Political Methodology

Committee on Concepts and Methods

Working Paper Series

26

February 2010

Causal Explanation and Multi-Method Research in the Social Sciences

David Kuehn

University of Heidelberg (kuehn@uni-heidelberg.de)

&

Ingo Rohlfing

University of Cologne (rohlfing@wiso.uni-koeln.de)

C&M
The Committee on Concepts and Methods
www.concepts-methods.org

IPSA
International Political Science Association
www.ipsa.org

CIDE
Research and Teaching in the Social Sciences
www.cide.edu

The C&M working paper series are published by the Committee on Concepts and Methods (C&M), the Research Committee No. 1 of the International Political Science Association (IPSA), hosted at CIDE in Mexico City. C&M working papers are meant to share work in progress in a timely way before formal publication. Authors bear full responsibility for the content of their contributions.

The Committee on Concepts and Methods (C&M) promotes conceptual and methodological discussion in political science. It provides a forum of debate between methodological schools who otherwise tend to conduct their deliberations on separate tables. It publishes two series of working papers: "Political Concepts" and "Political Methodology."

Political Concepts contains work of excellence on political concepts and political language. It seeks to include innovative contributions to concept analysis, language usage, concept operationalization, and measurement.

Political Methodology contains work of excellence on methods and methodology in the study of politics. It invites innovative work on fundamental questions of research design, the construction and evaluation of empirical evidence, theory building and theory testing. The series welcomes, and hopes to foster, contributions that cut across conventional methodological divides, as between quantitative and qualitative methods, or between interpretative and observational approaches.

Submissions. All papers are subject to review by either a member of the Editorial Board or an external reviewer. Only English-language papers can be admitted. Authors interested in including their work in the C&M Series may seek initial endorsement by one editorial board member. Alternatively, they may send their paper to wps[at]concepts-methods.org.

The C&M webpage offers full access to past working papers.

www.concepts-methods.org

1. Introduction

After a long and often fiercely fought debate over the respective values of quantitative and qualitative methods in the social sciences, multi-method research (MMR) is emerging as a new methodological paradigm. While there are numerous variants of MMR (Capoccia and Freeden, 2006), the current debate is mostly centered on the combination of a cross-case method, usually regression analysis or, though decidedly less often, Qualitative Comparative Analysis (QCA), with qualitative case studies (e.g. Bennett, 2002; Lieberman, 2005; Rihoux, 2006; Rohlfing, 2008).[1] The combination of cross-case and within-case methods is supposed to compensate the inherent weaknesses of one method by the respective strength of the other method (e.g. Achen, 2005; Collier, Brady and Seawright, 2004; Coppedge, 1999; Levy, 2008; Wolf, 2010). The cross-case method is employed to detect regularities on the cross-case level, but suffers from the problem that a systematic cross-case pattern does not equate causation. This is the rationale for process tracing that is performed to discern the causal process in place between X and Y. Case studies in turn are very limited in producing valid cross-case inferences, which in turn renders it beneficial to supplement them with a cross-case technique. The integration of quantitative and qualitative methods that follows these lines is presented as a "third way" that is supposedly superior vis-à-vis the employment of one method alone (Brady, Collier and Seawright, 2006; McKeown, 1999; Tarrow, 1995).

[1] We appreciate that case-studies can be qualitative and quantitative and that the number of observations can be very high (Gerring, 2004). Usually, however, case studies are qualitative and rely on "causal process observations" (Collier, Brady and Seawright, 2004) and it is this type of case study we refer to in the following when speaking of case studies and small-n research.

The debate on MMR meets with the ongoing discussion about standards of good causal inference and explanation (Bennett, 2003; Gerring, 2005; Mahoney, 2008; Mooney Marini and Singer, 1988). One view usually associated with quantitative research and forcefully articulated by King et al. (1994) states that causal inference means establishing the *causal effect* of a given cause on the phenomenon of interest (Morgan and Winship 2007). The elevation of a causal effect, that is, a cross-case regularity, as the standard of causal inference has been criticized by qualitative scholars because of the purported black-boxing of the causal processes connecting cause to effect (Abbott, 1998; Collier et al., 2004; McKeown, 1999).[2] These critics argue that a satisfying scientific explanation should theorize and empirically discern an uninterrupted process that links the purported cause and the outcome (Demetriou, 2009; Elster, 1998; Hedström and Swedberg 1998; Salmon, 1998; Yee, 1996).

While the two views on causality and explanation were pitted against each other for some time, their compatibility has been increasingly emphasized recently (e.g. Morgan and Winship 2007; Steel 2004; Woodward 2003; Waldner 2007; King and Powell 2008). It is now widely acknowledged in the philosophy of science as well as in the social sciences that one should strive for *causal explanations* that include propositions of both causal effects and causal processes (Cartwright, 2004).[3] Since the former are empirically assessed on the cross-case level and the

---

[2] A more general critique of the quantitative worldview comes from scholars rooted in realist, constructivist, interpretivist and post-structuralist traditions who generally deny the possibility that causal regularities are in place and can be searched for (open system perspective) {Kurki, 2007 #3923}. We do not further address this criticism in our discussion of MMR, since the combination of a large-n, cross-case method as a tool of causal inference presumes a belief in regularities and the possibility of meaningful comparison between cases (closed system perspective).

[3] To be fair, this latter claim remains to some degree controversial, even among case study researchers. Some authors argue that causal process hypotheses must be subjected to empirical evaluation in order to provide a

latter on the within-case level, it is evident that there is a strong sense of complementarity between the formulation of causal explanations and their development and testing through MMR.

Somewhat surprisingly, the actual potential of MMR to promote the formulation of causal explanations has received very little attention so far. There is agreement that causal explanations should be formulated and tested empirically, but it is unclear what this general advice means in practice. We argue that the salient dimension to be discussed in light of the plea for causal explanations and MMR concerns the notion of causality and the distinction between *determinism* and *probabilism* in particular. Of course, deterministic and probabilistic causality have been addressed in the methods literature before (e.g. Adcock, 2002; Bennett, 1999; Goertz, 2005; Goldthorpe, 1997a, 1997b; Lieberson, 1991). These discussions, however, have mainly focused on case studies and their ability to cope with probabilism for cross-case inferences (e.g. Munck, 2005). What is missing is a systematic elaboration of what determinism and probabilism imply for causal explanations and for the cross-case *and* within-case level. For example, it is an open question of what we can learn from our knowledge about cross-case level regularities about within-case level relationships (and vice versa) in a probabilistic world.

A rigorous discussion of the relationship between causal explanation and deterministic and probabilistic causality should therefore be the next step in the debate on multi-method

firm basis for causal explanation (e.g., George and Bennett 2005). Others agree that making statements about causal processes is intrinsic to causal theorizing, but hold that an empirical evaluation is not necessary or rather difficult (Gerring 2009). Since our paper is mainly concerned with MMR, which explicitly aims at an empirical analysis of propositions referring to the cross-case and the within-case level, we do not have to dig deeper into this debate. Furthermore, our discussion does not need to take position on the related question of whether causal inference in process tracing follows the same logic as cross-case analyses or not (Gerring 2009). As will become clear later on, our arguments on MMR and process tracing are independent of the logic of generating within-case inferences.

research. This is what we aim to provide in the first section of our paper. We elaborate in a step-wise fashion what determinism, probabilism, and randomness imply for the cross-case and within-case level. This discussion is necessary in order to move on beyond the general consensus that causal explanations are desirable (Mahoney, 2008). Equally important, it provides the ground for the proper development and testing of causal explanations through MMR. If there is no explicit understanding of how to distinguish determinism from probabilism and the latter from randomness in cross-case and within-case analyses, there is no basis for making informed causal inferences.

We argue that the opportunity to make a meaningful connection of the cross-case method and the within-case part is severely reduced if the causal relationship is probabilistic. The cross-case method describes general patterns between variables in a larger set of cases. This, however, does not allow for drawing conclusions about what to expect in one case or a small number of cases sampled from the population. At the same time, the number of cases examined in process tracing usually is too small to extend insights developed from within-case evidence to all cases under analysis in the cross-case part. Furthermore, we show that well-known strategies to enhance the inferential leverage of case study research, like the choice of “crucial cases”, do not solve the problems that stem from probabilistic causality in MMR. Consequently, we conclude that the causal inferences generated through MMR are considerably less certain than the current state of the debate suggests.

Our paper is organized as follows. In the next section, we address the difference between deterministic and probabilistic causal explanations in regards of the cross-case and within-case level. Building on this conceptual framework, we evaluate the potential of MMR to empirically investigate these two types of causal explanations in section three. In the conclusion, we summarize our argument and provide some tentative advice on how to interpret these findings.

2. Determinism, Probabilism, Randomness, and Causal Explanations

The integration of cross-case and within-case theorizing in causal explanations is to be welcomed. The two levels of analysis are complementary and lead to richer social science theories and more complete knowledge of the social world than cross-case and within-case theorizing alone. The importance of fully specified explanations notwithstanding, there has only been little attention to the core concept which rests at the heart of such endeavors: the notion of causality itself. Particularly, the distinction between deterministic and probabilistic causal relationships has not received sufficient scrutiny. Of course, different notions of causality have been addressed in the methodological dispute between quantitative and qualitative researchers before (e.g. Adcock, 2002; Bennett, 1999; Goldthorpe, 1997a; Lieberson, 1991; Munck, 2005). Up till now, however, these topics have not been systematically elaborated in the literature on causal explanations or MMR.

