Qualitative Comparative Analysis (QCA) and Fuzzy-Sets: Agenda for a Research Approach and a Data Analysis Technique

Claudius Wagemann[a] **and Carsten Q. Schneider**[b]

[a] *Istituto italiano di scienze umane, Palazzo Strozzi, Piazza degli Strozzi, 50123 Florence, Italy*
claudius.wagemann@sumitalia.it

[b] *Department of Political Science and Center for the Study of Imperfections in Democracies, Central European University, Nador utca 9, 1051 Budapest, Hungary*
schneiderc@ceu.hu

Abstract

"Qualitative Comparative Analysis" (QCA) is an increasingly applied methodological tool in comparative social sciences. It is well suited for the analysis of causally complex claims framed in terms of necessity and sufficiency. This article presents the epistemology of QCA and discusses its applicability to social science research questions. It also illustrates some of the features that have recently been added to this set of methodological tools. This article is best read in close conjunction with Schneider and Wagemann's "Standards of Good QCA Practice," the next paper in this journal issue.

Over the last years, "Qualitative Comparative Analysis" (QCA) has drawn increasing attention within the social sciences. In the late 1980s, this methodological family was introduced to a wider public by the American social scientist Charles Ragin (1987). Since then, QCA has been continuously modified, extended, and adapted to the needs that arise in empirical social research (Ragin 2000; Ragin 2008; Schneider and Wagemann 2007;

Rihoux and Ragin 2008).[1] As the applicability of QCA to empirical social scientific research questions was increasing, QCA became more diffused within the discipline. Today it is applied to a wide range of research topics in different social science disciplines (see www.compasss.org for an overview). While gaining increasing attention and acceptance within US academy, QCA is nonetheless still far more prominent among European scholars, who more actively teach, develop, write (also in other languages than English), and apply QCA.

One can only speculate about the reasons for this continental divide in the reception of QCA. One possibility is that the European academy, being more fragmented and characterized by less strong methodological cleavages than its American counterpart, has given more space to those scholars who understand QCA as a methodological "third way" between "qualitative" and "quantitative" methods. This had also been Ragin's initial thought, since the sub-title of his 1987 book announced QCA as "moving beyond qualitative and quantitative strategies." As such, QCA also contributes to a debate which unfolded with much power only some years after the publication of Ragin's first book (for this debate, see the seminal contributions by King, Keohane, and Verba 1994 and Brady and Collier 2004).

In any event, the "Q" in the acronym "QCA" stands for "qualitative" and indicates the firm grounding of QCA in the qualitative research tradition. But in QCA, qualitative research can be extended far beyond just a handful of national cases or regional studies. In fact, contrary to widespread wisdom, a full-blown QCA cannot be meaningfully applied to just two, four or eight particular cases (Berg-Schlosser, De Meur, Rihoux, and Ragin 2008:4f.; Schneider and Wagemann 2007:271; Wagemann 2007). Rather, its full potential it unfolds in studies based on a mid-sized N. This is precisely why QCA is an appealing alternative for those scholars whose number of cases, and also research interest or hypotheses, do not permit either a standard statistical analysis or a classical comparative case study approach.

QCA still remains subject to continuous development and upgrading. In this article, we present the current "state of the art" of QCA because evidence indicates that not enough applicants are familiar with all details of the most updated versions. We start with a differentiation, between

---

[1] For more on the history of QCA, see Berg-Schlosser et al. 2008:3ff.

QCA as a research approach and QCA as an analytical technique (part 2). We then describe basic principles of QCA (part 3) and discuss its episte-mology (part 4). After this, we present different variants of QCA (part 5) and introduce the current methodological and technical QCA agenda (part 6). We conclude with the observation that although QCA is quickly growing and steadily being improved, much uncertainly exists about how effectively to perform a high-quality QCA. In the next article in this jour-nal issue we suggest a list of criteria for a "good" QCA.

QCA as Research Approach and Analytical Technique

The nature of QCA can be – and, in fact, is – interpreted in two different but complementary ways: as a research approach writ large or as a data analysis technique (Ragin 1987; Berg-Schlosser et al. 2008). The full poten-tial of QCA rests in the acknowledgement of its double nature, that is, that it is neither just another way of analyzing data nor just another way of describing standard qualitative research practices.

The interpretation of *QCA as a research approach* mainly refers to the iterative process of data collection as part of the process of: moving "between ideas and evidence" (Ragin 1994:76; Ragin 2004:126); model specification; taking a holistic view of cases (Fiss 2007:1180); case selec-tion and re-conceptualization of conditions and outcome; and a specific view regarding causality. As regards the latter, QCA enables a researcher to deal with causal complexity, understood as causation which is equifinal, conjunctural, and asymmetric (for this, see below, Berg-Schlosser et al. 2008:8, and Schneider 2008:chapter 5). Furthermore, unlike most quan-titative techniques, QCA does *not rely on a set of* assumptions, such as: permanent causality; uniformity of causal effects; and unit homogeneity (Berg-Schlosser et al. 2008:9).

Quite obviously, all these aspects of QCA reflect its "qualitative roots," which are already reflected in its name.[2] Indeed, in traditional qualitative comparative research, it is common: to exclude and/or add cases from anal-ysis during an ongoing research process; to re-code values for certain cases;

---

[2] It is interesting to note that the French speaking tradition does not seem to share fully this perspective; it translates QCA as "AQQC" ('*Analyse quali-quantitative comparée*') (De Meur and Rihoux 2002).

or to re-conceptualize entire variables if preliminary empirical evidence suggests these updates of research design. By contrast, in quantitative, statistically oriented research, most of these practices are considered problematic, to say the least. Likewise, complex causality – a starting assumption in QCA – is the rule, not the exception, in qualitatively oriented work.

