mobility. This provided us with some confidence on the underlying dataset and its ability to represent a large section of the British social sciences—that is, of its capacity to extend claims beyond the dataset.

The issue of change in knowledge required additional information. Computational text analysis offered a means for approximating how knowledge changed by tracking the contributions of scholars over time both individually and in relation to their respective epistemic fields and organizational units. By looking at how texts changed over time, we should be able to say whether the types of questions, concerns, and vocabularies of disciplines were altered, connecting these to research evaluations' role in academic labor markets.

We identified abstracts rather than citation networks as means for tracing shifts in knowledge over time. While various works have looked at citation networks to identify changes in the organization of scholarly fields (Godin 2006), we were interested in pursuing an approach that would be more attentive to how scholars used words in their own intellectual products rather than referencing the work of others. Citation data was used in our model mostly as a proxy of productivity. In particular, abstracts were classified using Latent Dirichlet Allocation topic models to identify patterns of change within the corpus over time.

Topic models have gained increasing traction within sociology since their introduction to the study of cultural formations almost a decade ago (Bohr and Dunlap 2018; DiMaggio, Nag, and Blei 2013; Mohr and Bogdanov 2013; Roose, Roose, and Daenekindt 2018). Since then, they have grown to become methods *du jour* (Bail 2015), and have expanded in reach and complexity in other domains of computational social science (Roberts et al. 2013, 2014). It would not be an overstatement to say that topic models are one of the most successful methodological innovations associated to the rise of computational social science, given their ability to classify and code large collections of text relatively seamlessly—they offer, in a way, the forms of scalability that are desired, and promised, by computational research designs. Given their relatively low costs and barriers of entry, topic models are now routinely used to study constructs varying from meaning (Goldenstein and Poschmann 2019) and the dynamics of social movements (Lindstedt 2019) to organizational frames (Pardo-Guerra 2020) and research paradigms (Lockhart 2020).

In the context of our project, the Latent Dirichlet Allocation developed by Blei and coauthors (2003) was used to classifying the abstracts in our corpus (DiMaggio et al. 2013). As discussed by others, the challenges of topic modeling involve balancing statistical and semantic fit: model selection often involves inspecting model outputs and comparing these to the data via domain expertise (Marshall 2013; Mützel 2015; Nelson et al. 2018). In our case, this led to identify a 40-topic solution as adequate for capturing the disciplinary diversity of each field. This was partly a theory-driven decision insofar as it reflected a

representation of social scientific fields that would allow for comparisons across time: it was sufficiently granular to distinguish shifts in topical concerns in the corpus that bore some relation to disciplinary subfields. In economics, for example, a 40-topic solution allowed distinguishing between abstracts focused on formal statistical discussions (“test, estim, distribut, asymptot”) and practical discussions on time-series modelling (“test, rate, run, cointegr”). Solutions with a smaller number of topics lost this variation, while those with larger numbers produced less intelligible results.

The purpose of topic models was not primarily to classify the corpus but to provide a means for reducing the dimensionality of the data. This evokes the extended case method’s logic of spanning across scales: while computational techniques can be used to expand our abilities to classify digital objects, they can also be used to produce metrics that identify features of a process and that would be difficult to observe locally ((Legewie and Schaeffer 2016; Leung 2014; Stoltz and Taylor 2019; Taylor and Stoltz 2020). Topic models allowed producing two key metrics. The first consisted of a measure of *similarity* between a scholar and her peers, which consisted of calculating the topical distributions for all the publications produced over a period by organizational unit b , excluding those produced by scholar a , and comparing these to the topic distributions for scholar a over the same period to estimate their cosine similarities. This first measure provides a sense of whether a scholar was unique in her department or was topically similar to what others were publishing.

The second metric involved using the topical distance between organizational units a and b to construct a graph representing all departments in each field. Within this graph, large distances represented thematic divergences while small ones represented convergence. To capture structural position in the field (which we associated to a measure of departmental *typicality*), we estimated the betweenness centrality for each department in this topical network using Python’s NetworkX library (Hagberg, Swart, and S Chult 2008; Leydesdorff 2007; White and Borgatti 1994).

Calculated over each period, these two measures of disciplinary organization (similarity and typicality) were used in our full career model to test for the effects of research assessments on the dynamics of academic careers. The model evaluated the probability of changing jobs between evaluation periods in relation to both established demographic variables, individual productivity, and institutional prestige and these two metrics relative to the organization of disciplinary fields. To control for idiosyncratic variations, the random effects models we tested included fixed effects for periods and disciplines to account for variability across time and fields. Both the metrics and the models are presented in the appendix. What matters for the purpose of this article is that the point estimates associated to similarity and typicality were associated with higher rates of mobility. In other words, our initial computational findings showed that scholars that are more similar to their colleagues and those working in less typical institutions are more likely to change jobs between evaluation periods. Further analysis suggested that the moves tend to occur in a specific direction: from less typical to more typical settings, and from situations of more similarity to those of less.