This is a serious shortcoming for two related reasons. From a conceptual perspective, the meaning of deterministic and probabilistic causal *explanations* has thus far not been systematically defined. Existing methodological discussions on probabilism center on the question if small-n studies are suitable to make probabilistic cross-case inferences (Bennett, 2003; Goldstone, 1997; Lieberson, 1991). Although it is widely agreed that the unique selling point of case studies lies on the level of causal processes, the problem and meaning of probabilism on the within-case level has been ignored so far. This neglect means, second, that there are currently no criteria for evaluating the validity of causal inferences based on within-case level evidence and, to the extent that they are constitutive for them, causal explanations in general. We acknowledge that many philosophical treatments and the overwhelming majority of social scientists do not subscribe to a deterministic understanding of causality and assume probabilistic causality instead (Suppes 1970; Humphreys 1989; Reiss 2009; Bennett, 2003;

Bollen, Entwisle and Alderson, 1993; Gerring, 2005; King et al., 1994, chap. 3; Lieberson, 1991; Pearl, 2000; Salmon, 1998, chap. 2). Regardless of the individual understanding of causality, however, any discussion of these topics require a thorough understanding of determinism and probabilism in causal explanations. In order to address these issues and to provide criteria for assessing the nature of causal explanations, we define deterministic and probabilistic causal explanations on the cross-case and within-case level in the remainder of this section. In each section, we start with the cross-case level and consider the within-case level next.

3.1. Determinism

On a general level, a causal relationship is deterministic if there is an invariant link between cause and effect in a specific context (Adcock, 2002; Salmon, 1998, chap. 2; Hoefer 2008; Suppes 1999). The qualification “in a specific context” refers to the need of providing some scope conditions delineating the field in which the relationship is expected to hold, for it is unlikely that any cause-effect relationship is in place across all time and space. For the cross-case level, determinism means that the value of Y must be fully predictable as a function of X for each case in the specified population. We call this the *invariant-effect element* of a deterministic causal explanation.[4] For instance, a deterministic reading of the democratic peace thesis would state that a dyad of two democratic states (X) always displays peaceful relations (Y).[5]

The invariant-effect component only captures the cross-case regularity of the relationship between cause and effect. Therefore, it is insufficient to fully account for a determinist causal explanation which must be complemented with an elaboration of what determinism means on the within-case level. A causal relationship qualifies as deterministic on the within-case level if a causal process links cause and effect in *each* case of the specified population on the within-case level. We term this the *full-connectedness element* of a deterministic causal explanation. With respect to our democratic peace example, a deterministic within-case claim implies that the democratic quality of a country accounts for peace in *every* dyad by a causal process, for instance through the working of norms of peaceful conflict resolution.

Concerning the within-case level, there is a debate on how causal inferences are generated in process tracing. Some scholars argue that it does not suffice to specify and empirically observe *any* causal process linking X and Y that can be reasonably subsumed under the theory in question. Instead, it is recommended to formulate a *specific* sequence of events and intermediate steps in advance of the empirical analysis, which must then be exactly observed in the within-case analysis. This technique is dubbed pattern-matching, for the theorized pattern of intervening

---

[4] The notion 'causal effect' does not imply any assumption regarding the "type" of causality being as either covariational or set-relational relationship (i.e., necessity and/or sufficiency) (Mahoney et al. 2009). We are agnostic on this dimension of causality as our arguments on the cross-case component of causal explanations apply to both. For the discussion of MMR, this entails that the arguments we present hold true for regression analysis and QCA, which embody the two different causal worldviews.

[5] For simplicity, we use dichotomous causes for our examples. The same argument holds for continuous causes.

steps is compared to the empirical process which actually happened in a given case (George and Bennett, 2005, chap. 10; Hall, 2003). As such, pattern-matching is a more demanding technique than ordinary process tracing in which one collects causal-process observations and subsumes them under one or multiple theories in the end. Since all our arguments on the within-case level apply to ordinary processes and process tracing, they naturally extend to the more demanding case of pattern-matching. So, all arguments we raise regarding the more relaxed view on causal processes and within-case analysis, will necessarily hold for the more demanding variant, also.

Conceptually, an invariant effect and full connectedness are individually necessary and jointly sufficient attributes of a deterministic causal explanation (cf. Goertz, 2006, chap. 1). The causal explanation cannot be framed as deterministic if there is an invariant effect, but not a causal process in each case because one necessary requirement is not met. Similarly, the observation of causal processes without an invariant effect does not qualify as determinism, either. The fact that invariant effect and full connectedness are both individually necessary elements of our definition of a deterministic causal explanation has implications for situations where one only theorizes about the cross-case or the within-case level. Whenever a deterministic causal effect is hypothesized to be in place, this necessarily includes the, albeit implicit, assumption that there is a within-case process running from cause and effect in every unit in the population. An invariant effect without full connectedness is useful for prediction, but does not meet the requirements of a deterministic causal explanation. Similarly, a deterministic within-case hypothesis inevitably implies a deterministic cross-case argument. It may be interesting to conjecture and find favorable empirical evidence that X is always linked to Y in a specific way. If there is no deterministic causal effect, however, one is not dealing with a deterministic causal explanation because of the inability to perfectly predict the scores of Y given our knowledge of X.

Our understanding of determinism deviates from other definitions that exclusively focus on the cross-case level. Clark et al. (2006, 313) equate determinism with X ensuring the presence of Y. Besides that they focus on a specific type of cross-case relationship (that is, causal sufficiency), this definition ignores the within-case level that is crucial for us. Similarly, the understanding of determinism as "everything that happens has a cause or causes and could not have happened differently unless something in the cause or causes had also been different" (Carr 1961, cited from Adcock 2002: 2)[6] refers to the cross-case condition, since altered scores on the preconditions will lead to different causes on the outcome. While the within-case condition is not explicitly referred to, however, it is reasonable to infer from Adcock's definition a within-case dimension of deterministic causality, in that he implicitly includes a procedural element. Yet it is important to stress that a deterministic hypothesis necessarily implies a within-case component.

3.2. Probabilism

A relationship between cause and effect is probabilistic if there is an invariant, yet systematic link between X and Y. On the cross-case level, probabilism means that it is not possible to predict the score of Y for *each* case on the basis of X. However, it is still feasible to predict the frequency with which Y occurs in a specified *set* of cases, given the score on X (Salmon, 1998, chap. 6).

---

[6] Adcock (2002) discusses two further varieties of determinism. One captures the reductionist idea that a given effect is produced by one or a very limited number of causes, which then are assumed to be the sole determinants of the outcome. This understanding of deterministic causation rather catches the idea of explanatory parsimony or monocausal explanation than the invariant regularity of cause and effect which could easily apply to multicausal or conjectural cause-effect-relations. The other conception of determinism refers to the level of analysis, mainly the structure/agency divide. In this view, a causal account is called determinist if intention and choice do not play a role in explaining social phenomena. We do not see that this understanding is substantively distinct from the one proposed by us and shared by many researchers.

We dub this the *systematic-effect* element of a probabilistic causal explanation. Returning to the democratic peace example, this would mean that we would expect that, say, ninety percent of all democratic dyads enjoy peaceful relations.

While this cross-case aspect of probabilism is well understood, it has not been systematically discussed for the within-case level (e.g. Goldthorpe, 2001; Bennett, 2003; Goldstone, 1997). Corresponding to our discussion on determinism, we define within-case probabilism as the presence of a causal process in some, but not all cases of the specified population. As is the case for the cross-case level, it is only possible to make statements about the presence of processes in a larger *set* of cases, but not for individual cases, which we call the *systematic-connectedness* condition of a probabilistic causal explanation.

It is to note that the precise reason for probabilism on the cross-case and/or within-case level is not important in the context of causal explanations. In principle, probabilism can be *complexity-induced* and *ontological* (Bennett, 2003; King et al., 1994, chap. 3; Salmon, 1998, chap. 2). Complexity-induced probabilism denotes that the world is inherently deterministic, but looks probabilistic because of insufficiencies in data and/or data processing capacities. Ontological probabilism, on the other hand, captures the belief that there is some degree of inherent irregularity in the world which cannot be explained. These two views on probabilism are ontologically different and rest on distinct assumptions concerning the feasibility of deterministic explanations (Bennett, 2003). In practice, however, they are observationally equivalent because it is impossible to tell in a given situation whether the observed irregularity is due to complexity or ontology (King et al. 1994). From a conceptual perspective on causal explanations it does not matter why there is an imperfect regularity on the cross-case and within-case level.[7] It is simply the presence of a probabilistic causal effect or probabilistic process that suffices to qualify an explanation as probabilistic. Conceptually, therefore, the systematic-effect and systematic-connectedness component are individually necessary and jointly sufficient for qualifying a causal explanation as probabilistic.

Differentiating between determinism and probabilism is, in principle, easy, because it is unproblematic to separate an invariant causal effect from a varying one and to discern whether causal processes are present in all cases or a subset of them. The more difficult question concerns the distinction between probabilistic causality and *randomness*. It is therefore necessary to elaborate on what 'systematic' means on the cross-case and within-case level.

