

**OPPOSITION IN TRANSITION. DOES UNITY BRING DICTATORS DOWN?
A QCA RE-ANALYSIS OF HOWARD&ROESSLER'S STUDY OF LIBERALIZING
ELECTORAL OUTCOMES***

CARSTEN Q. SCHNEIDER

PhD Program Director, Assistant Professor
Department of Political Science
Central European University

Email: schneiderc@ceu.hu

http://www.personal.ceu.hu/departs/personal/Carsten_Schneider/

*I am indebted to Marc Morje Howard and Philipp G. Roessler
who generously provided me access to their data.

November 2007

Draft - Comments are welcome

Paper prepared for joined Panel of the Working Groups “Demokratieforschung” and “Methoden der Vergleichenden Politikwissenschaft”. This Panel will be held at the Conference “Stand und Zukunft der Vergleichenden Politikwissenschaft”, organized by different sections of the Deutsche Vereinigung für Politische Wissenschaft (DVPW), November 9th to 11th 2007, Delmenhorst, Germany.

Abstract

This paper investigates the question: *Can QCA help to generate additional analytical insights when applied to a data set that already has been care- and skillfully analyzed with an appropriate statistical technique?* In order to answer this methodological question, I use recently published data on explaining the occurrence of ‘liberalizing electoral outcome’ (Howard & Roessler, 2006), a specific form of political regime change, which receives growing attention in the literature. I provide arguments and evidence that the application of QCA can uncover analytically relevant data patterns that would remain hidden if only standard qualitative and quantitative data analysis techniques were used. As a consequence, in many research situations, researchers could and should consider the use of QCA in order to increase inferential leverage.

1 Introduction

Over the last years, the social scientific discipline is living through a period of revived interest in empirical methodological issues. Perhaps starting with the book published by King, Keohane, & Verba, 1994, a series of other publications (Behnke, Gschwend, Schindler, & Schnapp, 2006, Brady & Collier, 2004, George & Bennett, 2005, Gerring, 2001, Goertz, 2006b, Goertz & Starr, 2003, Jahn, 2006, Mahoney & Rueschemeyer, 2003, Pierson, 2004), conferences, and new activities in professional associations (such as the Qualitative Methods section at APSA and panels at ECPR events) enriched our understanding of the diversity and relative strengths and weaknesses of different approaches to analyzing data in the social sciences. Within this broader methodological awareness, one specific methodological approach received particular attention. Charles Ragin's 'Qualitative Comparative Analysis' (QCA), first introduced in a book in the late 1980s (Ragin, 1987), then extended to fuzzy set QCA in a book publication (Ragin, 2000), and further elaborated in several journal articles (Ragin & Sonnett, 2004, Ragin, 2006a, Ragin, 2006b, Ragin, forthcoming, Rihoux & Ragin, 2007, and Schneider & Wagemann, 2007, just to mention a few) has been subject to several discussions addressing issues as diverse as its similarity and dissimilarity to already existing statistical techniques, such as logistic regression (Grendstad, 2007); its underlying assumptions compared to those in linear regression (Seawright, 2004 and 2005, Ragin, 2005, and Achen, 2005), the limitations of using dichotomous data (Goldthorpe, 1997 and, more general on dichotomies, Bollen, 1993 and Collier & Adcock, 1999), the process of model specification (Amenta & Poulsen, 1994), QCA's capacity to distinguish random from real data (Liebersohn, 2004, Marx, 2006), or the usefulness of set theory and the notions of necessary (Goertz & Starr, 2003) and sufficient conditions (Ragin, 2000) for social science research.

In this paper, I will not and cannot deal with all these important and fundamental methodological issues. Instead, I focus on the relatively modest question: *Can QCA help to generate additional analytical insights when applied to a data set that already has been care- and skillfully analyzed with an appropriate statistical technique?* If the answer to this question is no, then most of the above mentioned more fundamental issues do not matter much. For even if QCA passes all the hurdles of being an acceptable method for analyzing empirical social scientific data in principle, if this method does not yield, in practice, any new and interesting insights that usually remain undetected with currently available qualitative and quantitative data analysis techniques, then not much reason is left for (learning and) applying QCA. In other

words, the proof of the pudding ultimately rests in demonstrating QCA's usefulness in the research practice.

In order to shed light on the main question, I make use of data on the occurrence of so-called 'liberalizing electoral outcomes' (Howard & Roessler, 2006).

I proceed in the following way. I first reflect on the issue that the choice of method must primarily be guided by some (rough) theoretical expectations. Here I focus on those aspects that are distinctive for QCA compared to most existing data analysis tools, that is, the notions of necessity and sufficiency, equifinality, asymmetric and conjunctural causation. Then I present the analysis by Howard & Roessler, 2006 showing what kind of variables, data, and method they use, which cases they analyze, what their findings are, what kind of potentially useful information their method could still contribute and, ultimately, why a re-analysis of this particular study with QCA seems fruitful. In a third section, I then analyze the data with QCA. Following the standards of good practice, I first analyze the occurrence of the outcome LEO and then its non-occurrence. Within each of these two separate analyses, I first investigate whether there are any necessary conditions for the outcome and after that, what are the sufficient conditions for the outcome. During the QCA analysis, I focus on two substantive questions, both stemming from the perspective taken by the Howard and Roessler's article. First, I analyze the role of the explanatory variable put into the center of attention by these authors - the formation of an opposition coalition – once different assumptions about the complexity of the underlying causal relationships are made. And, second, I discuss in what sense case study of the Kenyan 2002 elections does and does not count as evidence for Howard and Roessler's claim that LEO is crucially driven by the formation of opposition coalitions. I conclude with a brief summary.

2 The Choice of Methods – Some Reflections

As such, no empirical data analysis method is superior to any other. Instead, the choice of the most adequate method is contingent on features of the specific research design. Characteristics of this research, such as the number of cases, the form and quality of the data, or the nature of the outcome to be explained play a crucial role in choosing the appropriate method. The most important rationale for the choice of a method is, however, (a) whether it enables researchers to learn more about the phenomenon of interest and, thus, ultimately, about the cases that display that phenomenon and (b) whether it allows for the development (and/or 'testing') of plausible, interesting, and theory-informed hunches¹ about the causes driving the outcome. Since each

¹ I use the term hunch rather than hypothesis to underline some of the shortcomings of most of our social science theories: their predictive power is low, the causal mechanisms opaque and alternative mechanisms are often not

empirical method, by default, makes assumptions about the underlying nature of causality and since different data analysis techniques make different assumptions, careful attention must be paid to both the exact content of the hunches to be analyzed and the assumptions embedded in the empirical method chosen. Both – theoretical claims and methodologically induced assumptions – must be as compatible as possible, something which Peter Hall, 2003 put into the apt expression of ‘aligning ontology and methodology’.

2.1 *Features of Causal Complexity*

I will focus on the issue of causal complexity, a crucial term when thinking about the fit between causal theories and hypotheses, on the one hand, and empirical methods, on the other. Causal complexity can be understood in different, not necessary mutually exclusive ways (Braumoeller, 2003). One form of causal complexity unfolds through the interplay of relevant factors *over time*. Aspects such as time, timing, sequencing, or feedback loops (Abbott, 2001, Pierson, 2004) are commonly dealt with in research based on the intensive analysis of few cases, often under the label of process tracing (George & Bennett, 2005). The kind of causal complexity I am focusing on in this paper is of different, *static*, nature, though. It goes under the label of conjunctural causation, equifinality, and asymmetry (Ragin, 1987 and Ragin, 2000, Lieberman, 1985). *Conjunctural causation* can be defined as the situation in which one single condition unfolds its impact on the outcome to be explained only when combined with one or more other condition(s).² *Equifinality*, in turn, adds to this complexity by allowing for the possibility that different conjunctions can produce the same outcome. *Causal asymmetry* is present when a given conjunction contributes to explaining the presence of an outcome Y but, at the same time, is irrelevant for explaining the absence of that outcome.

A powerful way of framing conjunctural causation, equifinality, and causal asymmetry consists in the concepts of *necessity* and *sufficiency*. In case a condition (or a combination of conditions) is found to be sufficient but not necessary for an outcome, we implicitly acknowledge that there must be other sufficient conditions for that outcome, i.e. we acknowledge that there is equifinality. We also acknowledge that the same (combination of) conditions have a different, if any, causal relation to explaining the *absence* of the outcome, that is, we are incorporating causal asymmetry.

Necessity and sufficiency, in turn, have a close relationship to set theory. If we say that a condition X is sufficient for Y, we can reformulate this claim and say that all elements (i.e., cases) with the characteristic X are a subset of all elements (cases) with the characteristic Y. For

mutually exclusive, the assumptions the theories are based on are highly contested, the hypotheses are not uniquely explaining the outcome, i.e. there are more than one plausible account (model).

instance, the claim (hypothesis) that democraticness (X) causes richness (Y) is set theoretic in nature. This statement formulates the expectation that the set of democratic countries (X) is a subset of the set of rich countries (Y). From such a set-theoretically framed relation we can claim – provided we also have theoretical arguments - that X is sufficient, but not necessary for Y. This has some important implications as to what kind of patterns we expect to find in the data. Most importantly, there are other conditions for richness than democraticness since not all rich countries are democratic - just think of Saudi Arabia. Hence, cases like Saudi Arabia, which are rich but not democratic, do not contradict the statement that democraticness is sufficient for richness, because that statement only generates expectations about the value of the outcome Y (richness) for those cases that are also democratic, not for those cases that are not democratic (such as Saudi Arabia). As a consequence, when we test whether democraticness is a sufficient condition for richness non-democratic cases might be rich or not rich - we simply do not have any expectations and thus do not have to care because for this sufficiency claim cases such as Saudi Arabia are logically irrelevant.

In contrast, if X is said to be necessary for Y, then the set of cases with X is a superset of the set of cases with Y (Ragin, 2000). To continue the hypothetical example from above, if we expect that democraticness is necessary for richness, then any case such as Saudi Arabia, which is not democratic but, nevertheless, rich, does violate the expectation that democraticness is necessary for richness. However, at the same time, it is no violation of this statement if we find cases that are democratic but not rich (e.g. India or Mongolia), for the statement of necessity only generates expectations on the value of the causal condition for those cases that display the outcome Y (rich cases). Therefore, whether or not cases without the outcome also display the necessary condition is simply irrelevant when testing for necessity.³

In order to define what causal complexity is, it is helpful to pin down what it is not: the assumption that causes exert their effect (i.e. that the independent variables cause the dependent variables) in a linear, additive, and unifinal manner. Ragin (2006a) summarizes the attempt at estimating each independent variable's effect on the outcome under the label 'net effect thinking'. These are the assumptions based on which those standard statistical techniques operate that are most frequently employed in the current comparative social scientific literature.⁴

² In the framework of statistical analysis, this is usually called 'interaction effect', which is similar, but not identical to an 'intersection', the adequate term in set theory based methods such as QCA.

³ As a consequence, when testing for necessity, one must select only cases that show the outcome to be explained. Such a case selection violates two deeply engrained and often boldly stated textbook wisdoms: one is selecting on the dependent variable and there is no variance on that variable. Obviously, such a practice is a 'sin' if the aim of a study is to understand variation and if the methods used are designed to explain variation by means of correlations. However, not all studies share these aims, such as, for instance, the analysis of necessary (but not sufficient) conditions, which is common in qualitative (small N) research.

⁴ Obviously, there are more advanced statistical techniques that circumvent many, but not all of these assumptions, such as Structural Equation Models or Generalized Linear Models. These techniques, however, require research conditions that more often than not are absent in macro-comparative social research, most prominently the relatively low number of cases.

Data analyses techniques based on this epistemology are very powerful tools for summarizing complex data in a parsimonious way.

As such, neither the assumption of causal simplicity nor that of complexity can claim general superiority.⁵ Both have their strengths and weaknesses. Assuming simplicity allows for deriving parsimonious models from rather complex data while the assumption of complexity usually enables the researcher to pay more tribute to different classes of cases within their population – both valuable aims of social inquiry. On the downside, the methodologically induced assumption of simplicity runs the risk of generating over-simplified representations that are not only very much detached from the cases and data patterns that underlie the analysis, but which often also just present caricatures of the theories they claim to test (Munck, 2001). In turn, the starting assumption of complexity runs the risk of individualizing each and every single case without much progress towards generalization and with significant difficulties in theorizing (even *ex post*) that empirical complexity. The empirical analyses below exemplify these strengths and weaknesses.

