Assessing the Quality of Qualitative Comparative Analysis (QCA)

–

Evaluation, Improvement, Application

Von der Philosophischen Fakultät
der Gottfried Wilhelm Leibniz Universität Hannover
zur Erlangung des Grades eines
Doktors der Philosophie (Dr. phil.) genehmigte
Dissertation

von Jonas Buche

geboren am 21.08.1982 in Rostock

2017

Referentin: **Prof. Dr. Monika Jungbauer-Gans**
(Leibniz Universität Hannover &
Deutsches Zentrum für Hochschul- und Wissenschaftsforschung)

Korreferentin: **Prof. Dr. Christiane Gross**
(Julius-Maximilians-Universität Würzburg)

Tag der mündlichen Prüfung: 18.05.2017

Danksagung

Ich danke meinen Doktormüttern und meiner Familie.

Abstract/Zusammenfassung

Starting from the observation that the number of empirical applications of the set theoretic social science method Qualitative Comparative Analysis (QCA) founded by Charles C. Ragin in 1987 raised considerably over the last ten years, the cumulative dissertation asks: *"What is the concept of quality of QCA and how is it implemented in QCA applications?"* Based on a systematic conceptualization and operationalization of the "quality of QCA" the research question is answered in three steps. First, a thorough *evaluation* of the quality of QCA in empirical applications is conducted. It covers 139 articles published between 2006 and 2016 in peer reviewed journals from three scientific areas, namely sociology, political science, and business research. Second, with the aim to deliver an *improvement* of the quality of QCA the formulas for the calculation of consistency and coverage (Ragin 2006) are critically discussed and updated by the exclusion of cases that are irrelevant for the respective set relation. Finally, third, a best practice *application* of QCA is delivered that takes into account all criteria for a high quality QCA as conceptionalized in the beginning.

Ausgehend von der Beobachtung, dass die Anzahl empirischer Anwendungen der 1987 von Charles C. Ragin präsentierten mengentheoretischen Methode Qualitative Comparative Analysis (QCA) in den letzten zehn Jahren merklich angestiegen ist, stellt die vorliegende kumulative Dissertation die Frage: „*Wie lässt sich die Qualität von QCA konzeptualisieren und wie wird sie in Anwendungen von QCA umgesetzt?*" Die Forschungsfrage wird auf der Grundlage einer systematischen Konzeptualisierung und Operationalisierung der „Qualität von QCA" in drei Schritten beantwortet. Erstens wird eine gründliche *Evaluation* der Qualität von QCA in empirischen Anwendungen durchgeführt. Diese umfasst 139 zwischen 2006 und 2016 in Zeitschriften mit peer review publizierte Artikel aus den drei Bereichen Soziologie, Politikwissenschaft und Business Research. Mit dem Ziel einer *Weiterentwicklung* der Qualität von QCA werden, zweitens, die Formeln zur Berechnung von Konsistenz und Abdeckung kritisch diskutiert und durch den Ausschluss von für die jeweilige Mengenbeziehung irrelevanten Fällen angepasst. Schließlich wird, drittens, eine best practice *Anwendung* von QCA vorgestellt, die sämtliche eingangs skizzierten Kriterien des Konzepts der Qualität von QCA beachtet.

Keywords/Schlagworte

Qualitative Comparative Analysis (QCA); Quality of QCA; Evaluation of QCA Applications; Parameters of Fit in QCA; Consistency & Coverage

Qualitative Comparative Analysis (QCA); Qualität von QCA; Evaluation von QCA-Anwendungen; Gütemaße in QCA; Konsistenz & Abdeckung

Table of Content

List of tables

List of figures

0.

Zusammenfassung der Dissertation in deutscher Sprache

Summary of the doctoral thesis in German language

(Jonas Buche)

Qualitative Comparative Analysis (QCA) bezeichnet eine mengentheoretische Methode zur vergleichenden Fallanalyse, die von dem Soziologen Charles C. Ragin (1987, 2000, 2008) in den 1980er Jahren entwickelt wurde und sich mittlerweile im sozialwissenschaftlichen Methodenkanon etabliert hat. In den letzten Jahren ist einerseits ein deutlicher Zuwachs an publizierten Anwendungen von QCA in hochrangigen Zeitschriften zu verzeichnen (Compasss 2016). Andererseits verschaffen sich auch kritische Stimmen stärker Gehör, die anhand von Beispielstudien aufzeigen, welche methodischen Schwierigkeiten mit einer QCA einhergehen können (z.B. Hug 2013; Krogslund et al. 2015; Seawright 2005: 24).

Die kumulative Dissertation setzt sich zum Ziel, die Qualität von QCA als Methode sowie die Qualität von QCA Anwendungen nicht nur empirisch zu erheben, sondern auch zu evaluieren und zu verbessern. Aus diesem Grund werden insgesamt fünf thematisch zusammenhängende Fachartikel gemeinsam mit einer gesonderten Abhandlung, welche den thematischen Zusammenhang darlegt, als Dissertationsschrift vorgelegt.

Die Arbeit ist wie folgt aufgebaut: Die einführende Rahmenschrift (Kapitel 1) stellt einerseits das inhaltliche Vorgehen der fünf Fachartikel vor und enthält darüber hinaus eine umfangreiche Konzeptformierung der „Qualität von QCA-Anwendungen“. Insgesamt werden fünf Arbeitsschritte, die bei der Durchführung einer QCA wichtig sind, präsentiert und jeweils mit Indikatoren unterlegt.

Zwei der fünf Fachartikel (Kapitel 2 & 3) enthalten sodann Evaluationen von insgesamt 138 Anwendungen von QCA, die in der Soziologie (Buche & Siewert 2015) und der Business- und Managementliteratur (Wagemann et al. 2015) zu verorten sind. Zusätzlich wird eine einzelne Studie aus dem Bereich der Politikwissenschaft vertieft evaluiert und darüber hinaus repliziert (Kapitel 4, Buche et al. 2016). Die Evaluation aller Studien erfolgt jeweils anhand des zuvor spezifizierten Konzepts und zeigt im Wesentlichen, dass die Qualität der untersuchten Anwendungen von QCA den festgelegten Gütekriterien in weiten Teilen nicht zu entsprechen vermag.

Der vierte Artikel (Kapitel 5, Buche 2016) enthält eine Weiterentwicklung der methodischen Qualität von QCA. Ausgehend von der Unterscheidung von relevanten und irrelevanten Fällen in der Analyse von Mengenbeziehungen wird der Einfluss irrelevanter Fälle zu vermeiden gesucht und entsprechende Formeln zur Berechnung von relevanter Konsistenz und Abdeckung von Mengenbeziehungen erstellt. Zusätzlich wird eine weitere Formel präsentiert, die es ermöglicht, den verzerrenden Einfluss schiefer Verteilungen von Mitgliedern und Nicht-Mitgliedern in Mengen auf die Analyse von Mengenbeziehungen zu erkennen.

Der fünfte und letzte Artikel (Kapitel 6, Buche 2017) umfasst schließlich eine eigene thematische Anwendung von QCA aus dem Bereich der europaorientierten Parlamentsforschung. Es handelt sich um eine Analyse auf der Individualebene, die Interaktionsmustern von schwedischen Reichstagabgeordneten in nationalen und europäischen Sachfragen in vergleichender Perspektive untersucht. Bezugnehmend auf das im Rahmenpapier erstellte Konzept der Qualität von QCA werden sämtliche Erfordernisse an eine Analyse von hoher Qualität umgesetzt.

Im abschließenden Kapitel (7) wird die übergreifende Frage nach der methodischen Qualität von QCA und ihrer Anwendungen anhand der Ergebnisse aus den fünf Artikeln beantwortet.

1.

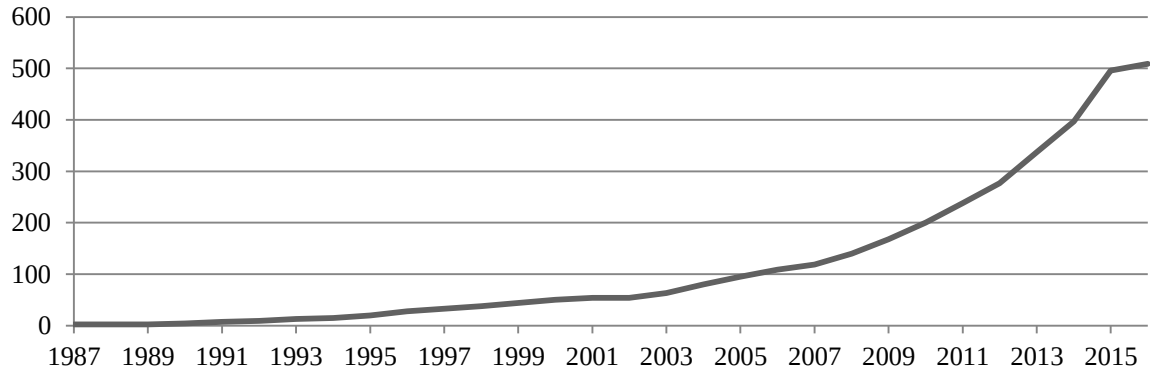
Conceptual Framework and Proceeding of the Cumulative Dissertation

(Jonas Buche)

1.1 Introduction

Qualitative Comparative Analysis (QCA) has become an important part of the methodological tool kit in the social sciences over the last 30 years. With his seminal contribution "The Comparative Method. Moving Beyond Qualitative and Quantitative Strategies", the American sociologist Charles C. Ragin (1987) laid the foundation of a new methodological perspective on social phenomena. Based on Boolean algebra QCA combines logic operation and set theoretic thinking with the aim of "extending the logic of case studies to comparative analysis" (Mahoney 2010: 133). Since Ragin's first textbook, QCA has been further developed and refined (Ragin 2000, 2006, 2008; Rihoux & Ragin 2009; Schneider & Wagemann 2007, 2010, 2012, 2013; Rohlfing 2015a/b), combined with other methods (Schneider & Rohlfing 2013, 2014; Rohlfing & Schneider 2013) and, finally, applied to the study of social reality. Especially in the last decade, the number of empirical applications of QCA raised considerably. While from 1987 to 2006 approximately 100 studies using QCA have been published in peer reviewed articles, this number has more than quintupled since then (see figure 1.1). Just in 2015, approximately 100 studies have been recognized by the bibliographic database of the compasss (2016) network. In other words, the "Ragin revolution" (Vaisey 2009) does not just continue but accelerates.

Figure 1.1 Number of articles published in peer-reviewed journals, since 1987



Source: Compasss 2016.

However, the rising number of applications does neither entail information about the quality of QCA as a method nor about the quality of empirical applications of QCA – except that all these studies underwent a peer-review process. Instead it strikes the eye that with the increasing number of applications critical articles against QCA raised considerably over the last years as well. For instance, some scholars criticized QCA to be prone to measurement error (Hug 2013) or confirmation bias (Krogslund et al. 2015) while others point out the "centrality

of untested assumption" (Seawright 2005: 24). As a reference, these researchers use several example applications of QCA substantiating their claims. Nevertheless, in his textbook on social science methodology John Gerring (2012: 350) states about QCA that "the potential utility of a method should be differentiated from its actual employment". Put differently, methodologically questionable applications of QCA should not be used as a reference for QCA not being a valuable method *per se*. Although this statement could be read as a call for a lively debate on both the quality of QCA as such *and* the quality of QCA applications, there are no systematic analyses on the latter so far. There are descriptive overviews on the use of QCA between 1984 and 2011 in general (Rihoux et al. 2013) or in specific fields like public policy research (Rihoux et al. 2009) and organization research (Marx et al. 2013). But methodological evaluations on the quality of QCA-based research are limited to two studies: On the one hand, there is an article on 19 QCA applications published in peer reviewed journals in the field of comparative welfare-state research (Emmenegger et al. 2013). Although the authors characterize this field as least likely for low quality QCAs they discreetly conclude that "many scholars can make more out of using QCA" (Emmenegger et al. 2013: 190). On the other hand, there is a conference paper on 21 applications from a variety of fields such as sociology, comparative politics, and international relations (Mello 2012). Although all of these studies have undergone a peer review process Mello detects several methodological problems and concludes that "in order to utilize the full potential of QCA as a method, researchers need to reconsider its underlying assumptions" (Mello 2012: 1). In the following I argue that not only but also because of the accelerating publication rates in peer reviewed journals there is a huge need for a systematic evaluation of the actual methodological quality of QCA applications. Consequently, I pose the question:

"What is the concept of quality of QCA and how is it implemented in QCA applications?"

In order to answer this question I, first, need to clarify the concept of 'quality of QCA' in detail. Already in 2010, Schneider and Wagemann (2010) proposed standards of good practice in QCA defining "what a "good" QCA-based research entails" (Schneider & Wagemann 2010: 1; emphasis by the authors). In the introductory chapter of this cumulative dissertation I will, on the one hand, display the basics of QCA and guidelines for a good QCA based on the 26 proposals Schneider and Wagemann (2010) updated toward a "receipt for a good QCA" in later work (Schneider & Wagemann 2012: 275 ff.). On the other hand, I will form a three-level-concept of the "quality of QCA" by identifying and operationalizing five research steps at the secondary level that every QCA-based research needs to be engaged in.

Second, I need to deliver a thorough evaluation of the quality of QCA in empirical applications, i.e. whether *actual practices* comply with the suggested *best practices* or not. This is conducted in chapters 2 to 4 of the cumulative dissertation that draw on three co-authored scientific articles:

> Buche, Jonas, and Markus B. Siewert. 2015. "Qualitative Comparative Analysis (QCA) in der Soziologie – Perspektiven, Potentiale und Anwendungsbereiche." *Zeitschrift für Soziologie* 44 (6):386-406.
>
> Wagemann, Claudius, Jonas Buche, and Markus B. Siewert. 2016. "QCA and Business Research: Work in Progress or a Consolidated Agenda?" *Journal of Business Research* 69 (7):2531–2540.
>
> Buche, Antje, Jonas Buche und Markus B. Siewert. 2016. "Fuzzy Logic or Fuzzy Application? A Response to Stockemer's "Fuzzy Set or Fuzzy Logic?"" *European Political Science* 15 (2): 359-378.