The statistical models allowed constructing specific empirical statements about the social sciences in Britain that were then used in an ethnographic context. The first statement was on the possible effects of the evaluations on more similar scholars and those working in less typical institutions (represented in Figure 2). The second statement derived from the trends of topical similarity across institutions which, as shown in Figure 3, concerns a decreasing diversity in the organizational forms taken by UK social science departments. Specifically, calculating the average topical distance between institutions across periods shows a decreasing trend that is suggestive of departments becoming “more similar” over time.

FIGURE 3: Associations between similarity and typicality at means for economics. The data shows a process of convergence where higher rates of mobility occur for those in specific structural conditions that are penalized by the research evaluations.

In a more traditional research design, generalizable claims could have resulted from these two statements. If the dataset is an adequate representation of the British social sciences, our study could have stopped at this point. This would have forestalled any

additional interpretative analysis of what these patterns meant for British academics. For this, an ethnographic approach was required, which would put to the test the statements produced from the quantitative models to assess their relevance to informants. The findings of our model were incorporated to an oral history protocol of British academics that sought to make sense of the structural forces at play in shaping careers through standardized research evaluations. This is a crucial methodological difference between grounded computational theory, computational ethnography, and the extended computational case method: the logic of using these computational techniques wasn't deductive (or inferential) but *demonstrative*: while we integrated textual and quantitative data into a career mobility model, the objective wasn't only estimating the effect of textual variables on rates of job changes but to use these to visualize patterns in the field in a way that could be discussed with our ethnographic informants.

FIGURE 4. Average topical distance between disciplines, 1986-2014

Oral histories allow making sense of the career patterns observed in our longitudinal dataset, scaling our findings into larger theoretical claims. Oral histories offer, in particular, a unique opportunity for academics to make sense of their organizational practices through a guided, situated conversation (Laslett 1999; Sandino 2006). The personal narratives we collected in two rounds of face-to-face conversations focused on the career experiences of social scientists and their research practices. Interviewees were selected through a convenience sample informed by the results of our quantitative model, capturing scholars at institutions that were both highly typical and highly atypical in their respective disciplines across recent evaluation periods. The conversations often started with the same prompts: How did you come into academia? How did you become aware of research evaluations? How have these shaped your research strategies? Results from the computational models were shared with respondents as a way of eliciting conversations about potential experiences of convergence in their careers.

Extending the computational results into the field and making them part of the conversational dynamics of oral history interviews permitted evaluating the extent to which existing theoretical accounts exaggerated aspects of knowledge production yet failed to identify other relevant features. For example, a longstanding assumption in the literature on the quantification of work is that evaluations are external interventions on scientific practices operating as bureaucratic constraints over which researchers have little agency (Feldman and Sandoval 2018; Hammarfelt, Rijcke, and Rushforth 2016; Lizotte 2021; Rijcke et al. 2016; Sayer 2014). Traditional labor mobility models agree with such finding, stressing work environment and fit as motivators of shifts rather than more overt pecuniary considerations (Long 1978).

Prompted by the computational findings, our conversations with scholars problematized this naïve model, highlighting points of agency and intervention where academics had opportunities to shape their research priorities and publication strategies. Most of our respondents, for example, did not see evaluations as impinging on the topics of their work—which contrasts with the findings of the models that suggest some topical convergence. They did, however, note that the evaluations created incentives for

submitting work for high-visibility journals, which required changing how they framed their research. Peter, a political scientist, noted

“It’s obvious to me that if I want to keep my job I have to do so” “It wouldn’t change the content of what I do but it does change in some ways the *framing* of what I do. Not always for the worse, I would say—even though, in general, I don’t agree with the exercise. You know, I probably should make more effort to talk to people that don’t agree with my premises. And if [...] part of that means going to a more generalist disciplinary [journal], like trying to get into [American Political Science Review] or something like that [...] if it fails, then at least it makes you think big. It’s not always that bad.”

In making sense of increasing homogeneity in the field (one of our two computational statements), respondents noted organizational rather than external incentives in publishing that were tied to the operationalization of research evaluations by local managers. Mark, a sociologist from London, noted that while the evaluations may make some forms of scholarship less attractive to administrators, what really mattered in the daily lives of academics was how they created forms of organizational surveillance that created internal mechanisms of punishment and reward. Every year, he noted,

a committee [of administrators] evaluates the quality of your publications. I have a cynical attitude towards them. Many times, the folks in the committees have no idea of what they are reading or doing. They also do it very fast. Five minutes per paper, looking at where it was published. We have an appeal procedure that allows us to check if the paper was properly judged. And we know it matters because it is a tool for increments and promotions.