The differentiation between systematic, albeit non-invariant causality and randomness requires specifying a benchmark. Regarding the cross-case level, it is established practice to say that an observed causal effect reflects a probabilistic causal relationship if it is large enough to be unlikely the result of pure chance. Correspondingly, separating probabilism from randomness on the within-case level requires the specification of the number of cases in the population in which a causal process must be present so as to render it sufficiently unlikely that it is due to randomness. While it is common practice in cross-case analyses to specify such a threshold, it is an issue that has not been systematically discussed for the within-case level. However, once one accepts that a causal explanation entails a cross-case and a within-case component, it is straightforward to demand the theoretical explication of a criterion distinguishing probabilism from randomness on both levels. Empirically, it must be additionally taken into account that it is

[7] As we will argue below, however, the two types of probabilism are more important to consider from a methodological perspective, especially when performing MMR.

highly unlikely to ever observe a causal effect that is perfectly random or to discern no theoretically meaningful causal process at all in a single case (an issue to be discussed in detail below). At the margin, it is easy to dismiss a causal explanation when one has, for example, a causal process in only one case out of 100 because one process does not represent a within-case regularity. But how to interpret the finding of 60 or 70 processes in that same sample? This hypothetical example exemplifies the necessity of specifying a benchmark separating probabilism from randomness that should be theoretically defined.

The importance of delineating probabilistic causality from non-systematicness on the level of processes is at the core of a recent exchange between Rosato (2003; 2005) and Slantchev et al. (2005) about the democratic peace research program. Rosato (2003) argues on the basis of a number of case studies and supplementary quantitative evidence that none of the theorized processes proposed to explain democratic peace has empirical substance. Slantchev et al. (2005) reply that the causal process hypotheses of democratic peace theorists are not deterministic, but probabilistic. Thus, a couple of case studies were not sufficient to cast serious doubt on democratic peace theory. In their eyes, a within-case hypothesis should be considered robust if it can make sense of ten percent of all cases. Rosato (2005) refutes this argument and claims ten percent of all cases too low a threshold for a theory to pass a test. In light of our argument, this disagreement can be interpreted as a result of different—and insufficiently explicated—benchmarks of what constitutes probabilism on the within-case level.

Our arguments on causal explanations are summarized in Table 1. Every causal explanation includes an argument concerning causal effects and causal processes. A causal claim is deterministic if it states an invariant causal effect *and* stipulates that a causal process runs from X to Y in each case. Both elements are individually necessary and jointly sufficient for determinism. A differentiation between probabilism and randomness calls for the specification of an appropriate benchmark on theoretical grounds in advance of the empirical analysis. Passing this benchmark on the level of effects *and* processes is individually and jointly sufficient for an explanation to qualify as probabilistic. If either condition is not met – either because the effect or the number of processes are likely to be the product of chance – the empirical relationship in question is non-systematic and therefore does not qualify as a causal relationship.

Table 1: Characteristics of different types of causal explanations

Hypothesized causal relationship	**Characteristic**		
	Cross-case		**Within-case**
Deterministic	Invariant effect	**and**	Process in each case
Probabilistic	Systematic effect	**and**	Process in a certain share of cases
Random	Random effect	**or**	No process in a certain share of cases

3. Causal Explanations and Multi-method Research Designs

In this section, we discuss MMR in light of our theoretical reflections on determinism, probabilism and significance. We are aware that any method, cross-case and within-case alike, rests on specific assumptions and is confronted with a range of problems (see for example Freedman, 1991; Kittel, 2006; Rohlfing, 2008; Wolf, 2010). These issues of course have to be taken into account when performing MMR. In the following, however, we limit our discussion of methods to their *general* suitability for the assessment of deterministic and probabilistic causal explanations and only touch on their shortcomings if they are of immediate relevance for our topic. As we explained in the introduction, the most popular MMR combine regression analysis

(Lieberman, 2003; Lynch, 2002; Wolf, 2010), and, to a lesser degree, QCA with process tracing. Because of this, our treatment of cross-case methods will be limited to these two large-n techniques.

3.1. Determinism and Probabilism in Regression and QCA

Regression analysis and QCA are employed in MMR for the same reasons as in standard, single-method research. The goal is to discern regularities over a large number of cases in order to infer causal relationships on the cross-case level. Both cross-case methods are capable of determining deterministic and probabilistic causality. Concerning determinism in regression analysis, one would expect to find a function which perfectly maps X into Y across all observations. In practice, regression analyses yield correlations which are less than perfect and spot a certain share of unexplained variance (Clark, Gilligan and Golder, 2006). The basic causal assumption behind the mathematical functions underlying the regression model, e.g. $Y = \alpha + \beta X$, however, is inherently deterministic in nature. Given high-quality data, reliable measurement instruments and the absence of idiosyncratic confounders, nothing in the methodological and epistemological foundations of regression analysis prohibits the identification of a deterministic empirical relationship in principle. The same is true for QCA, under which we subsume the three variants crisp-set QCA (csQCA), fuzzy-set QCA (fsQCA), and multi-value QCA (mvQCA) (Rihoux and Ragin, 2008).[8]

Turning to probabilism, regression models incorporate variance in the causal relationship by including an error term into the equation (e.g., $Y = \alpha + \beta X + e$). Reconsidering the distinction between ontological and complexity-induced probabilism, however, the error term only captures

[8] The three variants rest on the same logic of inference and only differ with respect to the measurement of conditions.

the former. This is different for measurement error and omitted variables as the two main sources of complexity-induced probabilism. The effects of measurement error on the regression output depend on whether it is systematic or non-systematic and occurs in the dependent or independent variable. Depending on the concrete constellation, the consequences are inefficiency at best and bias and inconsistency in the worst case (Gujarati, 2004: 524-528; Rabinovich, 2000). It seems safe to assume that all social science data exhibits some degree of non-systematic measurement error and sometimes even systematic error that should be taken into account. For example, Benoit et al. (2009) convincingly argue that the widely used Comparative Manifesto Project data (Budge et al., 2001; Klingemann, et al., 2006) exhibits random measurement error and that one can estimate a simulation-extrapolation model in order to assess its implications on OLS results. One means to address systematic measurement error is the estimation of an instrumental variables regression in which the variable suffering from the error is replaced by one without systematic mismeasurement (Dunning, 2008; Wooldridge, 2003: Chap. 15).[9]

Omitted variable bias is a well-known and hotly debated issue in the social sciences (Achen, 2002, 2005a) and there are various instruments for dealing with this problem (e.g. Clarke, 2005; Fox, 1991). For example, one can visually inspect whether the distribution of the residuals is no-random, which is an indicator for model misspecification. In addition, tests for mis-specification like RESET can to inspect if X has a non-linear effect on Y. Thus, complexity-induced probabilism creates problems for the statistical analysis of cross-case effects, but there are means to control for their presence and diminish their influence.[10]

---

[9] See Dunning (2008) for problems of instrumental variable regression.

[10] A third option that has been rarely considered so far is that the world is a mix of complexity-induced and ontological probabilism (Lieberson and Horwich, 2008). It may be that some variables are deterministically linked to

Concerning QCA, recent innovations have led to the development of parameters that capture probabilistic set-relations, in particular the concept of *consistency* (Ragin, 2006). Consistency captures the degree to which the cases at hand are consistent with a specific set-relationship. For example, if X is present in ten cases, but Y is only given in eight out of these ten, the consistency of a sufficient relation relationship is .8.[11] However, QCA currently lacks tools with which the different problems stemming from complexity-induced and ontological probabilism can be handled. Regarding complexity-induced probabilism, Seawright (2005) shows that the omission of a variable has similar implications for QCA as it has for regression results, since the QCA solution depends heavily on the variables included into the analysis *a priori*. However, different to regression, there is no equivalent to the visual inspection of the distribution of residuals or specification tests. It has been proposed to interpret a solution's consistency score as a proxy measure for the probability of having omitted a condition. The lower the consistency, the more likely it is that a condition is missing whose inclusion would improve the consistency score. However, this is a questionable strategy because it ultimately boils down to post-hoc fitting the solution to cross-case data. The regression approach toward omitted variables does not suffer from this problem because it draws on the distribution of the residuals and not on the R^2 (or whatever measure is used to assess a model's explanatory power) (King, 1991).

At present, QCA is equally prone to suffering from measurement error. In short, measurement error means that the observed scores of cases are different from their true score.

---

each other, but that ontological probabilism introduces some noise that makes it impossible to unambiguously discern the deterministic relationship. Similarly to ontological probabilism, this is not a problem for regression analysis.

[11] Presuming csQCA and dichotomous conditions. The calculation is different for fsQCA.

This implies that the place of a case in the truth table may be different from its location in the "true distribution" in the real world. Whether cases populate different rows in a truth table depends on the extent of the measurement error. *Ceteris paribus*, the probability of misclassification increases with increasing measurement error because cases become more likely to take another score on a condition and thus change rows in the truth table. Since the solution one derives from QCA hinges on the distribution of cases in the truth table, it is apparent that both random and systematic measurement error undermine the validity of the solutions derived from this technique.[12] Due to the lack of specification tests, the scores must be changed manually on a case-by-case basis and it must be evaluated to what extent the QCA solution is sensitive to measurement error. In comparison with regression analysis, QCA thus seems at present less well equipped to deal with the consequences of complexity-induced and ontological probabilism. In total, the implication is that the presence of probabilism may undermine the value of cross-case techniques to separate probabilism from randomness on the cross-case level. As we mentioned before, there is a whole list of problems that produce the same adverse effects, like the failure to control for serial correlation in regression analysis. However, the special problem of probabilism is that the very reason for which cross-case methods are applied, to discern a probabilistic causal effect, may render this goal infeasible.