These three studies cover evaluations of 139 articles published between 2006 and 2016 in peer reviewed journals from three scientific areas, namely sociology, political science, and business research.

With the aim of improving the quality of QCA, third, I extend the debate to the discussion on the 'parameters of fit' (Ragin 2006). The measures coverage and consistency can be seen as the major criteria for the explanatory power of QCA in general and the assessment of quality of set relations in particular. In chapter 5 of the cumulative dissertation I argue that both the formulas for consistency and coverage of both sufficiency and necessity might largely be driven by irrelevant cases. By that, they contradict the notion of asymmetry which can be seen as one major argument in favor of QCA (Ragin 2008, Schneider/Wagemann 2012). Next to highlighting the potential problems, I offer a possible update of these formulas that allows to avoid the influence of irrelevant cases on the assessment of consistency and coverage of set relations. Moreover, I discuss the implications of the updated formulas on uncovering skewed membership distributions and offer a formula for 'non-skewedness' in QCA. Chapter 5 bases on the single authored article "Relevant Consistency and Relevant Coverage in Fuzzy Set Qualitative Comparative Analysis (fsQCA)" which has not yet been published.

Finally fourth, I apply QCA to a specific topic myself. Chapter 6 of the cumulative dissertation works with a mix of qualitative and quantitative data collected by the author. It takes into account all guidelines of the 'standards of good practice' explained in the introductory part and avoids the pitfalls and weaknesses identified in the evaluation of published studies. This chapter draws on the single authored scientific article:

> Buche, Jonas. 2017. "Europeanization of Legislative-executive Relations at the Micro Level – Under Which Conditions Do Swedish MPs Interact with Ministerial Officials?" *COMPASSS Working Paper Series* 2017-87.

To conclude, the idea of the individual articles and the dissertation as a whole is not to speculate about reasons for the low quality of QCA applications. Instead, the idea is to systematically conceptualize the quality of QCA, to empirically evaluate the quality of QCA, to mathematically improve the quality of QCA, and to thoroughly apply a 'best practice' QCA of high quality. The introductory chapter demonstrates why the five scientific articles included into the cumulative dissertation just jointly fulfill its aim to assess the quality of QCA and, thus, need to be seen as one piece of research. Chapter 1 proceeds as follows. In section 1.2 the theoretical foundations, the variants, and the actual analytical technique of QCA are presented. The conceptualization of 'quality of QCA' is conducted in section 1.3. Out of the presentation and discussion of guidelines and standards of good practices in QCA the concept of quality of QCA is formed. This can be used for the subsequent evaluation of empirical applications. Additionally, section 1.3.3 offers a detailed examination of the operationalization of the concept's indicators. In section 1.4 the five papers included in the cumulative dissertation are presented in more detail. Moreover, my contributions to the respective co-authored articles are made clearly distinguishable and assessable as the individual scientific achievement. The final section concludes.

1.2 Theoretical and methodological basics of QCA

A systematic conceptualization of the quality of QCA requires not only a thorough understanding of how QCA works as a method but also of the theoretical foundation underlying it. Following the differentiation of Ragin (2000, 2008; see also Schneider & Wagemann 2010, 2012) between *QCA as a research approach* and *QCA as an analytic tool* subsequently both the set theoretic underpinning and the technique of QCA are presented.

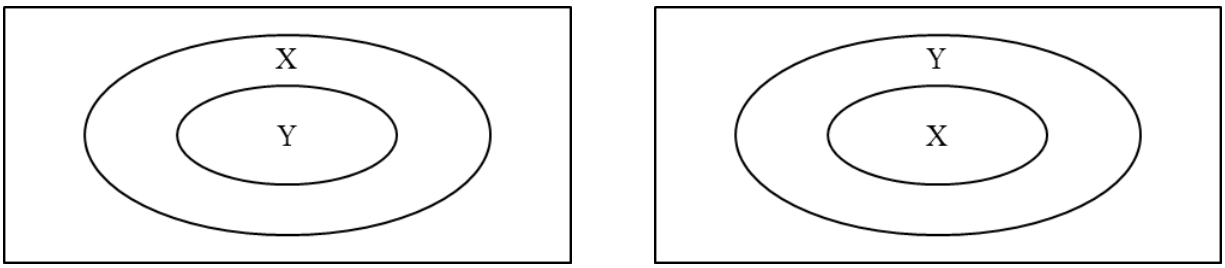
1.2.1 QCA as a Research Approach

Qualitative Comparative Analysis is a comparatively new way of systematic case comparisons in the social sciences. In its early days it has been promoted as a "third way" between the variable-oriented quantitative and the case-oriented qualitative research paradigm (Ragin 1987: 16-17; 34-68). In his later work, Ragin (2000, 2008) spelled out the importance of a clear case orientation over all of the analytical steps of QCA (see Schneider & Wagemann 2012; Blatter & Haverland 2012; Goertz & Mahoney 2012; Thiem et al. 2016). The positioning becomes obvious in the acronym QCA with the Q standing for *qualitative*. However, QCA aims at cross-case comparisons and the "possibility of formulating statements about broad, cross-case patterns" (Ragin 2000: 120). This is achieved by accentuating the set-theoretic basis of his comparative method. Sets are the main building blocks of the analysis,

i.e. cases are seen as configurations of memberships in purposively defined sets. Moreover, set relations of necessity and sufficiency are assumed to describe and explain complex social phenomena (Ragin 2008: 29-68; Schneider & Wagemann 2012: 56-90).

If one, for instance, is interested in explanations for the set of "people with a high life satisfaction" one possible explanatory factor could be the set of "people with excellent health". In order to analyze whether an excellent health (condition X) is necessary for a high life satisfaction (outcome Y) one has to check if whenever Y is present, X is present, too. Put differently, the necessity statement implies that there cannot be a highly satisfied person that is not displaying an excellent health status. Taking a set theoretic perspective, this means that the outcome is a subset of the condition set or, the other way around, the condition is a superset of the outcome set (see figure 1.2, left hand side).

Figure 1.2 Venn-diagrams for necessary and sufficient conditions



Necessary condition *Sufficient condition*

On the contrary, a condition is sufficient for an outcome if whenever the condition is present the outcome is present as well. To stick to the hypothetic example again; if X is sufficient for Y there cannot be any person with excellent health that is not highly satisfied with his or her life. In set theoretic terms, all healthy people are a subset of satisfied people (see figure 1.2, right hand side). While this simple example explains the set theoretic foundation of QCA, it does not take into account cases as configurations of set memberships. This changes if one adds another explanatory condition, for instance next to the set of "people with excellent health" (A) the set of "people in partnership" (B). Out of these two conditions, four configurations or ideal types appear: "healthy person in partnership" (A*B)[1], "healthy person not in partnership" (A*~B), "not healthy person in partnership" (~A*B), and "not healthy person not in partnership" (~A*~B). Adding another condition such as "high socio-economic status" (C) would lead to eight logically possible configurations.

[1] As QCA is based on Boolean algebra, the multiplication sign "*" equals the logical AND, the addition sign "+" equals the logical OR, and the tilde "~" implies the absence of a set. Moreover, arrows signal either necessity "←" or sufficiency "→" set relations.

Ragin (2000: 120) calls this the "set-theoretic foundation for diversity-oriented research". The importance of diversity is not limited to the understanding of cases as different configurations of set memberships. Rather, first, diversity orientation also applies for the explanatory power of configurations explaining an outcome. Instead of the average net effect of single variables on a specific outcome, QCA takes into account configurations of explanatory conditions jointly being necessary or sufficient for an outcome. Necessary combinations of conditions are so called SUIN conditions. SUIN is the acronym for "sufficient but unnecessary part of a factor that is insufficient but necessary for the result" (Mahoney et al. 2009). As an example, the combination (*union*) of living in partnership OR a high socio-economic status could be a necessary condition for a high life satisfaction. In Boolean terms the solution would look like this: B+C ← Y. Condition B and C are mutually substitutable necessary conditions, i.e. SUIN conditions. If a person does not live in partnership, this single – according to this example – must have a high socio-economic status in order to be able to achieve high life satisfaction.[2] For sufficiency on the contrary, an INUS condition is an "insufficient but necessary part of a condition which is itself unnecessary but sufficient for the result" (Mackie 1974: 62). Following the example again, a possible sufficient configuration (*intersection*) could be described like this: A*B → Y. This means that a high life satisfaction could be achieved by the combination of excellent health AND living in a partnership. A and B are INUS conditions, because each of them is a necessary part of a configuration which in turn is sufficient for the outcome.

Moreover second, QCA offers the opportunity to detect equifinal solutions. Instead of one explanatory model that covers the whole sample several mutually non-exclusive configurations of conditions might explain the occurrence of a certain outcome for different (clusters of) cases. For instance, next to the configuration of excellent health AND partnership (A*B) the combination of being a single AND having a high socio-economic status (~B*C) could also be detected as sufficient for a high life satisfaction: A*B + ~B*C → Y.

Finally, third, due to its set theoretic foundation QCA deals with asymmetric hypothesis about social phenomena. This clearly differs from linear hypothesis such as "the healthier a person (A), the higher its life satisfaction (Y)" which implies "the lower A, the less Y" as well. Instead, QCA applies set theoretic hypothesis like "if A, then Y". From the hypothetical sufficiency statement that whenever a person is member in the set of people with excellent health it will also be in the set of people with a high life satisfaction (A → Y), one cannot draw any

[2] More precisely, the logical OR in QCA is inclusive which means that A+B needs to be read as A or B or both.

[3] Note that the 0.5 anchor itself is not allocated to any of the cases since it defines the phase shift from mem-

conclusion on whether the outcome is present or not if the condition is absent (Ragin 2008: 176-187; Schneider & Wagemann 2012: 76-89). The asymmetry of set relations can also be exemplified with the condition "living in partnership" (B) in the hypothetical equifinal solution term A*B + ~B*C → Y. While in the first configuration (A*B) partnership in its presence is sufficient for the outcome, in the second configuration (~B*C) the absence of partnership is sufficient. In short, configurational thinking, equifinality, and asymmetry are the main pillars of QCA as a diversity-oriented comparative research approach.

1.2.2 Variants of QCA

In principle one can distinguish three variants of QCA; crisp set, fuzzy set, and multi value QCA. In the first place, Ragin (1987) developed QCA for dichotomous crisp sets (csQCA) that separate cases into members and non-members of sets. For example the proposed condition set C would include people with high socio-economic status (membership score 1) and exclude people without high socio-economic status (membership score 0) indicating the difference in kind between the two groups. However, the need of dichotomization and the enormous loss of information coming along with it raised a great deal of criticism (for an overview, see De Meur et al. 2009). As a result, Ragin (2000) transferred the idea of fuzzy sets (Zadeh 1965, 1969; Klir et al. 1997) from informatics to social science methodology. Fuzzy sets (fsQCA) enable researchers to differentiate gradations of set membership (scores between 0.5 and 1) and non-memberships (scores 0 and 0.5; see also Schneider & Wagemann 2012: 24-31; Ragin 2008: 29-34). Still, the dichotomy persists and the set membership score of 0.5 as the "point of maximum ambiguity" (Ragin 2008: 30) defines the transition point between membership and non-membership. But in addition to the difference in kind, fuzzy sets allow for differences in degree by allocating gradual (non-)memberships. Thus, the loss of information among different people with a "more-than-not" membership in the set of high socio economic status (membership scores above 0.5) can be avoided. The same is true for the group of non-members (scores below 0.5). The third variant to award set memberships to cases has been developed by Cronqvist (2005; Cronqvist & Berg-Schlosser 2009). Multi-value QCA (mvQCA; see comments by Schneider & Wagemann 2012: 255 ff.; Thiem 2013; Vink & Van Vliet 2013) allows to transfer multi-nominal concepts into set memberships (scores 0, 1, 2, 3, etc.). Here, several differences in kind can be displayed. For instance, the membership in the set "partnership" could take the score 0 indicating to be a single, score 1 to be in a partnership but not married, score 2 to be married, score 3 to be divorced, and so forth. However, the empirical importance of mvQCA is rather limited so far. While csQCA and fsQCA have

been utilized in approximately 250 publications in peer reviewed journals each, the use of mvQCA did not exceed 13 applications until September 2016 (Compasss 2016).

Table 1.1 Crisp and fuzzy set memberships of example "life satisfaction"

Cases	Set membership scores			
	A	B	C	Y
Tom	1	1	1	1
Jerry	1	1	0	1
Tina	0	0	1	1
Linda	1	0	1	1
Toby	0	0	1	1
Bob	0	0	0	0
Larry	1	0	0	0
Rob	0	1	0	0
Berta	0	1	1	0

Crisp set data

Cases	Set membership scores			
	A	B	C	Y
Tom	0.8	1	0.6	1
Jerry	0.6	0.6	0	0.8
Tina	0.2	0	0.8	0.6
Linda	1	0.2	1	1
Toby	0.4	0.4	0.8	0.6
Bob	0	0.4	0	0
Larry	1	0.2	0.2	0.4
Rob	0.2	0.8	0.4	0.2
Berta	0	1	0.8	0

Fuzzy set data

Notes: Condition A: set of people with excellent health, condition B: set of people in partnership, condition C: set of people with high socio-economic status, condition Y: set of people with high life satisfaction

Table 1.1 refers to the hypothetical example discussed above and displays nine imaginary cases' set memberships in crisp sets and fuzzy sets. Whenever the crisp membership score is 1 on the left hand side indicating a full membership in the respective set, the fuzzy score on the right hand side is higher than 0.5 also indicting a (more-than-not) membership. All crisp memberships scores 0 are transferred into fuzzy increments of non-membership (scores below 0.5).[3] Thus, regarding the difference in kind the information in both tables is the same. As pointed out above, the fuzzy membership scores additionally allow for differences in degree of (non-)memberships.