These micro-historical accounts didn’t prove or disprove the findings from the computational models but complemented their claims, showing how evaluations come to matter practically in academic life. They provided, furthermore, important negative cases that allowed putting to the test naïve models of academic marketization and added weight to our interpretations of quantitative patterns. For example, rather than describing evaluations as external impositions (a trope that is central to the existing literature), some scholars underlined the importance of participating in these intrusive evaluations: for some informants, for example, such exercises are necessary sources of legitimacy that

sustain the UK's higher education sector as a public service. For others, the evaluations were even presented as "important exercises", as Carl, a London-based economist argued, because they allowed signaling the worth of their institutions to funders and administrators, despite not having the prestige of old, high-status universities like Oxford and Cambridge. Such organizational dynamics are simply invisible within the original corpus and the existing theoretical accounts (hence their qualities as negative cases), requiring the iterative scaling of findings between the computational and ethnographic fields.

4.3 Scaffolding

Establishing the robustness of ethnographic claims requires moving back and forth between the local scale and structural forces. What does this look like in the extended computational case method? One example of this is provided by research designs that, considering local ethnographic observations, extend these through additional evidentiary scaffolds.

While providing surprising insights about organizational mechanisms that linked research evaluations to the epistemic strategies and practices of scholars, our oral histories also suggested supplementary analyses of the original corpus that could further qualify our claims. Some informants, for example, raised questions about what was being highlighted by the computational patterns: were these shifts topical, or did they represent changes in meaning? Changes in topics were also made sense of by informants through structural transformations in the governance of universities and discipline-wide realignments around research trends, both of which are not immediately related to research evaluations. This called for a more detailed analysis of *semantic* changes in the corpus, rather than just of topical shifts in the four social science disciplines. Additional evidence of convergence in the use of key terms—which would represent similar shifts in the structure of knowledge—could thus bolster the evidence presented by the computational models and elaborated through the oral histories.

To assess changes in the semantic context of the terms used by social scientists in the abstracts of their articles, we turned to word embeddings. Word embeddings are a growing technique of text analysis that allows reducing a large universe of concepts into low-dimensional dense vectors (Kozłowski, Taddy, and Evans 2019; Taylor and Stoltz 2020). In this transformation, word embeddings preserve semantic similarity as spatial proximity, mirroring the distributional hypothesis of language (Kozłowski et al. 2019) and allowing for studying how linguistic units evolve over time in their context of use (Bizzoni et al. 2019; Hamilton, Leskovec, and Jurafsky 2016; Szymanski 2017).

Word embeddings allowed identifying important contextual changes for concepts in the different fields. First, using the word2vec implementation and segmenting our corpus by periods of evaluation, we estimated both changes in the immediate semantic universe of key terms identified by the first iteration of topic models as well as subsets of terms that had experienced the greatest shifts with respect to others in the corpus. This last operation involved calculating a weighted Kendall tau for the cosine similarities between every term across consecutive periods to find those that had the highest level of change. Through this technique, we could ‘navigate’ the corpus as a process of linguistic shift and confirm our initial hypothesis about changes in the structure of meaning of disciplinary language.

An extension of word embeddings (doc2vec) also provided additional evidence of disciplinary changes. Built on top of a word embedding model, doc2vec transforms texts into comparable, dense vectors. In this transformation, proximity equates semantic similarity. Unlike word embedding models, document embedding models can discriminate between different texts even when they are mostly identifiable by ambiguous words. The word “culture”, for example, is important for several sociological subfields, yet it performs different roles in each: the ‘culture’ of an economic or political sociologist does not fully match the ‘culture’ of a cultural sociologist (Liu et al. 2018). Document embedding models allow identifying these different contexts of use by not only taking into account a representation of language but also of sub-units of texts like sentences and paragraphs (Haj-Yahia, Sieg, and Deleris 2019; Navigli and Martelli 2019)—hence providing a semantic-aware means for evaluating textual similarity.

Calculated using the texts produced by institutions over each period, the doc2vec models allowed estimating the similarity between the work of scholars across different departments and evaluating their evolution over time. Consistent with the findings from the topic models, and despite a growing number of academics in each field which suggests growing diversity, we identified a secular increase in the semantic similarity of papers between institutions, providing further evidence of epistemic convergence in disciplines (Figure 4). This extension of questions generated by informants in the ethnographic field back into computational analyses constituted an additional scaffold for our claims that provided further confidence in our overall findings.

FIGURE 4. Semantic similarity of all texts in economics, political science, and sociology, 1996-2014.

The result of these interventions was a refinement of existing structural theories of academic labor. Expanding on a literature that presents academic work as captured by neoliberal institutions and marketized forms of management, our research strategy was capable of discerning mechanisms used by scholars to make sense of their organizational environments. Rather than emphasizing *only* the computational results, our findings stressed the importance of vocational logics identified by scholars in interpreting and dealing with the evaluation of their work (Pardo-Guerra 2022). Standardized forms of work quantification were read not as external impositions as much as ‘games’ that had to be played as part of the professional academic field. Distinct organizational cultures mediate, with different intensities, foci of attention, and logics of practice, how evaluations are interpreted and made sensible and legitimate. For some institutions, they create environments where the cognitive and economic costs of searching and pursuing other positions is acceptable; for others, evaluations simply do not matter because of how administrators and scholars are buffered from the pecuniary logics the exercise. The results of this study are suggestive: rather than acting as singular pressure points, the effect of research evaluations depends on the organizational ecology of the field and the variability of circumstances, resources and cultures across institutions, features that can only be examined through an extended case method, attentive to structural forces and trends.

