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

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* 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

2. Determinism, Probabilism, Randomness, and Causal Explanations

The integration of cross-case and within-case theorizing in causal explanations is to be

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 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 medery

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

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 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.

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 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

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

(or

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

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 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.

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.

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

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 li7444 1(i)-0.1(kel)6.1(o[-0.1(7w)5(ae)6.1(ue Xfi)-6)-0.1(7ay)-5.7(Y. Whi)-4.6(e0)5.1it on-suorostc rty.1(u)-1-5

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

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 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.

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

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.

Goldstone, Jack A. (1997): Methodological Issu

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.

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):987-1

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.