1.2.3 QCA as an Analytical Tool

The "analytic moment" (Ragin 2000) displays the actual data analysis technique of QCA. Although it exposes the potential of QCA as an analytical tool for the examination of set relations, it must be separated from QCA as a research approach (see section 1.3; see also Ragin 2000, 2008; Schneider & Wagemann 2010, 2012). The analysis of necessity and sufficiency are two separate analytical steps. Software tools like fsQCA (latest version Ragin/Davey 2014), TOSMANA (Cronqvist 2016), and several packages for the R environment („QCA3" by Huang 2014, „QCA" by Duşa & Thiem 2014, „cna" by Ambuehl et al. 2015, "QCA(GUI)

[3] Note that the 0.5 anchor itself is not allocated to any of the cases since it defines the phase shift from membership to non-membership and vice versa. Thus, a case with a membership score of 0.5 is neither a member nor a non-member in the respective set.

by Duşa 2016", and "QCApro" by Thiem 2016) offer appropriate help. However, the rather low complexity of the crisp set example data in table 1.1 allow for an analysis by hand.

Remember, a condition is necessary if the outcome cannot occur without the condition being present. Put differently, one has to examine if any (combination of) condition(s) is also present whenever the outcome is present. Turning the attention to the imaginary cases in table 1.1, five cases are members of the outcome set (Tom, Jerry, Tina, Linda, and Toby) but no single condition is also present over all of them. Nevertheless, whenever the outcome is present either condition B (see cases Tom and Jerry) OR condition C (see cases Tina, Linda, and Toby) is present. Thus, as pointed out above the combination of either being in partnership OR having a high socio-economic status is necessary for a high life satisfaction in this hypothetical example: B+C←Y.[4]

For the analysis of sufficiency, QCA makes use of so called truth tables (see the truth table for the hypothetical example in table 1.2.). Truth tables display all logically possible combinations of conditions. Thus, every truth table row is one ideal typical configuration of all conditions; for instance the combination of the presence of A AND the presence of B AND the absence of C: A*B*~C (truth table row 2 in table 1.2). For the analysis, every empirical case is allocated to its ideal type, i.e. the respective truth table row (see table 1.2, column 6).

Table 1.2 Truth table of crisp set example "life satisfaction"

Truth table row	Truth value of condition in configuration			Truth value of sufficiency	Cases covered by configuration
	A	B	C		
1	1	1	1	1	Tom
2	1	1	0	1	Jerry
3	1	0	1	1	Linda
4	1	0	0	0	Larry
5	0	1	1	0	Berta
6	0	1	0	0	Rob
7	0	0	1	1	Tina, Toby
8	0	0	0	0	Bob

Since the hypothetical explanatory model covers three explanatory conditions and every condition can be either present or absent (truth value in configuration, see column 2, 3, and 4), there are eight possible logical combinations of conditions (truth table rows). The last column covers the empirical cases allocated to their respective ideal type configuration. Since two cases (Tina and Toby) fit the same ideal type, they are assigned to truth table row 7 (configu-

[4] The same is true for the combination of condition A+~B. See for a detailed discussion section 1.3.3, step III.

ration ~A*~B*C). Finally, the column "truth value of sufficiency" displays whether the given configuration is sufficient for the outcome (score 1) or not (score 0). To be clear, the truth table is not the analysis of sufficiency itself. Rather, it serves as a tool for structuring and processing the empirical data in QCA; it "is at the core of QCA, both in the understanding of it as an approach and as a technique" (Schneider & Wagemann 2012: 91). The actual analysis of the information on sufficiency contained in a truth table is the reduction of complexity according to the algorithm by Quine and McCluskey (Ragin 1987: ch. 6; see for a different minimization algorithm e.g. Baumgartner 2009; Baumgartner & Thiem 2015). Given the truth value of sufficiency (column 5), in a first step one can easily see that four configurations are sufficient for the outcome (i.e. truth table rows 1-3 and 7). In Boolean algebra the solution term is the following: A*B*C + A*B*~C + A*~B*C + ~A*~B*C → Y.[5] Apparently, the information contained in that Boolean term is highly complex. The reduction takes place in a second step by "matching similar conjunction" (Schneider & Wagemann 2012: 105ff.). As a comparison of the first (A*B*C) and the second (A*B*~C) sufficient truth table row in the Boolean solution reveals, the combination of conditions A AND B is sufficient for the outcome regardless of the presence or absence of condition C. Thus, the information on C or ~C is superfluous, and by excluding it the solution term gets less complex (A*B). Likewise, the other two sufficient truth table rows (A*~B*C + ~A*~B*C) uncover that the combination of ~B AND C is sufficient for Y irrespective the presence or absence of A. Again, matching these similar conjunctions by excluding superfluous information makes the Boolean solution term more parsimonious (~B*C). To be accurate, also the first (A*B*C) and the third (A*~B*C) truth table row could be logically minimized. Regardless the presence or absence of condition B, the configuration A*C alone is sufficient for Y as well. To conclude, by matching similar conjunctions the complexity of sufficiency statements has been minimized from the union of four entire truth table rows (so called *primitive expressions* since the display the most complex, i.e. primitive statement of sufficiency) to three reduced conjunctions of only two conditions each (so called *prime implicants* since they cannot be reduced any further).

[5] This reads as the combination of "the presence of excellent health AND the presence of partnership AND the presence of high socio-economic status" OR the combination of "the presence of excellent health AND the presence of partnership AND the absence of high socio-economic status" OR the combination of "the presence of excellent health AND the absence of partnership AND the presence of high socio-economic status" OR "the absence of excellent health AND the absence of partnership AND the presence of high socio-economic status" is sufficient for "the presence of high life satisfaction".

However, in order to possibly uncover an even more parsimonious sufficiency statement the Quine-McCluskey-algorithm in a third step tries to detect redundant prime implicants. Put differently, it looks for the minimum number of prime implicants that still cover the entire sufficiency information of all primitive expressions, i.e. sufficient truth table rows. This is conducted by making use of a prime implicant chart (see table 1.3).

Table 1.3 Prime implicants chart of crisp set example "life satisfaction"

		Primitive expressions / sufficient truth table rows			
		A*B*C	A*B*~C	A*~B*C	~A*~B*C
Prime implicants	A*B	X	X		
	~B*C			X	X
	A*C	X		X	

Prime implicant charts contain information on both primitive expressions (columns) and prime implicants (rows). Moreover, they display the sub- and superset-relations between the two.[6] For instance, the prime implicant A*B is a superset of the primitive expressions A*B*C and A*B*~C and, consequently, covers the sufficiency information of both. In the chart this is indicated by the X. Likewise, prime implicant ~B*C covers the primitive expressions A*~B*C and ~A*~B*C. All four sufficient truth table rows from the hypothetical truth table (table 1.2) are already covered by the two configurations A*B and ~B*C. Thus although the third prime implicants A*C also covers two primitive expressions, this additional information is considered to be redundant for the analysis aiming at a parsimonious solution.[7] To conclude, by using the Quine-McCluskey algorithm, the highly complex solution term consisting of four sufficient combinations (A*B*C + A*B*~C + A*~B*C + ~A*~B*C → Y) of three conditions each can be narrowed down to only two configurations containing only two conditions each (A*B + ~B*C → Y).

1.3 Conceptualization of the Quality of QCA

After the short theoretical and technical introduction to QCA in the following I identify five basic research steps of every applied QCA and connect them with the code of good standard

[6] Put differently, prime implicant charts display the result of the minimization process by matching similar conjunctions.

[7] It goes without saying that a mechanic reduction of complexity might be counterproductive when searching for theoretically proper sufficient configurations. The search for the minimum number of prime implicants might hide alternative explanations for social phenomena (see for the problem of model ambiguities Thiem 2014a).

in QCA (Schneider & Wagemann 2010) to a comprehensive concept of "quality of QCA". Moreover, I discuss the operationalization of the concept by explaining the respective indicators used throughout the evaluation.

1.3.1 Code of Conduct in QCA Applications

Both the set theoretic foundation of QCA as a research approach and the mathematical basis of QCA as an analytic tool are relatively easy to understand. But even the hypothetical example uncovers a couple of difficulties in the actual application of QCA. To start with, the allocation of set memberships to cases, a process that is called set calibration in QCA, poses questions. For instance, how can one define the (fuzzy) membership score of a person in a set such as "people with high life satisfaction". Moreover, there might be more than two or three potential explanatory factors that need to be included into the analysis. But how many conditions should or can be integrated into the explanatory model – and how many cases? Furthermore, in the example all statements of necessity and sufficiency have been deterministic. But the analysis of social reality more often than not reveals deviances from perfect subset-superset relations. How to deal with contradictory cases that are members of the same configuration of conditions but differ in their membership in the outcome set? For instance, the imaginary cases Toby and Tina are both not excellent healthy singles with a high socio-economic status (configuration ~A*~B*C; see table 1.1). If one of them would be unsatisfied with his or her life, the researcher could no longer decide whether the configuration ~A*~B*C is sufficient for Y or not – since half of the empirical evidence confirms sufficiency while the other half disproves. This problem of inconsistencies in sufficiency (and necessity) is even more present in fuzzy set QCA (see Schneider & Wagemann 2012: ch. 5). Finally, the diversity-orientation of QCA asks for a large variety of configurations of conditions in order to describe set relations. In contrast to the hypothetical example with ideal data, in social science reality not all logically possible configurations might always be covered by empirical cases. How to deal with this limited empirical diversity?

Because questions like these arise during the application of QCA, Schneider and Wagemann (2010: 1) felt the need to provide a "guideline for authors, reviewers, and readers of QCA" covering the most problematic issues.[8] They arrange 26 proposals (from A to Z) according to the differentiation of QCA as a research approach and QCA as a tool for the analytic mo-

[8] See for the first versions of the list in their German textbook on QCA (Schneider & Wagemann 2007: 266ff.) and Schneider & Wagemann 2008. For an update see Schneider & Wagemann 2012: 275ff. See for a similar approach Rihoux & Ragin 2009.

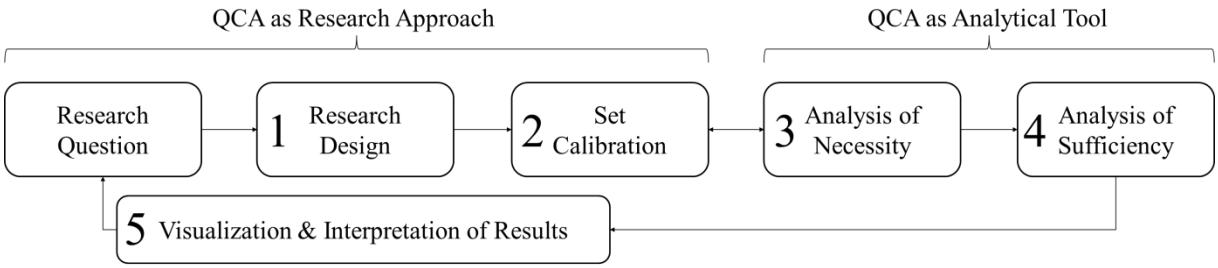
ment.[9] The phase before to the analytical moment concentrates on research design questions such as the number of and selection criteria for cases and conditions. Schneider and Wagemann (2010) emphasize the justification of case (non-)selection (*proposal d*), and the need of theoretical and/or empirical knowledge for the selection of a moderate number of conditions (*proposals e and f*). Moreover, they stress the importance of familiarity with cases (*proposal c*). This combination of in-depth case knowledge and detailed theoretical and conceptual knowledge is the basis for transparent and convenient calibration (*proposal g*). During the analytical moment the authors suggest to analyze necessity and sufficiency in consecutive steps (*proposal i*) and to check for contradictions prior to the actual analysis of sufficiency (*proposal j*). Still, as (slight) deviances of set relation mostly cannot be avoided they call for transparent (*proposal m*) and justified (*proposal l*) decisions on the level of inconsistencies included into the explanatory model. The same is true for the treatment of truth table rows without empirical evidence, so called logical remainder rows (*proposal n*). After the analytical moment users of QCA should consider the case-orientation of QCA (*proposals r and s*) and carefully interpret the Boolean solution terms (*proposals t, u, and v*). Finally, transparency also applies to the availability of research material which makes the results reproducible, for instance a matrix covering the not yet calibrated raw data (*proposal w*), the truth table (*proposal x*), the parameters of fit (*proposal z*), and the solution formulas (*proposal y*).

1.3.2 Concept of Quality of QCA

Schneider and Wagemann (2010: 3) state that "while some of these proposals might seem obvious, evidence from research reality suggests that this is not the case for everybody." This, on the one hand, corresponds to the finding by Mello (2012: 1) that "in order to utilize the full potential of QCA as a method, researchers need to reconsider its underlying assumptions". On the other hand, it reveals that a systematic evaluation of the actual quality of QCA application has to consider both the methodological basics outlined above as "QCA as a research approach" and the genuine analysis of set relations as outlined in the section on "QCA as an analytical tool". Of course, decisions in one phase have an impact on the other phase as well. But based on the differentiation of QCA as approach and analytic tool one can identify five steps of every QCA-based research needs to be engaged in (see figure 1.3).

[9] Note that not all the alphabet is considered since not all proposals are important for the conceptualization of the quality of QCA but regard, for instance, a call for mixed methods designs (proposal b).