Discussion

In considering how to integrate computational methods into sociological research, scholars are presented with various options. Most commonly, the emergence of new means for analyzing and processing digital data have involved an augmentation of statistical modeling and pattern recognition. In this paper we advocate a different approach which, inspired by Burawoy’s extended case method, creates opportunities for using computational findings within a quasi-ethnographic design to tackle structural questions directly motivated by theoretical concerns.

Our proposal differs from other computationally augmented research designs in important ways. Notably, the extended computational case method diverges from approaches like computational grounded theory in three important respects:

First, while sharing CGT's interest in the affordances of computer assisted text analysis methods, ECCM is less concerned about the augmentation of the researcher's throughput (that is, the capacity to scale up analysis by considering larger corpuses) than with the ability to discern puzzling findings using computational techniques that can be discussed in the field. In our study of career transitions in British academia, computational methods were relevant as means for generating both a large longitudinal dataset of labor mobility that is otherwise unavailable and identifying shifts in academic markets that are tied to specific textual measures. Because emphasis is theoretical, certain forms of work—such as corpus construction or data “cleaning”—acquire more prominent roles in the ECCM as ways of delimiting empirical instances through theoretical questions. “Counting” words without *ex ante* interpretations and data manipulations (Lee and Martin 2015) is simply impossible because data is always material, and corpora are always necessarily constructed (they result from particular decisions of inclusion and exclusion, data formatting standards, and inscribed institutional histories that limit the usability of any formal maps we may produce from them) (Piper 2020; Tufekci 2014). The emphasis, then, is less on finding an ideal corpus that, like a model animal (Radford and Joseph 2020), serves as the basis for generalizing research and extending claims and more on *creating* a corpus that attends to particular theoretical questions and that can be explored in the field.

Second, ECCM joins other calls for integrating digital analysis into ethnographic designs (Geiger and Halfaker 2017), and computational mixed methods more generally (Abramson et al. 2018; Small 2011). Rather than circumscribing the analysis to the corpus and extending claims uniquely from it, ECCM calls for scaling narratives between and across ethnographic and computational domains. Computational methods certainly afford observing patterns that would be otherwise invisible (think, for instance, of the visualizations produced through topic modeling or word embeddings and which show changes in the way words are used over time), yet in our methodological approach, these

changes only acquire theoretical significance when tested against the observations of ethnographic informants. This form of scaling allows testing the verisimilitude of patterns observed computationally in the context of ethnographic conversations. In this, the ECCM breaks away from the CGT in extending claims by taking them outside of the corpus and the coding strategy.

Third, in establishing the robustness and validity of its claims, CGT puts to the test its classifications on the corpus through a process of testing and verification, as does computational ethnography. By contrast, and consistent with its logic of scaling, the ECCM extends claims by establishing scaffolds across domains that serve as additional points for elaborating theoretical claims. This involves a logic of reflexive corroboration rather than testing: ethnographic observations can inspire, for example, the adoption of further forms of data analysis that are meaningful for assessing the stability of certain empirical claims and concatenating these into general, structural statements. This occurs iteratively and across forms of evidence rather than within the corpus, providing additional evidentiary basis to findings and observations.

The strategy that we propose here serves as a general framework for integrating computational methods and interpretative, ethnographic designs. Importantly, the extended computational case study allows tackling structural theoretical arguments without committing to either the primacy of the raw data (as is assumed in grounded theory) or the generalizability of statistical claims (as occurs in traditional verificationist frameworks). This issue of integrating computational methods into ethnographic designs is of increasing relevance, given the expansion of computational methods in our profession. An indicator of the need for such frameworks is given by how qualitative techniques are now being used in the aid of verificationist frameworks. Consider the well-known Fragile Families Challenge (Salganik et al. 2019), which implicitly assumes that complex social phenomena can be rendered legible and predictable with sufficient data and computational models: recent iterations of this project seek to integrate qualitative interviews into the research design, not as a mechanism for building up claims or incorporating ethnographic depth, but as a means for finding the “dark matter” (that is, novel unmeasured factors)

otherwise invisible in the original corpus. Our proposal breaks from this by echoing computational grounded theory's attentiveness to close reading and human-centered interpretation while offering means for assessing general theoretical claims. By bringing the strategies of the extended case method to bear on mixed computational studies, we hope to complement and expand the methodological repertoire of contemporary scholars.

Author's note

A modified, anonymized version of the dataset used to produce figures 3 and 4 is publicly available at <https://osf.io/4jg8b/> along with the associated code for the random effects panel through which mobility probabilities were predicted. The data excludes variables that would allow for a rapid identification of individual UK academics, leading to slightly modified results. Specifically, the data does not exclude individuals who moved into non-academic positions. This added level of anonymity is consistent with expectations of anonymity discussed with informants and that are standard practice in data-handling in the United Kingdom.