From all that, one should derive the conclusion that the adequacy of one or the other epistemology depends on the theories and hunches about the nature of the causal phenomenon observed and the research aim. If theories and/or expectations indicate that one can safely assume that one and the same phenomenon that is observed across many cases (say, the presence of democratic political regimes around the world) is indeed the product of one single condition, then one should employ methods that assume unifinality, additivity, and maybe also linearity. A wide range of (sophisticated) statistical data analysis techniques is available for that. If, instead, one has indications that the subject under study is better understood in terms of necessary and sufficient conditions and, by virtue of that, by equifinality, conjunctural causation, and asymmetry, that is, by causal complexity, then the choice of method should be different. Likewise, if the research interest rests more in assessing the causal role of (single) variables rather than in explaining (groups of) cases, then statistical techniques are the right choice, if not, the choice of method should probably be different.

2.2 Features of QCA

Since necessity and sufficiency can be rooted in set theory and formal algebra and since these concepts can adequately be expressed with the notation rules of Boolean algebra, any method based on set theory and Boolean algebra seems to be a promising tool for dealing with causal complexity. Qualitative Comparative Analysis (QCA) and its different variants (most importantly fuzzy set QCA (fsQCA) but also multi value QCA (mvQCA)) is such a method.

All QCA variants make use of the formal notation and operations of Boolean algebra and the data these methods process is formulated in terms of membership of cases in sets. As a consequence, all QCA variants generate findings that are characterized by precisely the features of complex causal statements outlined above. A hypothetical example shall clarify this point. Let A, B, C, D, and E be the causal conditions and Y the outcome. Furthermore, let capital letters denote the presence of a condition or outcome and small letters their negation; the + signs indicates the logical OR (the union of the two sets to the left and the right of this sign); the * sign indicates the logical AND (the intersection of the two sets to the left and right of this sign); and the \rightarrow sign indicates that the set of conditions to its left implies the set to its right, which is usually interpreted in terms of sufficiency (the set of conditions to the left are sufficient for the outcome to the right).⁶

Let the outcome of a QCA analysis look as follows.⁷

$$\begin{array}{lll} \text{For the occurrence of Y:} & A*D + C*E + a*B & \rightarrow Y \\ \text{For the occurrence of non-Y:} & B*D + a*c & \rightarrow y \end{array}$$

Such a solution formula contains all the elements of causal complexity as defined in this paper. First, it is equifinal, as is indicated by the logical operator OR. The outcome Y is produced through the condition A*D OR the condition C*E OR the condition a*B. Second, it is conjunctural, as is indicated by the logical operator AND. For instance, it is not A alone that produces Y but only when it is combined with condition D. Condition A, thus, is an INUS conditions because it is an *insufficient but necessary part of a condition which is itself unnecessary but sufficient for the result* (Mackie, 1974: 62; Goertz, 2003a: 68).⁸ Furthermore, under certain circumstances, that is, when combined with the condition B, it is the *absence* of condition A that contributes to producing the outcome Y. Finally, the solution is asymmetric because the solution formula for explaining the outcome Y is not simply the opposite, or negation, of the solution formula for explaining the non-occurrence of Y (denoted as y).

As mentioned, QCA is based on set-theoretic notions. This is important to note because most verbally formulated social science theories (i.e. all but purely formal ‘theories’) are almost exclusively of a set theoretic nature. Just recall the above example of the verbal statement that ‘democracies are rich’, which can be translated as ‘democratic case area a subset of rich countries’. As shown, the set theoretic nature of verbally defined relations between causally

⁵ An increasing number of scholars, though, argues that the burden of proof should be on the shoulders of those who assume causal simplicity and not on those who assume causal complexity (Coppedge, 2002: 3).

⁶ See Ragin, 1987 and 2000, Ragin & Rihoux, 2004, Schneider & Grofman, 2006, and Schneider & Wagemann, 2007 for notation in QCA and potential sources of confusion.

⁷ Contrary to frequently formulated critiques, QCA does not require a deterministic understanding of causality, nor does the data have to be dichotomous when fsQCA is used. Hence, the following solution formulas can be generated based on non-dichotomous data and they might not explain 100% of the outcome. See Ragin, 1987 and Schneider & Wagemann, 2007 for an extensive description of the QCA algorithm that leads from the ‘raw’ data to such solution formulas.

⁸ Such INUS conditions are particularly, if not impossible, to detect with standard statistical techniques.

relevant conditions (e.g. democracy) and the outcome (e.g. richness) strongly suggest the use of the concepts of necessity and sufficiency, and, by virtue of that, the notion of causal complexity in terms of asymmetric, equifinal, and conjunctural relations.⁹

Standard statistical techniques, such as regression analysis, are not designed for dealing with the type of causal complexity outlined above. To start with, the data handled with these techniques is not of set theoretic nature. This means that it is hardly possible to interpret the results in terms of necessity and sufficiency. In addition, the fact that many of these techniques rely on correlations instead of set relations (as QCA does) for relating variables also means that regression techniques assume symmetric causal relations between the independent variables and the dependent variable.¹⁰ Hence, the implicit assumption of symmetry is built into regression techniques through the use of the most common correlation coefficients, which are all symmetric measures.¹¹ Some recent attempts to develop statistical tools for dealing with the kind of set theory based causal complexity show (a) that, in principle, is possible to come up with such statistical procedures but also that such procedures (b) will look much different from current dominant practices and (c) will not escape the need for a (very) large N (see Braumoeller, 2003, Braumoeller & Goertz, 2003, Eliason & Stryker, 2005).

Further arguments for considering the application of QCA are a medium number of cases to be compared (20-60); a stronger interest of the researcher in understanding the cases and the reasons why some of them display the outcome while others do not, without, at the same time, losing sight of more general (and generalizable) patterns as it often happens in classical small N comparisons. QCA, if executed properly, can combine inductive and deductive elements and it can rest on both hypothesis 'testing'¹² and hypothesis formulating parts.

To sum up, any data analysis technique rests on a set of assumptions. Thus, the choice of method predetermines the kind of empirical findings one can generate with it. From this follows that the choice of methods should be adapted both to the theoretical propositions at hand, and the analytic aims and research strategies pursued. If theory indicates the presence of causal complexity in terms of necessity and/or sufficiency, if in-depth case knowledge (for some or all cases) exists, if the number of cases is not large, and if case studies are planned in order to foster

⁹ See Goertz (2003b) who counts more than 150 hypotheses about necessary conditions formulated in recent social science literature, many of which do not explicitly use the term necessity but mean it while others do say necessity but do not mean it.

¹⁰ This is expressed through the use of the =-sign in regression equations. See Schneider & Grofman, 2006 on that point.

¹¹ Also, think of another instance of assumed symmetry which is expressed by the fact that in logistic regression the probability for the non-occurrence of the outcome is simply calculated as 1 minus the occurrence of the outcome. There are, of course, asymmetric statistical measures of association, but the point is that they are not (yet) used in mainstream social science statistical data analysis.

¹² In QCA, hypothesis testing cannot follow the same logic as in inferential statistics for the simple reason that QCA is not an inferential statistic and, thus, does not make use of the notion of probability theory and distributions. This, however, does not mean that hypotheses cannot be tested in QCA at all. For some ideas of how to do this, see Schneider & Wagemann, 2007, 118-122.

the causal inference, then the use of QCA might be as justifiable as the use small N comparisons based on Mill's methods and/or the use of regression analysis.

3 Howard & Roessler, 2006: Liberalizing Electoral Outcome

A recently published article by Marc Howard and Philip Roessler shows virtually all characteristics that warrant the application of QCA. Yet, the authors use only logistic regression combined with a single case narrative. The question I explore in the following is: Can we learn more from the data if QCA is applied to it?

In this section, I provide the basic features of the study by Howard & Roessler, 2006 and present their main findings. I also present some additional analytic perspectives on the data that provide further arguments for why the application of QCA to this particular data set can potentially yield additional insights on the topic of liberalizing electoral outcomes. The analysis of this data has been published in an article in the *American Journal of Political Science*, one of the leading journals in the profession, and a conference paper version of it has won the 2006 Best Convention Paper Award of APSA's 'Comparative Democratization' section. Hence, the scientific and analytic merits of this work are beyond doubt. In fact, I have chosen this article precisely because it is very good. In case it can be demonstrated that the application of a new method (QCA) to a data set already well-analyzed with statistical techniques (logistic regression) followed by an insightful case study (see below) yields additional analytically relevant insights, then this should be even more so the case in less well executed mixed-methods designs (Lieberman, 2005, for a critique see Rohlfing, forthcoming). From all this follows that the aim of this re-analysis is not to argue that the authors could have done better in their statistical and case study analyses. Instead, the re-analysis should bring to the fore how different methods predetermine certain views on the underlying evidence.

3.1 Features and Findings of the Study

The main research question the authors try to answer is: Under which conditions do elections in a competitive authoritarian regime turn out to be of a liberalizing electoral outcome (LEO)? The dependent variable is the occurrence of LEO, operationalized as the a simultaneous increase of 1 point in Freedom House's scale of individual freedom (see Freedom House, 2006) and a three point increase in Jagers' and Gurr's Polity index (Jagers & Gurr, 1995) "in the year after the election, compared to the year before" (p. 369).

The unit of analysis in Howard and Roessler's paper are national elections in competitive authoritarian regimes (Levitzky & Way, 2002). All together, they include 50 of these elections

(cases) in their study, which took place in 31 different countries around the globe between 1990 and 2002.

In their cross-national study they test nine hypotheses. The variables are listed in Table 1. The dependent variable LEO and 5 out of 9 independent variables are operationalized as dichotomies. For the remaining 4 non-dichotomous independent variables arguments can be provided that a dichotomous representation corresponds to the notion of the underlying concepts to be measured at least to the same extent as the interval level indicators do. More than half of their variables are operationalized as dichotomies. In Appendix 1, I describe each variable and provide detailed rationales for dichotomization.

3.2 Cross-National Findings

For their cross-national analysis, Howard & Roessler, 2006 use logistic regression, a statistical technique designed for analyzing dichotomous dependent variables. This technique estimates for each independent variable the marginal effects on the probability that the outcome occurs.

Howard & Roessler, 2006 find that out of their 9 independent variables none of the structural factors shows a statistically significant effect on the outcome LEO. Only three actor-based and strategic factors (as opposed to contextual factors) are significant: incumbent turnover and opposition mobilization are significant at the 5% level, while the factor opposition coalition is significant at the 1% level.

By using a very useful graphical display (p. 376) they demonstrate that opposition coalition has by far the largest substantive impact on the occurrence of LEO. Everything else being constant, the formation of an opposition coalition increases the probability of the occurrence of a liberalizing electoral outcome by over 80%. This means that in ‘cases’ that are described by the average value on the non-dichotomous variables and by the minimum value of the dichotomous variables, the fact that an opposition coalition is formed increases the likelihood of the occurrence of LEO by 80%. In the same ‘fictitious case’¹³ incumbent turnover does increase the probability by only less than 40% and opposition coalition, the third statistically significant factor, by only about 20% with each standard deviation increase from their respective means.

¹³ If the data is a random sample from a well-defined population, one can infer that the effect of opposition coalition, on average and in all (also unobserved) cases, is the one found by Howard and Roessler. If the cases studied are not a random sample and/or if there is no well-defined population, then one should be careful with such an interpretation. Below I will discuss the role of limited diversity – that is logically possible combinations of cases for which no empirical evidence is at hand, or, unobserved cases – for drawing causal inference.

Table 1: Explaining Liberalizing Electoral Outcomes: Logistic Regression Analysis

	<i>Coefficients and Standard Errors</i>
Opposition Coalition	7.72** (3.03)
Opposition Mobilization	.91* (.40)
Incumbent Turnover	3.15* (1.51)
Economic Growth	.33 (.24)
Foreign Direct Investment	-.10 (.31)
Foreign Aid	.008 (.02)
Parliamentarism	-3.07 (2.18)
Regime Openness	1.04 (.99)
Prior Liberalizing Change	-1.38 (1.73)
Constant	-1.33 (5.24)
N	50

Note: Table entries are regression coefficients, with standard errors in parentheses. The dependent variable is whether or not a country experienced a liberalizing electoral outcome (LEO).
*p < .05; **p = .01

All in all, these results suggest that opposition coalition is the primary cause for LEO. Imagine Western democracy promoters who read this article: are they well advised to put their money and effort into fostering coalitions between diverse opposition movements in autocratic regimes? In the following I provide arguments and evidence that this conclusion rests on many untested assumptions and is, thus, hazardous.