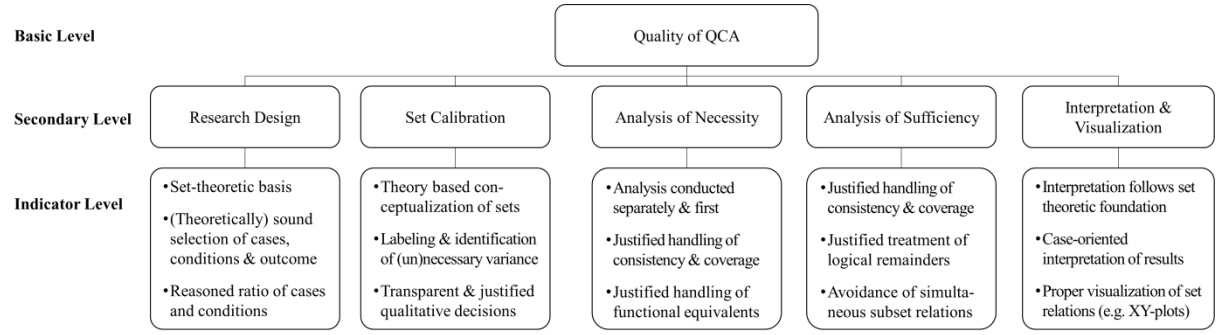
Figure 1.3 Research steps of QCA as an approach and as an analytic tool



Source: own illustration.

These five steps are the basis for the concept of "quality of QCA" formed and applied in the following. According to concept formation literature by Goertz (2006a: ch. 2), the quality of QCA can be conceptualized in three levels: basic level, secondary level, and indicator level (see figure 1.4).

Figure 1.4 Three-level-concept of Quality of QCA



Source: own illustration according to Goertz 2006: 50.

While the five research steps compose the secondary level, at the indicator level the most important proposals of Schneider and Wagemann (2010) are echoed. In the next subsection, the indicator level is discussed for every research step in more detail.

1.3.3 Operationalizing the Concept of Quality of QCA

In order to evaluate the quality of QCA in applications the different parts of the concept have to be operationalized. In other words, the indicators for the respective research steps provided in figure 1.4 need to be at hand in a measurable format. While for the two more technical steps 'analysis of necessity and sufficiency' it is easier to identify quantifiable indicators, for 'research design', 'calibration', and 'interpretation' the indicators discussed in the following strongly relies on each other and, consequently, must be discussed in connection.

Step I – research design

In his book on "Constructing Social Research" Ragin (1994: 26) defines a research design as "a plan for collecting and analyzing evidence that will make it possible for the investigator to answer whatever questions he or she has posed". Obviously this broad claim is not limited to studies making use of QCA. However, a research design that builds on this comparative method comes along with some requisites.

First, as outlined above QCA is deeply rooted in set theory and raises hypothesis that state (mostly) asymmetric and configurational set relations of necessity and/or sufficiency. This strongly differs from covariational hypothesis as applied in regressional analytic methods (see Thiem et al. 2016). Thus next to an appropriate terminology not mixing together the dissimilar fundamentals of these two research approaches, a high quality QCA-based research design will relate its plan for collecting and analyzing evidence to this theoretical foundation. More precisely, the use of QCA as one (or even the most) valuable tool for answering a certain question will not be motivated by the number of cases but by set theory, i.e. the benefits of configurational thinking, equifinality, and asymmetry for the analysis of social phenomena.

To be very clear, the number of cases itself is of no interest for the evaluation of the quality of a QCA-based research design. Ragin (2000: 25) suggests that QCA is especially useful for studies with a mid-sized *n*, i.e. between 10 and 50 cases. This traces back to the understanding of QCA as a third way between case studies and statistics, since a mid-sized *n* might both exceed the complexity of (comparative) case study research and might not allow robust statistical analyses. However, next to that this argument should not be used as a justification for the application of QCA in general, the mid-sized *n* argument is asymmetric. While it defines QCA to be very useful for a medium *n*, it does not say anything about the usability of QCA for analyses with less than 10 or more than 50 cases. Empirically, most applications of QCA deal with a medium number of cases (see chapter 2 and 3). But one can also find QCAs dealing with less than 10 (e.g. Achilov & Shaykhutdinov 2013; Da Roit & Weicht 2013) or even several thousands of cases (for instance Amenta et al. 2009; Glaesser & Cooper 2011, 2012a, 2012b). More important conceptually, every *n* comes along with strength and difficulties a high quality QCA research design has to deal with. To start with, while in QCAs with a very small *n* the familiarity with cases is very high they might lack empirical diversity which leads to the problem that many truth table rows will not be covered by cases; i.e. the number of logical remainders raises. In contrast, while large-*n* QCAs might offer (but do not guarantee) greater diversity, they come along with a lack of detailed case knowledge that Ragin (2000,

2008) emphasizes. Heterogeneity among cases does not only have a limiting effect on the number of logical remainders. It also increases the probability of inconsistent set relations, i.e. that cases with the same configuration differ in their membership in the outcome set (see for the discussion of consistency and coverage the following steps 3 and 4, and chapter 5). Finally, a medium number of cases might either combine empirical diversity, low levels of inconsistencies and profound case knowledge – or fall victim to all of these problems. To conclude, based on the sheer number of cases included the quality of a research design cannot be judged. Instead, any evaluation of the quality of QCA needs to consider the 'number of cases' when dealing with the 'case selection strategy', the 'number & selection of conditions' included into the model, and the 'case-oriented interpretation' of the results, respectively.

Thus second, the selection of cases is a very important factor for the evaluation of the quality of QCA regarding the research design. Also for the selection of cases different good practices are possible. Given a certain population of units several types of probability or non-probability sampling might be as convincing as including the whole population into the analysis – as long as the case selection strategy is transparent and justified. To get to the core of the matter, findings of a QCA in principle are not generalizable since set relations are examined only for those cases included into the analysis. Adding cases will (sometimes only slightly) change the results, and even removing cases will influence the empirical importance of certain findings or might even change them significantly. Thus, a meaningful case selection that takes into account both the population and scope conditions important for the specific research question is essential for a high quality QCA research design.

Third, also the number of conditions included into the explanatory model is an essential part of the quality of a QCA research design and needs to be evaluated in combination with the number of cases. As outlined above (see section 1.2.1), in QCA cases are seen as configurations of (non-)memberships in purposively defined sets. When three conditions are included into the analysis, every case is either (partial) member or (partial) non-member in each of the three conditions which generates eight different configurations ($3^2=8$). Every additional condition doubles the number of logically possible configurations. This is important to be aware of since the number of configurations quickly exceeds the number of cases. If one, for instance, includes five conditions ($5^2=32$) into an analysis on the EU member states (N=28; as of Nov 2016) there will be at least four configuration (32-28=4) without empirical evidence. This has enormous effects on the construction (and analysis) of the truth table (compare section 1.2.3) since the decision whether a configuration is sufficient for the outcome or not cannot be based on empirical information by default. Given the problem of limited empirical di-

versity, i.e. several EU member states might display the same configuration of memberships and non-memberships, probably even more than four truth table rows will not be covered by cases. However, while the latter problem is an empirical one since cases might cluster, the former one goes back to the research design. To conclude, a high quality QCA-based research design provides a reasoned ratio of number of conditions and cases.

Finally fourth, also the selection of conditions and the outcome needs to be justified in a high quality research design. Again, this claim is not limited to studies using QCA since research designs will always need appropriate theoretical underpinnings of *explanans* and *explanandum* in order to convince potential readers. However, the selection of conditions displays a central feature of QCA, namely the iterative "back and forth between theory and empirical evidence" (Ragin 2008: 78). While a proper theoretical foundation of the respective condition is necessary for the inclusion into the model, the decision for this explanatory factor might be subject to change in a later research step due to the (lack of) empirical importance for the selected cases. Put differently, at any point of the research process Ragin (2000: ch. 2) explicitly calls for an iterative adaption of the explanatory model in order to come up with a sufficiently specified truth table that offers the best possible analysis of the social phenomena under scrutiny. Still, this process needs to be comprehensible, i.e. both transparent and reasonable.

To sum up, the evaluation of the quality of a QCA-based research design can be operationalized by the set-theoretic foundation of the research plan, the ratio of cases and conditions, and the (theoretically) reasoned selection of cases and conditions.

Step II – set calibration

The assignment of set membership scores to single cases is called calibration. During this research step empirical data need to be (meaningfully) transformed into sets. To be clear, the importance of set calibration cannot be overestimated and it should be seen as the centerpiece of every QCA. Or as Ragin (2008: 104) puts it: "Set-theoretic analysis without careful calibration of set membership is an exercise in futility." Interested in the quality of calibration, though, the following criteria can be utilized for its evaluation.

First, as set calibration is always an interpretation of empirical data by the researcher, it must offer a strong link between theory and data. Put differently, a proper calibration bases on a clear concept formation (Goertz 2006a; Adcock & Collier 2001). According to Goertz (2006a: 5) concepts are "theories about the fundamental constitutive elements of a phenomenon." As such concepts both outline the theoretical underpinning of phenomena and reveal the connection to measurement and data collection. This is also suggested by Ragin's (2008: 78) idea of

an “iterative back and forth between ideas and evidence”. Any concept underlying set calibration does not only build upon theory, but also includes knowledge about selected cases and the empirical data at hand. Out of the interaction of theory and empirical data purposeful allocations of set membership scores are possible. Thus, the quality of calibration, in the first instance, can be evaluated against the conceptual connection of theory and empirical data. A high quality calibration will strongly focus on this interaction.

Moreover, sets are not variables. The awareness of the differences can be evaluated by considering two major principles of set calibration: the idea of phase shifts (Ragin 2008: 73) and, in fsQCA, the distinction between necessary and unnecessary variance (Ragin 2008: 83).

To start with, second, the idea of phase shifts in calibration points to the fact that sets in the first place differentiate members from non-members (see section 1.2.2). More precisely, the calibration of sets requires the conceptually based definition of a phase shifting point from which on cases turn from members to non-members or vice versa, i.e. the 0.5-anchor. For illustration Ragin (2008: 72) takes the example of the Celsius temperature scale as it is purposively calibrated according to the volume of water. Here, phase shifts take place at 0°C and at 100°C and both of them clearly demarcate qualitative breakpoints. It goes without saying that these clear cut qualitative breakpoints are (mostly) not at hand when applying the ideas of phase shifts to the assessment of social phenomena. Thus, in order define the “point of maximum ambiguity” (Ragin 2008: 30) purposefully, sets need to be labeled. For instance, in the hypothetical example presented above I am interested in the outcome set of “people with a *high* life satisfaction”. In contrast to the variable “general life satisfaction” that for example the German Socio Economic Panel Study (SOEP 2016) collects on the basis of peoples’ individual self-attribution on a 0 (low) to 10 (high) scale, the label *high* defines the target set. The change of variable’s symmetric conceptual poles “*high* and *low*” to set theoretic asymmetric poles “*high* and *non-high*“ life satisfaction enables the researcher to assigns the phase shift on a conceptual, not a numerical basis. For instance, cases with a *medium* life satisfaction such as values 4, 5, or 6 on the SOEP indicator clearly cannot be (more than not) members of the target set of “people with a *high* life satisfaction”. Moreover, if one changed the set’s label into “people with a *very, very high* life satisfaction” not even persons with a self-ascribed general life satisfaction value of 7 or 8 might be appropriate members of this set. In more general terms however, regardless whether dichotomous crisp sets or fuzzy sets are utilized, the conceptual underpinning of the phase shift decision is central to set calibration. During evaluation the labeling of conditions as well as the non-allocation of the 0.5 anchor to any case can be used as indicators for a high quality calibration.

Third, whenever calibration goes beyond crisp sets the distinction of necessary and unnecessary variance is an important factor for the quality of calibration. Next to the decision on the transition point between membership and non-membership described above, calibration of fuzzy sets requires to assign another two qualitative breaking points, i.e. *full* membership (score 1) and *full* non-membership (score 0). Again closely connected to the conceptually defined label of the respective set, empirical variation in the underlying variable might be deemed as unnecessary. Thus, during calibration empirical variance might be truncated in a way that the target set reflects the underlying concept indicated by the chosen label. Taking the example set of people with a *high* life satisfaction again, it might conceptually be unnecessary to distinguish (too many) different instances of non-high life satisfaction. Thereby, during calibration all persons with an individual self-attribution of general life satisfaction between 0 and 5 (according to the SOEP variable) could be regarded as *full non-members*, i.e. assigned a membership score of 0 in the set of people with a *high* life satisfaction. By that decision, the relevant variance for the calibrated set differs remarkably from the original variable's variance the calibration bases on. To conclude, in addition to a meaningful label the identification of conceptually relevant and irrelevant variation of variables also signals a high quality set calibration.

Table 1.4 displays fictitious raw data for the nine cases from the example (compare table 1.1) according to a hypothetical self-ascription of general life satisfaction as of SOEP (2016). Furthermore, it offers two fundamentally different calibration options (calibration I and II) for the example outcome set "people with a *high* life satisfaction". The options differ according to the decisions on the qualitative anchors, i.e. full (non-)membership (scores 1 and 0) and the crossover-point (score 0.5). Calibration I can hardly be identified as a calibration as explained above but is more a simple transformation of the variable's concept into numbers between 1 and 0. By that, for instance, case 'Rob' is neither a member nor a member of the set as indicated by the score 0.5. In contrast, calibration II takes into consideration the chosen label '*high*' life satisfaction in the way explained above. Thus, the anchor for the phase shift between membership and non-membership changes from the mathematical middle category (SOEP-value 5) applied in calibration I to a conceptual breaking point defined between the SOEP-values 6 and 7 at 6.5. Furthermore, calibration II truncates the irrelevant variation by setting the new anchor for full membership at SOEP-value 9 and the full non-membership at SOEP-value 5 "so that the resulting membership scores faithfully reflect the target set's label" (Ragin 2008: 84). For instance, case 'Rob' is a non-member of the set now.