References

- Abramo, Giovanni, Ciriaco Andrea D'Angelo, and Alessandro Caprasecca. 2009. "Allocative Efficiency in Public Research Funding: Can Bibliometrics Help?" *Research Policy* 38(1):206–15.
- Abramson, Corey M., Jacqueline Joslyn, Katharine A. Rendle, Sarah B. Garrett, and Daniel Dohan. 2018. "The Promises of Computational Ethnography: Improving Transparency, Replicability, and Validity for Realist Approaches to Ethnographic Analysis." *Ethnography* 19(2):254–84.
- Azoulay, Pierre, Joshua Graff-Zivin, Brian Uzzi, Dashun Wang, Heidi Williams, James A. Evans, Ginger Zhe Jin, Susan Feng Lu, Benjamin F. Jones, and Katy Börner. 2018. "Toward a More Scientific Science." *Science* 361(6408):1194–97.
- Baden, Christian, Christian Pipal, Martijn Schoonvelde, and Mariken AC G. van der Velden. 2021. "Three Gaps in Computational Text Analysis Methods for Social Sciences: A Research Agenda." *Communication Methods and Measures* 1–18.
- Bail, Christopher A. 2015. "Lost in a Random Forest: Using Big Data to Study Rare Events." *Big Data & Society* 2(2):2053951715604333.

- Barnes, Barry. 2013. *Scientific Knowledge and Sociological Theory*. Routledge.
- Ben-David, Joseph, and Randall Collins. 1966. "Social Factors in the Origins of a New Science: The Case of Psychology." *American Sociological Review* 451–65.
- Ben-David, Joseph, and Teresa A. Sullivan. 1975. "Sociology of Science." *Annual Review of Sociology* 1(1):203–22.
- Berman, Elizabeth Popp, and Daniel Hirschman. 2018. *The Sociology of Quantification: Where Are We Now?* SAGE Publications Sage CA: Los Angeles, CA.
- Bizzoni, Yuri, Stefania Degaetano-Ortlieb, Katrin Menzel, Pauline Krielke, and Elke Teich. 2019. "Grammar and Meaning: Analysing the Topology of Diachronic Word Embeddings." Pp. 175–85 in *Proceedings of the 1st International Workshop on Computational Approaches to Historical Language Change*.
- Bjerre-Nielsen, Andreas, and Kristoffer Lind Glavind. 2022. "Ethnographic Data in the Age of Big Data: How to Compare and Combine." *Big Data & Society* 9(1):20539517211069892.
- Blei, David M., Andrew Y. Ng, and Michael I. Jordan. 2003. "Latent Dirichlet Allocation." *Journal of Machine Learning Research* 3(Jan):993–1022.
- Boellstorff, Tom. 2013. "Making Big Data, in Theory." *First Monday* 18(10).
- Bohr, Jeremiah, and Riley E. Dunlap. 2018. "Key Topics in Environmental Sociology, 1990–2014: Results from a Computational Text Analysis." *Environmental Sociology* 4(2):181–95.
- Burawoy, Michael. 1998a. "The Extended Case Method." *Sociological Theory* 16(1):4–33.
- Burawoy, Michael. 2009. *The Extended Case Method*. University of California Press.
- Burrows, Roger, and Mike Savage. 2014. "After the Crisis? Big Data and the Methodological Challenges of Empirical Sociology." *Big Data & Society* 1(1):2053951714540280.
- Campagnolo, Gian Marco. 2020. *Social Data Science Xennials: Between Analogue and Digital Social Research*. Springer Nature.
- Cetina, Karin Knorr. 2009. *Epistemic Cultures: How the Sciences Make Knowledge*. Harvard University Press.
- Decoteau, Claire Laurier. 2017. "The AART of Ethnography: A Critical Realist Explanatory Research Model." *Journal for the Theory of Social Behaviour* 47(1):58–82.
- Deterding, Nicole M., and Mary C. Waters. 2021. "Flexible Coding of In-Depth Interviews: A Twenty-First-Century Approach." *Sociological Methods & Research* 50(2):708–39.