[12] We acknowledge that there are techniques with which one can assess whether the observed distribution of cases in a 2x2 table is likely to be the result of measurement error or systematic effects (Braumoeller and Goertz, 2000; Ragin, 2000). Yet these tests do not apply to truth tables and therefore do not help with respect to the problem that we point out.

4.2. Determinism and Probabilism in Case Studies

Despite attempts to employ quantitative methods for comparative case study designs (e.g. Abadie, Diamond and Hainmueller, forthcoming) and to statistically estimate models on within-case causal processes (e.g. Box-Steffensmeier and Jones, 1997; Galtung, 1970; Glynn and Quinn, 2007; Goldthorpe, 2001; Imai, Keele and Yamamoto, 2008; Pötter and Blossfeld, 2001), regression techniques are usually deemed inadequate for the analysis of causal processes. Instead, qualitative case study designs are held to be superior for the identification of causal processes and to elucidate whether a causal effect can be attributed to a causal link connecting cause to effect (Bennett and Elman, 2006; George and Bennett, 2005, ch. 10). Two reasons are brought forth for this judgment. First, case studies do not depend on data-set observations for making causal inferences, but rely on causal-process observations (Brady et al. 2004). They are therefore held to be a perfect match to deal with the variety of non-comparable observations one is likely to encounter when examining the empirical implications of causal process hypotheses (King and Powell, 2008). Second, process tracing is supposedly suited to uncover omitted variables and spurious correlations, two important problems which arguably cannot be *sufficiently* dealt with by large-n cross-case analyses.[13] These two rationales make process tracing the ideal method to complement cross-case analyses and provide for explanatory leverage in MMR, which is acknowledged by quantitative and qualitative researchers alike (e.g. Achen, 2005b; Brady, Collier and Seawright, 2006; Lieberman, 2005).

---

[13] In addition, case studies are praised for assessing and improving concept validity and the measurement of variables (Adcock and Collier, 2001; Coppedge, 1999). While, of course, all causal inferences and explanations depend on adequate and reliable measures, the problem is not connected to the inherent logic of causal reasoning which we will discuss in the remainder of this treatment. Hence, we focus only on the two issues named in the text.

This optimism concerning the inferential potential of the case study part of MMR stands in stark contrast to the long history of outspoken criticism of case studies as single-method designs (e.g. Beck, 2006; Goldstone, 1997; Lieberson, 1991). It is true that the systematic combination of cross-case analyses and process tracing might ameliorate some of the critical issues of single-method case studies. For one, the cross-case analysis provides for a systematic foundation to select the cases to be chosen in detail (Bäck and Dumont, 2007; Lieberman, 2005; Shively, 2006). This is an important contribution since all sampling rules and selection techniques must rest on some analysis of the larger population and the cross-case analysis provides an adequate formal technique to map the population (Gerring, 2007a, chap. 5; Tarrow, 1995). Second, case studies in MMR designs are somewhat less prone to suffer from the classic "degrees-of-freedom" problem (Campbell, 1975; Lieberson, 1991). Since the identification of causal regularities on the *macro*-level is provided by the cross-case method, the case study's inherent inability to discriminate between competing macro-level hypotheses is irrelevant.

Being relieved from making cross-case inferences, the question is to what degree process tracing is appropriate for assessing deterministic and probabilistic propositions on the within-case level and to account for the deficiencies of large-n techniques. Again starting with determinism, it is regularly stated that case studies can be used for testing deterministic propositions. Goertz (2003), for instance, argues that small-n methods are adequate for testing necessary condition hypotheses as these can be refuted by a single deviant case. However, the same is not true for corroborating deterministic hypotheses because one would need to trace the processes in *all* cases of a given population, which directly follows from the nature of deterministic causal explanations. The *full-connectedness condition* of a deterministic causal explanation requires that X is linked to Y by a causal process in each case. While not impossible in principle, considering the high demands process tracing puts on the quality of data and depth of analysis, a satisfactory empirical evaluation of the full-connectedness condition is usually unattainable. The best circumstances for testing deterministic causal explanations in MMR designs might arguably hold in Comparative Historical Analysis (CHA), where populations rarely exceed ten to twenty cases (Mahoney and Rueschemeyer, 2003). In most MMR designs, however, dozens or hundreds of cases are analyzed which cannot be all made subject to process tracing.

While the problem of identifying the full-connectedness condition of deterministic causality is hampered only by these practical concerns, the possible existence of probabilistic causality poses a more fundamental problem for the case study part of MMR designs. On first sight, case studies seem to be well-equipped to deal with the complexity-induced variant of probabilism. Indeed, the supposed ability of process tracing to circumvent the variation stemming from measurement error and omitted variables is a common argument in the small-n literature (Bennett and Elman, 2006). We are more skeptical in this regard because the discussion of case studies and measurement error is usually concerned with conceptual validity on the *cross-case level* only, that is, the conceptualization and measurement of the data-set observations due to non-systematic errors, misspecified concepts or inappropriate indicators (Adcock and Collier, 2001). Case studies seem to provide a powerful tool to evaluate and refine the measurement instruments used to capture the empirical realities of the dependent and independent variables. As we have discussed before, however, a case study embedded in MMR does not provide much inferential leverage for the cross-case method and its main contribution to the value of a given explanation rests on its ability to adequately assess the causal processes on the within-case level.[14] The

[14] Furthermore, even if the case study method finds that in the cross-case method the stochastic variance can be attributed to measurement error in some cases, this does not provide any information about the cross-case

possibility of committing measurement error on that level of analysis is, in turn, largely neglected in the small-n literature. Equivalent to cross-case mismeasurement of the causes, measurement error on the within-case level might just as well derive from subsuming a certain process observation under a wrong concept or interpreting some piece of empirical evidence wrongly. This danger might in fact be even more prevalent than for the cross-case variable because of the highly disparate nature of causal process observations and the problems of formalizing the concepts on which these observations are identified, interpreted and coded.[15] To give an example, in his case study of the Tokyo round of trade harmonization in the 1980s, Grieco (1990) derives multiple observable implications from two competing theories, neo-realism and neo-liberalism, and claims that the empirical evidence is more in accord with the neo-realist explanation.[16] While the empirical within-case evidence that he provides supports the neo-realist argument, it is not conclusive as the process observations he cites can equally well be subsumed under the liberalist hypothesis that trade negotiations are actually shaped by domestic politics and the interests of economic actors (e.g. Gilligan, 1997; Pahre, 1998; Pahre, 2008).

The problem of misinterpretation of process evidence is particularly salient for those within-case hypotheses which "bottom out" on the level of individual or collective actors and try

---

regularities and measurement validity in those cases which have not been subjected to the small-n analysis. This argument is equivalent to our further discussion, below.

[15] In fact, the literature on the correct conceptualization and operationalization of cross-case variables is highly developed (Adcock and Collier, 2001; Goertz, 2006; Sartori, 1984).

[16] In short, neo-realism says that countries are concerned about their survival. They seek relative gains in international cooperation since this increases strengthens the position relative to potential future enemies. Neo-liberalism stipulates that countries seek absolute gains and that they only worry about being cheated. Whether or not other countries gain more from cooperation is not important.

to explain social phenomena by these actors' intentions (Machamer, Darden, and Craver 2000). Njølstad (1990), for instance, shows that such problems are ubiquitous even in a well-researched area like US nuclear policy, which is replete with disagreements on how to interpret the supposedly well-documented interests and beliefs of even the most important actors and how these may have contributed to even the most critical and obvious events and developments. Process tracing as such is, hence, not beyond the problems of measurement error and the resulting implications for observing probabilism on the within-case level.

A similar rationale holds for the second reason of imperfect regularities, the omission of a variable. In the absence of homogenously defined and conceptualized variables on the within-case level, the "omitted variable" problem in process tracing presents itself as an "omitted evidence" problem. Since the available sources only catch a certain aspect of an event of interest and the researcher usually is not able to decide if she has collected a representative amount of evidence, the existence of one or more highly important pieces of evidence may go undetected (Thies, 2002). Furthermore, sources for within-case evidence are susceptible to bias (Lustick, 1996), at least as much so as data-set observations. Secondary sources, e.g., historical books, usually adhere to a certain historiography and focus on some aspects of the empirical phenomenon while leaving others aside. Official documents may only contain information that the respective authors want to disseminate to the public, while the real motivations are not documented (or classified and therefore inaccessible for the researcher). The content of newspaper reports may be driven by a certain editorial policy and/or the information of political insiders pursuing their own goals. For a similar reason, the information gathered in expert interviews should be treated carefully (George and Bennett, 2005, chap. 5).