Table 1.4 Raw data and different calibrations of the example set "people with a high life satisfaction"

Cases	**Raw data**	**Calibration I** *Anchor for full membership:* ***10*** *Anchor for phase Shift:* ***5*** *Anchor for full non-membership:* ***0***			**Calibration II** *Anchor for full membership:* ***9*** *Anchor for phase Shift:* ***6.5*** *Anchor for full non-membership:* ***5***		
		Crisp sets	Qualitative Fuzzy sets	Direct Fuzzy sets	Crisp set	Qualitative Fuzzy set	Direct Fuzzy sets
Tom	10	1	1	0.95	1	1	0.99
Jerry	8	1	0.8	0.86	1	0.8	0.86
Tina	7	1	0.7	0.77	1	0.6	0.65
Linda	9	1	0.9	0.92	1	1	0.95
Toby	7	1	0.7	0.77	1	0.6	0.65
Bob	4	0	0.4	0.35	0	0	0.01
Larry	6	1	0.6	0.65	1	0.4	0.27
Rob	5	0.5	0.5	0.5	0	0.2	0.05
Berta	1	0	0.1	0.08	0	0	0

Notes: Raw data: hypothetical self-ascription of general life satisfaction according to SOEP (2016), condition Y: set of people with high life satisfaction,

Lastly, each of the two options I and II is divided into a crisp set calibration, and two different fuzzy calibrations, a 'qualitative' and the 'direct' one. The 'qualitative' fuzzy set calibration bases on another round of qualitative decisions on relevant variation within members and within non-members of cases. For instance, (ideally) based on conceptual or theoretical reasoning I allocate a higher membership (score: 0.8) to cases with the value 8 in the raw data than to cases with a general life satisfaction value of 7 (score: 0.6). Direct calibration, instead, bases on a mathematical transformation of interval-scale data by "using estimates of the log odds of full membership" (Ragin 2008: 87; see for a detailed explanation Ragin 2008: ch. 5). Direct fuzzy calibration is implemented in several software solutions (e.g. based on a logarithmic membership function Ragin/Davey 2014; see for different membership functions for calibration Thiem 2016) and allows very detailed differentiations of set membership score. While the direct calibration might be considered as mathematically more exact, from a more qualitative standpoint very fine grained set membership differences might be hardly distinguishable. For instance, one might question the qualitative difference between Tom's (score 0.99) and Linda's (score 0.95) membership in the example outcome set "people with a *high* life satisfaction" (see table 1.4). One way to combine 'qualitative' and 'direct' calibration is 'indirect' set calibration (Ragin 2008: 94 ff.). To conclude, fourth, given researchers' both conceptual and technical flexibility in set calibration, transparency and justification are major criteria for the evaluation of the quality of set calibration. Regarding transparency, a high quality QCA will make raw data and set data available as well as openly present the qualita-

tive decisions on the phase shift anchor and the full (non-)membership anchors. Besides, a high quality QCA will provide a detailed discussion and justification of these decisions since only this allows the respective reader to fully comprehend set calibration.

To sum up, the evaluation of the quality of a QCA-based set calibration can be operationalized by the strength of the conceptual link between theory and data, the handling of irrelevant variation, and the intersubjectivity of qualitative decisions indicated by open and transparent reasoning.

Step III – analysis of necessity

QCA as an analytical tool strongly emphasizes set relations of sufficiency rather than necessity. Given the centrality of truth tables and of logical minimization this "sufficiency bias" (Schneider & Wagemann 2012: 220) in QCA does not really surprise. However, the analysis of necessity is far from being superfluous. A necessary condition X is defined as whenever the outcome Y is present, X has to be present as well. Put differently, to find a necessary condition means to identify a factor X without which the social phenomena of interest Y simply cannot occur. Thus, the analysis of necessity is of great importance in QCA and needs to meet certain quality criteria.

First, the analysis of necessity is not part of the sufficiency test but needs to be conducted as a separate research step. In contrast to the minimization of truth tables the check for necessary conditions analytically bases on the data matrix only (compare section 1.2.3). Examining the necessity of a certain (conjunction of) condition(s) for the outcome means, in the first place, to check that there is no case that is a member in the outcome set without being a member in the necessary (conjunction of) condition(s). Mathematically, all cases' membership scores in Y must not be higher than their membership scores in the necessary condition. For instance, the analysis of necessity for the crisp set example data (see table 1.1) revealed that all cases with a membership in the set of people with high life satisfaction (Y) are either members in the set of people in partnership (B) or in the set of people with high socio-economic status (C); i.e. B+C←Y. In mathematical terms, all cases' membership scores in the necessary condition B+C are greater than or equal to Y; i.e. $(B+C)_i \geq Y_i$ (see table 1.5, left hand side).[10]

However second, analyses of real world social phenomena (almost) never discover deterministic necessity statements and the problem of inconsistent set relations needs to be faced. Es-

[10] In accordance to Boolean algebra, a case's membership score in the union B+C is the maximum of the respective single set memberships.

pecially when fuzzy sets are applied, deviances from perfect set relations are almost inevitable. In fuzzy set QCA, a condition is necessary if all cases' *partial* membership scores in the condition are greater than or equal to the *partial* membership in the outcome set. Using the hypothetical example again, table 1.5 displays case by case whether or not the necessity statement B+C←Y is consistent. As examined before (see section 1.2.3) the dichotomous crisp sets fulfill the 'greater than or equal to'-postulate. In contrast, using the more fine grained calibrated fuzzy set data two cases (Jerry & Larry) are less members in the condition B+C than in the outcome set (see table 1.5, right hand side).

Table 1.5 Analysis of necessity of example "life satisfaction" with crisp and fuzzy data

Cases	**Crisp set membership scores**		**Consistent i.e. B+C ≥ Y**	**Fuzzy set membership scores**		**Consistent, i.e. B+C ≥ Y**
	B+C	Y		B+C	Y	
Tom	1	1	✓	1	1	✓
Jerry	1	1	✓	0.6	0.8	✗
Tina	1	1	✓	0.8	0.6	✓
Linda	1	1	✓	1	1	✓
Toby	1	1	✓	0.8	0.6	✓
Bob	0	0	✓	0.4	0	✓
Larry	0	0	✓	0.2	0.4	✗
Rob	1	0	✓	0.8	0.2	✓
Berta	1	0	✓	1	0	✓
Consistency of necessity			1.00	Consistency of necessity		0.91
Coverage of necessity			0.67	Coverage of necessity		0.64

Notes: Condition B: set of people in partnership, condition C: set of people with high socio-economic status, condition Y: set of people with high life satisfaction

Whenever a case is less a member in the condition than in the outcome set, the statement of necessity needs to be questioned since – in a simplified manner – there is more (partial) outcome present than the necessary condition is able to explain. A decade ago, Charles Ragin (2006) engaged in this problem and introduced the measures of consistency and coverage for the evaluation of set relations in social research (see for a critical discussion chapter 5). The aim was to deliver "simple descriptive measures for evaluating the strength of the empirical support for theoretical arguments describing set relations" (Ragin 2006: 292). The consistency measure reveals how much the necessity claim deviates from a perfect set relation. The second to last row in table 1.5 exhibits that B+C←Y is not fully consistent but slightly inconsistent. A consistency score of 0.91 can be understood in a way that the necessity statement is true for 91% of the fuzzy outcome set. At the same time it also means that 9% of the fuzzy outcome set is wrongly explained if one says that the configuration B+C is necessary for Y. Note, however, that the consistency score does not entail information on the number of inconsistent cases. It only calculates the fuzzy deviance from a perfect superset-subset-relation.

Whether or not a researcher classifies a condition as (quasi-)necessary on the basis of a 0.91 consistency is, finally, another qualitative decision. While Schneider and Wagemann (2012: 143) claim that the consistency threshold for a necessary condition should not be lower than 0.9, the actual decision needs to be substantiated with good arguments.[11] Next to consistency, Ragin (2006) developed the measure coverage of necessity which reveals how non-trivial a necessity statement is (see for a detailed discussion see Goertz 2006b; Schneider & Wagemann 2012: 235; and chapter 5). For instance, a condition like the "set of human beings" would be a perfectly trivial necessary condition for the analysis of high life satisfaction. Regardless whether the outcome is present or absent, every case will be member of the condition. Thus, the "set of human beings" would also be necessary for the absence of the outcome since it is a constant. The last row in table 1.5 shows that the combination B+C is not a constant (7 of 9 cases are (more than not) members) but only about two thirds of the set is covered by the outcome set. However, regarding the quality of the analysis of necessity one can conclude that a high quality QCA application will provide an informed and transparent utilization of these measures.

Moreover third, it is advisable to check for necessity prior to sufficiency (see Schneider & Wagemann 2010) for at least two reasons. On the one hand, finding a necessary condition might be useful for the analysis of sufficiency, especially for the treatment of logical remainders (see research steps 'research design' and 'analysis of sufficiency'). Remember, if a condition is necessary for the outcome, there simply cannot be any sufficient configuration that does not include this condition. The very same is true for all those configurations that are not covered by cases, i.e. logical remainders. Thus, analyzing necessity first enables the researcher to exclude all those logical remainders from the truth table analysis that do not show the necessary condition (see for the so called 'enhanced standard analysis' Schneider & Wagemann 2012: 200 ff.). On the other hand, analyzing necessary conditions first prevents researchers to incorrectly deduce necessity claims from results on sufficient configurations (see for the so called 'false' necessary conditions Schneider & Wagemann 2012: 220 ff.). If, for example, every sufficient solution term contains one and the same condition, this does not automatically mean that this condition is necessary for the result. Instead, this might be due to inconsistent truth table rows or the treatment of logical remainders (see research step 'analysis of sufficiency'). Thus, QCAs of high quality will analyze necessity both separately and prior to sufficiency in order to avoid this risk.

[11] See for possible reasoning by graphical assistance research step V: 'interpretation and visualization'.

Fourth, and finally, the assessment of necessity also needs to be based on theoretical and conceptual reasoning. Specifically this means that necessary configurations should be contrasted with their theoretical benefit (Schneider & Wagemann 2012: 74 f.). Consider that the necessity statement B+C←Y means that without either a partner or a high socio-economic status no person in the data set will be highly satisfied with his or her life. While this finding technically is relatively robust, its theoretical benefit is rather limited since on the conceptual level it is hard to argue that 'partnership' and 'high socio-economic status' act as functional equivalents for a higher-order concept (Adcock & Collier 2001). The same is true for the configuration A+~B identified as a necessary condition in footnote 3. According to this finding either to be healthy OR *not* to be in partnership (or both) enables a person to be highly satisfied. However, claiming that either of these unions is a necessary condition for the outcome is conceptually not plausible since they are no functional equivalents for one higher-order concept. In contrast, for instance a union of the sets "people with high education degree" OR "people with high income" could serve as functional equivalents for the higher-order concept "people with high social status". Generally speaking, OR-combination of two or even more conditions might be identified as necessary relatively easy. Still, next to the analytical support a convincing theoretical argument needs to be available as well. A high quality analysis of necessity will either limit itself to the detection of single necessary conditions or bring up substantive arguments that the necessary union combines functional equivalents for one higher-order concept.

To sum up, the evaluation of the quality of the analysis of necessity in QCA can be operationalized by checking whether the analysis has been conducted at all, conducted separately and prior to the analysis of sufficiency, and conducted as well as understood correctly by utilizing the measures consistency and coverage of necessity and by tracing the theoretical benefit of the necessity claim.

Step IV – analysis of sufficiency

As outlined before the analysis of sufficiency is the analytical core of QCA. In contrast to necessity, the analysis of sufficiency starts off from the greatest possible complexity, i.e. truth table rows, and tries to reduce it by logical minimization. A sufficient configuration such as A*B is defined as whenever the condition is present, the outcome Y is present as well. In mathematical terms, all cases' membership scores in the sufficient condition are smaller than

or equal to Y; i.e. $(A*B)_i \leq Y_i$.[12] However, the analysis of sufficiency is not free of potential pitfalls that need to be addressed in any application, too. Thus, the treatment of these issues can fruitfully be utilized as indicators for the evaluation of the quality of sufficiency analyses in QCA.

First, also sufficient set relations are hardly ever perfect when social reality is analyzed and the measures of consistency and coverage of sufficiency needs to be treated with care. As explained above, the analysis of sufficiency starts with the construction of a truth table. Table 1.6 displays the truth table for the fuzzy set example.

Table 1.6 Truth table of fuzzy set example "life satisfaction"

Truth table row	Truth value of condition in configuration			Truth value of sufficiency	Consistency of sufficiency	Cases covered by configuration
	A	B	C			
1	1	1	1	1	1.00	Tom
2	1	1	0	1	1.00	Jerry
3	1	0	1	1	1.00	Linda
4	1	0	0	**?**	0.78	Larry
5	0	1	1	0	0.44	Berta
6	0	1	0	0	0.50	Rob
7	0	0	1	**?**	0.88	Tina, Toby
8	0	0	0	0	0.63	Bob

On the first sight it seems to be similar to the crisp set based truth table (table 1.2). It also consists of eight truth table rows each representing one ideal typical configuration of the presence and/or absence of the three conditions A, B, and C. Moreover, the last column covers the empirical cases allocated to their respective ideal type configuration and the column "truth value of sufficiency" displays whether the given configuration is sufficient for the outcome (score 1) or not (score 0). However, but a closer look reveals the column 'consistency of sufficiency' has been added. It presents information on how consistently sufficient the respective truth table row is for the outcome. In mathematical terms, the consistency score calculates the respective configuration's deviance from a perfect subset relation indicated by $X_i \leq Y_i$.[13] The consistency of sufficiency value reveals the truth table rows 1 to 3 are perfectly consistent (value: 1.00). Put differently, the configurations A*B*C + A*B*~C + A*~B*C are sufficient

[12] In accordance to Boolean algebra, a case's membership score in the intersection A*B is the minimum of the respective single set memberships.