3.3 Further cross-national information

Howard & Roessler, 2006 use their cross-national statistical analysis in order to show broad patterns of association and delegate the task of paying attention to specific cases to subsequent, non-comparative, in-depth case studies. This is a perfectly valid and commonly practiced approach for drawing causal inference. It, nevertheless, does not rule out that already during the cross-national analysis more attention to specific (groups of) cases could and should be paid. In the next chapter, I demonstrate that QCA is a more adequate cross-case method for sticking closer to case evidence than are common statistical techniques with their reliance on deviations from sample means. Before that, I now briefly highlight some possible ways of putting cases more into the center of attention even when using logistic regression.

First, in logistic regression the dependent variable is a dichotomy. It is therefore easy to calculate the percentage of cases that are correctly predicted by the estimated model. Here, three the pieces of information are relevant: the percentage of (in)correctly predicted cases (a) with LEO; (b) without LEO; and (c) percentage of overall correctly predicted cases.

Table 2: Classification Table, Logistic Regression

		LEO Predicted		Percentage Correct
		0	1	
LEO Observed	0	33	2	94.3
	1	4	11	73.3
Overall Percentage				88.0

Table 2 cross-tabulates the observed number of LEO with the number of LEO predicted by Howard and Roessler’s logistic regression model. If the model correctly predicted all cases, cases would only fall into the 0;0 and the 1;1 cells of the 2x2 matrix. This, however, is not the case. Out of the 15 cases with LEO, the model only correctly predicts 11. Put differently, over 25% of cases with LEO are not predicted by their model. The model works somewhat better for the non-occurrence of LEO. Here 33 out of 35 cases, or, about 94%, are correctly predicted. The overall percentage of correctly predicted cases (i.e. both LEO and Non-LEO) is 88%.

Second, the focus on case can be further increased by specifying *which* cases are incorrectly predicted by the model.

Table 3: Classification Table with Country Names, Logistic Regression

		LEO Predicted		Percentage Correct
		0	1	
LEO Observed	0	33	<i>Ivory Coast 2000</i> <i>Mexico 1994</i>	94.3
	1	<i>Armenia 1998</i> <i>Guatemala 1995</i> <i>Peru 2000</i> <i>Thailand 1992</i>	11	73.3
Overall Percentage				88.0

Table 3 reproduces the information contained in Table 2 and adds the country names in the cells for incorrectly predicted cases. The four cases with LEO that are not captured by the model are Armenia 1998, Guatemala 1995, Peru 2000, and Thailand 1992. In contrast, the model wrongly predicts the occurrence of LEO in Ivory Coast 2000 and Mexico 1994. This information on cases could be used to select cases for case studies and to further improve the model by looking at additional factors that are relevant for understanding why these cases - and not others - are not

predicted by the model.¹⁴ One important fact about these 6 cases that are mis-classified by the cross-national logistic regression model raises doubts on whether Howard and Roessler's conclusion is warranted that the formation of an opposition coalition truly has an outstanding causal impact on LEO: none of these mis-classified cases shows the formation of an opposition coalition (see truth table in Table 5). This implies two things: (a) their model predicts LEO for cases *without* opposition coalition that, in fact, do not have LEO (Ivory Coast 2000 and Mexico 1994) and (b) their model fails to predict LEO for cases that have experienced LEO but which did *not* show an opposition coalition (Armenia 1998, Guatemala 1995, Peru 2000, Thailand 1992).

Further evidence casting doubts on Howard and Roessler's praise of opposition coalition as the single most important causal factor for LEO is the following simple 2x2 table.

Table 4: Cross-Tabulation: LEO – Opposition Coalition

		OPPOSITION COALITION		Total
		0	1	
LEO	0	(a) 32	(b) 3	35
	1	(c) 7	(d) 8	15
Total		39	11	50

Table 4 shows a simple cross-tabulation of the dependent variable LEO and the main independent variable opposition coalition. We should not be so much concerned that the bivariate correlation between the two variables, while statistically significant, is rather modest with 0.495, for one of the reasons of running a multivariate regression is precisely to tease out the marginal effect of one variable net of the effects of all other relevant variables. What is more interesting to note from the perspective of causal complexity - and here especially the aspect of asymmetric causality - is the following fact: among the 15 cases in which the outcome LEO is observed, the distribution of cases with and without opposition coalition is almost even. Only 8 out of 15 cases with LEO also show opposition coalitions while 7 out of 15 manage to achieve LEO without any signs of opposition coalition. Such a 50-50 situation among cases *with* LEO implies that the relationship Howard and Roessler find between opposition coalition and LEO is mainly driven by the distribution of cases in the row of Table 4 that contains cases *without* LEO. Here the picture is clear: 33 out of 35 cases without LEO do not experience opposition coalition while only 2 out of 35 do have opposition coalition but not LEO.¹⁵

From a viewpoint of symmetric causality and from that of correlational analysis, there is nothing suspicious about the data pattern found in Table 4. The question and answer on what *fosters* LEO are identical to those on what *hinders* LEO. This is why cell (a) in Table 4 in which cases

¹⁴ See Seawright & Gerring, 2005 on different statistical rationales for selecting cases for subsequent case studies.

are gathered with both the independent and the dependent variable are *absent* (0) can count as evidence that the *presence* of the independent variable fosters the *presence* of the dependent variable. Whether or not the 0,0 cell should be used for drawing causal inference is subject to philosophical debates held under the label of the black raven paradox. For an application of this debate in the social sciences, see Seawright, 2002a, Seawright, 2002b, Braumoeller & Goertz, 2002, and Clarke, 2002).

If, however, we subscribe to the idea of asymmetric causal relationships, the notion of necessary and sufficient conditions, and set theoretic relations as a means of relating variables with each other, then exactly cell (a) in which by far the most cases fall and which drives the correlation between LEO and opposition coalition, becomes logically irrelevant. This 0;0 cell can contain as many cases as one is pleased in any test for necessity and sufficiency for the occurrence of LEO – it does not alter the interpretation of the set relation at all. What becomes relevant, instead is, first, the relation between cells (d) and (c) for the test whether opposition coalition is necessary for LEO¹⁶ and, second, the relation between cells (d) and (b) for the test whether it is sufficient¹⁷ (Ragin, 2006b, Goertz, 2006a).

To sum up, the inferences drawn by Howard and Roessler based on their methodological approach is valid, plausible, and insightful. By and large, and statistically speaking, the values for the variables opposition coalition and LEO are systematically related. Opposition coalition matters most – net of all the other factors – for the occurrence of LEO. This interpretation rests, however, on strong assumptions about the basic nature of the causal relationship and these assumptions are induced by the data analysis technique used rather than derived from the existing theory. In fact, some of their proto-theoretical arguments run counter to what they have to assume when performing their hypotheses tests with logistic regression. What their analysis does not show and what, in fact, is a blind spot induced by the method and its underlying assumptions, is that their model and, even more so, their pet variable, perform much better in explaining the non-occurrence of LEO rather than its occurrence. While this difference is nonsensical under the assumption of causal symmetry, it is far from trivial once we assume asymmetrical causal processes. More broadly speaking, a closer look at the data provides evidence that a different perspective on the data, one that is working under the assumption of causal complexity, might reveal a different and more differentiated causal picture. Below I will follow up on this point when I analyze the data with QCA.

¹⁵ This shows how crucial the selection of non-cases, i.e. cases in which LEO did *not* occur, is for analysing which factors contribute to the *occurrence* of LEO (on the intricacies of selecting negative cases and the impact on drawing causal inference, see Mahoney & Goertz, 2004).

¹⁶ If all cases fall into cell (b) and no case in cell (d), then the condition could be interpreted as necessary for the outcome, with cells (a) and (c) being irrelevant for this test of necessity.

¹⁷ If all cases fall into cell (b) and no case in cell (a), then the condition could be interpreted as sufficient for the outcome, with cells (d) and (c) being irrelevant for this test of sufficiency.

3.4 Rationale for Reanalysis

The study by Howard & Roessler, 2006 is a fine piece of careful, multi-method nested analysis published in one of the most respectable journals of the Political Science discipline. From this, it is clear that any attempt to reanalyze their data does not aim at showing that Howard & Roessler, 2006 are wrong in how they applied their methods, nor that their findings are irrelevant or flawed as such. Their study has a number of characteristics which suggest the application of QCA to their cross-national data.

First, in several occasions, the authors express the expectation that the effect of single variables will most likely depend on the value of other variables. Even if the authors come short of formulating a full-blown interactive theory – something missing in most social science analyses at large – their expectations of the underlying causal mechanisms make the incorporation of some conjunctural terms into the data analysis a reasonable choice. As already mentioned, based on just 50 cases the inclusion of (several) interaction terms into their logistic regression model is not feasible. With QCA, however, one is able to look at conjunctions.

Second, the majority of the concepts used are operationalized in a dichotomous way. QCA is designed for analyzing dichotomous data. Third, their number of cases is moderate. Although it is by now common practice to run (logistic) regression analyses with 50, or often even fewer cases, it is clear that the power of statistical data analysis techniques cannot be brought to its full bear with such a low number of cases. QCA is not based on the same notions and concepts from inferential statistics like logistic regression and, thus, many of the well-known assumptions that need to be met in order to run statistical analyses and to draw meaningful inference from a sample to a (hopefully well-defined) universe of cases, do not apply. In QCA, inference beyond the cases analyzed is often not the aim because all relevant cases are included in the study. If not, inference beyond the sample is based on substantive argumentation rather than probability theory.

Fourth, a medium-sized number of cases allows for acquiring more in-depth case knowledge and to focus more intensively on single (groups of) cases during the analysis and during the interpretation of the results rather than having them disappear behind coefficients and (often moderate) explanatory power of the overall model. In fact, the authors themselves acknowledge that “[...] statistical analysis is useful in drawing out broad patterns across countries, it has limited utility for specifying causal processes in more detail”(Howard & Roessler, 2006: 376). From this, one can derive the argument that the more a cross-case method yields case-sensitive results, the more it should be useful for subsequent case-intensive, ‘model-testing’ (Lieberman, 2005) analysis, simply because the analytic results from such cross-case methods provide for more detailed information as to which factors one needs to look at. QCA’s premium given to conjunctural causation and equifinality, together with its grounding in notions of set theory and

Boolean algebra allows for an unequivocal identification of each single case to one (or more) of the different paths leading to the outcome.¹⁸

Fifth, Howard & Roessler, 2006 make the point only briefly in their conclusions, but it is obviously only a small step from their analytic finding to policy advice: if, indeed, the formation of opposition coalition matters very much for bringing about LEO, then democracy promoters should learn this lesson and direct their funds and efforts at this particular issue. However, more often than not, the *context* in which certain policies are implemented is crucial for the success of that policy. For this reason, the ‘net effect thinking’ (Ragin, 2006a) that is inbuilt into statistical techniques might not generate the most useful policy advice. At the end, the 80% increase in the probability that LEO occurs if the opposition manages to agree on a coalition refers to a specific but also purely hypothetical case: it is the hypothetical country with average opposition mobilization, no incumbent turnover, average economic growth, average foreign direct investment, and average foreign aid, with a presidential system, average regime openness, and no prior liberalizing changes (Howard & Roessler, 2006: 375). Not only are there many more logically possible types of cases, but the ‘case’ for which the 80% increase in probability is calculated does neither exist in their data, nor, most likely, in reality. This is the phenomenon of *limited diversity* (Ragin, 2000 and also Schneider & Wagemann, 2007), which is omnipresent in comparative social research and which has several implications for drawing causal inferences. Therefore a more policy relevant approach to studying the effect of opposition coalition formation on the occurrence of LEO seems to be to observe how this factor combines with other analytically relevant factors in all those conjunctions for which empirical evidence is at hand, rather than calculating its average effect and holding everything else constant. More likely than not, opposition coalition formation, *sometimes* might be beneficial, but not always and more likely than not, it will not be beneficial on its own but only in (different) combinations with other factors. Again, the assumption of equifinal and conjunctural causation favored in QCA generates results that do not render themselves to be interpreted in terms of net effects of single variables regardless of the context (that is the values of the other factors) in which it is inserted. A final reason for reanalyzing the data with QCA is the laudable attempt by Howard & Roessler, 2006 to strengthen the inference drawn from their cross-case analysis with evidence generated by a case study. In their discussion of Kenya’s election in 2002 they credibly show that the formation of an opposition coalition, indeed, was decisive for bringing about a LEO. In their plausible account of what was going on in Kenya in 2002 they refer to some, but not all independent variables from their cross-national model. Following the authors, it seems to be crucially important for understanding the occurrence of LEO in Kenya 2002 that – unlike in the

¹⁸ The increased attention to outliers in regression analysis and their use for improving causal inference is certainly a welcome development.

elections in 1992 and 1997 - the incumbent president Moi did not run for elections again (condition 'incumbent turnover' = 1), that the governing party KANU lacked financial means for increasing pressure (p. 379), partly stemming from lack of economic growth (condition 'economic growth' = 0), partly from lack of foreign aid (p. 379, condition 'foreign aid' = 0).¹⁹ While the causal mechanism they depict is plausible, two questions as to how that mechanism in Kenya 2002 travels well to other cases – even to those that seem to confirm their cross-national finding by displaying both an opposition coalition and LEO – emerge. Both questions stem from the unclear status of how representative the Kenyan story is for other cases of successful LEO. As will become clear, the representation of the data in the form of a truth table – a key analytic device in QCA-based research – can help assess the question on how much we learn about other cases when analyzing Kenya 2002.