Based on limited within-case evidence, a researcher could fall for a type I error by inferring a causal process between X and Y, while in fact the evidence is flawed or there is some uncovered evidence that the uncovered process is more apparent than real. At the same time, there is a risk of erroneously accepting the null hypothesis when indeed there was some causal process which remained undetected because of inaccessible evidence. Either way, the presence or absence of convincing within-case evidence does not necessarily mean that the identified causal process (or its absence) is true. In sum, the fact that measurement error and omitted causes on the within-case level cannot be ruled out suggests that process tracing must be able to account for probabilism even if the causal relationship is assumed to be deterministic. This is obviously the case if the causal relation is theorized to be probabilistic in the first place.

The problems that we discussed so far pertain to internal validity, that is, to infer that the theorized process is in place or not on the basis of available evidence. Even more limiting, however, are the inherent problems of external validity in case studies. The need to generalize is inherent to MMR process tracing aiming to contribute to the development of probabilistic causal explanations because an integral component of the latter is that a process is given in a certain share of cases in the *population*. The ability of case studies to achieve this with a sufficient degree of certainty is disputable. Confidence in the generalizability of within-case inferences can hardly be robust, given that only a small fraction of cases can be studied in within-case analysis. Suppose, for example, a MMR study entails 25 cases in the large-n part and the cross-case analysis indicates a systematic effect. Suppose further that three of the 25 cases are chosen to be studied through process tracing. Finally, assume that in this sample of three cases, a causal process can be uncovered in two cases. What can one infer from this finding? For the reasons we explained above, answering this question first of all requires to specify the threshold, which we set here at 80 percent. This means that the criterion for distinguishing probabilism from randomness on the within-case level is that 80 percent of all cases in the population are assumed to have the process in place. Statistically, drawing cases from the population, inferring that a process is in place or not in these cases, and generalizing to the population follows the logic of hypergeometric distribution (the logic is analogous to drawing balls from an urn without placing them back before drawing the next one). Given the threshold, the probability of observing a process in two of three cases is about .40, which means that one cannot reject the hypothesis that the sample is drawn from a sample in which 80 percent of the cases have the process given. The importance of the threshold becomes apparent when considering that the probability for observing a process in one of three cases is about .10. Although it would not be satisfactory to perform such a within-case analysis where the process is absent more often than not, we still cannot reject the threshold at the conventional level of significance. Even more important is, however, that the probability of discerning a process in two out of three cases is .10 if there *is* such a process in only 20 percent of the cases in the population. These 20 percent are far off the benchmark of 80 percent, but it is not possible to reject either of the two thresholds by observing a process in two out of three cases. The opportunity to generalize from the sample to the population of course depends on all parameters, populations size, sample size, observed processes in the sample, and the benchmark. Yet this example, in which the parameters took rather favorable values, shows that process tracing is generally unable to reliably assess the systematic-connectedness component of a probabilistic causal explanation. This problem stems, of course, from the well-known "small-n problem" of case study research (Goldthorpe, 1997; Lieberson, 1991). It should be noted, however, that the traditional small-n problem refers to the cross-case level, that is, what one can infer from patterns of scores on the variables/conditions about the whole population. Our discussion of case studies, in contrast, particularly aims at the within-case level and process tracing, which is at the heart of qualitative case studies.

Some authors argue that MMR research is particularly suited to overcome this problem, because the large-n study provides a foundation from which to gauge the representativeness of a case in regard to the population. For regression analysis it has been proposed to analyze the residuals of cases in order to find *typical cases* (i.e. cases with small residual and which lie on or close to the regression line) and *deviant cases* (those whose residuals are larger and which lie farther off the regression line) (Eckstein, 1975; Lieberman, 2005; Lijphart, 1971; Seawright and Gerring, 2008). A similar technique has been recently suggested for *fuzzy set* QCA (Rohlfing and Schneider, 2009). While these selection criteria are sensible and reduce the problem of selection bias in case study research to some degree (Collier and Mahoney, 1996), they do not ameliorate the specific problems related to probabilistic causality. Typical cases are by definition representative of the larger population, but they are only representative on the cross-case level to the extent that a case is well-captured by a regression model or QCA solution. If one accepts that a cross-case pattern is not causation, the cross-case representativeness of the sampled cases does not provide any certainty concerning the representativeness of these cases on the within-case level. This, rather, has to be evaluated empirically, which, as we have discussed above, is does not provide much inferential leverage for the whole population given the unfavorable sample/population ratio.

A second case selection strategy that may be used to counter our critique chooses cases on he basis of theoretical expectations and prior empirical knowledge in order to enhance inferential power. The most prominent case selection techniques which build on such prior expectations are *most-likely* and *least-likely* tests (Eckstein, 1975; Gerring, 2007b; Lijphart, 1971). The argument behind the most-likely test is that if one does not find the specified process in a case where it is most probable to be present, it is unlikely that the process will be observed in other cases (which are themselves not analyzed). In regard to the least-likely test it is argued that if one finds a process in cases in which it is unlikely to be expected, it is safe to assume that in all other cases the process is likely to be present as well. A range of criticisms has been formulated regarding the usefulness of such most-likely and least-likely tests, for instance the weak state of theory in political science as well as the ambiguous nature and often contested interpretation of empirical evidence, which undermine the researcher's ability to adequately estimate the certainty of his prior expectations (cf. Sober 2002).[17]

Besides these caveats, however, our point remains that either way one cannot distinguish probabilism from randomness on the within-case level. This is because any inferences from tests on very few cases rest on untestable (and usually implicit) assumptions about the frequency with which non-systematic and omitted systematic variables are at work in the population. For instance, the failure of a most-likely test due to an idiosyncratic factor may be limited to the case examined. Alternatively, the case could be a member of a larger subset of cases in the population in which a non-systematic variable is interfering with the causal link between X and Y. In order to know which of the two possibilities holds, one first needs to form a theoretical expectation about how robust the process between X and Y is to the influence of non-systematic factors. In addition, it is necessary to know the frequency distribution of those idiosyncratic factors which have a strong enough impact to deteriorate the link between X and Y. This distribution, which is

[17] Recently, the logic behind these case selection strategies has been restated in terms of Bayesian probability (Bennett, 2008; Dion, 1998). Our criticism is unaffected by these intriguing efforts to formalize the underlying principles of selecting cases based on knowledge on priors for the following reasons: First, as Dion underscores, his Bayesian approach is limited to necessary conditions. Second, both authors are not concerned with generalizing from a sample to a population, but deal with *internal* validity. Matters of sample size and population size are not discussed by Dion or Bennett. Third, the number of cases is only small when the *a priori* confidence in the own hypotheses is high and confidence in the alternative hypothesis is low. Besides presuming strong theory, cases which meet these criteria are most-likely cases, which are theoretically uninteresting for empirically evaluating causal explanations because of the high probability to observe what one expects to observe (Eckstein, 1975).

observable only on the within-case level, however, is unknown because the sample size for within-case analysis is too small to make inferences on the population with certainty. It is therefore impossible to tell precisely how much a failed most-likely test should decrease our confidence regarding the tested within-case part of a causal explanation. The same logic, naturally, applies to most-likely tests that fail because of omitted variables.[18] If a researcher concludes on the basis of causal process observations that a variable should be added to the explanation, she implicitly assumes that this variable is also connected to Y in a significant share of the population. Again, this assumption is untestable so that there is no certainty about the weight one should attach to the failed most-likely test. Because of this, theory-driven case selection and most-likely/least-likely tests cannot compensate for the inherent inability of case studies to capture the within-case dimension of probabilistic causality.

4. Conclusion

There is a wide ranging agreement in the social sciences that the primary objective of scientific research should be the establishment of causal explanations; that is, stating the causes which produce the phenomenon of interest. This includes identifying the relevant causes and their effect on the outcome. In addition, explanation also entails the explication of the processes which link the purported cause to the effect. Since the observation of a cross-case pattern is not sufficient for determining causality, it is widely held that, "[s]patiotemporal continuity […] makes the critical

[18] Mismeasurement is only of secondary importance because the problems of how to interpret failed most-likely cases or past least-likely cases equally apply to the deteriorating influence of non-systematic and systematic variables on the link between X and Y. While it of course makes a difference whether a variable is classified as random or systematic, the methodological implications for the evaluation of failed tests are the same.

difference [...]. When we have provided spatiotemporally continuous connections between correlated events, we have fulfilled a major part of the demand for an explanation of the correlation" (Salmon, 1998, 113). Therefore, empirically sound causal explanations involve two distinct steps. For one, the relevant causes on the macro-level needs to be identified and their covariation with the outcome of interest needs to be established. This cross-case analysis must be complemented with within-case investigations in the next step to uncover these "spatiotemporally continuous connections between correlated events". Given these requirements, the combination of large-n cross-case techniques and small-n within-case analysis into a single MMR design seems particularly suited for providing robust empirical underpinnings of causal explanations.

While we agree that combining large-n methods and case studies is, in principle, a fruitful approach for developing causal explanations, we are skeptical about the actual degree to which MMR can deliver what it seems to promise in regard of producing and assessing causal explanations. Regarding the large-n part, regression analysis and QCA are suited for analyzing cross-case relationships. Similarly, we concur that case studies are appropriate for tracing the causal processes linking cause and effect in discrete cases. We are much less convinced, however, concerning their capacity to offer as much explanatory leverage as most of the existing MMR literature seems to put in them.