[13] The formula for the consistency is the following: $\text{Consistency}_{\text{sufficient conditions } (x_i \leq y_i)} = \frac{\sum_{i=1}^{I} min(X_i, Y_i)}{\sum_{i=1}^{I} (X_i)}$.
Of course, the consistency of sufficiency can also be calculated for truth table rows that base on crisp sets. However, a) the measure was not yet introduced in section 1.2.3, and b) all the configurations in the hypothetical crisp example are either perfectly consistent (value 1.00) or perfectly inconsistent (value 0.00).

for the outcome without any deviances since all empirical evidence supports the set relation. Hence, one can easily assign a '1' as 'truth value of consistency' indicating that the respective truth table rows are regarded as sufficient and need to be included into logical minimization. Such an easy decision is not at hand for the remaining configurations as the consistency scores reveal that they are not perfectly sufficient conditions. Unquestionably, in a deterministic understanding all (slightly) inconsistent truth tables must not be included into minimization since they are clearly not sufficient. But QCA is not deterministic per se and researches must be aware that the decision to exclude any inconsistent truth table row comes to a price; solution terms can only cover those cases that are covered by the configurations included to logical minimization. Put differently, if one decides to explain set relations only on the basis of perfectly consistent configurations than one will not be able to explain very much. The alternative is to include configurations that are (slightly) not consistently sufficient for the outcome. Accordingly, if one decides to explain more, than one will not be able to explain as good as before. In other words, consistency and coverage in fsQCA work against each other. This calls for another qualitative decision by the researcher on how much inconsistency should be allowed. While for instance Schneider and Wagemann (2010: 10) suggest that "no consistency values lower than 0.75 should be accepted" the actual decision is up to the researcher him- or herself. Regarding the example in table 1.6, the truth value of sufficiency has been decided by me for six out of eight truth table rows. Next to the perfectly consistent rows 1-3 that are regarded as sufficient for the outcome, the decision not to include them into logical minimization (indicated by a truth value of '0'; column five) has been taken for rows 5, 6, and 8. A consistency of 0.44, 0.5, and 0.63 respectively suggests that the empirical evidence just poorly supports the sufficiency claim. Still, whether or not to include rows 4 and 7 is not decided yet since the consistency scores are relatively high.

Table 1.7 displays the different results of the logical minimization when either no inconsistency is included (neither truth table row 4 nor 7 included; consistency cutoff: 1.00; solution term I), only slight inconsistency is included (truth table row 7 included; consistency cutoff: 0.88; solution term II), or even more inconsistency is included (truth table rows 4 and 7 included; consistency cutoff: 0.78; solution term III).

A single look at the results reveals that the decision on the consistency cutoff has tremendous effects on the analysis of sufficiency. Above all, the solution terms differ remarkably. While solution II equals the solution term from the crisp set analysis, term III is more general in comparison, i.e. a superset of term II. Next to being a rich single (~B*C) excellent health alone (A) is sufficient for high life satisfaction – no longer just in combination with being in

partnership (A*B). Instead, according to solution term I excellent healthy people in partnership or excellent healthy rich people are highly satisfied with their lives.[14]

Table 1.7 Analysis of sufficiency of fuzzy example "life satisfaction" with different consistency cut offs

Parameter of fit	**Solution term I**	**Solution term II**	**Solution term III**
	A*B + A*C → Y	A*B + ~B*C → Y	A + ~B*C → Y
Solution consistency	1.00	0.95	0.84
Solution coverage	0.74	0.83	0.91
Cases covered by solution	3 (Tom, Jerry, Linda)	5 (Tom, Jerry, Linda, Tina, Toby)	6 (Tom, Jerry, Linda, Tina, Toby, Larry)

Notes: Condition A: set of people with excellent health, condition B: set of people in partnership, condition C: set of people with high socio-economic status, condition Y: set of people with high life satisfaction

Moreover, the number of cases that are covered by the solution doubles from term I to term III since more empirical evidence is included. Lastly, the parameters of fit change as well. While the solution consistency decreases from term I to term III since truth table rows that (slightly) contradict the sufficiency claim are included, the solution coverage increases since more of the outcome set can be explained.[15] To summarize, results of the analysis of sufficiency are very dependent on decisions regarding the consistency cutoff in truth table analysis.[16] Considering that all three solution terms of the hypothetical example are in line with the claim by Schneider and Wagemann that only „consistency levels (well) above 0.75 are advisable" (Schneider & Wagemann 2012: 279), the problem becomes even more apparent. Thus, a high quality QCA will substantiate and justify the decision on the amount of inconsistency included into logical minimization.[17]

Second, next to the question of handling inconsistent truth table rows also the treatment of logical remainders appears to be a decisive part throughout the evaluation of the quality of sufficiency analyses in QCA. As brought up before, logical remainders or 'counterfactuals'

[14] Solution term I, in addition, provides a proper example for the risk of identifying 'false necessary conditions' from the analysis of sufficiency (see research step 'analysis of necessity' and Schneider & Wagemann 2010). Although condition A is part of every sufficient condition and by that seems to be necessary for Y, the separate analysis of necessity reveals that it is not (consistency: 0.78; coverage: 0.85).

[15] The comparison of 'solution coverage' and 'cases covered by solution' nicely displays that coverage value (as well as consistency) does not contain information on the number of cases that are covered by a solution term. In contrast, it means that the solution term covers between 74% (term I) and 91% (term III) of the outcome set.

[16] Note that some authors (and software tools) speak of inclusion score instead of consistency cutoff; see e.g. Thiem 2016.

[17] A set theoretic approach to interpret and review the amount of inconsistency is delivered in research step 'interpretation and visualization'.

appear whenever the number of logically possible configurations exceeds the empirical diversity of cases. In order to 'create' limited empirical diversity in the hypothetical example data at least one configuration needs to get rid of its empirical evidence. Thus, I change the membership of case Linda in the set A from score 1 to 0. In other words, Linda is no longer excellently healthy and by that displays no longer the configuration A*~B*C but the configuration ~A*~B~*C. As now no case is a member in the configuration A*~B*C anymore, this truth table row becomes a logical remainder (see table 1.8, row 3). When it comes to truth table analysis, all configurations that lack empirical evidence can (n)either be clearly regarded as sufficient (n)or as insufficient for the outcome as indicated by the question mark. However, their needs to be a decision on whether or not to include these 'empty' truth table rows into logical minimization, i.e. whether to assign a truth value of sufficiency of '1' or '0' in the truth table.

Table 1.8 Truth table of fuzzy set example "life satisfaction" with logical remainder

Truth table row	**Truth value of condition in configuration**			**Truth value of sufficiency**	**Consistency of sufficiency**	**Cases covered by configuration**
	A	**B**	**C**			
1	1	1	1	1	1.00	Tom
2	1	1	0	1	1.00	Jerry
3	1	0	1	**?**	-	-
4	1	0	0	0	0.78	Larry
5	0	1	1	0	0.50	Berta
6	0	1	0	0	0.50	Rob
7	0	0	1	1	0.92	Tina, Toby, Linda
8	0	0	0	0	0.63	Bob

In principle, there are three different approaches to handle logical remainders. The *conservative* strategy (Schneider & Wagemann 2012: 162) is based on the maxim that whenever there is no empirical evidence for a certain configuration one simply cannot say that this is a sufficient condition for the outcome. Thus, due to the lack of empirical instances for the respective configurations the conservative strategy is to decide that all logical remainder rows are excluded from the minimization process. Following this strategy, the life satisfaction example truth table row 3 would be assigned a truth value of sufficiency of '0'. It goes without saying that the conservative solution term has a solid reasoning by including only empirically observed configuration. Nevertheless, there are two more strategies that actually make use of counterfactuals. Including configurations even without empirical evidence into logical minimization is, always, a thought experiment that claims: if there was a case displaying this configuration, then it would show the outcome. But the genuine reason to include counterfactuals to the analysis of sufficiency is twofold. Technically, on the one hand, logical remainders

might simplify the analysis of sufficiency by extending the potential to 'match similar conjunctions' (see section 1.2.3). In other words, simplifying counterfactuals are those configurations that can be used for logical minimization to arrive at a less complex solution. For instance, if one would include the example counterfactual truth table row 3, i.e. A*~B*C, it could be matched with row 1 and row 7.[18] This purely technical argument substantiates the *most parsimonious* solution.[19] As the name suggest, this proceeding aims at the least complex solution term by including all simplifying logical remainders. Theoretically, on the other hand, counterfactuals might broaden the analytical power of the analysis. Put differently, thought experiments might improve the understanding of social phenomena in comparison to analyses that do base their knowledge on the diversity that was empirically covered by cases. Regarding the example logical remainder row 3, one could design a counterfactual thought experiment as follows:

Empirical observation:

> The conjunction in truth table row 7 (i.e. ~A*~B*C) is – based on a strong empirical basis – sufficient for the outcome Y (i.e. 'high life satisfaction').

Theoretical expectation:

> The presence of set A (i.e. 'excellent health') in theory is beneficial for the presence of outcome Y (i.e. 'high life satisfaction').

Counterfactual reasoning:

> If already the absence of A in combination with ~B*C is sufficient for Y, then a conjunction with the additional presence of 'A' with all other condition being constant (i.e. A * ~B*C) would also be sufficient for Y if such cases did exist.

On the basis of such a thought experiment, truth table row 3 could be assigned a '1' as truth of sufficiency and by that be included into the minimization process. In contrast to the most parsimonious solution, this decision is not (only) based on the technical argument that row 3 simplifies the solution term. Instead, it combines the ideas of the technical and the theoretical

---

[18] Matching truth table rows 3 and 1: A*~B*C + A*B*C → Y = A*C → Y
Matching truth table rows 3 and 7: A*~B*C + ~A*~B*C → Y = ~B*C → Y

[19] While many scholars formulate the risk to conflict theoretical consideration when treating all simplifying counterfactuals as sufficient for the outcome (see for so called "difficult counterfactuals" Schneider & Wagemann 2012: 175 ff.), some rate the most parsimonious solution as the only way to approach causality in QCA (Baumgartner 2009, 2013, 2015).

argument which is implemented in the *intermediate* solution term.[20] The three strategies' exemplified sufficient solution terms are displayed in table 1.9. All three base upon exactly the same empirical information and just differ in their treatment of logical remainders. Moreover, since the only counterfactual (row 3) was included both due to technical and theoretical reasons, the most parsimonious and intermediate solution for the hypothetical example are the same.

Table 1.9 Analysis of sufficiency of fuzzy example "life satisfaction" with different treatment of logical remainder

Parameter of fit	**Conservative solution**	**Most parsimonious solution**	**Intermediate solution**
	A*B + ~A*~B*C → Y	A*B + ~B*C → Y	A*B + ~B*C → Y
Solution consistency	0.95	0.95	0.95
Solution coverage	0.83	0.83	0.83
Cases covered by solution	5 (Tom, Jerry, Linda, Tina, Toby)	5 (Tom, Jerry, Linda, Tina, Toby)	5 (Tom, Jerry, Linda, Tina, Toby)

Notes: Condition A: set of people with excellent health, condition B: set of people in partnership, condition C: set of people with high socio-economic status, condition Y: set of people with high life satisfaction

To conclude, there are different strategies to deal with counterfactuals. Three of them, the conservative, the most parsimonious, and the intermediate solution term are implemented in the standard software tools and jointly form the "standard analysis" (see Ragin & Davey 2014). This means that all of them are reported by default at the end of any analysis of sufficiency. Thus, a high quality QCA both will substantiate and justify the treatment of logical remainders and will openly discuss the selected solution term in contrast to the two alternatives.

Third, while the treatment of inconsistencies and logical remainders has been broadly discussed at least within the QCA research community, the skewedness of set memberships is a rather unexplored problem of the analysis of sufficiency in QCA. So far, scholars primarily discussed the effects of skewedness with regard to necessity (see research step 'analysis of necessity above'; Goertz 2006b; but see Schneider & Wagemann 2012: 232 ff.). However, also for the analysis of sufficiency skewed memberships are highly problematic – if not even more. In a nutshell, skewed distributions of set memberships might lead to inaccurate inferences about set relations. In more detail, skewedness may result in simultaneous subset (and

[20] The intermediate solution also includes simplifying logical remainders to the sufficiency analysis, but only those that are in line with the theoretical expectations (see for the so called "easy counterfactuals" Schneider & Wagemann 2012: 168 ff.).

superset) relations of the presence and/or the absence of a certain condition X on the presence and/or absence of the outcome Y (Schneider/Wagemann 2012: 232). For instance, if (almost) all cases are members of a certain condition (i.e. $X_i > 0.5$) regardless whether the outcome is present or absent, than the presence of X will be identified as simultaneously sufficient for both the outcome and the non-outcome (i.e. X→Y and X→~Y). Likewise, if (almost) all cases are non-members of a certain condition (i.e. $X_i < 0.5$) than the absence of X will be identified as simultaneously sufficient for both the outcome and the non-outcome (i.e. ~X→Y and ~X→~Y). It goes without saying that simultaneous sufficiency statements about one condition for both the presence and the absence of a social phenomenon are at least incoherent if not untenable. However, to claim that a high quality QCA will avoid skewedness at all is beyond researchers' ability to notice as there is no suitable measure at hand so far. Instead, one has to differentiate skewedness of single sets and more or less complex configurations. Whenever single sets are skewed the problem can easily be related to inappropriate set calibration. For instance, Schneider and Wagemann (2012: 248) argue that skewedness in an outcome set Y empirically is not so much of a problem due to the Y-orientation in applied QCA. Since outcome sets usually consist of a single set only, skewed membership scores are very likely to be noticed. At least for a high quality calibration the same will apply to single condition sets. In contrast, especially throughout the analysis of sufficiency explanatory conditions might be very complex configurations (e.g. truth table rows). This hinders to detect skewedness in X even if the single sets are not skewed. Consequently, for complex configurations of sets a measure for the extent of skewedness is needed. Such a value could display problems of skewedness among complex configurations by identifying the level of empirical diversity included into the explanatory model (for a suggestion for the calculation of such a measure see chapter 5). For single conditions and the outcome set, instead, a high quality QCA will both avoid skewedness and be aware of the analytical consequences of skewedness, i.e. simultaneous subset/superset relations of conditions and (non-)outcome.