4 QCA Analysis of Liberalizing Electoral Outcomes

As should have become clear, QCA is a method that tries to bridge the gap between primarily case-oriented and dominantly variable-oriented approaches to understanding social phenomena. Hence, using QCA, there are many different perspectives from which one could analyze the data provided by Howard & Roessler, 2006. If the main aim of this article was purely substantive and consisted in generating new substantive insights on the (non-)occurrence of LEO and on some of the 50 cases under study, I would have to take all these different perspectives. That would require (lengthy and time consuming) redefinitions of cases, concepts, measurements, and models. The main aim of this paper, though, is a methodological one and the empirical data largely serves presentational purposes. Hence, in order to keep the task in a manageable size, I will follow the authors' main line of argument and focus on the question whether opposition coalition is, indeed, such a decisive factor for LEO and, if so, whether we can generate more knowledge on the question under which conditions this factor is relevant and under which ones it is not. In addition to observing the role of opposition coalition once we drop the assumption of causal simplicity, I focus on where the case of the Kenyan elections in 2002 is located in the equifinal and conjunctural paths towards LEO that are generated by using QCA and in which sense this can (not) confirm their cross-case evidence.

¹⁹ One can express a simplified version of their account of Kenya 2002 in Boolean terms as follows: $OPPCOAL * INCTURN * ecgrowth * foraid$
Below I will use Boolean logic to intersect their narrative with the results generated by QCA to show what is not said about the occurrence of LEO when just looking at Kenya 2002.

4.1 *Truth Table and Limited Diversity*

The basis for any QCA analysis is the information contained in a truth table. Table 5 presents the data by Howard & Roessler, 2006 in such a structure.²⁰ Unlike in common data matrices, in truth tables, the rows are not empirically observed cases (such as ‘elections’ in the present example) but all 2^k logically possible combinations of the k conditions specified. Each empirically observed case can then be attributed to one truth table row. Such a truth table is useful for many things. It not only shows which of the empirically observed cases fall into the same truth table row and are, thus, analytically identical, but it also makes the amount of limited diversity straightforwardly visible. All of the 2^k truth table rows that do not describe any empirical case are logical remainders. All logical remainders taken together describe the extent of limited diversity (Ragin, 2000, Schneider & Wagemann, 2006).

For instance, from Table 5 we learn that only 39 out of the 512 logically possible combinations bear empirical evidence. In other words, for $512 - 39 = 473$ logically possible combinations we do not know whether or not the outcome LEO would be produced because these logically possible ‘cases’ do not exist in our data. Limited diversity is the rule rather than the exception in comparative social research based on observational data – regardless of the data analysis techniques used (Ragin, 2000). Hence, in one way or the other this missing information influences our possibilities for drawing causal inference. While no method can circumvent the problem of limited diversity (for instance, by simply inventing these cases based on a set of assumptions), the best way of tackling this methodological problem is to be conscious about its existence and to have a set of possible treatments of these logical remainders among which the researcher can deliberately choose. QCA performs well on this criterion compared to virtually all other currently existing empirical research methods. In contrast, the standard interpretation of multivariate statistical models which usually starts with the ‘*ceteris paribus*’ or ‘holding everything else constant’ clause²¹ runs the danger of insinuating at least at the semantic level that information on all logically possible types of cases was available when estimating the coefficients while, in fact, it was not.²² In the present example, there are 244 logically possible combinations that involve opposition coalition and which do not empirically exist. There are, for instance, no cases that combine opposition coalition with economic growth ad regime openness, or with parliamentarism and prior liberalization.

²⁰ See Appendix 1 for information on how Howard and Roessler’s non-dichotomous variables were transformed into dichotomies.

²¹ For a critique at the idea of *ceteris paribus*, see Woodward, 2002 and Kittel, 2006.

²² The assumption of a randoml drawn sample on which regression analysis rests, mitigates the problem of limited diversity to some extent. In practice, the problems are (a) that often the sample is not a sample but represents the population, (b) it is not a random sample, or (c) even if it is randomly drawn, the population itself, to which one infers might be limited in its diversity without the researcher being aware of this fact.

Table 5: Truth Table – LEO and 9 Conditions

Row #	Opp. coalition	Opp.mobi- lization	Incumb. turnover	Econ. Growth	FDI	Foreign Aid	Parlia- menta	Regime openness	Prior Lib.	LEO	Empirical Cases
1	0	0	0	1	0	1	1	0	1	0	ALB96
2	0	1	0	1	0	1	1	0	0	0	ALB97
3	0	0	0	1	0	1	1	0	0	0	ALB01
4	0	0	0	1	0	1	0	0	1	0	ARM96,T OG98
5	0	0	1	1	0	1	0	0	0	1	ARM98
6	0	0	0	1	0	1	0	0	0	0	CEN99,G AB98,GUI 98, SRI94,UG A01
7	0	0	0	0	1	0	0	0	0	0	CHA01
8	0	0	1	0	0	1	0	0	0	0	IVO00,GN B99
9	0	0	0	1	0	0	0	0	0	0	CRO97
10	1	0	1	0	1	0	1	0	0	1	CRO00
11	0	0	1	1	0	0	0	0	0	1	DOM96,G UA95
12	1	0	0	1	0	1	0	0	1	1	GHA96
13	1	0	1	1	0	1	0	0	1	1	GHA00
14	0	0	1	1	0	1	0	0	1	0	GUA90
15	0	0	0	1	0	0	1	0	0	0	INS92,MA L90,NEP9 9
16	0	1	1	0	0	0	1	0	1	1	INS99
17	0	0	0	1	0	0	0	1	0	0	IRN93
18	0	0	1	1	0	0	0	1	0	0	IRN97
19	1	0	1	0	0	0	0	0	0	1	KEN02
20	0	0	0	1	1	0	1	0	1	0	MAL95
21	1	0	0	0	1	0	1	0	0	0	MAL99
22	0	0	1	0	0	0	0	0	0	0	MEX94,R US00
23	0	0	0	1	0	0	1	0	1	0	NEP94
24	1	0	0	0	0	1	0	0	0	1	NIC90
25	0	0	0	1	1	0	0	0	0	0	PER95
26	0	1	0	0	0	0	0	0	0	1	PER00
27	0	1	1	0	0	0	0	0	1	1	PER01
28	1	0	0	0	0	0	0	0	1	0	ROM92
29	1	0	0	1	0	0	0	0	1	1	ROM96
30	0	0	0	0	0	0	0	0	0	0	RUS996
31	0	0	0	0	0	1	0	0	0	0	SEN93
32	1	0	0	1	0	1	0	0	0	1	SEN00
33	0	0	0	1	1	0	1	0	0	0	SIN91,SIN 97
34	0	0	0	1	0	0	0	0	1	0	SRI99
35	0	0	1	1	0	0	1	0	0	1	THI92
36	0	1	0	1	0	0	1	1	0	0	YGS96
37	1	1	0	0	0	1	0	1	1	1	YGS00
38	0	0	1	0	1	1	0	0	0	0	ZAM01
39	1	0	0	0	0	0	0	1	0	0	ZIM02
40 - ... 512										?	-

Another observation is that most cases are analytically different from each other, which can be seen from the fact that most rows with empirical evidence contain just one case each.²³ There are

²³ The truth table does not contain any contradictory rows (Ragin, 1987: 113-118). These are truth table rows which contain cases that are identical in their combination of conditions (this is why they are in one and the same truth

only 6 rows with more than one case. Among those, row # 6 stands out by covering a total of five cases (Central African Republic 1999, Gabon 1998, Guinea 1998, Sri Lanka 1994, Uganda 2001).

In principle, a truth table can be interpreted as the most complex answer to the question which factors are sufficient for the outcome, for each truth table row is a statement of sufficiency for LEO (rows with LEO = 1) and for non-LEO (rows with LEO = 0)). In the present example, however, it is better to resort to Boolean algebra and the Quine-McClusky algorithm for minimizing the complexity of the truth statements displayed in Table 5. This is what I will do in the following section.

4.2 Analysis of the occurrence of LEO

The standards of good practice in QCA (see Ragin & Rihoux, 2004, Schneider & Wagemann, 2007 and Wagemann & Schneider, 2007) dictate that for the analysis of a given phenomenon one needs to perform the following analytic steps. First of all, one needs to search for necessary conditions separate from the analysis of sufficient conditions, with the analysis of necessity coming first. Then for the analysis of sufficient conditions, one needs to generate the solution formula *without* making simplifying assumptions about logical remainders, then *with* simplifying assumptions about these logically possible but empirical unobserved cases, and finally with only *easy counterfactuals* (Ragin & Sonnett, 2004). Each of these separate analyses represents the information contained in a truth table in a different way, but these different solution formulas are not in logical contradiction to each other. In contrast, they are all logically equivalent and only vary in their degree of complexity: those analyses based on simplifying assumptions yield less complex and those without such simplifying assumptions yielding more complex solution terms. Finally, due to the notion of asymmetric causality, each of the above mentioned analytic steps needs to be performed separately for the occurrence of the outcome LEO and then for its non-occurrence.

Analysis of necessary conditions for LEO

Table 6 displays two pieces of information for each of the 9 conditions and their complements (that is, their negation indicated by a \sim -sign). The first column ‘consistency’ displays the degree to which the given condition is consistent with the statement of being a necessary condition for the outcome LEO. A value of 1 indicates a perfect consistency (all cases with LEO also display the condition), while lower values indicate an increasing deviation from this perfect set relationship. Put differently, a consistency score of 1 in a necessary condition test means that all

table), but which differ in the value of the outcome (i.e. some cases display the outcome LEO while some others do not).

cases with the outcome also show the necessary condition, but the inverse is not required to hold.²⁴ Needless to point out that such a perfect subset relation will yield only a modestly strong correlation. In general, fairly high consistency values should be applied for the test of necessity (see Schneider & Wagemann, 2007 for arguments on that). The second column ‘coverage’ yields a numerical value of the empirical importance of necessary conditions. The higher the coverage value the more empirically important is the necessary condition. Low coverage values, in turn, indicate that the condition under question is a trivially necessary one (Ragin, 2006b, Goertz, 2006a).²⁵ It is important to note that the interpretation of the coverage scores is only relevant for those conditions that pass the threshold of consistency, similar to the practice of only interpreting those coefficients in a regression analysis that pass a certain threshold of significance.