For one, we cannot be sure that the observation of a process in one or a few cases really is a systematic feature of the causal relationships in question. Similarly, from the absence of a theoretically expected process in the small-n sample we cannot infer with a sufficient degree of confidence the conclusion that there is no causal relationship in the population. As much as correlation is not causation, no process does not mean no causation, either. As we have shown, these problems cannot be mitigated by consciously choosing the cases for within-case analysis based on the results of the large-n method or through theoretical expectations. Furthermore, these problems hold true regardless of whether probabilism is due to our inability to adequately measure the complexity of reality or whether it is an inherent feature of the social world.

In sum, within-case analysis on a few cases does not allow confident inferences regarding the existence and distribution of processes in the population. Case studies are therefore limited in what they can hope to contribute to the robustness of MMR results, that is, whether the observed cross-case relationship is indeed causal or spurious. This resonates with most of the philosophical literature on causality, where the distinction between singular and general explanations is made very clearly. As Ellery Eells points out, very little about what happens on the level of the single unit or case (“token”) can be inferred from aggregate-level (“type”) probabilistic causal claims, and very little can be learned about population-level probabilistic causal relations from case-level probabilistic causal claims (Eells, 2008).

What to do with these findings? How should we approach the within-case part of MMR designs if we agree with David Zuckerman’s testimony that we have “strong reasons to view the political world as containing nonlinear relationships among variables, probabilistic outcomes and structures, aperiodic systems, unpredictable phenomena, chance factors, and open-ended probabilities” (Zuckerman, 1997, 287)? One option would be to forgo the small-n part completely and to focus on the large-n analysis. And indeed, some authors argue that case studies cannot avoid the problems of generalization and one therefore should not aim to make inferences from single instances to a larger class of cases (cf. George and Bennett, 2005, 30-32; Gerring, 2007a, chap. 2). However, this suggestion is not a viable option as the within-case analysis is necessary in order to evaluate the within-case propositions of a causal explanation. As long as there is not enough comparable within-case data to run statistical or QCA analyses on an adequately large sample of cases, case study process tracing in a few cases will remain the method of choice for the within-case part of MMR designs—and, thus, the problems outlined above will remain.

Should, then, MMR be discarded? Clearly not. Even if the potential for inferences is limited, case studies do contribute to our knowledge about at least some of the cases. Knowing little is better than knowing nothing, after all, and as Paul Humphreys aptly elucidates, even partial causal explanations about which we cannot be sure that they hold for all cases are informative; they are neither false by necessity, nor do they hinder the accumulation of knowledge (Humphreys 1989). However, in order to realistically evaluate the chances of accumulation of knowledge through the combination of large-n methods and case-studies, we propose two tentative implications of our analysis.

First, concerning the problem of probabilism, the epistemological value of process observations on the within-case level must be reassessed. If they are not an adequate basis for making strong inferences, they can provide little more than *informative clues* about the veracity of the original causal proposition in regard to a larger number of cases. In this view, it might be more appropriate to think of within-case observations of nothing more as individual "pieces of information" (Collier, Brady and Seawright, 2004), that need to be supplemented with more evidence from other cases (Beck, 2006).

Second, it seems necessary to reconsider case selection rules in multi-method research. At present, the standard prescription for case selection is to choose cases according to their residuals (Gerring, 2007b; Lieberman, 2005; Seawright and Gerring, 2008). If there is little inferential value in observing a process (or not), there is little sense in selecting cases without knowing that there is a process (or not). In our eyes, a more appropriate selection strategy would be to select cases on the cross-case *and* the within-case level, that is, on the basis of the scores on X and Y and with the knowledge that a process is present. Having selected a case on the process, the goal of the within-case analysis, then, would be to test or search for genuine within-case implications of a theory. We believe that this is a valuable feature of process tracing which has its place in multi-method research despite the problems deriving from probabilism.

References

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller (forthcoming): Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of Calafornia's Tobacco Control Program. *Journal of the American Statistical Association*.

Abbott, Andrew (1998): The Causal Devolution. Sociological Methods & Research 27(2):148-181.

Achen, Christopher H. (2002): Toward a New Political Methodology: Microfoundations and Art. Annual Review of Political Science 5(1):423-450.

Achen, Christopher H. (2005): Two Cheers for Charles Ragin. Studies in Comparative International Development 40(1):27-32.

Achen, Christopher H. (2005a): Let's Put Garbage-Can Regressions and Garbage-Can Probits Where They Belong. Conflict Management & Peace Science 22(4):327-339.

Adcock, Robert (2002): "Determinism" and Comparative-Historical Analysis: Clarifying Concepts and Retrieving Past Insights. Paper presented at the annual meeting of the American Political Science Association 2002.

Adcock, Robert, and David Collier (2001): Measurement Validity: A Shared Standard for Qualitative and Quantitative Research. American Political Science Review 95(3):529-546.

Bäck, Hanna, and Patrick Dumont (2007): Combining Large-N and Small-N Strategies: The Way Forward in Coalition Research. West European Politics 30(3):467-501.

Beck, Nathaniel (2006): Is Causal-Process Observation an Oxymoron? Political Analysis 14(3):347-352.

Bennett, Andrew (1999): Causal Inference in Case Studies: From Mill's Methods to Causal Mechanisms. Paper presented at the American Political Science Association Conference 1999.

Bennett, Andrew (2003): Beyond Hempel and Back to Hume: Causal Mechanisms and Causal Explanations. Paper presented at the annual meeting of the American Political Science Association 2003.

Bennett, Andrew (2008): Process-Tracing: A Bayesian Perspective. In Oxford Handbook of Political Methodology, edited by Janet M. Box-Steffensmeier, Henry Brady, and David Collier, pp. 702-721. Oxford: Oxford University Press.

Bennett, Andrew, and Colin Elman (2006): Qualitative Research: Recent Developments in Case Study Methods. Annual Review of Political Science 9:455-476.

Benoit, Kenneth, Michael Laver, and Slava Mikhaylov (2009): Treating Words as Data with Error: Uncertainty in Text Statements of Policy Positions. American Journal of Political Science 53(2):495-513.

Bollen, Kenneth A., Barbara Entwisle, and Arthur S. Alderson (1993): Macrocomparative Research Methods. Annual Review of Sociology 19:321-325.

Box-Steffensmeier, Janet M., and Bradford S. Jones (1997): Time Is of the Essence: Event History Models in Political Science. American Journal of Political Science 41(4):1414-1461.
Brady, Henry E. (2003): Models of Causal Inference: Going Beyond the Neyman-Rubin-Holland Theory. Paper presented at the Annual Meetings of the Political Methodology Group 2002.
Brady, Henry E., and David Collier (2004): Rethinking Social Inquiry: Diverse Tools, Shared Standards. Lanham: Rowman & Littlefield.
Brady, Henry E., David Collier, and Jason Seawright (2006): Toward a Pluralistic Vision of Methodology. Political Analysis 14(3):353-368.
Braumoeller, Bear F., and Gary Goertz (2000): The Methodology of Necessary Conditions. American Journal of Political Science 44(4):844-858.
Budge, Ian, Hans-Dieter Klingemann, Andrea Volkens, Judith Bara, Eric Tanenbaum, Richard C. Fording, Derek J. Hearl, Hee Min Kim, Michael D. McDonald, and Silvia M. Mendes (2001): Mapping Policy Preferences: Estimates for Parties, Electors, and Governments, 1945-1998. Oxford: Oxford University Press.
Bunge, Mario (1997): Mechanism and Explanation. Philosophy of the Social Sciences 27(4):410-465.
Campbell, Donald T. (1975): Degrees of Freedom and the Case Study. Comparative Political Studies 8(2):178-193.
Capoccia, Giovanni C., and Michael Freeden (2006): Multi-Method Research in Comparative Politics and Political Theory. Committee on Concepts and Methods working paper series, no. 9.
Cartwright, Nancy (2004): From Causation to Explanation and Back. In The Future for Philosophy, edited by Brian Leiter, pp. 230-245. Oxford: Oxford University Press.
Cartwright, Nancy. 2009. Causality, Invariance, and Policy. In The Oxford Handbook of Philosophy of Economics, ed. Don Ross and Harold Kincaid, 410-423. Cambridge: Cambridge University Press.
Clark, William Roberts, Michael J. Gilligan, and Matt Golder (2006): A Simple Multivariate Test for Asymmetric Hypotheses. Political Analysis 14(3):311-331.
Clarke, Kevin A. (2005): The Phantom Menace: Omitted Variable Bias in Econometric Research. Conflict Management & Peace Science 22(4):341-352.
Collier, David, and James Mahoney (1996): Insights and Pitfalls: Selection Bias in Qualitative Research. World Politics 49(1):56-91.
Collier, David, Henry E. Brady, and Jason Seawright (2004): Sources of Leverage in Causal Inference: Toward an Alternative View of Methodology. In Rethinking Social Inquiry. Diverse Tools, Shared Standards, edited by Henry E. Brady, and David Collier, pp. 229-266. Lanham: Rowman & Littlefield.
Coppedge, Michael (1999): Thickening Thin Concepts and Theories - Combining Large N and Small in Comparative Politics. Comparative Politics 31(4):465-476.
Cox, Robert. 1995. Realism, Positivism and Historicism. In Approaches to World Order, ed. Robert Cox and Timothy Sinclair. New York: Cambridge University Press.
Demetriou, Chares (2009): The Realist Approach to Explanatory Mechanisms in Social Science: More Than a Heuristic? Philosophy of the Social Sciences, advance access.
Dessler, David (1991): Beyond Correlations: Toward a Causal Theory of War. International Studies Quarterly 35(3):337-355.