Fourth, and finally, it should be mentioned that the analysis of sufficient conjunctions that explain the absence of the outcome might be of great interest as well. Still, if conducted one needs to remember that set relations are asymmetric in nature. Thus, it is not possible to infer sufficient configuration for the absence of the outcome out of the sufficient solution term for its presence (compare section 1.2.1). For instance, just because the combination of excellent health AND living in partnership in our example is sufficient for high life satisfaction (A*B→Y) it is not the case that all sick singles (~A*~B) in the data set are unsatisfied with their lives (~Y). Mathematically, in contrast to linear algebra set relations base on Boolean

algebra. Given De Morgan's law the negation of the intersection of two sets is the same as the union of their negations, i.e. ~(A*B) = ~A+~B. Moreover, the negation of sufficiency is necessity. Together, a simple negation of the sufficiency statement for the presence of an outcome constitutes a necessity statement for the absence of the outcome, i.e. ~ (A*B→Y) = ~A+~B←~Y. Verbally, in order to be not highly satisfied with his or her life it is necessary for the hypothetical cases to either not to be excellent healthy OR not to be in partnership OR both. Note, however, that De Morgan's law cannot be applied whenever logical remainders have been utilized in the analysis of sufficiency. In short, a high quality QCA will conduct the analysis of sufficiency for the absence of the outcome separately (Schneider & Wagemann 2010: 12 f.). But still, any analysis of the negation of the outcome requires more than the understanding of Boolean algebra. One needs to conceptually examine what actually is of interest when ~Y is analyzed. If one, for instance, wants to detect configurations of conditions under which people turn out to be *not highly* satisfied with their lives this will work out perfectly fine. In contrast, if one wants to detect configurations of conditions under which people turn out to be *lowly* satisfied with their lives, this new target set needs a new calibration. Again, due to asymmetry in set calibration the negation of a set does not (always) depict its opposite pole. Maybe even the theoretical background and the respective conditions for the analysis of the non-outcome "*low* life satisfaction" needs to be changed. To conclude, a high quality QCA will discuss the (a)symmetry of set calibration prior to (separate) analyses of sufficiency for the presence and absence of the outcome.

To sum up, the evaluation of the quality of the analysis of sufficiency in QCA can be operationalized with various indicators. On the one hand, the choice and justification of consistency cutoffs for logical minimization as well as the treatment of logical remainders are important factors. A high quality QCA will avoid the mechanical utilization of standardized thresholds from the literature but argue on the basis of its own research design. In addition, a high quality QCA will (theoretically) justify the (non-)use of counterfactuals and select a solution term not because it is most easy to interpret but because it is in line with the conceptual foundation. Finally, a high quality QCA will avoid skewedness in single conditions, and – if conducted at all – analyze the sufficiency for the absence of the outcome separately.

Step V – interpretation and visualization

After the analyses of necessity and sufficiency, the empirical findings need to be properly interpreted. A proper interpretation of empirical results always needs to relate findings to the theoretical and methodological substance of the respective study. This is not different in QCA

but, again, comes along with some QCA-specific requisites. Relating empirical findings and the theoretical foundation in QCA primarily means to consider configurational thinking, equifinality, and asymmetry as well as a strong case orientation throughout interpretation. Interpretation and visualization are regarded as one research step in the concept of quality of QCA since the visualization of results with appropriate tools might be of added value for the (correct) interpretation of (non-perfect) set relations.

First, taking configurational thinking seriously means to interpret necessary and/or sufficient combinations of conditions in their complexity. In other words, a high quality QCA will not interpret the importance of single INUS-conditions within a sufficient configuration separated from the other conditions. As the analysis revealed the respective conditions only jointly are sufficient for the outcome.[21]

Likewise second, taking equifinality seriously means to interpret the solution term's single sufficient configurations (or path) not 'against' each other. Any interpretation of a solution term that tries to differentiate the importance of sufficient paths should proceed with care since there are at least four different understandings of 'importance' possible. On the one hand, even the term 'empirical importance' is ambiguous because the path covering the largest amount of cases might be different from the path covering the largest part of the outcome set. For instance, the solution term II from the analysis of sufficiency of the fuzzy example "life satisfaction" (table 1.7) consists of two paths: A*B + ~B*C → Y. Examining the parameter coverage of sufficiency for the respective path reveals that A*B alone covers approximately 52%[22] of the outcome set (consistency score: 0.92). The very same is true for path ~B*C, it also has a raw coverage of 0.52 of the outcome set (consistency score: 1). Consequently, regarding empirical importance in the sense of coverage of sufficiency, A*B and ~B*C are equally important. However, looking for the number of cases covered by the respective path exposes that A*B covers the three cases Tina, Linda, and Toby, while ~B*C covers only two cases, Tom and Jerry. As a proper visualization in this regard a detailed table including consistency and coverage scores, but also the number of cases covered by the single paths might be used since it supports the double-edged interpretative effort. On the other hand, theoretical importance is ambiguous as well. Theoretical importance might either de-

[21] The same applies to SUIN-conditions in necessary configurations.

[22] The calculation of the coverage value of single path differentiates raw and unique coverage. While raw coverage displays how much of the outcome is covered by the path, unique coverage exposes how much of the outcome is covered *exclusively* by this path. Here, both paths' raw coverage is 0.52, the unique coverage is 0.3 for both of them.

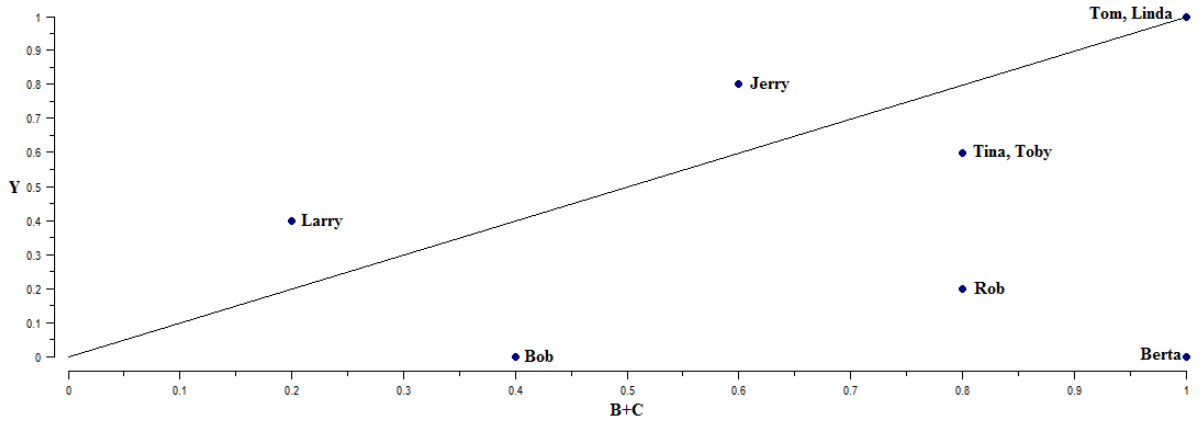
scribe the ability of a path to provide a theoretically new (and maybe path-breaking) explanation for the occurrence of a certain phenomenon – or be a reassurance that existing theories are still valid. To conclude, a high quality QCA will be very precise when interpreting the added value of the individual paths for the solution term.

Third, taking asymmetry seriously in the interpretation of empirical results is associated both with analytical issues and calibration issues. On the analytical side, any interpretation needs to take into account the separate sufficiency checks for the presence and absence of the outcome (as outlined in research step IV). Both analytical steps also need to be interpreted separately and must not go beyond the scope of the respective findings. With regards to calibration, interpretation needs to correspond to the asymmetry of concepts as indicated by the labeled target sets. This applies whenever the absence of a certain condition is an SUIN or INUS-condition in a necessary and/or sufficient configuration. For instance, cases displaying the configuration ~A*~B*~C in the hypothetical example do not need to be interpreted as 'sick' (~A) and 'poor' (~C) 'singles' (~B). Due to asymmetric calibration, they might be people of 'very good but not excellent health' (also covered by ~A), people with an 'average level of income' (also covered by ~C), and people with 'a lover they do not call their partner' (probably also covered by ~B). In short, a high quality QCA will keep in mind these labels and carefully limit the interpretation of empirical findings to the analytical and conceptual scope as determined by asymmetry.

Finally fourth, taking QCA's case orientation seriously during interpretation means bringing the case back in. Remember that according to Mahoney (2010: 133) QCA aims at "extending the logic of case studies to comparative analysis". Thus, detailed knowledge on cases as the constitutive research objects is essential for any interpretation of empirical results. Thus, interpretation must not remain on the level of the solution term as a whole or even the respective sufficient paths. Instead, the individual cases as well as the diversity among them covered by such a path need to be interpreted and discussed in detail. On the one hand this means to see which cases are explained by a certain necessary and/or sufficient configuration. On the other hand, cases contradicting a set relation claim (i.e. deviant cases) are at least as interesting for a proper interpretation as onlier cases since they might challenge theory or simply terminate the (deterministic) strength of an otherwise consistently necessary and/or sufficient condition. To conclude, a high quality interpretation in QCA will offer a detailed discussion on the relation between empirical findings and cases. Here, another form of visualization of set relations is strongly advisable; the graphical visualization with XY-plots.

XY-plots provide a both easy to create and easy to interpret visualization tool for the quality of set relations. For the frequent user, XY-plots reveal the extent of inconsistency as well as of coverage at first sight. Moreover, XY-plots help to identify those cases that for instance are responsible for a low consistency or coverage of a configuration. As an example, figure 1.5 displays the XY-plot for the (quasi-)necessary condition B+C for the presence of Y.

Figure 1.5 XY-plot for the analysis of necessity of example "life satisfaction" with fuzzy data



As outlined in research step III, conditions are necessary if the outcome cannot occur without the condition being present. In fuzzy set, the (partial) membership in the necessary condition needs to be greater than or equal to the (partial) membership in the outcome set. If this applies to BC ← Y then all cases need to be located in the lower triangular, i.e. directly at or below the diagonal showing X=Y. But this is not the case; both Larry and Jerry (compare figure 1.5 and table 1.5) are located above the diagonal and by that lower the consistency of necessity slightly to a value of 0.913 (see table 1.5). The analytical interpretation of B+C as necessary conditions could proceed as follows:

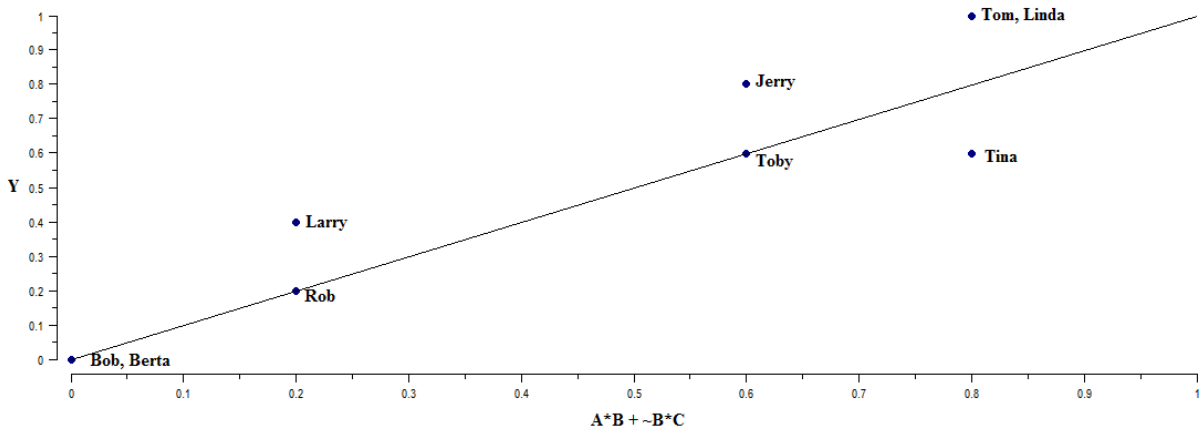
> From an analytical standpoint, the configuration B+C is quasi-necessary since the inconsistency caused by two cases Larry and Jerry is negligible. Case Larry is neither a member in the configuration B+C (membership score: 0.2) nor in the outcome set Y (membership score: 0.3). As such, he is an irrelevant case for the statement that the presence of B+C is necessary for the presence of Y (compare chapter 5). Case Jerry, instead, is both a member in the configuration B+C (membership score: 0.6) and in the outcome set Y (membership score: 0.8). As such, he supports the necessity claim in general, even if the partial presence of B+C is not fully explaining the partial presence of Y. To conclude, as there are no truly contradictory cases to the statement of necessity, B+C might be regarded as necessary.

Remember, however, that from a conceptual standpoint the theoretical benefit of combination B+C is unclear since there are no substantive arguments that B and C might act as functional equivalents for one higher-order concept (see research step III). However, now it is up to the

researcher to decide whether B+C not only analytically can be interpreted as quasi-necessary, but whether it also is conceptually appropriate.

Moreover, as an example for the visualization of sufficient conditions, figure 1.6 displays the XY-plot for the statement A*B + ~B*C → Y.