QCA as an analytical technique refers to the so-called "analytical moment" (Ragin 2000) which sets in when the cases have already been selected and all conditions as well as the outcome have already been calibrated (on calibration, see Ragin 2008 and below). The main goal of this step – this "moment" – is to find empirical patterns in the data. This is usually achieved with the help of specific software packages. Currently, there are four programs suitable for different variants of QCA (for an introduction to the different variants of QCA, see below):

- the fsQCA software by Ragin and collaborators for csQCA and fsQCA (Ragin, Drass, and Davey 2006);
- Tosmana for csQCA and mvQCA(Cronqvist 2006);
- a package in R for csQCA (Dusa 2007); and
- a command in Stata for csQCA and fsQCA (Longest and Vaisey 2008).

This software aspect of QCA shows some similarities to quantitative, variable-oriented techniques of data analysis, such as, for example, regression analysis in its various forms. Despite these affinities, however, QCA cannot and should not be reduced to just another data analysis technique. That is, we think it is wrong to give primacy to the analytical moment and thus interpret the case-oriented, qualitative aspect of QCA as a second-rank feature. If QCA is limited to the analytical moment, one risks misunderstanding the nature of QCA and falling short of employing its full analytic strength. One particular risk would be to apply purely *quantitative* standards of analysis when evaluating QCA-based research. Only if we open our mind to the more qualitatively oriented research design aspects of QCA, can we see clearly the need for applying different evaluation standards.

Basic Principles of QCA

The general goal of a QCA is to support the researcher in the attempt to arrive at a meaningful interpretation of the (causal) patterns displayed by the cases under examination. Thus, QCA is predominantly oriented towards understanding cases. In turn, cases are perceived holistically as configurations of analytically relevant characteristics. These properties – chosen on the basis of the theoretical literature and sociological imagination – and the common interest of all QCA analyses is to explain the relation between one case property defined as the outcome and other properties defined as conditions.

In QCA, "relation" is conceptualized as set relation and not, as in standard statistics, as "cor-relation." For QCA, sub-sets and super-sets (and, at later stages, unions and intersections of sets) are cornerstones for developing causal claims. These set relations are related to the ideas of necessity and sufficiency of conditions (for the analytical connection between set theory and necessity/sufficiency, see Ragin 2000 and Schneider and Wagemann 2007:31ff.). This is what makes QCA such a powerful tool for examining (combinations of) conditions that are sufficient and/or necessary for a given outcome. This aim is strikingly different from that in standard statistical analyses, which are geared towards detecting co-variations of variables and estimating net effects (Ragin 2008:176ff.).[3]

A set-theoretic foundation gives rise to another interesting argument in favor of QCA. Most verbally formulated social scientific theories can be interpreted in terms of set relations between conditions and an outcome. Since set relations can be translated into necessity and sufficiency relations, we can find a huge amount of social science theories which generate hypotheses on necessary and/or sufficient conditions (for an impressive collection of necessary conditions, see Goertz 2003). If this proves true, a method like QCA is more adequate for investigating set-theoretic hypotheses than many other data analysis techniques that are not anchored in set theory.

---

[3] This does not exclude statistical instruments from also being helpful in examining necessary and/or sufficient conditions. See Dion 2003 for possibilities offered through the Bayesian approach, and Braumoeller 2003 and Braumoeller and Goertz 2003 for a discussion of other statistical approaches in this area. However, the currently dominant standard correlation-based statistical techniques do not do the best job.

As an example of a QCA solution formula, let us assume that we wanted to work out the sufficient (and later necessary) conditions for the "stabilization of a democracy" (Y). On the basis of consulting theoretical and empirical literature on this issue, we have identified as potential conditions "a developed economy" (A), "a homogeneous society without major social differences" (B), and "the dominance of a clan" (C).[4] The result of an analysis of the sufficient conditions could yield that

a) a simultaneous presence of a far developed economy and a homogeneous society without any notable differences is a sufficient condition for the stability of a democracy (in a more abstract formulation: the simultaneous presence of both characteristics logically implies the outcome of a stable democracy); and that

b) as a non-exclusionary alternative to this, the absence of a dominant clan is a sufficient condition for the stability of a democracy; that is, wherever no dominant clan can be observed, we can find a stable democracy.

This result shows that two alternative sufficient conditions ("paths") exist. These paths are not mutually exclusive. It is possible that one and the same case follows both the first path (simultaneous presence of a far developed economy and a homogeneous society without any notable social differences) and the second path (absence of a dominant clan). However, this QCA solution term also contains the possibility that cases can reveal a stable democracy despite having a dominant clan. This is true for those cases that are characterized by a developed economy *and*, at the same time, a homogeneous society without major social differences.

In QCA, such a result is usually represented with a standardized formal notation. In our example, the result would be written as

$$AB + c \rightarrow Y \tag{1}$$

---

[4] All examples are intended only for methodological illustration, not to contribute more directly to substantial discussions.