- 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.
- Edelmann, Achim, Tom Wolff, Danielle Montagne, and Christopher A. Bail. 2020. "Computational Social Science and Sociology." *Annual Review of Sociology* 46:61–81.
- Espeland, Wendy Nelson, and Mitchell L. Stevens. 2008. "A Sociology of Quantification." *European Journal of Sociology/Archives Européennes de Sociologie* 49(3):401–36.
- 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/>).
- Feldman, Zeena, and Marisol Sandoval. 2018. "Metric Power and the Academic Self: Neoliberalism, Knowledge and Resistance in the British University." *TripleC: Communication, Capitalism & Critique* 16(1):214–33.
- Flick, Uwe. 2004. "Triangulation in Qualitative Research." *A Companion to Qualitative Research* 3:178–83.
- Foster, Jacob G., Andrey Rzhetsky, and James A. Evans. 2015. "Tradition and Innovation in Scientists' Research Strategies." *American Sociological Review* 80(5):875–908.
- Garfinkel, Harold. 1963. "A Conception of and Experiments with 'Trust' as a Condition of Concerted Stable Actions." *The Production of Reality: Essays and Readings on Social Interaction* 381–92.
- Garfinkel, Harold. 1967. "Ethnomethodology." *Englewood Cliffs*.
- Geiger, R. Stuart, and Aaron Halfaker. 2017. "Operationalizing Conflict and Cooperation between Automated Software Agents in Wikipedia: A Replication and Expansion of 'even Good Bots Fight'." *Proceedings of the ACM on Human-Computer Interaction* 1(CSCW):1–33.
- Geuna, Aldo, and Ben R. Martin. 2003. "University Research Evaluation and Funding: An International Comparison." *Minerva* 41(4):277–304.
- Glaser, Barney G., and Anselm L. Strauss. 2017. *Discovery of Grounded Theory: Strategies for Qualitative Research*. Routledge.

- Gläser, Jochen, and Grit Laudel. 2016. "Governing Science: How Science Policy Shapes Research Content." *European Journal of Sociology/Archives Européennes de Sociologie* 57(1):117–68.
- Godin, Benoît. 2006. "On the Origins of Bibliometrics." *Scientometrics* 68(1):109–33.
- Goldenstein, Jan, and Philipp Poschmann. 2019. "Analyzing Meaning in Big Data: Performing a Map Analysis Using Grammatical Parsing and Topic Modeling." *Sociological Methodology* 49(1):83–131.
- Grigoropoulou, Nikolitsa, and Mario L. Small. 2022. "The Data Revolution in Social Science Needs Qualitative Research." *Nature Human Behaviour* 1–3.
- Hagberg, Aric, Pieter Swart, and Daniel S Chult. 2008. *Exploring Network Structure, Dynamics, and Function Using NetworkX*. Los Alamos National Lab.(LANL), Los Alamos, NM (United States).
- Haj-Yahia, Zied, Adrien Sieg, and Léa A. Deleris. 2019. "Towards Unsupervised Text Classification Leveraging Experts and Word Embeddings." Pp. 371–79 in *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*.
- Hamilton, William L., Jure Leskovec, and Dan Jurafsky. 2016. "Diachronic Word Embeddings Reveal Statistical Laws of Semantic Change." *ArXiv Preprint ArXiv:1605.09096*.
- Hammarfelt, B. M. S., S. de Rijcke, and Alexander D. Rushforth. 2016. "Quantified Academic Selves: The Gamification of Science through Social Networking Services." *Information Research* 21(2):eSM1.
- 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.
- Hesse, Mary. 2020. *The Structure of Scientific Inference*. University of California Press.
- Hicks, Diana. 2012. "Performance-Based University Research Funding Systems." *Research Policy* 41(2):251–61.
- Jick, Todd D. 1979. "Mixing Qualitative and Quantitative Methods: Triangulation in Action." *Administrative Science Quarterly* 24(4):602–11.
- Karell, Daniel, and Michael Freedman. 2019. "Rhetorics of Radicalism." *American Sociological Review* 84(4):726–53.
- Knigge, Antonie, Ineke Maas, Marco HD Van Leeuwen, and Kees Mandemakers. 2014. "Status Attainment of Siblings during Modernization." *American Sociological Review* 79(3):549–74.

- Kozlowski, Austin C., Matt Taddy, and James A. Evans. 2019. "The Geometry of Culture: Analyzing the Meanings of Class through Word Embeddings." *American Sociological Review* 84(5):905–49.
- Kuhn, Thomas S. 1957. *The Copernican Revolution: Planetary Astronomy in the Development of Western Thought*. Vol. 16. Harvard University Press.
- Laslett, Barbara. 1999. "Personal Narratives as Sociology." *Contemporary Sociology* 28(4):391–401.
- Latour, Bruno. 1993. *The Pasteurization of France*. Harvard University Press.
- Lazer, David, Alex Pentland, Lada Adamic, Sinan Aral, Albert-Laszlo Barabasi, Devon Brewer, Nicholas Christakis, Noshir Contractor, James Fowler, and Myron Gutmann. 2009. "Social Science. Computational Social Science." *Science (New York, NY)* 323(5915):721–23.
- Leahey, Erin. 2007. "Not by Productivity Alone: How Visibility and Specialization Contribute to Academic Earnings." *American Sociological Review* 72(4):533–61.
- Lee, Monica, and John Levi Martin. 2015. "Coding, Counting and Cultural Cartography." *American Journal of Cultural Sociology* 3(1):1–33.
- Legewie, Joscha, and Merlin Schaeffer. 2016. "Contested Boundaries: Explaining Where Ethnoracial Diversity Provokes Neighborhood Conflict." *American Journal of Sociology* 122(1):125–61.
- Leung, Ming D. 2014. "Dilettante or Renaissance Person? How the Order of Job Experiences Affects Hiring in an External Labor Market." *American Sociological Review* 79(1):136–58.
- Lewis, Kevin, and Jason Kaufman. 2018. "The Conversion of Cultural Tastes into Social Network Ties." *American Journal of Sociology* 123(6):1684–1742.
- 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):1303–19.
- Li, Zhuofan, Daniel Dohan, and Corey M. Abramson. 2021. "Qualitative Coding in the Computational Era: A Hybrid Approach to Improve Reliability and Reduce Effort for Coding Ethnographic Interviews." *Socius* 7:23780231211062344.
- Lichterhan, Paul, and Isaac Ariail Reed. 2015. "Theory and Contrastive Explanation in Ethnography." *Sociological Methods & Research* 44(4):585–635.
- Lindstedt, Nathan C. 2019. "Structural Topic Modeling for Social Scientists: A Brief Case Study with Social Movement Studies Literature, 2005–2017." *Social Currents* 6(4):307–18.