Table 6: Test of Necessity: Outcome LEO

Conditions tested:	Consistency	Coverage
opp._coa	0.53	n.r.
~opp._coa	0.47	n.r.
inc._tur	0.60	n.r.
~inc._tur	0.40	n.r.
Parliame	0.20	n.r.
~parliame	0.80	n.r.
prior_li	0.40	n.r.
~prior_li	0.60	n.r.
oppmob_d	0.27	n.r.
~oppmob_d	0.73	n.r.
ecgrow_d	0.53	n.r.
~ecgrow_d	0.47	n.r.
fdi_d	0.07	n.r.
~fdi_d	0.93	0.33
foraid_d	0.40	n.r.
~foraid_d	0.60	n.r.
regope_d	0.07	n.r.
~regope_d	0.93	0.31

n.r. = not relevant because consistency value too low

No condition is a 100% consistent necessary condition for LEO. Two conditions, however, display high consistency values: the *absence* of foreign direct investment and the *absence* of regime openness. In substantive terms, especially the second condition seems counter-intuitive. It seems difficult (but not impossible) to find theoretically grounded arguments why the absence of regime openness should be necessary for LEO and not, if anything, the presence of regime openness. A look at the coverage values of these two conditions shows, however, that they are both trivially necessary and therefore do not deserve any further substantive interpretation as necessary conditions. Trivialness of necessary conditions occurs, when, empirically, virtually all

²⁴ The more general formulation that is also applicable to fuzzy set QCA is that across all cases the scores in the condition X are higher than or equal to their scores in the outcome Y (Ragin, 2000). In short: $X_i \geq Y_i$.

²⁵ As an obvious example for a trivial necessary condition, think of oxigyn being necessary for the outbreak of war.

cases in the data set do show this trait.²⁶ And, in fact, the information in the truth table reveals that there is hardly any variation on those two variables: almost all cases score 0 on foreign direct investment and regime openness.²⁷

If we look at the main variable of Howard & Roessler, 2006, we see that opposition coalition shows a consistency value as a necessary condition for LEO of 0.53. Opposition coalition, thus, is far from necessary for producing LEO. This finding is very much in line with Howard and Roessler's careful interpretation of opposition coalition being far from determining the outcome LEO.²⁸

Analysis of sufficient conditions for LEO

The complete solution formulas for the occurrence of LEO can be found in Appendix 2. In order to facilitate the discussion that interests us most in this paper (Is the role of opposition coalition in bringing about LEO really that crucial once we allow for causal complexity?), Table 7 summarizes the results based on four criteria: (a) the number of paths towards the outcome; (b) the number of times, opposition coalition is sufficient on its own; (c) the number of times, opposition coalition is part of a sufficient path; (d) the number of times, the *negation* of opposition coalition forms part of a sufficient path. Each dimension will be applied to the three different solution formulas: the most complex based on no simplifying assumptions, the most parsimonious based on computer-generated simplifying assumptions and a medium-complex solution formula, based only on easy simplifying assumptions.

The first point to note is that the more parsimonious solutions contain a lower number of different paths towards the outcome. This is not surprising since the number of paths (and of conditions involved in each path) represents the very definition of parsimony. More interesting is the second row which shows that in none of the three different analyses of the truth table, opposition coalition is a sufficient condition on its own. This indicates that we cannot find empirical evidence for the claim that it is this factor that leads to LEO in any of the countries under study – not even if we allow the computer to make as many simplifying assumptions as possible. Instead, it needs to be combined with different other factors in order to exert its causal effect on the occurrence of LEO.

²⁶ Notice that if there was absolutely no variation, such fully trivial necessary conditions could be reinterpreted as scope conditions of an analysis (on scope conditions, see Walker & Cohen, 1985).

²⁷ Out of the 50 cases, there are 45 with regime openness = 0 and 42 with FDI = 0. A third variable without much variation is opposition mobilization (44 cases = 0). Of course, that lack of variation is partly due to the dichotomization of the ordinal/interval scale variables. It remains, however, an empirical fact that there are very few cases with substantively high prior liberalization and with FDI, respectively, and these few cases are those that received a score of 1 in my dichotomization.

²⁸ Below, we will see, however, that the *absence* of opposition coalitions can be interpreted as a necessary for the *absence* of LEO – logically speaking quite a different statement.

Table 7: Summary Sufficiency Analysis, Outcome LEO

Row		Solution formulas		
		Most complex Solution without simplifying assumptions	Intermediate Solution only with easy simplifying assumptions	Most parsimonious Solution with simplifying assumptions
1	Number of paths towards LEO	10	7 ^b	5 ^a
2	Number of times Opposition Coalition is sufficient on its own	-	-	-
3	Number of times Opposition Coalition is part of paths	6	3	3
4	Number of times Negation of opposition coalition is part of paths	4	1	-

^a: based on selection of 1 prime implicant out of 2 logically redundant

^b: based on selection of 3 prime implicants out of 9 logically redundant

As row three shows, the factor opposition coalition forms part of many of the different paths towards LEO. For instance, in the analysis that is based on not a single simplifying assumption and which, thus, is closest to the observable empirical information on the cases, opposition coalition appears in 6 out of 10 different paths. This indicates the influential role this factor plays in bringing about LEO, even though only as an INUS condition and thus together in a complex interplay with other conditions. However, other factors, such as incumbent turnover (see solution formulas in Appendix 2) appear almost as frequently in sufficient conjunctions for LEO as opposition coalition does. They could thus deserve the same analytic attention as opposition coalition gets.

What is more, as the first column in row 4 of Table 7 shows, in 4 out of 10 paths towards LEO it is the absence of opposition coalition that, combined with other factors, produces LEO. Even if we include only easy counterfactuals when minimizing the truth table information (row 4, column 2, in Table 7), we still obtain one path towards the outcome that contains the absence of opposition coalition as part of a sufficient conjunction. Only when we go for maximum parsimony by allowing the computer program to simulate outcome values for potentially 473 logical remainders do we get a result that features opposition coalition only in its presence as part of 3 out of 5 paths towards LEO.

The most parsimonious QCA solution formula, while also not displaying opposition coalition as the overriding causal force, at least does not contain the disturbing information that for certain cases the absence of opposition coalition is involved in bringing about LEO. This ‘nice’ solution formula is, however, based on a large amount of fully a-theoretical, computer algorithm generated assumptions about the empirically non-existing 473 logical remainders. More precisely, the software assumes the outcome of 1 for 306 out of 473 logical remainders cases. In other words, if, based on a QCA analysis of Howard and Roessler’s data we still want to conclude that forming an opposition crucially matters more than any other factor, then we have

to be aware that this claim rests on many weak (if not non-sensical) simplifying assumptions about logically possible types of cases.

In sum, from the evidence shown in Table 7, we can infer that opposition coalition plays an important role in understanding the occurrence of LEO – but from a QCA perspective this important role is not that it explains, on average and everything else being the same, more than other factors. Instead, its importance stems from the observation that most paths towards LEO – but not all – require the presence of this factor in combination with other conditions. Opposition coalition, then, can be seen as an important INUS condition. Allowing for causal complexity in terms of conjunctural causation and equifinality renders the focus on the causal role of single conditions almost meaningless and replaces this mono-causal perspective with the attempt at revealing the contexts in which opposition coalition – and sometimes even its absence – produces the outcome.

4.3 Analysis of the occurrence of NOT-LEO

As already mentioned several times above, one essential feature of causal complexity in terms of necessity, sufficiency, conjunctural causation, and equifinality is the issue of asymmetry. Put briefly, the (combinations of) factors responsible for the occurrence of an outcome and those for its negation/absence are expected to be not simply their mirror images. Put more simply, even if X causes Y, it does not follow from that not-X causes not-Y. Unlike most statistical techniques, set-theory based QCA requires the separate analysis of the occurrence and the non-occurrence of the outcome. I thus analyze the outcome NOT-LEO separately and, again, focus on the causal role opposition coalition and the negation of this factor plays.

Analysis of necessary conditions for NOT-LEO

The results of the necessary condition test for NOT-LEO are reported in Table 8. It is revealed that 2 out of the 18 (9 conditions and their respective complements) reach an acceptable consistency level of higher than 0.9 as a necessary condition for LEO²⁹: the *absence* of opposition mobilization and the absence of an opposition coalition. Unlike in the necessity analysis for the outcome LEO, this time, at least one of the two consistently necessary conditions is not trivial. The coverage value for the absence of opposition coalition is a stunning 0.82.³⁰

²⁹ The consistency value of 0.91 is the result of dividing the number of cases with not-opposition coalition *and* not-LEO (n = 32) by the sum of all cases with not-LEO (35).

³⁰ The coverage value results from dividing the sum of all cases with not-opposition coalition *and* not-LEO (n = 32) by the sum of all cases with not-opposition coalition (n = 39). The coverage value for opposition mobilization is also still quite high at 0.75.

A more detailed look at the data reveals which cases deviate from the pattern of not-opposition coalition being necessary for not-LEO. As Table 4 above already showed, there are three cases that do not have LEO but have opposition coalition (Malaysia 1999, Romania 1992, and Zimbabwe 2002), thus providing evidence that contradict the statement that not-opposition coalition is consistently necessary for not-LEO and thus lower the consistency value from 1 (perfect consistency) to 0.85. Overall, however, cases with not-LEO tend to also have not-opposition coalition.

Table 8: Test of Necessity: Outcome NOT-LEO

Conditions tested	Consistency	Coverage
opp._coa	0.09	n.r.
~opp._coa	0.91	0.82
inc._tur	0.20	n.r.
~inc._tur	0.80	n.r.
Parliame	0.34	n.r.
~parliame	0.66	n.r.
prior_li	0.23	n.r.
~prior_li	0.77	n.r.
oppmob_d	0.06	n.r.
~oppmob_d	0.94	0.75
ecgrow_d	0.69	n.r.
~ecgrow_d	0.31	n.r.
fdi_d	0.20	n.r.
~fdi_d	0.80	n.r.
foraid_d	0.43	n.r.
~foraid_d	0.57	n.r.
regope_d	0.11	n.r.
~regope_d	0.89	n.r.

n.r. = not relevant

Analysis of sufficient conditions for NOT-LEO

Table 9 follows the same structure and logic as Table 7. For the three different sufficiency analysis (most complex, intermediate, most parsimonious), I display the total number of paths, the number of times not-opposition coalition is sufficient on its own, and in how many paths this condition and its negation is included. The full versions of the three different solution formulas for the non-occurrence of LEO can be found in Appendix 3.

First of all, we notice that the complexity of the most complex solution formula is greater than it was when we analyzed the outcome LEO. A total of 12 paths are leading to NOT-LEO. This is more than the 10 paths leading to LEO (see Table 6) but consider that there are also more truth table rows with the outcome NOT-LEO (LEO = 0) than with the outcome LEO (LEO = 1, see Table 5). In other words, in the most complex version of the solution formulas, we need 10 paths to cover 15 cases of LEO but only 12 paths to cover 35 cases of NOT-LEO. The main portion of

cases is covered/explained by one path. This sufficient conjunction alone applies to 15 cases.³¹ In substantive terms this means that the cases without LEO are analytically much more similar than the cases with LEO. If we allow the computer to compute simplifying assumptions, the complexity is reduced to just 5 paths. Here, again, the majority of cases – 22 - follow one path.³²

Table 9: Summary Sufficiency Analysis, Outcome NOT-LEO

Row		Solution formulas		
		Most complex Solution without simplifying assumptions	Intermediate Solution with only simplifying assumptions	Most parsimony Solution with simplifying assumptions
1	Number of paths towards LEO	12	11	5 ^a
2	Number of times NOT- Opposition Coalition is sufficient on its own	-	-	-
3	Number of times Opposition Coalition is part of paths	3	3	-
4	Number of times Negation of opposition coalition is part of paths	9	8	3

^a: based on selection of 2 prime implicants out of 4 logically redundant

In none of the sufficiency analyses does the *absence* of opposition coalition alone produce the outcome NOT-LEO, nor does, as expected, the presence of opposition coalition. Furthermore, we see that while the factor ‘*absence* of opposition coalition’ is present in a majority of paths leading towards the outcome NOT-LEO, it is not part of all the sufficient conjunctions. What is more, some paths towards NOT-LEO contain the *presence* of opposition coalition (rows # 21, 28, 39 in Table 5 covering the cases Malaysia 1999, Romania 1992, and Zimbabwe 2002). In other words, the factor spelled out by Howard and Roessler (Howard & Roessler, 2006) as the main driving force for the occurrence of LEO, in fact, seems to play a causal role in explaining the non –occurrence of LEO in some cases.

As was the case in the analysis of the occurrence of LEO, this rather troubling empirical evidence can only be ‘masked’ once we opt for maximum parsimony by allowing the computer to make simplifying assumptions about the 473 logical remainders: in the most parsimonious solution term, no path towards the non-occurrence of LEO contains the presence of opposition coalition anymore. Notice, however, that this most parsimonious solution is based on 219 simplifying assumptions, most of which are highly unlikely to pass any serious plausibility test.