"AB" stands for the combination of the conditions A and B.[5] The plus sign stands for the logical OR and indicates the two alternative paths to the outcome Y. This might be confusing at first, since linear algebra, as taught from elementary school onwards, uses a plus sign for "and;" however, it is interpreted as an "or" in the algebras on which QCA techniques are based, namely, Boolean and fuzzy algebra (see below).

A small letter for the condition C indicates that not C itself but rather its negation (the absence of a dominant clan) is a sufficient condition for the outcome. The arrow pointing to Y indicates that the expression to its left-hand side logically implies the expression to its right-hand side (Ragin and Rihoux 2004), that is, that the expression to its left is a subset of the expression to its right. In available case-theoretical arguments, this empirical subset relation can be causally interpreted in terms of sufficiency.[6]

In the process of analyzing data with QCA, it is the rule rather than the exception that a single condition is neither necessary nor sufficient, yet plays a crucial causal role. Thus, consider the condition A in the following example

$$AB + c \rightarrow Y \tag{1}$$

Such conditions are called "INUS conditions." INUS stands for *"insufficient but <u>n</u>ecessary part of a condition which is itself <u>u</u>nnecessary but <u>s</u>ufficient for the result"* (Mackie 1974:62; Goertz 2003:68; Mahoney 2008). Condition A alone is not sufficient, but it is a necessary component of the (combined) condition AB – which itself is not necessary but only sufficient for Y. This means that starting from a general focus on set-theoretical relationships QCA enables a researcher to model complex causal relations in such a way that also those factors are identified as causally relevant that are alone neither sufficient nor necessary.

In QCA, conditions do not compete against each other in a race for explaining more of the variation in an outcome (Ragin 2003:8; Ragin 2006a; Ragin 2008:chap 10). Rather, different paths – as discussed above – can be

---

[5] An alternative notation is "A*B," where the * sign represents the logical AND or the intersection of the sets A and B, respectively.

[6] In everyday language, we are tempted to say that the combination AB + c "leads to" Y. However, we recommend to avoid such wording, since the inverse arrow cannot be substituted with "leads to" in the case of necessary conditions (see below).

equivalent[7] ("equifinal") alternatives for one another. This even goes so far that in a solution formula for sufficient conditions, such as

$$AB + aC \rightarrow Y \tag{2}$$

a condition (in the example the condition A) takes on different roles. That is, in order to explain Y, it has to be *present* if combined with the condition B and *absent* if combined with the condition C. Such a situation is an example of a clear break from configurational theory with a predominant linear paradigm (Fiss 2007:1181).

To this analysis of sufficient conditions, an analysis of necessary conditions has to be added. As far as the example presented above is concerned, there is no condition which can be discovered in all the alternative paths leading to the result. This suggests that no condition is necessary. If any condition were necessary, then it would have to be contained in all "paths" causing the outcome.

If, in another hypothetical example, we had found a necessary condition X, the following formal notation would be appropriate:

$$X \leftarrow Y \tag{3}$$

Note that the inverse direction of the arrow does not suggest any causal mechanism. Y does not "lead to" or "cause" X. The arrow represents a logical implication: it says that, wherever we find Y, we will also find X. Here is the definition of a necessary condition based on a subset relationship between condition X (the superset) and the outcome Y (the subset).

Causal Complexity in QCA

As we now see, this discussion of technical aspects of QCA (whose use of software, formulas, and even numbers and letters might remind us of statistical techniques) is closely connected to a very specific understanding of causal complexity. First, the fact that usually more than one sufficient

---

[7] This logical equivalence of paths towards an outcome does not rule out the possibility of assessing their (different) degrees of empirical importance, usually achieved through the coverage measures (see Ragin 2006b and below).

condition is discovered as an explanation for an outcome points to the presence of *equifinality* (see above and Katz and Kahn 1978:30 for an early notion of this phenomenon). QCA is based on the assumption that "many roads lead to Rome." This is in sharp contrast to the unifinal perspective of many statistical techniques, among them the usually applied additive and linear regression models.

In a regression equation of the type

$$y = a + b_1x_1 + b_2x_2 + b_3x_3 + \ldots + e \qquad (4)$$

only one way exists for how an outcome is produced, namely, the way described in the additive regression equation. No matter how the values are distributed among the variables, the linear combination on which the estimation is based is always the same. Contrary to this, a solution for sufficient conditions in QCA shows the different paths which count as causal *alternatives* for an outcome. This is indicated through the logical OR (+).

Linked to this, *conjunctural causation* becomes evident in the combination of various conditions through the logical AND (*). A condition exerts its effect not alone, but only in combination with one or more other conditions. Here, too, we see a sharp contrast to the practice of standard statistical techniques. Even if the specification of interaction terms in multivariate regression analysis is becoming more popular and the understanding of how (not) to interpret the coefficients is improving (Braumoeller 2004; Kam and Franzese 2007), there are limits with regard to the number and complexity of interactions that can be specified and interpreted in statistical analyses (Fiss 2007:1182). Third-order interaction terms are already quite rare; and fourth order interaction terms are virtually absent in the social science literature. Yet, QCA commonly produces just this type of conjunction of four (or more) conditions.

Related to this, it is important to point out that in QCA multi-collinearity is neither a technical nor an epistemological problem (Scharpf 2000: 59). As with most other qualitative research approaches, so QCA also reckons and knows to handle the fact that social phenomena occur in clusters. Not all logically possible types of a phenomenon manifest themselves empirically. With QCA, "missing" types of cases can be detected easily and the scope of the causal claim restricted accordingly.