Dion, Douglas (1998): Evidence and Inference in the Comparative Case Study. Comparative Politics 30(2):127-145.

Dunning, Thad (2008): Model Specification in Instrumental-Variables Regression. *Political Analysis* 16(3):290-302.

Eckstein, Harry (1975): Case Study and Theory in Political Science. In Strategies of Inquiry. Handbook of Political Science, Vol. 7, edited by Fred I. Greenstein, and Nelson W. Polsby, pp. 79-137. Reading, Mass.: Addison-Wesley.

Eells, Ellery (2008): Probabilistic Causality. Cambridge: Cambridge University Press.

Elster, Jon (1998): A Plea for Mechanisms. In Social Mechanisms. An Analytical Approach to Social Theory., edited by Peter Hedström, and Richard Swedberg, pp. 45-73. Cambridge: Cambridge University Press.

Elster, Jon. 2007. Explaining Social Behavior: More Nuts and Bolts for the Social Sciences. Cambridge: Cambridge University Press.

Falkner, Gerda (2007): Time to Discuss: Data to Crunch or Problems to Solve? A Rejoinder to Robert Thomson. West European Politics 30(5):1009-1021.

Fox, John (1991): Regression Diagnostics. Newbury Park: Sage.

Freedman, David A. (1991): Statistical Models and Shoe Leather. Sociological Methodology 21:291-313.

Galtung, Johan (1970): Diachronic Correlation, Process Analysis and Causal Analysis: The Quest for Diachronic, Nomothetic Social Science. Quality & Quantity 4(1):55-94.

George, Alexander L., and Andrew Bennett (2005): Case Studies and Theory Development in the Social Sciences. Cambridge, Mass.: MIT Press.

Gerring, John (2004): What Is a Case Study and What Is It Good For? American Political Science Review 98(2):341-354.

Gerring, John (2005): Causation: A Unified Framework for the Social Sciences. Journal of Theoretical Politics 17(2):163-198.

Gerring, John (2007a): The Case Study Method: Principles and Practices. Cambridge: Cambridge University Press.

Gerring, John (2007b): Is There a (Viable) Crucial-Case Method? Comparative Political Studies 40(3):231-253.

Gerring, John (2008): The Mechanismic Worldview: Thinking inside the Box. British Journal of Political Science 38(1):161-179.

Gerring, John (forthcoming): Causal Mechanisms: Yes, But... Comparative Political Studies.

Gilligan, Michael J. (1997): Lobbying as a Private Good with Intra-Industry Trade. International Studies Quarterly 41(3):455-474.

Glynn, Adam, and Kevin Quinn (2007): Non-Parametric Mechanisms and Causal Modeling. working paper.

Goertz, Gary (2003b): Cause, Correlation, and Necessary Conditions. In Necessary Conditions. Theory, Methodology, and Applications, edited by Gary Goertz, and Harvey Starr, pp. 47-65. Lanham: Rowman & Littlefield.

Goertz, Gary (2005): Necessary Condition Hypotheses as Deterministic or Probabilistic: Does It Matter? APSA Qualitative Methods Newsletter 3(1):22-27.

Goertz, Gary (2006): Social Science Concepts: A User's Guide. Princeton: Princeton University Press.

Goertz, Gary, and Harvey Starr (2003): Introduction: Necessary Condition Logics, Research Design, and Theory. In Necessary Conditions. Theory, Methodology, and Applications, edited by Gary Goertz, and Harvey Starr, pp. 1-24. Lanham: Rowman & Littlefield.

Goldstone, Jack A. (1997): Methodological Issues in Comparative Macrosociology. Comparative Social Research 16:107-120.
Goldthorpe, John H. (1997a): Current Issues in Comparative Macrosociology: A Debate on Methodological Issues. Comparative Social Research 16:1-26.
Goldthorpe, John H. (1997b): A Response to the Commentaries. Comparative Social Research 16:121-132.
Goldthorpe, John H. (2001): Causation, Statistics, and Sociology. European Sociological Review 17(1):1-20.
Grieco, Joseph M. (1990): Cooperation among Nations: Europe, America, and Non-Tariff Barriers to Trade. Ithaca: Cornell University Press.
Gujarati, Damodar N. (2004): Basic Econometrics. Toronto: McGraw-Hill.
Hall, Peter A. (2003): Aligning Ontology and Methodology in Comparative Research. In Comparative Historical Analysis in the Social Sciences, edited by James Mahoney, and Dietrich Rueschemeyer, pp. 373-404. Cambridge: Cambridge University Press.
Hall, Peter A. (2006): Systematic Process Analysis: When and How to Use It. European Management Review 3:24-31.
Hedström, Peter, and Richard Swedberg (1998): Social Mechanisms: An Analytical Approach to Social Theory. Cambridge: Cambridge University Press.
Hedström, Peter. 2008. Studying Mechanisms to Strengthen Causal Inferences in Quantitative Research. In The Oxford Handbook of Political Methodology, ed. Janet M. Box-Steffensmeier, Henry E. Brady, and David Collier, 319- 335. Oxford: Oxford University Press.
Hempel, Carl G. 1965. Aspects of Scientific Explanation and Other Essays in the Philosophy of Science. New York: Free Press.
Hitchcock, Christopher. 2004. Do All and Only Causes Raise the Probabilities of Effects? In Causation and Counterfactuals, ed. John Collins, Ned Hall, and L.A. Paul. Cambridge: MIT Press.
Hoefer, Carl. 2008. Causal Determinism. In The Stanford Encyclopedia of Philosophy (Winter 2008 Edition). http://plato.stanford.edu/archives/win2008/entries/determinism-causal/.
Holland, Paul (1986): Statistics and Causal Inferences. Journal of the American Statistical Association 81(396):945-960.
Humphreys, Paul. 1989. The Chances of Explanation: Causal Explanation in the Social, Medical, and Physical Sciences. Princeton: Princeton University Press.
Imai, Kosuke, Luke Keele, and Teppei Yamamoto (2008): Identification and Inference in Causal Mediation Analysis. working paper.
Kemp, Stephen, and John Holmwood (2003): Realism, Regularity and Social Explanation. Journal for the Theory of Social Behaviour 33(2):165-187.
King, Gary (1991): Truth Is Stranger Than Prediction, More Questionable Than Causal Inference. American Journal of Political Science 35(4):1047-1053.
King, Gary, and Eleanor Neff Powell (2008): How Not to Lie without Statistics. Paper prepared for delivery at the 2008 Annual Meeting of the American Political Science Association.
King, Gary, Robert O. Keohane, and Sidney Verba (1994): Designing Social Inquiry: Scientific Inference in Qualitative Research. Princeton: Princeton University Press.
Kittel, Bernhard (2006): A Crazy Methodology? On the Limits of Macro-Quantitative Social Science Research. International Sociology 21(5):647-677.

Klingemann, Hans-Dieter, Andrea Volkens, Judith Bara, Ian Budge, and Michael MacDonald (2006): Mapping Policy Preferences Ii. Estimates for Parties, Electors, and Governments in Eastern Europe, the European Union and the OECD, 1990-2003. Oxford: Oxford University Press.

Kurki, Milja (2007): Causation in International Relations: Reclaiming Causal Analysis. Cambridge: Cambridge University Press.

Kurki, Milja. 2008. Causation in International Relations: Reclaiming Causal Analysis. Cambridge: Cambridge University Press.

Levy, Jack S. (2008): Case Studies: Types, Designs, and Logics of Inference. Conflict Management and Peace Science 25(1):1-18.

Lieberman, Evan S. (2003): Race and Regionalism in the Politics of Taxation in Brazil and South Africa. Cambridge: Cambridge University Press.

Lieberman, Evan S. (2005): Nested Analysis as a Mixed-Method Strategy for Comparative Research. American Political Science Review 99(3):435-452.

Lieberson, Stanley (1991): Small Ns and Big Conclusions: An Examination of the Reasoning in Comparative Studies Based on a Small Number of Cases. Social Forces 70(2):307-320.

Lieberson, Stanley (2001): Book Review: Fuzzy-Set Social Science. Contemporary Sociology 30(4):331-334.

Lieberson, Stanley , and Joel Horwich (2008): Implication Analysis: A Pragmatic Proposal for Linking Theory and Data in the Social Sciences. Sociological Methodology 38(1):1-50.

Lijphart, Arend (1971): Comparative Politics and the Comparative Method. American Political Science Review 65(3):682-693.

Little, Daniel. 1998. Microfoundations, Method, and Causation: On the Philosophy of the Social Sciences. New Brunswick: Transaction.

Lustick, Ian S. (1996): History, Historiography, and Political Science: Multiple Historical Records and the Problem of Selection Bias. American Political Science Review 90(3):605-618.

Lynch, Julia (2002): The Age of Welfare: Citizens, Clients, and Generations in the Development of the Welfare State. University of California, Berkeley: Ph.D. thesis.

Machamer, Peter, Lindley Darden, and Carl F. Craver. 2000. Thinking about Mechanisms. Philosophy of Science 67(1): 1-25.