- Liu, Qian, He-Yan Huang, Yang Gao, Xiaochi Wei, Yuxin Tian, and Luyang Liu. 2018. "Task-Oriented Word Embedding for Text Classification." Pp. 2023–32 in *Proceedings of the 27th international conference on computational linguistics*.
- Lizotte, Mathieu. 2021. "If You Do Not Deign to Quantify, Someone Else Will Do It for You: In Support of a Balanced Approach to the Evaluation of Science." *Social Science Information* 05390184211021364.
- Lockhart, Jeffrey W. 2020. "Paradigms of Sex Research and Women in STEM."
- 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* 703–22.
- Marshall, Emily A. 2013. "Defining Population Problems: Using Topic Models for Cross-National Comparison of Disciplinary Development." *Poetics* 41(6):701–24.
- McFarland, Daniel A., Kevin Lewis, and Amir Goldberg. 2016. "Sociology in the Era of Big Data: The Ascent of Forensic Social Science." *The American Sociologist* 47(1):12–35.
- McMahan, Peter, and James Evans. 2018. "Ambiguity and Engagement." *American Journal of Sociology* 124(3):860–912.
- Mennicken, Andrea, and Wendy Nelson Espeland. 2019. "What's New with Numbers? Sociological Approaches to the Study of Quantification." *Annual Review of Sociology* 45.
- Mische, Ann. 2009. *Partisan Publics*. Princeton University Press.
- Mohr, John W., Christopher A. Bail, Margaret Frye, Jennifer C. Lena, Omar Lizardo, Terence E. McDonnell, Ann Mische, Iddo Tavory, and Frederick F. Wherry. 2020. *Measuring Culture*. Columbia University Press.
- Mohr, John W., and Petko Bogdanov. 2013. *Introduction—Topic Models: What They Are and Why They Matter*. Elsevier.
- Molina, Mario, and Filiz Garip. 2019. "Machine Learning for Sociology." *Annual Review of Sociology* 45:27–45.
- Mützel, Sophie. 2015. "Facing Big Data: Making Sociology Relevant." *Big Data & Society* 2(2):2053951715599179.
- Navigli, Roberto, and Federico Martelli. 2019. "An Overview of Word and Sense Similarity." *Natural Language Engineering* 25(6):693–714.

- Nelson, Laura K. 2020a. "Computational Grounded Theory: A Methodological Framework." *Sociological Methods & Research* 49(1):3–42.
- Nelson, Laura K. 2020b. "Computational Grounded Theory: A Methodological Framework." *Sociological Methods & Research* 49(1):3–42.
- Nelson, Laura K. 2021a. "Cycles of Conflict, a Century of Continuity: The Impact of Persistent Place-Based Political Logics on Social Movement Strategy." *American Journal of Sociology* 127(1):1–59.
- Nelson, Laura K. 2021b. "Leveraging the Alignment between Machine Learning and Intersectionality: Using Word Embeddings to Measure Intersectional Experiences of the Nineteenth Century US South." *Poetics* 101539.
- Nelson, Laura K., Derek Burk, Marcel Knudsen, and Leslie McCall. 2018. "The Future of Coding: A Comparison of Hand-Coding and Three Types of Computer-Assisted Text Analysis Methods." *Sociological Methods & Research* 0049124118769114.
- Nelson, Nicole C. 2013. "Modeling Mouse, Human, and Discipline: Epistemic Scaffolds in Animal Behavior Genetics." *Social Studies of Science* 43(1):3–29.
- Pacewicz, Josh. 2020. "What Can You Do with a Single Case? How to Think about Ethnographic Case Selection like a Historical Sociologist." *Sociological Methods & Research* 0049124119901213.
- Pardo-Guerra, Juan Pablo. 2020. "Where Are the Market Devices? Exploring the Links among Regulation, Markets, and Technology at the Securities and Exchange Commission, 1935–2010." *Theory and Society* 1–32.
- Petersen, Alexander Michael, Santo Fortunato, Raj K. Pan, Kimmo Kaski, Orion Penner, Armando Rungi, Massimo Riccaboni, H. Eugene Stanley, and Fabio Pammolli. 2014. "Reputation and Impact in Academic Careers." *Proceedings of the National Academy of Sciences* 111(43):15316–21.
- Piper, Andrew. 2020. *Can We Be Wrong? The Problem of Textual Evidence in a Time of Data*. Cambridge University Press.
- Radford, Jason, and Kenneth Joseph. 2020. "Theory In, Theory Out: The Uses of Social Theory in Machine Learning for Social Science." *ArXiv Preprint ArXiv:2001.03203*.
- Ragin, Charles C., and Howard Saul Becker. 1992. *What Is a Case?: Exploring the Foundations of Social Inquiry*. Cambridge university press.
- 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):1954–58.