Finally, the concept of *asymmetric causality* (Lieberson 1985) is important when evaluating the potential of QCA for social science research. This

means that the explanation of the presence of a phenomenon – say, the existence of a welfare state – does not imply that the inverted explanation automatically accounts for the absence of the same phenomenon, that is, the non-existence of a welfare state. Quite different from most statistical procedures, which are based on correlation measures and thus assume a symmetric relation of the correlated variables, QCA links conditions and an outcome through asymmetric set-theoretical relations. Thus, the presence of a phenomenon and its absence are examined in two separate analyses. Sometimes, the results of this can be surprising and can lead to a better understanding of the phenomenon under examination. This is in clear contrast to standard quantitative research, where negative outcomes (or negative conditions) are similar to the positive outcomes (or positive conditions) and only the sign +/- differs. As a consequence of this, an explanation for a positive outcome is simply the inverse of the explanation for a negative outcome.[8]

The simultaneous incorporation of three elements of causal complexity – equifinality, conjunctural causation, asymmetric causal relationships – makes QCA different from standard statistical techniques.[9] This difference goes back to the fundamental point that quantitative methods are based on the use of correlations, not of subset relations, when examining causally relevant relations between the independent and dependent variables. In this, the aim is usually to identify the most powerful predictor for explaining variance in the dependent variable and to make probability statements for the generalization ("significance tests") from (a hopefully representative) sample to a (hopefully well defined) underlying population.

All of this makes standard statistical techniques a powerful set of tools for summarizing complex data into parsimonious equations and in filtering out the "net-effect" (Ragin 2008) of independent variables. But these

---

[8] The results of (logistic) regression, for instance, do not substantively change if the scale of the dependent variable is inverted.

[9] Of course, there are other dimensions of causal complexity than equifinality, conjunctural causality, and asymmetric causality. Above all, the time dimension has to be mentioned. It is often of decisive importance for the presence or absence of an outcome when and/or in which sequence certain factors appear. Although QCA focuses on these central questions of qualitative research, because of its close contact with the cases under examination, the formal integration of time, timing, and sequencing into the QCA algorithm is making only slow progress (Caren and Panofsky 2005; for a critique, see Ragin and Strand 2008).

techniques rest on the starting assumption that social phenomena are driven by unifinality, additivity, and symmetry. This explains why it is difficult to model equifinal, conjunctural, and asymmetric set relations in terms of sufficiency and necessity using correlation-based statistical techniques (Fiss 2007:1190).

Of course, QCA's propensity to detect certain types of complex causal structures is only an asset if there are good (theoretical) reasons to believe that a phenomenon under study is driven by such a causal structure. No method is *per se* superior. Rather, its usefulness is determined by its fit to the research problem at hand.

Certainly, the type of causal complexity which we have mentioned throughout the text also exists both in standardized statistical techniques and in (historical and non-historical) single-case studies. But it is much more restricted in these methods: in statistics, modeling causal complexity often goes hand in hand with technical problems (such as, as mentioned, the integration of interaction effects and the related loss of degrees of freedom, or the phenomenon of multicollinearity). With regard to equifinality, "[s]tandard regression methods are essentially unable to take equifinality into account" (Fiss 2007:1182).[10] In comparative case studies, causal complexity often leads to idiosyncratic explanations for single cases, and thus is achieved at the expense of generalizability of the results beyond the case(s) under examination.

While QCA resembles Mill's methods (Berg-Schlosser et al. 2008:2, 17; Mill 1865; Mahoney 2003), we deem it important to point out that QCA overcomes several important shortcomings of this commonly applied framework in the comparative social sciences (for critiques of Mill's methods, see also Lieberson 1991; Mahoney 2000; or Schneider and Wagemann 2007:73–77). Three of the most consequential flaws of Mill's methods are the following. First, hardly ever is a single condition found to be sufficient for *all* cases under examination. Instead, empirical and research reality most of the time reveals that conditions are sufficient in combination with other conditions ("conjunctural causation"), and that there are more of those causal combinations leading to the same outcome ("equifinal causa-

---

[10] For recent attempts to make statistical techniques more amenable for analysis of causal complexity, see, e.g. Braumoeller 2003 or Braumoeller and Goertz 2003. These approaches usually set high requirements for data quality and yet still have difficulties in simultaneously tackling all features of causal complexity.

tion"). Mill's methods, by and large, are not suitable for analyzing such complex causal relations. Second, it is impossible to apply Mill's methods in a meaningful way once the number of cases reaches a commonly encountered medium-sized N. Third, the application of Mill's methods makes it difficult to acknowledge and analyze the pervasive and decisive problem of "limited diversity" (see Schneider and Wagemann 2007:109ff. and Schneider and Wagemann 2006). Here also, QCA – thanks to its use of truth tables – goes further.

Variants of QCA

Many of those aspects that relate to QCA as a data analysis technique have undergone numerous modifications over the last years. One result of this is that there are different variants of QCA today. In order to avoid terminological confusion, Rihoux and Ragin (2008) introduced the term "Configurational Comparative Methods" (CCM), which comprises all types of QCA.

The original version (Ragin 1987) is referred to as "crisp set QCA" (csQCA). In set theory, a "crisp set" is one in which an element is either a member of the set or it is not (Klir, Clair, and Yuan 1997:48). In-between cases with partial memberships do not exist. For instance, the element "Tuesday" is a member in the set of "weekdays," whereas the element "finger nail" is not a member of this set. Translated into social science applications, csQCA requires conditions and outcomes to be either present or absent. While this binary structure of the data makes it possible to apply Boolean algebra, it also presents a notable, often criticized shortcoming of csQCA (e.g. Goldthorpe 1997).