³¹ The path is: opp. coalition * opp mobiliz. * inc. turnover * ECON. GROWTH * fdi * regime openness → NOT-LEO

The cases covered by this path are:
ALB96+ALB01+ARM96, TOG98+CEN99, GAB98, GUI98, SRI94, UGA01+CRO97+INS92, MAL90, NEP99+NEP94+SRI99.

³² The path is: opp. coalition * inc. turnover * ECON. GROWTH → NOT-LEO

The cases covered by this path are:

Hence, if this most parsimonious solution term is used for the substantive interpretations that the presence of opposition coalition is not linked to the non-occurrence of LEO, we must be aware that this rests on many questionable assumptions.³³

4.4 *What can we learn from the case of Kenya?*

As mentioned above, Howard and Roessler discuss the case of Kenya 2002 to foster the plausibility of their claim that, in fact, opposition coalition plays a crucial role in bringing about LEO. Guiding questions for this chapter follow the standard questions that should be asked about case studies in mixed methods (small-N – large-N) designs: Is their case well chosen? How representative is it? And, does the Kenyan evidence really foster the cross-national claim? In their account of Kenya 2002, they refer to four of their nine independent variables: the presence of an opposition coalition and of an incumbent turnover and the lack of economic growth and foreign aid. The first point to mention is that these are not the statistically significant variables. Hence, the question arises why are only a subset of their nine variables from the cross-national analysis looked and why those four and not others? In order to argue that the plausible analytic story told about Kenya is also applicable to and insightful for other cases, one must argue that variation on all the other variables that had been introduced as potentially relevant factors for LEO in the cross-national model, but which are not addressed in Kenya, are irrelevant. The best way of supporting such an argument is to provide empirical evidence that, in fact, cases with *different* values on those conditions that are not addressed in the analysis of Kenya, but the *same* values on those that are addressed in the Kenyan analysis, also show the outcome LEO. This, however, is not the case and a truth table representation of the data is very helpful in revealing this empirical fact. As can be seen in Table 5, only one case scores the same as Kenya 2002 on the four conditions discussed in the Kenyan case study: it is Croatia 2000 (truth table row 10). The valid question then is: Which broader causal pattern is Kenya 2002 (and Croatia 2000) supposed to stand for and whether opposition coalition is playing the same causal role in different cases?

Second, by choosing a case that scores positive on both the expected condition (opposition coalition) and the outcome (LEO), they create a perfect setting for teasing out the causal mechanism linking these two variables, which is a perfectly valid and effective strategy (on strategies of case selection, see Lijphart, 1971 and 1975, Yin, 1994, Dion, 1998, Gerring, 2004,

ALB96+ALB97+ALB01+ARM96,TOG98+CEN99,GAB98,GUI98,SRI94,UGA01+CRO97+INS92,MAL90,NEP99+IRN93+MAL95+NEP94+PER95+SIN91,SIN97+SRI99+YGS96.

³³ In fact, it is even highly likely that several contradictory simplifying assumptions (Vanderborgh & Yamasaki, 2003, Schneider & Wagemann, 2007) were made by the computer in the analysis of the occurrence of LEO and then its non-occurrence. In substance, this means that for arriving at the most parsimonious solutions the computer assumes for *one and the same* logical remainder in the analysis of LEO that it would produce the outcome LEO and in the analysis of NOT-LEO that it would not produce the outcome NOT-LEO.

George & Bennett, 2005, Lieberman, 2005, Seawright & Gerring, 2005, or Rohlfing, forthcoming). It contains the danger, though, to overlook that the phenomenon of LEO might result from equifinal causal processes and that not all paths towards LEO necessarily involve the formation of an opposition coalition. And, in fact, as shown above, quite many cases (7 out of 15) achieved LEO *without* having an opposition coalition. Again, the tool of a truth table can help to find analytically fruitful cases for a matched comparison with Kenya 2002 and for further refining and strengthening the argument about the causal role played by opposition coalitions.

Third, another interpretation of the choice of Kenya 2002 could be that it follows the least likely case scenario (Gerring, 2007). The information in Table 5 reveals that Kenya, apart from having an opposition coalition and an incumbent turnover, does not display any of the other 7 conditions that are expected to contribute to the outcome LEO.³⁴ Since almost no other factor was present, the formation of an opposition coalition can more safely be regarded as the crucial factor for bringing about LEO. Also in this interpretation it remains important to note that there is no other single case that looks, if not identical, than at least similar to Kenya. If we define similarity as two cases that share the same values on all but one of the causal conditions and that are identical in the outcome value (just as it is done in Mill's method of a difference (Mill, 1865) or Przeworski and Teune's (1970) most similar system design), then we find no such case in the data set. Again, then, the question is how much we learn about the general causal role of opposition coalition when looking at Kenya 2002. Most probably other causal mechanisms are in place in other cases and opposition coalition is only sometimes involved.

Fourth, the authors are careful enough to point out that their causal claim remains a probabilistic one and that, therefore, they do not expect that the formation of opposition coalition always lead to LEO. In order to demonstrate their point, they refer to the case of Zimbabwe's elections in 2002 in which there was a coalition of opposition forces but, nevertheless, the incumbent president Robert Mugabe managed to hold on to political power at the cost of increased state pressure, economic decline, and international isolation (Howard & Roessler, 2006: fn 31).³⁵ . While such cases do not necessarily contradict the probabilistic statement that, overall, opposition coalition contributes to the occurrence of LEO, the identification and closer examination of these 'unexpected cases' can potentially yield analytically fruitful results. In addition to this, I think some further reflections can be made on the observation that opposition coalition and LEO do not perfectly correlate.

³⁴ All but the condition 'executive format' for which no clear directional expectation exists, are coded such that a 1 value indicates the hypothesized positive effect on the occurrence of LEO.

³⁵ To be precise, this is true for three cases: Mali 1999, Romania 1992, and Zimbabwe 2002. While in the case of Zimbabwe 2002 it seems plausible to argue that an opposition coalition was not enough to lead to LEO because it triggered severe counter-measures by the incumbent regime, the same rationale is far less convincing for Romania 1992, where larger scale and severe human rights violations did not occur.

Fifth, a non-deterministic relationship between opposition coalition and LEO also entails that there are cases that do not have an opposition coalition yet arrive at the outcome LEO. As shown above (see Table 4) this holds for a total of 7 cases (Armenia 1998, the Dominican Republic 1996, Guatemala 1995, Indonesia 1999, Peru 2000, Peru 2001, and Thailand 1992). Studying at least one of these cases could strengthen the inference by Howard and Roessler that what ultimately matters are strategic decisions by actors (and here, most importantly, the decision to unite in one opposition coalition) and not the structural context conditions. Maybe, upon closer examination one finds that there was a functional equivalent (Przeworski & Teune, 1970, Adcock & Collier, 2001) to an opposition coalition in the above mentioned cases.³⁶ Or, closer analyses of these unexpected cases of LEO would show that what looked like a LEO, in fact, was not because, for instance, all of these cases went back to some sort of hybrid regime much quicker than any of the other cases classified as LEO.³⁷

In sum, both types of rows (no opposition coalition with LEO and opposition coalition without LEO) represent empirical evidence that the core causal claim by Howard and Roessler of opposition coalition producing LEO is more complex. Howard and Roessler acknowledge this complexity by explicitly stating that this relation is not deterministic, which is correct but probably only partially a solution for also QCA results do not have to be deterministic. Where a QCA approach fundamentally differs from standard statistical approaches is in the treatment of this complexity. Rather than delegating this complexity into a lower explanatory power, the aim is to specify under which different conditions opposition coalition produces LEO and under which conditions it does not.

5 Conclusions

This paper had a primarily methodological focus and centered on the question: Can one safely assume that the application of QCA in addition to standard qualitative and quantitative techniques does not contain the potentials for generating additional analytic insights? The tentative answer based on the QCA analysis of Howard and Roessler's data on LEO is that one should probably not assume that QCA is a redundant data analysis technique.

³⁶ For instance, if the causal impact of an opposition coalition on LEO rests on the causal mechanism of signaling a credible alternative to the incumbent and her policies, then such a credible shift could probably also be produced by a radical redirection in policy and personnel of the ruling autocrats. Of course, any attempt to come up with theoretically fruitful causal mechanisms is seriously hampered by the fact that the dependent variable LEO is measured in a very coarse, if not opaque way. Remember, the authors use a combination of different increases on the Freedom House and the Polity scales, regardless at which levels of the scales these point increases occur. For more general critiques at the Freedom House and Polity scales see, for instance, Munck & Verkuilen, 2002 or McHenry, 2000).

³⁷ The fact that Peru appears as a case of LEO both in 2000 and in 2001 might indicate in such a direction at least for this case.

In this paper, I started out by highlighting the fact that any empirical data analysis technique operates on (often times hidden) assumptions about the nature of causality. It should, thus, not come as a surprise that the choice of method determines the *kind* of results one obtains from *one and the same* empirical information. Certain methods give priority to causal parsimony over causal complexity while others do the opposite. It is hardly surprising, thus, that the results of a QCA analysis differ from those generated by standard statistical techniques such as (logistic) regression. The argument made in the present paper is that in certain research situations, the application of QCA can reveal insightful patterns in the data that otherwise would remain hidden and that, often times, are more in line with the theory-guided expectations And research aims. As made clear several times in this paper, the aim of the QCA re-analysis was not to verify the results produced by Howard and Roessler, nor to criticize any of their analyses.

Space limitations prevented me from going into further details of the causes behind LEO. But one message should have been conveyed. The role of opposition coalition in bringing about LEO is a complex one. Only when very many simplifying assumptions are allowed for – either in QCA or in logistic regression – can the empirical information at hand be summarized such that opposition coalition alone and net of all the other relevant factors, is playing a crucial role. Put differently, in order to attribute a exceptional causal role to opposition coalition, we must make a series of far reaching assumptions that are methodologically induced rather than backed by substantive arguments. More specifically, we must assume that each factor has the same effect across all cases; that the context does not matter; that the cases observed are a random sample of a well-defined population; that the causes for LEO automatically explain not-LEO; and that there is just one path towards LEO. Whether or not such an extensive use of assumptions is analytically desirable ultimately depends on the intended goals of the analysis. The more simplifying assumptions are made - and they are made by default in standard statistical techniques, while in QCA one can opt not to make them -, the more parsimonious and theoretically digestible, but also the more detached from the case evidence will be the results. In contrast, if less, or even not a single, simplifying assumptions are made, the result stays more true to the empirical case evidence at hand. Suppose the aim of ths research on LEO was to generate policy relevant knowledge, useful for (external) democracy promoters. More likely than not, the overriding importance attributed to opposition coalition via regression analysis would lead to less useful policy advice than the more differentiated picture generated with QCA. In QCA, not only is the effect of single variables put into context, but it is also possible to specify the (groups of) individual cases that follow the different causal configurations.