- 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):161–69.
- Roberts, Margaret E., Brandon M. Stewart, Dustin Tingley, and Edoardo M. Airoidi. 2013. "The Structural Topic Model and Applied Social Science." Pp. 1–20 in *Advances in neural information processing systems workshop on topic models: computation, application, and evaluation*. Vol. 4. Harrahs and Harveys, Lake Tahoe.
- Roberts, Margaret E., Brandon M. Stewart, Dustin Tingley, Christopher Lucas, Jetson Leder-Luis, Shana Kushner Gadarian, Bethany Albertson, and David G. Rand. 2014. "Structural Topic Models for Open-Ended Survey Responses." *American Journal of Political Science* 58(4):1064–82.
- Roose, Henk, Willem Roose, and Stijn Daenekindt. 2018. "Trends in Contemporary Art Discourse: Using Topic Models to Analyze 25 Years of Professional Art Criticism." *Cultural Sociology* 12(3):303–24.
- Rosenfeld, Rachel A. 1992. "Job Mobility and Career Processes." *Annual Review of Sociology* 18(1):39–61.
- Salganik, Matthew J., Ian Lundberg, Alexander T. Kindel, and Sara McLanahan. 2019. "Introduction to the Special Collection on the Fragile Families Challenge." *Socius* 5:2378023119871580.
- Sandino, Linda. 2006. *Oral Histories and Design: Objects and Subjects*. Oxford University Press.
- Sayer, Derek. 2014. *Rank Hypocrisies: The Insult of the REF*. Sage.
- Schritt, Jannik. 2019. "An Ethnography of Public Events: Reformulating the Extended Case Method in Contemporary Social Theory." *Ethnography* 1466138119891446.
- Seim, Josh. 2021. "Participant Observation, Observant Participation, and Hybrid Ethnography." *Sociological Methods & Research* 0049124120986209.
- Small, Mario Luis. 2009. "How Many Cases Do I Need?' On Science and the Logic of Case Selection in Field-Based Research." *Ethnography* 10(1):5–38.
- Small, Mario Luis. 2011. "How to Conduct a Mixed Methods Study: Recent Trends in a Rapidly Growing Literature." *Annual Review of Sociology* 37:57–86.
- Stoltz, Dustin S., and Marshall A. Taylor. 2019. "Textual Spanning: Finding Discursive Holes in Text Networks." *Socius* 5:2378023119827674.

- Szymanski, Terrence. 2017. "Temporal Word Analogies: Identifying Lexical Replacement with Diachronic Word Embeddings." Pp. 448–53 in *Proceedings of the 55th annual meeting of the association for computational linguistics (volume 2: short papers)*.
- Tavory, Iddo, and Stefan Timmermans. 2014. *Abductive Analysis: Theorizing Qualitative Research*. University of Chicago Press.
- Taylor, Marshall A., and Dustin S. Stoltz. 2020. "Integrating Semantic Directions with Concept Mover's Distance to Measure Binary Concept Engagement." *Journal of Computational Social Science* 1–12.
- Teplitskiy, Misha, Eamon Duede, Michael Menietti, and Karim R. Lakhani. 2020. "Status Drives How We Cite: Evidence from Thousands of Authors." *ArXiv E-Prints* arXiv-2002.
- Timmermans, Stefan, and Iddo Tavory. 2012. "Theory Construction in Qualitative Research: From Grounded Theory to Abductive Analysis." *Sociological Theory* 30(3):167–86.
- Tufekci, Zeynep. 2014. "Big Questions for Social Media Big Data: Representativeness, Validity and Other Methodological Pitfalls." in *Eighth international AAAI conference on weblogs and social media*.
- Van Maanen, John. 2011. *Tales of the Field: On Writing Ethnography*. University of Chicago Press.
- Van Velsen, Jaap. 1979. "The Extended-Case Method and Situational Analysis." Pp. 129–49 in *The craft of social anthropology*. Elsevier.
- White, Douglas R., and Stephen P. Borgatti. 1994. "Betweenness Centrality Measures for Directed Graphs." *Social Networks* 16(4):335–46.
- Wolfinger, Nicholas H. 2002. "On Writing Fieldnotes: Collection Strategies and Background Expectancies." *Qualitative Research* 2(1):85–93.