As one reaction to the shortcomings, limitations, and critiques of csQCA, Ragin developed "fuzzy set QCA" (fsQCA, Ragin 2000 and 2008). In fuzzy sets, elements can have differing degrees of membership in sets. These degrees vary between full membership and full non-membership. This flexibility enables social scientists not only to make qualitative differences between countries that are, say, democratic and those that are not. It also permits them to differentiate quantitatively the degree to which countries are democracies and non-democracies, respectively. While the categorization of cases is in line with most qualitative reasoning, the possibility for gradation corresponds better to quantitative social scientific thinking about

social reality. Obviously, precision, discipline, and transparency of the codification rules – the so-called calibration of fuzzy sets – are indispensable (see Ragin 2000; Ragin 2008b and Schneider and Wagemann 2007:180ff. and below).

fsQCA operates on fuzzy algebra (Klir et al. 1997:73ff.; Kosko 1993; Smithson and Verkuilen 2006; Zadeh 1965 and 1968; for the application of fuzzy algebra, see Schneider and Wagemann 2007:220ff.). Fuzzy algebra can be seen as a general version of Boolean algebra. csQCA, therefore, is a special case of fsQCA. All rules and algorithms developed for fsQCA can also be used for csQCA, but not *vice versa.* All concepts such as necessity, sufficiency, equifinality, conjunctural causation, and even technical details such as truth tables are common to csQCA and fsQCA.

A further variant of CCM is the so-called multi-value QCA (mvQCA) (Cronqvist 2005; Cronqvist and Berg-Schlosser 2008). It allows researchers to work with multinomial concepts (such as different political party families). While in certain research situations this extension seems to offer a neat solution, some shortcomings and unresolved issues with regard to mvQCA need to be mentioned (see Schneider and Wagemann 2007:262ff. and Vink and Van Vliet 2009).

First, contrary to common claims, mvQCA notably increases the problem of "limited diversity." By allowing single conditions to have more than two values (presence and absence), more logically possible combinations of case properties are created and this decreases the likelihood that all of them are represented by empirical cases. The research-practical reverberation of this is that in mvQCA only a few conditions should be multinomial. The majority of them still need to be dichotomous. In addition, these few multinomial conditions should have only few categories, preferably not more than three or maximum four. Of course, this does not go much beyond the possibilities which the dichotomous version csQCA offers.

Second, fuzzy sets are not allowed to be used in mvQCA. Arguably, fuzzy sets contain more information than both crisp and multi-value sets, and not being able to include them in a QCA – even if a researcher has them at hand – can be seen as a step back. Third, the outcome in mvQCA must be dichotomous. No multinomial or fuzzy set outcomes are allowed.

Fourth, as Vink and Van Vliet (2009) point out, most mvQCA applications seem to be at odds with some of the set-theoretic foundations of

QCA. This is particularly true when continuous variables are calibrated into multi-value conditions with three (or more) categories, and the set-theoretic status of the middle categories remains unclear. Finally, mvQCA remains particularly vague with regard to the status and analysis of necessary conditions. In sum, while mvQCA seems to offer an intuitive improvement over csQCA, important theoretical and practical concerns limit its applicability.

The Current QCA Agenda

QCA is a relatively young method. It is still under development. Analytic features are added, new software modules developed, and innovative forms of graphical representation suggested. In the following we address some of the most important recent developments.

One important milestone in the development of the method has been the introduction of consistency and coverage measures (Ragin 2006b; Goertz 2006; Schneider and Wagemann 2007:86ff.). These parameters can be seen as a response to the critique that QCA – mostly due to its roots in formal logic and set theory – is a deterministic method, not properly suited for the analysis of notoriously noisy social science data. Consistency measures provide a numeric measure in how far the empirical data support set theoretic statement that a (combination of) condition(s) is sufficient and/or necessary. The coverage parameters, in turn, evaluate how much of the outcome is explained by every single path and by the overall solution term.

Already earlier, Ragin (2000) had proposed binomial tests, benchmark proportions, and fuzzy adjustments in order to deal with the issue of misfit between logical statements and the underlying data, due to measurement error, imprecise theories, or stochastic social phenomena (see also Dion 2003). One drawback of introducing elements of inferential statistics is that one is buying into the usual assumptions on which standard statistical techniques rely – something QCA usually attempts to overcome (Seawright 2004; Seawright 2005; Ragin 2005).

One further recent development is an increasing emphasis on the analysis of *necessary* conditions (see especially Goertz and Starr 2003). There are several reasons why, so far, analysis of necessary conditions has been

considered of secondary importance by the majority of QCA applicants. First, QCA, by default, produces sufficient conditions,[11] but not automatically also always necessary conditions – even if they are present in the data. Second, daily language invites the researcher to formulate sufficient conditions as the only explanation for a given outcome: "X *leads to* Y" seems to leave no room for additional explanatory factors. And, third, it is tempting but almost always wrong to assume that from a solution formula indicating sufficient conditions, such as

$$AB + AC \rightarrow Y \tag{5}$$

the presence of necessary conditions can be unequivocally inferred by simply factoring out condition A and writing

$$A(B + C) \rightarrow Y \tag{6}$$

The interpretation of A as a necessary condition would only be correct under very specific and rare empirical conditions. The truth table from which the solution formula has been derived would have to show no signs of limited diversity, that is, all logically possible combinations of conditions would need to be represented by empirical cases. In addition, the truth table would be free of contradictory, or inconsistent, rows.