Reference List

- Abbott, A. (2001). *Time matters. on theory and method*. Chicago and London: The University of Chicago Press.
- Achen, C.H. (2005). Two cheers for Charles Ragin. *Studies in Comparative International Development* 40, 1, 27-32.
- Adcock, R., & Collier, D. (2001). Measurement validity: a shared standard for qualitative and quantitative research. *American Political Science Review* 95, 3, 529-546.
- Amenta, E., & Poulsen, J.D. (1994). Where to begin. a survey of five approaches to selecting independent variables for qualitative comparative analysis. *Sociological Methods & Research* 23, 1, 22-53.
- Behnke, J., Gschwend, T., Schindler, D., & Schnapp, K.-U. (eds.) (2006). *Methoden der Politikwissenschaft: neuere quantitative und qualitative Analyseverfahren*. Baden-Baden: Nomos.
- Bollen, K.A. (1993). Liberal democracy: validity and method factors in cross-national measures. *American Journal of Political Science* 37, 4, 1207-1230.
- Brady, H.E., & Collier, D. (eds.) (2004). *Rethinking social inquiry: diverse tools, shared standards*. Lanham: Rowman&Littlefield.
- Braumoeller, B.F. (2003). Causal complexity and the study of politics. *Political Analysis* 11, 3, 209-233.
- Braumoeller, B.F., & Goertz, G. (2002). Watching your posterior: sampling assumptions, falsification, and necessary conditions. *Political Analysis* 10, 2, 198-203.
- Braumoeller, B.F., & Goertz, G. (2003). The statistical methodology of necessary conditions. In G. Goertz & Starr, H. (Eds.), *Necessary conditions: theory, methodology, and applications* . (pp. 197-223). Lanham, Md., Boulder:Rowman & Littlefield.
- Clarke, K.A. (2002). The reverend and the ravens: a response to 'Testing for necessary and/or sufficient Causation: which cases are relevant?'. *Political Analysis* 10, 2, 194-197.
- Collier, D., & Adcock, R. (1999). Democracy and dichotomies: a pragmatic approach to choices about concepts. *Annual Review of Political Science* 2, 537-565.
- Coppedge, M. (2002). *Theory building and hypothesis testing: large- vs. small-N research on democratization*. Paper prepared for presentation at the Annual Meeting of the Midwest Political Science Association, Chicago, Illinois, April 25-27
- Cronqvist, L. (2005). *Introduction to Multi-Value Qualitative Comparative Analysis (MVQCA)*. Marburg. Internet: <http://www.tosmana.net>.
- Dion, D. (1998). Evidence and inference in the comparative case study. *Comparative Politics* 30, 127-145.
- Eliason, S.R., & Stryker, R. Goodness-of-fit tests and descriptive measures in fuzzy-set analysis , unpublished manuscript. (un pub)
- George, A.L., & Bennett, A. (2005). *Case studies and theory development in the social sciences*. Cambridge, Mass.: MIT Press.
- Gerring, J. (2001). *Social science methodology. a critical framework*. Cambridge: Cambridge University Press.
- Gerring, J. (2004). What is a case study and what is it good for? *American Political Science Review* 98, 2, 341-354.
- Gerring, J. (2007). Is there a (viable) crucial-case method? *Comparative Political Studies* 40, 3, 231-253.
- Goertz, G. (2003a). *Assessing the importance of necessary or sufficient conditions in fuzzy set social science*. Compass Working Paper, WP 2003-3
- Goertz, G. (2003b). The substantive importance of necessary condition hypotheses. In G. Goertz & Starr, H. (Eds.), *Necessary conditions*. (pp. 65-94). Lanham:Rowman & Littlefield Publishers.

- Goertz, G. (2006a). Assessing the trivialness, relevance, and relative importance of necessary and sufficient conditions in social science. *Studies in Comparative International Development* 41, 2,
- Goertz, G. (2006b). *Social science concepts. a user's guide*. Princeton : Princeton University Press.
- Goertz, G., & Starr, H. (eds.) (2003). *Necessary conditions: theory, methodology, and applications* . Lanham, Md., Boulder: Rowman & Littlefield.
- Goldthorpe, J.H. (1997). Current issues in comparative macrosociology: a debate on methodological issues. *Comparative Social Research* 16, 1-26.
- Grendstad, G. (2007). Causal complexity and party preference. *European Journal of Political Research* 46, 1, 121-149.
- Hall, P.A. (2003). Aligning ontology and methodology in comparative politics. In J. Mahoney & Rueschemeyer, D. (Eds.), *Comparative historical analysis in the social sciences*. (pp. 373-404). Cambridge/New York: Cambridge University Press.
- Howard, M.M., & Roessler, P.G. (2006). Liberalizing electoral outcomes in competitive authoritarian regimes. *American Journal of Political Science* 50, 2, 365-381.
- Jagers, K., & Gurr, T.R. (1995). Tracking the third wave with the polity III data. *Journal of Peace Research* 32, 4, 469-482.
- Jahn, D. (2006). *Einfuehrung in die vergleichende Politikwissenschaft*. Wiesbaden: VS Verlag.
- King, G., Keohane, R.O., & Verba, S. (1994). *Designing social inquiry. scientific inference in qualitative research*. Princeton: Princeton University Press.
- Kittel, B. (2006). A crazy methodology? on the limits of macroquantitative social science research. *International Sociology* 21 , 5, 647-677.
- Levitzky, S., & Way, L. (2002). The rise of competitive authoritarianism. *Journal of Democracy* 13, 2, 51-65.
- Lieberman, E.S. (2005). Nested analysis as a mixed-research strategy for comparative research. *American Political Science Review* 99, 3, 435-451.
- Lieberson, S. (1985). *Making it count: the improvement of social research and theory*. Berkeley: University of California Press.
- Lieberson, S. (2004). Comments on the use and utility of QCA. *Qualitative Methods. Newsletter of the American Political Science Association Organized Section on Qualitative Methods* 2, 2, 13-14.
- Lijphart, A. (1971). Comparative politics and comparative method. *American Political Science Review* 65, 3, 682-693.
- Lijphart, A. (1975). The comparable-cases strategy in comparative research. *Comparative Political Studies* 8, 2, 158-177.
- Mackie, J.L. (1974). *The cement of the universe*. Oxford: Oxford University Press.
- Mahoney, J. & Goertz, G. (2004). *The possibility principle and case selection: choosing negative cases in comparative analysis*. Compass Working Paper, WP 2004-14
- Mahoney, J., & Rueschemeyer, D. (eds.) (2003). *Comparative historical analysis in the social sciences*. Cambridge/New York: Cambridge University Press.
- Marx, A. (2006). *Towards more robust model specification in QCA. results from a methodological experiment*. Compass Working Paper, WP 2006-43
- McHenry, D.E.Jr. (2000). Quantitative measures of democracy in Africa: an assessment. *Democratization* 7, 2, 168-185.
- Mill, J.S. (1865). Of the four methods of experimental inquiry. In J. S. Mill (Ed.), *System of logic, ratiocinative and*

inductive, vol. 1, book 3, chap. VIII, sixth edition. (pp. 427-450).

Munck, G.L. (2001). The regime question. theory building in democracy studies. *World Politics* 54, October, 119-144.

Munck, G.L., & Verkuilen, J. (2002). Conceptualizing and measuring democracy: evaluating alternative indices. *Comparative Political Studies* 35, 1, 5-33.

Pierson, P. (2004). *Politics in time: history, institutions, and social analysis*. Princeton N.J.: Princeton University Press.

Przeworski, A., & Teune, H. (1970). *The logic of comparative social inquiry*. New York, London, Toronto, Sydney: Wiley Interscience.

Ragin, C.C. (1987). *The comparative method. moving beyond qualitative and quantitative strategies*. Berkeley/Los Angeles/London: University of California Press.

Ragin, C.C. (2000). *Fuzzy-set social science*. Chicago: University of Chicago Press.

Ragin, C.C. (2005). Core versus tangential assumptions in comparative research. *Studies in Comparative International Development* 40, 1, 33-38.

Ragin, C.C. (2006a). The limitations of net-effect thinking. In B. Rihoux & Grimm, H. (Eds.), *Innovative comparative methods for policy analysis. beyond the quantitative-qualitative divide*. (pp. 13-41). New York: Springer.

Ragin, C.C. (2006b). Set relations in social research: evaluating their consistency and coverage. *Political Analysis* 14, 3, 291-310.

Ragin, C.C. (forthcoming). Fuzzy sets: calibration versus measurement. In D. Collier, Brady, H., & Box-Steffensmeier, J. (Eds.), *Methodology volume of Oxford handbooks of political science*.

Ragin, C.C., & Rihoux, B. (2004). Qualitative Comparative Analysis (QCA): state of the art and prospects. *Qualitative Methods. Newsletter of the American Political Science Association Organized Section on Qualitative Methods* 2, 2, 3-12.

Ragin, C.C., & Sonnett, J. (2004). Between complexity and parsimony: limited diversity, counterfactual cases, and comparative analysis. In S. Kropp & Minkenberg, M. (Eds.), *Vergleichen in der Politikwissenschaft*. Wiesbaden: VS Verlag für Sozialwissenschaften.

Rihoux, B., & Ragin, C.C. (eds.) (2007). *Configurational comparative methods*. Thousand Oaks/London: Sage.

Rohlfing, I. (forthcoming). Pitfalls and Principles of Nested Analysis in Comparative Research. *Comparative Political Studies*

Schneider, C. Q. & Grofman, B. (2006). *It might look like a regression ... but it's not! An intuitive approach to the presentation of QCA and fs/QCA results*. Compass Working Paper, WP2006-39

Schneider, C.Q., & Wagemann, C. (2006). Reducing complexity in Qualitative Comparative Analysis (QCA): remote and proximate factors and the consolidation of democracy. *European Journal of Political Research* 45, 5, 751-786.

Schneider, C.Q., & Wagemann, C. (2007). *Qualitative Comparative Analysis (QCA) und Fuzzy Sets. Ein Lehrbuch für Anwender und alle, die es werden wollen.*: Verlag Barbara Budrich.

Seawright, J. (2002a). Testing for necessary and/or sufficient causation: which cases are relevant? *Political Analysis* 10, 2, 178-193.

Seawright, J. (2002b). What counts as evidence? prior probabilities, posterior distributions, and causal inference. *Political Analysis* 10, 2, 204-207.

Seawright, J. (2004). Qualitative Comparative Analysis vis-a-vis regression. *Qualitative Methods. Newsletter of the American Political Science Association Organized Section on Qualitative Methods* 2, 2, 14-17.

Seawright, J. (2005). Assumptions, causal inference, and the goals of QCA. *Studies in Comparative International Development* 40, 1, 39-42.

Seawright, J. & Gerring, J. (2005). *Case-selection techniques in case study research: a menu of qualitative and quantitative options*. unpublished manuscript, Berkeley and Boston

Vanderborght, Y. & Yamasaki, S. (2003). *The problem of contradictory simplifying assumptions in Qualitative Comparative Analysis (QCA)*. Paper presented at the ECPR General Conference, Marburg, 18-21 September

Wagemann, C. & Schneider, C. Q. (2007). *Standards of good practice in Qualitative Comparative Analysis (QCA) and fuzzy sets*. Compass Working Paper, WP forthcoming

Walker, H.A., & Cohen, B.P. (1985). Scope statements: imperatives for evaluating theory. *American Sociological Review* 50, 3, 288-301.

Woodward, J. (2002). There is no such thing as a ceteris paribus law. *Erkenntnis* 57, 303-328.

Yin, R.K. (1994). *Case study Research. design and methods*. (2. ed.). Thousand Oaks, London, New Delhi: Sage Publications.

Appendix 1: Description of variables and thresholds for dichotomizing non-dichotomous variables

Their first variable is ‘opposition coalition’. It is hypothesized that if opposition groups manage to unite and agree on a common strategy, program, and candidate for a national election, they can generate LEO. This is expected to be the case because opposition coalition (a) manage to take away votes from the incumbent, (b) prevent incumbents from playing opposition groups against each other, (c) increase the risk and costs of repression for the incumbent, and (d) make it more attractive and less dangerous for voters to vote against the incumbent (Howard & Roessler, 2006: 371). The variable ‘opposition coalition’ is operationalized as a dichotomy with 1 indicating the presence of an opposition coalition.

Second, they include the variable ‘opposition mobilization’ and hypothesize that “a higher level of mobilization will be positively associated with the occurrence of a LEO” (Howard & Roessler, 2006: 372). This variable is operationalized through the average number of peaceful public gathering of opposition groups in the years before and during the elections. While this operationalization yields a continuous measure, their hypothesis can be read in a way that makes a dichotomization an equally appropriate representation of the underlying condition and its expected causal effect on LEO. Howard & Roessler, 2006 do not say ‘the higher the opposition mobilization, the more likely will be LEO’, but rather that a higher level of mobilization exerts an effect on LEO. Hence, the authors’ hypothesis, indeed, makes a statement about a set relation between high levels of opposition mobilization and the occurrence of LEO and not of a correlation between the variables ‘opposition mobilization’ and LEO. Above I contrasted set relational with correlational hypotheses and showed that they generate different expectations about the underlying data patterns. If we apply these different expectations to the current example, one can argue two things: First, higher levels of mobilization are juxtaposed to all other possible levels of mobilization (medium, low, etc.), which justifies a dichotomization of this concept. Second, since, in essence the claim is that high opposition mobilization is sufficient for LEO, we only have expectations about the value for LEO when ‘mobilization’ is high, but when mobilization is not high, LEO can either occur or not occur. From this follows that we do not assume that opposition mobilization and LEO are related in a monotonic and symmetric fashion (‘the more of x, the more of y and vice versa, the less of x, the less of y). I therefore recode this variable with 1 indicating high mobilization and 0 not-high mobilization. The following threshold is chosen:

Opposition mobilization (average number of peaceful opposition gatherings of at least 100 people in year prior to elections and election year):

Lower than 3.5 = 0

Higher than 3.5 = 1

Their third independent variable is ‘incumbent turnover’. In some elections, the incumbent presents herself for reelections while in others the competitive authoritarian ruling elites are presenting a new candidate to the electorate. The authors hypothesize that LEO are more likely if the incumbent is not running again. This variable is already operationalized as a dichotomy in the original study with 1 indicating that the incumbent does not run again.