Mahoney, James (1999): Nominal, Ordinal, and Narrative Appraisal in Macrocausal Analysis. American Journal of Sociology 104(4):1154-1196.

Mahoney, James (2004). Reflections on Fuzzy Set/QCA. Qualitative Methods. Newsletter of the American Political Science Association Organized Section on Qualitative Methods and Multi-Method Research 2(2):17-22.

Mahoney, James (2004): Comparative-Historical Methodology. Annual Review of Sociology 30:81-101.

Mahoney, James (2008): Toward a Unified Theory of Causality. Comparative Political Studies 41(4-5):412-436.

Mahoney, James, and Dietrich Rueschemeyer (2003): Comparative Historical Analysis in the Social Sciences. Cambridge: Cambridge University Press.

Mahoney, James, Erin Kimball, and Kendra L. Koivu (2009): The Logic of Historical Explanation in the Social Sciences. Comparative Political Studies 42(1):114-146.

Marchionni, Caterina (2008): Explanatory Pluralism and Complementarity: From Autonomy to Integration. Philosophy of the Social Sciences 38(3):314-333.

Marx, Axel (2006): Towards More Robust Model Specification in QCA: Results from a Methodological Experiment. COMPASSS Working Paper.
Mayntz, Renate (2004): Mechanisms in the Analysis of Social Macro-Phenomena. Philosophy of the Social Sciences 34(2):237-259.
McKeown, Timothy J. (1999): Case Studies and the Statistical World View. International Organization 53(1):161-190.
Mooney Marini, Magret, and Burton Singer (1988): Causality in the Social Sciences. Sociological Methodology 18:347-409.
Morgan, Stephen L., and Christopher Winship. 2007. Counterfactuals and Causal Inference: Methods and Principles for Social Research. Cambridge: Cambridge University Press.
Munck, Gerardo L. (2005): Ten Fallacies About Qualitative Research. Newsletter of the APSA Organized Section on Qualitative Methods and Multi-Method Research 3(1):2-5.
Munck, Gerardo L., and Richard Snyder (2007): Debating the Direction of Comparative Politics: An Analysis of Leading Journals. Comparative Political Studies 40(1):5-31.
Njølstad, Olav (1990): Learning from History? Case Studies and Limits to Theory-Building. In Arms Races: Technological and Political Dynamics, edited by Nils Petter Gleditsch, and Olav Njølstad, pp. 220-246. Oslo: Sage.
Pahre, Robert (1998): Reactions and Reciprocity: Tariffs and Trade Liberalization from 1815 to 1914. Journal of Conflict Resolution 42(4):467-492.
Pahre, Robert (2008): Politics and Trade Cooperation in the Nineteenth Century: The "Agreeable Customs" of 1815–1914. Cambridge: Cambridge University Press.
Pearl, Judea (2000): Causality: Models, Reasoning, and Inference. Cambridge: Cambridge University Press, 1. ed.
Pinto, Pablo M., and Jeffrey F. Timmons (2005): The Political Determinants of Economic Performance: Political Competition and the Sources of Growth. Comparative Political Studies 38(1):26-50.
Pötter, Ulrich, and Hans-Peter Blossfeld (2001): Causal Inference from Series of Events. European Sociological Review 17(1):21-32.
Rabinovich, Semyon G. (2000): Measurement Errors and Uncertainties: Theory and Practice. New York: AIP Press.
Ragin, Charles C. (1987): The Comparative Method: Moving Beyond Quantitative and Qualitative Strategies. Berkeley: University of Berkeley Press.
Ragin, Charles C. (2000): Fuzzy-Set Social Science. Chicago: University of Chicago Press.
Ragin, Charles C. (2005): Core Versus Tangential Assumptions in Comparative Research. Studies in Comparative International Development 40(1):33-38.
Ragin, Charles C. (2006): Set Relations in Social Research: Evaluating Their Consistency and Coverage. Political Analysis 14(3):291-310.
Reiss, Julian. 2009. Causation in the Social Sciences: Evidence, Inference, and Purpose. Philosophy of the Social Sciences 39(1): 20-40.
Rihoux, Benoit (2006): Qualitative Comparative Analysis (QCA) and Related Systematic Comparative Methods: Recent Advances and Remaining Challenges for Social Science Research. International Sociology 21(5):679-706.
Rihoux, Benoit, and Charles Ragin (2008): Configurational Comparative Methods. Thousand Oaks: Sage.
Rohlfing, Ingo (2008): What You See and What You Get: Pitfalls and Principles of Nested Analysis in Comparative Research. Comparative Political Studies 41(11):1492-1514.

Schneider, Carsten Q. and Ingo Rohlfing (2009): Explaining Party Competition: A Multi-Method Approach Combining Fuzzy Set Qualitative Comparative Analysis (fsQCA) and Case Studies. General Conference of the European Consortium for Political Research, Potsdam, Sep 10, 2009

Rosato, Sebastian (2003): The Flawed Logic of Democratic Peace Theory. American Political Science Review 97(4):585-602.

Rosato, Sebastian (2005): Explaining the Democratic Peace. American Political Science Review 99(3):467-472.

Rubin, Donald B. (2008): For Objective Causal Inference, Design Trumps Analysis. Annals of Applied Statistics 2(3):808-840.

Ruggie, John Gerard. 1999. What Makes the World Hang Together? Neo-Utilitarianism and the Social Constructivist Challenge. In Exploration and Contestation in the Study of World Politics, ed. Peter J. Katzenstein, Robert O. Keohane, and Stephen D. Krasner, 215-246. Cambridge: MIT Press.

Salmon, Wesley C. (1998): Causality and Explanation. New York: Oxford University Press.

Sartori, Giovanni (1984): Guidelines for Concept Analysis. In Social Science Concepts: A Systematic Analysis, edited by Giovanni Sartori, pp. 15-85. Beverly Hills, Calif.: Sage Publications.

Seawright, Jason (2005): Qualitative Comparative Analysis Vis-À-Vis Regression. Studies in Comparative International Development 40(1):3-26.

Seawright, Jason, and John Gerring (2008): Case-Selection Techniques in Case Study Research: A Menu of Qualitative and Quantitative Options. Political Research Quarterly 61(2):294-308.

Shively, W. Phillips (2006): Case Selection: Insights from Rethinking Social Inquiry. Political Analysis 14(3):344-347.

Simon, Herbert A. (1954): Spurious Correlation: A Causal Interpretation. Journal of the American Statistical Association 49(267):467-479.

Slantchev, Branislav L., Anna Alexandrova, and Erik Gartzke (2005): Probabilistic Causality, Selection Bias, and the Logic of the Democratic Peace. American Political Science Review 99(3):459-462.

Sober, Elliott (2002): Bayesianism – Its Scope and Limits. In Bayesianism, ed. Richard Swinburne. London: British Academy.

Steel, Daniel. (2004): Social Mechanisms and Causal Inference. Philosophy of the Social Sciences 34(1):55-78.

Stinchcombe, Arthur L. (1991): The Conditions of Fruitfulness of Theorizing About Mechanisms in Social-Science. Philosophy of the Social Sciences 21(3):367-388.

Suppes, Patrick (1970): A Probabilistic Theory of Causality. Acta philosophica Fennica 24. Amsterdam: North-Holland Publishing Company.

Suppes, Patrick (1999): The Noninvariance of Deterministic Causal Models. Synthese 121(1/2): 181-198.

Tarrow, Sidney (1995): Bridging the Quantitative-Qualitative Divide in Political-Science. American Political Science Review 89(2):471-474.

Thies, Cameron G. (2002): A Pragmatic Guide to Qualitative Historical Analysis in the Study of International Relations. International Studies Perspectives 3(4):351-372.

Thomson, Robert (2007): Time to Comply: National Responses to Six EU Labour Market Directives Revisited. West European Politics 30(5):987-1008.

Van Evera, Stephen (1997): Guide to Methods for Students of Political Science. Ithaca: Cornell University Press.

Waldner, David (2007): Transforming Inferences into Explanations: Lessons from the Study of Mass Extinctions. In Theory and evidence in comparative politics and international relations, ed. Richard Ned Lebow and Mark I. Lichbach, pp. 145-177. 1. ed. Basingstoke: Palgrave Macmillan.

Wolf, Frieder (2010): Enlightened Eclecticism or Hazardous Hodgepodge? Mixed Methods and Triangulation Strategies in Comparative Public Policy Research. Journal of Multi-Method Research.

Woodward, James (2003): Making Things Happen: A Theory of Causal Explanation. Oxford: Oxford University Press.

Woodward, James (2009): Scientific Explanation. In The Stanford Encyclopedia of Philosophy (Fall 2009 Edition). http://plato.stanford.edu/entries/scientific-explanation/.

Wooldridge, Jeffrey M. (2003): Introductory Econometrics: A Modern Approach. Cincinnati, Ohio: South-Western College Pub.

Yee, Albert S. (1996): The Causal Effects of Ideas on Policies. International Organization 50(1):69-108.

Zuckerman, Alan S. (1997): Reformulating Explanatory Standards and Advancing Theory in Comparative Politics. In Comparative Politics: Rationality, Culture and Structure, edited by Mark Irving Lichbach, and Alan S. Zuckerman, pp. 277-310. New York: Cambridge University Press.