But research practice tells that social science data hardly ever is this neat and, thus, these two requirements are hardly ever met (for a more detailed discussion, see Schneider and Wagemann 2007:63, 112ff.). Yet, by now clear guidelines exist for analysis of necessary conditions, facilitated by new menu items in the computer software fsQCA 2.0 (Ragin et al. 2006).

Another recent development concerns the practice of assigning fuzzy set membership scores to individual cases, the so-called calibration of fuzzy sets (for this step of the research process see Ragin 2000 and 2008:chap 4–5). Meaningful calibration of (fuzzy) sets is key to a successful QCA. In one of the new proposals about calibration (the "direct method of calibration"), Ragin (2008) starts from the idea that the qualitative "anchors"

---

[11] The reason for this is that QCA is based on analysis of truth tables. Each truth table row, in turn, is itself a statement of sufficiency.

of 0 (full non-membership), 1 (full membership) and 0.5 (point of indifference) are derived on the basis of theoretical information. In a next step, fuzzy set membership scores in-between these anchors are assigned with the help of a logistical (and not a proportional) function; this converts values on quantitative scales into fuzzy scores. By imposing qualitative anchors, this procedure maintains the qualitative and theory-guided character of calibrating fuzzy set while also making use of the more fine-grained information contained in the (quantitative) data. The "indirect method of calibration," in turn, starts out from more qualitatively defined thresholds and establishes a logit distribution with the help of a STATA syntax (Ragin 2008).

Both calibration strategies can only be applied if quantitative data is at hand. Even then, the calibration of fuzzy scores still depends on the type of data and a researcher's theoretical knowledge.

Researchers working with QCA will unavoidably be forced to engage with the omnipresent phenomenon of limited diversity of social science data. This becomes straightforwardly visible once the data is sorted into a truth table and not all logically possible combinations (the truth table rows) contain empirical cases. Such scarcity of data – in statistical research often referred to as empty cells – poses problems for drawing causal inference regardless of the data analysis technique used. In QCA, however, some progress is made in handling limited diversity. The standard procedure consisted in either making no assumptions about the outcome value of these logical remainders, which produces the "most complex solution," or in letting a computer algorithm generate assumptions about the outcome value such that the "most parsimonious solution" is produced (Schneider and Wagemann 2007:101ff.).

Beyond this, Ragin and Sonnett 2004 have recently proposed using "easy counterfactuals," that is, assumptions about the outcome values only for those logical remainders for which strong theoretical expectations exist. The advantage of this theory-guided treatment of logical remainders is that it does not delegate the task to a computer algorithm or method-induced assumptions while, at the same time, it avoids overtly complex solution formulas. Other contributions to tackling the problem of limited diversity comprise: warnings that contradictory simplifying assumptions should be avoided (Vanderborght and Yamasaki 2003); or reducing the

number of truth table rows by dividing QCA into two (or more) analytic steps (Schneider and Wagemann 2006; Schneider 2008).[12]

Other recent developments concentrate on progress regarding the presentation of QCA. This includes both improvements of software packages functions and suggestions on the notation and graphical representation of results generated with QCA (Schneider and Grofman 2006). Since QCA is conceptualized as a method at the interface between qualitative case-oriented and quantitative variable-oriented research, presentation and interpretation of its results have to reflect the information both on cases and on variables. It is therefore not enough to just present the solution formula, for this refers only to the relation between sets and, as a result, underemphasizes QCA's case perspective. For achieving the latter, different instruments can be applied, such as truth tables, Venn diagrams, X-Y plots and dendograms (Schneider and Grofman 2006).

Conclusion

This contribution contains an overview of the state of the art of Qualitative Comparative Analysis. We have differentiated between two essential aspects of QCA, namely, as a systematic, case-oriented approach and as a data analysis technique. The logic and epistemology of QCA are inspired by the qualitative social science research tradition. The technical issues involved when analyzing the data show resemblance to quantitative approaches. We have clarified that QCA can fully unfold its potentials only in a skilful dialogue and combination of these two phases of a research process.

As shown, causal complexity plays a central role in QCA. Causal complexity in QCA refers to the concepts of equifinal, conjunctural, and asymmetric causality. These concepts, in turn, can be modelled with the terms of necessary and sufficient conditions and be grounded in formal set theory. Such a set-theoretic reasoning is implicitly present in many social science hypotheses. Usually, it is not possible to model the same causal relations

---

[12)] In this approach, the level of causality is further differentiated by distinguishing between "remote" and "proximate" causal factors. Thus, two-step approaches do not only contribute to a solution of the problem of limited diversity, but also account for a better integration of formal solution and theoretical hypotheses. In fact, different levels of causal effects are typical for most social science theories.

with standard statistical techniques. This does not mean that one method is superior to the other but instead it points to the necessity of choosing a methodological approach according to the theories at hand. In many research situations in the comparative social sciences, QCA offers an appropriate systematic alternative to the quantitative and qualitative paradigms.

We have also demonstrated how QCA – over time – has been differentiated into different variants, in an attempt to better correspond to research practical needs when analyzing social science data. Some features of QCA, however, are still works in progress. We have pointed to new ways of: handling the calibration of fuzzy sets; dealing with limited diversity; and dealing with contradictory or inconsistent truth table rows.