In addition to these strictly political variables, Howard & Roessler, 2006 introduce a series of plausible structural conditions for LEO. They justify this inclusion of structural features not only by arguing that one needs to control for these factors. At different points of their argument they also claim that different structural conditions can create different incentives for actors and their strategies for bringing about (or hindering) LEO. This is equivalent to saying that they expect one and the same actor-based factor (opposition coalition, opposition mobilization, incumbent turnover) to have different effects on LEO depending on the structural context in which these factors are embedded. Above, this causal structure was discussed under the label of equifinality and conjunctural causation. In statistical terms, the authors’ wording insinuates the presence of second (or even higher) order interaction effects. In their empirical model they fail to test (any) of these interactions. Most probably this is due to purely methodological constraints rather than substantive arguments, i.e. the already low power of their main effects model (9 independent variables and only 50 cases). The structural features Howard & Roessler, 2006 specify are subdivided into economic, global, international, and political institutional factors.

‘Economic crisis’: They expect that low or negative economic growth increases the likelihood of LEO: “[...] a crisis helps to tilt the balance of power in favor of the opposition and weaken the bargaining power of the incumbents” (p. 372). This variable is operationalized as the average growth of a country during the two years preceding the elections. Again, while this yields a continuous numerical measure the verbal hypothesis can be used for justifying a dichotomization and for expecting that economic growth and LEO are not necessarily highly correlated but rather set related. The expectation formulated by the authors only refers to low and negative growth rates and thus generates no expectations on whether or not LEO occur when this variable takes very high numerical values. I recode this variable into a dichotomy with 1 indicating low or negative growth. The following threshold is chosen:

Economic growth (percentage change of GDP):
Lower than 3.5 = 0
Higher than 3.5 = 1

‘Connection to the West’: The expectation is that countries with greater connection to the West have higher chances for experiencing LEO. This variable is operationalized as the country’s average level of foreign direct investments as a percentage of its GDP and the average foreign aid per capita during the two years preceding the elections. When using QCA – at least in its

crisp set version (csQCA) - the data for both the conditions and the outcome must be dichotomous.³⁸ And this is why this and other variables in the original study by Howard and Roessler are dichotomized. Beyond this purely methodological need for dichotomization, also some substantive justifications can be given: I think it is as (im)plausible to assume that the effect of connection to the West increases (decreases) in a linear fashion with each Euro of FDI as it is to claim that what matters is whether or not a country has close ties to the West. In other words, although a dichotomization of the continuous variable comes at the cost of a loss of information, it is far from straightforward that all the *numerical* variation contained in these two variables corresponds to some *theoretically* meaningful variation. A careful calibration of the empirical information into a dichotomous concept would take into account this correspondence between theoretical and numerical variation (on calibration, see Ragin, forthcoming). I transform both variables into dichotomies with 1 indicating high connectedness with the West. The thresholds are the following:

Foreign direct investment (as percentage of GDP):

Lower than 3.5 = 0
Higher than 0.5 = 1

Foreign aid (level of aid per capita):

Lower than 25 = 0
Higher than 25 = 1

‘Executive format’: Here the authors look at whether a country at the time of the elections has a presidential or a parliamentary system. This variable is a dichotomy with parliamentarism (1) and presidentialism (0). The authors follow the literature in which no clear expectations on the direction of the causal effect is expressed: “[...] the theoretical expectations can point in either direction” (Howard & Roessler, 2006: 374). It all seems to depend on the configuration of other factors within which an executive format operates. This is, again, a hint that Howard & Roessler, 2006 expect the outcome LEO to be the result of complex causal structures, an expectation not picked up in their cross-national statistical analysis that can, however, be analyzed with QCA.

‘Regime openness’: The expectation is that LEO are more likely to occur in political regimes that prior to elections signal their openness by an increased respect for civil rights because, and this is important to note, “[...] one would expect the opposition to find it easier to mobilize and organize for an electoral contest against the incumbent” (Howard & Roessler, 2006: 374). In other words, the variable ‘regime openness’ alone is not expected to have any impact on the occurrence of LEO. According to the authors, it must be combined with the other conditions

³⁸ One way out of the need for strict dichotomization would be the use of fuzzy set QCA (fsQCA). In fsQCA, conditions can either be dichotomous or take any value between 0 and 1. The outcome, however, must be a fuzzy set. This rules out the application of fsQCA to the present data where the outcome Leo is operationalized as a clear-cut dichotomy. Of course, one could (and probably better should) think of Leo as a matter of degree and operationalize it as a fuzzy set, a task that I postpone to the future. Another option would be to use multi-value

‘opposition coalition’ and ‘opposition mobilization’. Clearly, this is yet another indication that the authors do not rule out the possibility of conjunctural causation. The variable regime openness is operationalized as the average Freedom House civil liberty score during the two years prior to the elections. Again, one can argue that a dichotomy expresses the concept as adequately as a metric measure does because we only have expectations about high levels of regime openness, not so much about intermediate or low levels. I therefore transform the data with 1 indicating prior regime openness. The following threshold is used:

Regime openness (Freedom House civil liberties, reversed scale):
Lower than 5.1 = 0
Higher than 5.1 = 1

‘Regime openness change’: The hypothesis is that LEO is the result of a culmination of endogenous liberalization that unfolded in the years and decades prior to the elections under investigation. If so, other, more actor-centered factors would be the result of these broader trends and should, thus, have little own explanatory power for LEO. The variable is operationalized as a dichotomy with cases having experienced a regime openness change (operationalized as the increase in their Freedom House political rights score in the year prior to the election with regard to their score 5 years before) receiving a score of 1.

Appendix 2: Three solution formulas for outcome LEO

Most complex solution formula

(No simplifying assumptions about logical remainders)

$opp_coa*INC_TUR*prior_li*oppmob_d*ECGROW_D*fdi_d*foraid_d*regope_d+$
 $opp_coa*INC_TUR*parliame*prior_li*oppmob_d*ECGROW_D*fdi_d*regope_d+$
 $OPP_COA*inc_tur*parliame*prior_li*oppmob_d*fdi_d*FORAID_D*regope_d+$
 $opp_coa*INC_TUR*PRIOR_LI*OPPMOB_D*ecgrow_d*fdi_d*foraid_d*regope_d+$
 $OPP_COA*inc_tur*parliame*PRIOR_LI*oppmob_d*ECGROW_D*fdi_d*regope_d+$
 $OPP_COA*parliame*PRIOR_LI*oppmob_d*ECGROW_D*fdi_d*FORAID_D*regope_d+$
 $opp_coa*inc_tur*parliame*prior_li*OPPMOB_D*ecgrow_d*fdi_d*foraid_d*regope_d+$
 $OPP_COA*INC_TUR*parliame*prior_li*oppmob_d*ecgrow_d*fdi_d*foraid_d*regope_d+$
 $OPP_COA*INC_TUR*PARLIAME*prior_li*oppmob_d*ecgrow_d*FDI_D*foraid_d*regope_d+$
 $OPP_COA*inc_tur*parliame*PRIOR_LI*OPPMOB_D*ecgrow_d*fdi_d*FORAID_D*REGOPE_D$

solution coverage: 1.000000

solution consistency: 1.000000

Most parsimonious solution formula

(Computer generated simplifying assumptions on potentially all logical remainders)

OPP MOBILIZ. * econ. growth +	
(INS99+PER00+PER01+YGS00)	
OPP. COALITION * INC. TURNOVER +	
(CRO00+GHA00+KEN02)	
OPP. COALITION * ECON. GROWTH +	
(GHA96+GHA00+ROM96+SEN00)	
OPP. COALITION * FOREIGN AID +	
(GHA96+GHA00+NIC90+SEN00+YGS00)	
INC. TURNOVER * ECON. GROWTH * regime openness * prior lib change	
(ARM98+DOM96,GUA95+THI92)	

solution coverage: 1.000000

solution consistency: 1.000000

Number of simplifying assumptions: 306

Intermediate complex solution

(‘Easy’ counterfactuals based on directional expectations)

$parliame*OPPMOB_D*ecgrow_d+$
 $OPP_COA*INC_TUR*prior_li*foraid_d+$
 $INC_TUR*prior_li*ECGROW_D*fdi_d*regope_d+$
 $OPP_COA*parliame*prior_li*regope_d+$
 $INC_TUR*OPPMOB_D*ecgrow_d*fdi_d+$
 $OPP_COA*parliame*ECGROW_D*regope_d$

solution coverage: 1.000000

solution consistency: 1.000000

Appendix 3: Three solution formulas for outcome NOT-LEO

Most complex solution formula (No simplifying assumptions about logical remainders)

opp._coa*inc._tur*oppmob_d*ECGROW_D*fdi_d*regope_d+
 opp._coa*parliame*prior_li*oppmob_d*ecgrow_d*fdi_d*regope_d+
 opp._coa*inc._tur*parliame*prior_li*oppmob_d*foraid_d*regope_d+
 opp._coa*inc._tur*PARLIAME*oppmob_d*ECGROW_D*foraid_d*regope_d+
 opp._coa*INC._TUR*parliame*prior_li*oppmob_d*ecgrow_d*FORAID_D*regope_d+
 opp._coa*parliame*prior_li*oppmob_d*ECGROW_D*fdi_d*foraid_d*REGOPE_D+
 opp._coa*inc._tur*PARLIAME*prior_li*ECGROW_D*fdi_d*FORAID_D*regope_d+
 opp._coa*parliame*PRIOR_LI*oppmob_d*ECGROW_D*fdi_d*FORAID_D*regope_d+
 OPP._COA*inc._tur*parliame*PRIOR_LI*oppmob_d*ecgrow_d*fdi_d*foraid_d*regope_d+
 OPP._COA*inc._tur*parliame*prior_li*oppmob_d*ecgrow_d*fdi_d*foraid_d*REGOPE_D+
 OPP._COA*inc._tur*PARLIAME*prior_li*oppmob_d*ecgrow_d*FDI_D*foraid_d*regope_d+
 opp._coa*inc._tur*PARLIAME*prior_li*OPPMOB_D*ECGROW_D*fdi_d*foraid_d*REGOPE_D

solution coverage: 1.000000
 solution consistency: 1.000000

Most parsimonious solution formula (Computer generated simplifying assumptions on potentially all logical remainders)

foreign aid * REGIME OPENNESS + (IRN93+IRN97+YGS96+ZIM02)
opp. coalition * opp mobiliz. * econ. growth + (CHA01+IVO00,GNB99+MEX94,RUS00+RUS996+SEN93+ZAM01)
opp. coalition * opp mobiliz. * PRIOR LIB CHANGE + (ALB96+ARM96,TOG98+GUA90+MAL95+NEP94+SRI99)
opp. coalition * inc. turnover * ECON. GROWTH + (ALB96+ALB97+ALB01+ARM96,TOG98+CEN99,GAB98,GUI98,SRI94,UGA01+CRO97+INS92, MAL90,NEP99+IRN93+MAL95+NEP94+PER95+SIN91,SIN97+SRI99+YGS96)
opp mobiliz. * inc. turnover * econ. growth * foreign aid (CHA01+MAL99+ROM92+RUS996+ZIM02)

solution coverage: 1.000000
 solution consistency: 1.000000
 Number of simplifying assumptions: 219

Intermediate complex solution ('Easy' counterfactuals based on directional expectations)

opp._coa*inc._tur*oppmob_d*ECGROW_D*regope_d+
 opp._coa*inc._tur*PARLIAME*OPPMOB_D*ECGROW_D+
 opp._coa*parliame*PRIOR_LI*ECGROW_D*FORAID_D+
 opp._coa*parliame*prior_li*oppmob_d*ecgrow_d*regope_d+
 inc._tur*PARLIAME*oppmob_d*ecgrow_d*FDI_D*regope_d+
 inc._tur*parliame*PRIOR_LI*ecgrow_d*fdi_d*foraid_d*regope_d+
 opp._coa*parliame*prior_li*oppmob_d*ECGROW_D*foraid_d*REGOPE_D+
 inc._tur*parliame*prior_li*oppmob_d*ecgrow_d*fdi_d*foraid_d*REGOPE_D

solution coverage: 1.000000
 solution consistency: 1.000000