The development of new features, together with a rising number of applications of QCA, increases the need for the formulation and refinement of standards of good QCA practice. While the majority of social scientists knows (or should know) how to read and write statistical analyses the same cannot be taken for granted with their understanding of how this applies to QCA. Gaps in understanding here create the dangers that QCA is applied in a vague and superficial way or, even if applied correctly, is misunderstood by readers. This neither produces reliable research results nor contributes to the recognition of the novelty and potential benefits of QCA as an additional social science research tool. In the next paper we present such a "standard of good practice."

References

Berg-Schlosser, Dirk, Gisele De Meur, Benoit Rihoux and Charles C. Ragin. 2008. "Qualitative Comparative Analysis (QCA) As an Approach." Pp. 1–18 in *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques*, edited by Benoît Rihoux and Charles C. Ragin. Thousand Oaks/London: Sage.

Brady, Henry E. and David Collier, eds. 2004. *Rethinking Social Inquiry: Diverse Tools, Shared Standards*. Lanham: Rowman & Littlefield.

Braumoeller, Bear F. 2003. "Causal Complexity and the Study of Politics." *Political Analysis* 11(3):209–33.

——. 2004. "Hypothesis Testing and Multiplicative Interaction Terms." *International Organization* 58:807–820.

Braumoeller, Bear F. and Gary Goertz. 2003. "The Statistical Methodology of Necessary Conditions." Pp. 197–223 in *Necessary Conditions: Theory, Methodology, and Applications*, edited by Gary Goertz and Harvey Starr. Lanham, MD, Boulder: Rowman & Littlefield.

Caren, Neal and Aaron Panofsky. 2005. "TQCA. a Technique for Adding Temporality to Qualitative Comparative Analysis." *Sociological Methods & Research* 34(2):147–72.

Cronqvist, Lasse. 2005. *Introduction to Multi-Value Qualitative Comparative Analysis (MVQCA)*. Marburg. Internet: http://www.tosmana.net.

——. 2006. *Tosmana – Tool for Small-N Analysis [SE Version 1.25]*. Marburg. Internet: http://www.tosmana.net.

Cronqvist, Lasse and Dirk Berg-Schlosser. 2008. "Multi-Value QCA (MvQCA)." Pp. 69–86 in *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques*, edited by Benoit Rihoux and Charles C. Ragin. Thousand Oaks/London: Sage.

De Meur, Gisele and Benoit Rihoux. 2002. *L'Analyse quali-quantitative comparée. Approche, techniques et applications en sciences humaines*. Louvain-La-Neuve: Bruylant Academia.

Dion, Douglas. 2003. "Evidence and Inference in the Comparative Case Study." Pp. 95–112 in *Necessary Conditions: Theory, Methodology, and Applications*, edited by Gary Goertz and Harvey Starr. Lanham, MD, Boulder: Rowman & Littlefield.

Dusa, Adrian. 2007. *QCA Graphical User Interface Manual*. Compasss Working Paper, WP 2007–50.

Fiss, Peer. 2007. "A Set-Theoretic Approach to Organizational Configurations." *Academy of Management Review* 32(4):1180–1198.

Goertz, Gary. 2003. "The Substantive Importance of Necessary Condition Hypotheses." Pp. 65–94 in *Necessary Conditions*, edited by Gary Goertz and Harvey Starr. Lanham: Rowman & Littlefield Publishers.

——. 2006. "Assessing the Trivialness, Relevance, and Relative Importance of Necessary and Sufficient Conditions in Social Science." *Studies in Comparative International Development* 41(2):88–109.

Goertz, Gary and Harvey Starr, eds. 2003. *Necessary Conditions: Theory, Methodology, and Applications*. Lanham, MD, Boulder: Rowman & Littlefield.

Goldthorpe, John H. 1997. "Current Issues in Comparative Macrosociology: a Debate on Methodological Issues." *Comparative Social Research* 16:1–26.

Kam, Cindy D. and Robert J. Jr. Franzese. 2007. *Modeling and Interpreting Interactive Hypotheses in Regression Analysis*. Ann Arbor: University of Michigan Press.

Katz, D. and R. L. Kahn. 1978. *The Social Psychology of Organizations*. New York: Riley.

King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry. Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.

Klir, George J., Ute St. Clair, and Bo Yuan. 1997. *Fuzzy Set Theory. Foundations and Applications*. Upper Saddle River, NJ: Prentice Hall PTR.

Kosko, Bart. 1993. *Fuzzy Thinking: the New Science of Fuzzy Logic*. New York: Hyperion.

Lieberson, Stanley. 1985. *Making It Count: the Improvement of Social Research and Theory*. Berkeley: University of California Press.

——. 1991. "Small N's and Big Conclusions: An Examination of the Reasoning in Comparative Studies Based on a Small Number of Cases." Pp. 105–18 in *What Is a Case? Exploring the Foundations of Social Inquiry*, edited by Charles C. Ragin and Howard S. Becker. Cambridge: Cambridge University Press.

Longest, Kyle C. and Stephen Vaisey. 2008. "Fuzzy: a Program for Performing Qualitative Comparative Analysis (QCA) in Stata." *The Stata Journal* 8(1):79–104.

Mackie, John L. 1974. *The Cement of the Universe*. Oxford: Oxford University Press.

Mahoney, James. 2000. "Strategies of Causal Inference in Small-N Analysis." *Sociological Methods & Research* 28(4):387–424.

——. 2003. "Strategies of Causal Assessment in Comparative Historical Analysis." Pp. 337–72 in *Comparative Historical Analysis in the Social Sciences*, edited by James Mahoney and Dietrich Rueschemeyer. Cambridge: Cambridge University Press.

——. 2008. "Toward a Unified Theory of Causality." *Comparative Political Studies* 20 (forthcoming).

Mill, John S. 1865. "Of the Four Methods of Experimental Inquiry." Pp. 427–50 in *System of Logic, Ratiocinative and Inductive, Vol. 1, Book 3, Chap. VIII, Sixth Edition*, edited by John S. Mill.

Ragin, Charles C. 1987. *The Comparative Method. Moving Beyond Qualitative and Quantitative Strategies*. Berkeley/Los Angeles/London: University of California Press.

——. 1994. *Constructing Social Research. the Unity and Diversity of Method*. Thousand Oaks: Pine Forge Press.

——. 2000. *Fuzzy-Set Social Science*. Chicago: University of Chicago Press.

——. 2003. *Making Comparative Analysis Count*. Compasss Working Paper, WP 2003–6.

——. 2004. "Turning the Tables: How Case-Oriented Research Challenges Variable-Oriented Research." Pp. 123–38 in *Rethinking Social Inquiry: Diverse Tools, Shared Standards*, edited by Henry E. Brady and David Collier. Lanham: Rowman & Littlefield.

——. 2005. "Core Versus Tangential Assumptions in Comparative Research." *Studies in Comparative International Development* 40(1):33–38.

——. 2006a. "The Limitations of Net-Effect Thinking." Pp. 13–41 in *Innovative Comparative Methods for Policy Analysis. Beyond the Quantitative-Qualitative Divide*, edited by Benoit Rihoux and Heike Grimm. New York: Springer.

——. 2006b. "Set Relations in Social Research: Evaluating Their Consistency and Coverage." *Political Analysis* 14(3):291–310.

——. 2008. *Redesigning Social Inquiry: Set Relations in Social Research*. Chicago: University of Chicago Press.

Ragin, Charles C. and Benoit Rihoux. 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, Charles C. and John Sonnett. 2004. "Between Complexity and Parsimony: Limited Diversity, Counterfactual Cases, and Comparative Analysis." *Vergleichen in der Politikwissenschaft*, edited by Sabine Kropp and Michael Minkenberg. Wiesbaden: VS Verlag für Sozialwissenschaften.

Ragin, Charles C. and Sarah Strand. 2008. "Using Qualitative Comparative Analysis to Study Causal Order. Comment on Caren and Panofsky (2005)." *Sociological Methods & Research* 36(4):431–41.

Ragin, Charles C., Drass, Kriss A., and Davey, Sean. 2006. *Fuzzy-Set/Qualitative Comparative Analysis 2.0*. Tucson, Arizona: Department of Sociology, University of Arizona.

Rihoux, Benoit and Charles C. Ragin, eds. 2008. *Configurational Comparative Methods. Qualitative Comparative Analysis (QCA) and Related Techniques*. Thousand Oaks/London: Sage.

Scharpf, Fritz W. 2000. *Interaktionsformen. Akteurzentrierter Institutionalismus in der Politikforschung*. Opladen: Leske+Budrich.

Schneider, Carsten Q. 2008. *The Consolidation of Democracy: Comparing Europe and Latin America*. London: Routledge.

Schneider, Carsten Q. and Grofman, Bernard. 2006. *It Might Look Like a Regression… but It's Not! An Intuitive Approach to the Presentation of QCA and Fs/QCA Results*. Compasss Working Paper, WP2006–39.

Schneider, Carsten Q. and Claudius Wagemann. 2006. "Reducing Complexity in Qualitative Comparative Analysis (QCA): Remote and Proximate Factors and the Consolidation of Democracy." *European Journal of Political Research* 45(5):751–86.

——. 2007. *Qualitative Comparative Analysis (QCA) und Fuzzy Sets. Ein Lehrbuch für Anwender und alle, die es werden wollen*. Opladen and Farmington Hills: Verlag Barbara Budrich.

Seawright, Jason. 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.

——. 2005. "Assumptions, Causal Inference, and the Goals of QCA." *Studies in Comparative International Development* 40(1):39–42.

Smithson, Michael and Jay Verkuilen. 2006. *Fuzzy Set Theory: Applications in the Social Sciences*. Thousand Oaks: Sage Publications.

Vanderborght, Yannick and Yamasaki, Sakura. 2003. *The Problem of Contradictory Simplifying Assumptions in Qualitative Comparative Analysis (QCA)*. Paper presented at the ECPR General Conference, Marburg, 18–21 September.

Vink, Maarten and Olaf Van Vliet. 2009. "Not Quite Crisp, Not Yet Fuzzy? Assessing the Potentials and Pitfalls of Multi-value QCA." *Field Methods* 21(3):265–289.

Wagemann, Claudius. 2007. "Qualitative Comparative Analysis und Policy-Forschung." *Die Zukunft der Policy-Forschung*, edited by Frank Janning and Katrin Toens. Wiesbaden: VS Verlag für Sozial-wissenschaften.

Zadeh, Lofti A. 1965. "Fuzzy Sets." *Information and Control* 8:338–53.

——. 1968. "Fuzzy Algorithms." *Information and Control* 12:99–102.