Figure 1.6 XY-plot for the analysis of sufficiency of example "life satisfaction" with fuzzy data



A condition is sufficient if whenever the condition is present the outcome will be present, too. In fuzzy set, the (partial) membership in the sufficient condition needs to be smaller than or equal to the (partial) membership in the outcome set (X≤Y). If this applies to the complex configuration A*B + ~B*C then all cases need to be located in the upper triangular, i.e. directly at or above the diagonal showing X=Y. As figure 1.6 displays, this is almost perfectly the case. Only case Tina is located below the diagonal which causes a slight inconsistency (consistency score: 0.95; compare solution term II in table 1.7). The analytical interpretation of A*B + ~B*C as sufficient conditions could proceed as follows:

> From an analytical standpoint, the configuration A*B + ~B*C is almost perfectly consistent sufficient since the inconsistency caused by the case Tina is negligible. Case Tina is both a member in the configuration (membership score: 0.8) and in the outcome set Y (membership score: 0.6). As such, she supports the sufficiency claim in general, even if the partial presence of A*B + ~B*C is not fully explaining the partial presence of Y. To conclude, as there is no truly contradictory case to the statement of sufficiency, A*B + ~B*C might be regarded as sufficient for the outcome.

To sum up, the evaluation of the quality of the interpretation of empirical findings in QCA can be operationalized with regard to the methodological foundation of QCA. A high quality QCA interpretation will pay attention to configurational thinking, equifinality, and asymmetry. Moreover, high quality QCA interpretation will bring the cases back in. Finally, a high quality QCA interpretation will underpin its argument with appropriate forms of visualization. The use of XY-plots is especially recommendable for the interpretation of set relations since

they allow for a case-oriented depiction of set relations and help to identify inconsistencies as well as problems regarding coverage and skewedness.

For a comprehensive presentation of all indicators identified for all five research steps see the three-level-concept of the quality of QCA in table 1.4.

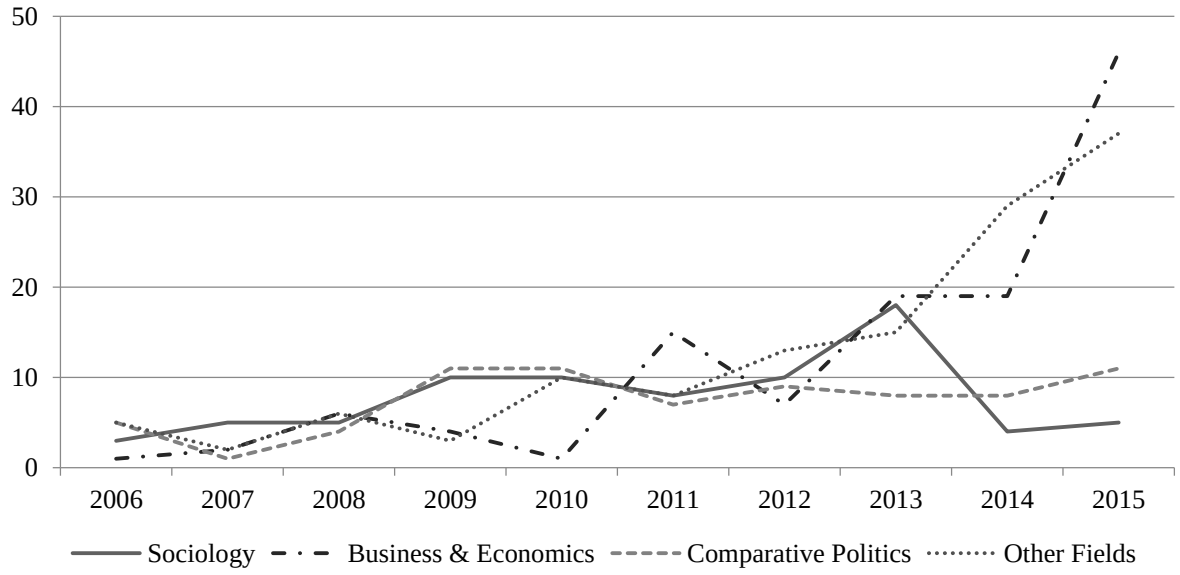
1.4 Relevance of the Scientific Articles

In this section the relevance of the scientific articles entailed in this cumulative dissertation are discussed. Starting off from the concept of the quality of QCA outlined above, the added value of the three articles evaluating the quality of QCA applications in sociology, business research, and political science as well as of the methodological improvement and the application article are examined. As three articles of this cumulative dissertation are co-authored, my individual contribution is made clearly distinguishable and assessable as an individual scientific achievement, too.

1.4.1 Evaluating the Quality of QCA

Ever since the development of Qualitative Comparative Analysis in 1987, comparative macro sociology and comparative politics have been the major scientific areas for QCA applications (Rihoux et al. 2013: 177).

Figure 1.7 Number of articles published in peer-reviewed journals by scientific area, since 2006



Source: Compasss 2016.

However, over the last years especially scholars from the field of business and economics used QCA to a larger extent. Today, business and economics, comparative politics, and sociology constitute the most active scientific fields of QCA applications (see figure 1.7). Thus, assessing the quality of published QCAs it seems likely to deal with empirical applications that stem from these fields. This cumulative dissertation comprises evaluations of 139 peer-reviewed journal articles in the following pieces:

Qualitative Comparative Analysis (QCA) in der Soziologie – Perspektiven, Potentiale und Anwendungsbereiche

The article "Qualitative Comparative Analysis (QCA) in der Soziologie – Perspektiven, Potentiale und Anwendungsbereiche" (*English title: Qualitative Comparative Analysis (QCA) and Sociology – Perspectives, Potential, and Areas of Application*) by Jonas Buche and Markus B. Siewert highlights perspectives and potentials which QCA offers to address sociological research questions. It reviews and evaluates seventy-seven publications in the field of sociology using QCA. Next to a broad overview on topics and fields of applications within sociology it examines crucial steps and potential pitfalls by exemplifying typical practices and best practices (Buche & Siewert 2015). By that, it evaluates all five research steps identified and operationalized by the concept formation in this introduction (see section 1.3). The article was written in collaboration with Markus Siewert (Goethe University Frankfurt) and includes six chapters:

1. *Was es heißt, soziale Phänomene mit QCA zu betrachten /*
   What it means to look at social phenomena with QCA
2. *Einleitung /* Introduction
3. *QCA in der Soziologie: Eine Rundschau /* QCA in Sociology: A Review
4. *Beispielstudien: Analyseschritte und potentielle Fallstricke einer QCA /*
   Sample studies: analytical steps and potential pitfalls in QCA
5. *Aktuelle Anwendungsbeispiele von QCA in der Soziologie /*
   Current application examples of QCA in sociology
6. *Abschließende Bemerkungen zum Mehrwert von QCA für die Soziologie /*
   Concluding remarks on the added value of QCA for sociology

My contribution to this publication was drafting and, together with Markus Siewert, finalizing the empirical analysis of the evaluated studies (i.e. chapter 3-5). Additionally, I took part in writing the introduction and conclusion.

First, I was the author of the review section on QCA applications in sociology. This includes both the analysis of the studies' topics and research questions, the research design, the level of

analysis (micro-, meso- and macro-sociology) and the respective data basis (quantitative aggregated indicators, surveys, interview data, qualitative analyses, ethnographical studies, etc.). Moreover, I reviewed the use of the variants of QCA (crisp set QCA, fuzzy set QCA), the number of conditions, and the number of cases included to the analyses. Secondly, using a large number of published example studies, I evaluated the realization of four analytical steps of a QCA, i.e. calibration, analysis of necessary, analysis of sufficiency, and the treatment of consistency and coverage. The idea of this approach was to uncover potential pitfalls in the application of QCA, to document them with practical examples, and to provide guidelines and advices on how to avoid these problems. Thirdly, it was my task to present three best practice QCA applications from the field of sociology, one at the micro, meso and macro level each.

The article has been published in the journal "Zeitschrift für Soziologie" which had an impact factor of 0.809 in the year of publication (Thomson Reuters 2016).

QCA and Business Research: Work in Progress or a Consolidated Agenda?

The article "QCA and Business Research: Work in Progress or a Consolidated Agenda?" by Claudius Wagemann, Jonas Buche, and Markus B. Siewert is based on a rigorous evaluation of sixty-one published, peer-reviewed journal articles applying QCA within the field of business and management studies. Like the first article, it relates the concept of the quality of QCA to its actual application and demonstrates minor and major shortcoming in the studies under scrutiny. In addition, the article presents several suggestions on how to overcome or avoid them in the first place (Wagemann et al. 2015). The article was written jointly with Claudius Wagemann and Markus Siewert (both Goethe University Frankfurt) and includes the following five chapters:

1. Introduction: QCA meets business research
2. QCA: approaching an acronym
3. QCA in business research: a birds-eye view
4. Does QCA in business research avoid common pitfalls?
5. QCA and business research: how big is the intersection?

My contribution to this publication was mainly the overview on the 61 peer-reviewed articles included into the evaluation of QCA in business research (i.e. chapter 3). I also took part in writing the introduction and conclusion.

First, I highlighted the enormous development of QCA in business research, which today constitutes the most active research area and produces the highest number of QCA studies per year. Secondly, the overview of QCA applications encompassed research topics and research

questions, the levels of analysis, the numbers of cases and conditions, and the different approaches to set calibration. Finally, I have extracted five facets from the overview that characterize the use of QCA in business and management literature.

The article has been published in the "Journal of Business Research" which had an impact factor of 2.219 in the year of publication (Thomson Reuters 2016).

Fuzzy Logic or Fuzzy Application? A Response to Stockemer's "Fuzzy Set or Fuzzy Logic?

The article "Fuzzy Logic or Fuzzy Application? A Response to Stockemer's 'Fuzzy Set or Fuzzy Logic?" assesses the quality of one specific application of QCA within the field of comparative politics. In contrast to the first two papers, its aim is not to give a broad overview on published applications but to demonstrate the consequences of a misinformed application throughout the various research steps of fsQCA (and OLS regression) in detail. Going beyond the scope of a pure evaluation of the selected study by Stockemer (2013), the third article offers a replication of the analysis of both methods and systematically discusses the calibration process, the analysis of necessary and sufficient conditions, and the interpretation of the results (Buche et al. 2016). The article was written jointly with Antje Buche (University of Erlangen-Nuremberg) and Markus Siewert (Goethe University Frankfurt) and consists of four chapters:

1. Introduction
2. Misunderstanding QCA
3. Misinformed application
    a. Debating case selection and operationalization
    b. Replicating the analysis
    c. Assessing the interpretation
4. Conclusion

My contribution to this publication is essentially the presentation of the uninformed use of QCA in Stockemer (2013) by means of research design, operationalization, and calibration as well as the replication and interpretation of the analysis (chapter 3). I also participated in writing the introduction and conclusion.

First, I unfold and discuss the problems both regarding content and methodology that result from case selection, the high number of explanatory factors, and the inappropriate calibration of the underlying data. Yet, my contribution is limited to those sections of the paper that deal with QCA, not including those that analyze Stockemer's similarly problematic utilization of OLS regression. Secondly, I conceptually prepared the replication of Stockemer's study. This

essentially involved collecting arguments for the re-calibration of the most problematic conditions on "statutory women's quota" and "electoral system type". Thirdly, I replicated the analysis of necessity and sufficiency by Stockemer (2013), provided a counterfactual argumentation for the inclusion logical remainders, and finally interpreted the results in an appropriate set-theoretic way. Here, next to the interpretation with regards to content I discussed the methodological implications of Stockemer's original selection of cases and conditions as well as the skewed distribution of the data.

The article has been published in "European Political Science" which had an impact factor of 0.553 in the year of publication (Thomson Reuters 2016).

1.4.2 Improving the Quality of QCA

While the first three articles offer an empirical evaluation of the application of the concept of quality of QCA, the fourth article goes beyond the concept as laid out above and engages in the improvement of the quality of QCA. It does so by interrelating the conceptual and the mathematical foundation of QCA in the following piece:

Relevant Consistency and Relevant Coverage in Fuzzy Set Qualitative Comparative Analysis (fsQCA)

The single-authored article "Relevant Consistency and Relevant Coverage in Fuzzy Set Qualitative Comparative Analysis (fsQCA)" provides an improvement of the currently most important measures of the quality of set relations, namely consistency and coverage. As previously mentioned, Charles Ragin (2006: 292) introduced consistency and coverage as "simple descriptive measures for evaluating the strength of the empirical support for theoretical arguments describing set relations". The importance of these parameters for the validity and reliability of fuzzy set relation cannot be overestimated. Put differently, the validity and reliability of results from almost all studies applying fsQCA prior to the development of these parameters might be regarded as 'uncertain'. However, the formulas for consistency and coverage values have the major shortcoming that they always include all cases into the analysis. Thus, cases where a potentially sufficient condition is absent might strongly influence the statement that the presence of this condition is sufficient for the outcome or not. The same applies to the consistency of necessity. Here, cases where the outcome is absent influence the strength of the empirical support for the claim that the outcome cannot be present without the condition being present. In short, all the formulas presented by Ragin (2006) rely on cases that are irrelevant for the respective set relation. This strongly contradicts the statement of asymmetry

which can be seen as one major argument in favor of QCA (see section 1.2.1). To frame it harsh: the mathematical foundation of the parameters of fit as developed by Ragin (2006) does not fit the conceptual foundation of QCA.

In this article of the cumulative dissertation I, first, identify different types of relevant and irrelevant cases and their effects on the consistency of sufficiency. Second, I propose a formula for 'relevant consistency' and apply it to artificial data. Inspired by Anscombe's quartet (Anscombe 1973) I use four datasets that have nearly identical statistical properties, but appear very different when graphed in XY-plots. While the Ragin formula – due to the influence of irrelevant cases – produces the same consistency scores regardless of the presence or absence of "true logical contradictory cases" (Schneider & Wagemann 2012: 127), the 'relevant consistency' displays major differences. Third, I argue that the effect of irrelevant cases is not limited to the formula of consistency of sufficiency. Instead, the influence of irrelevant cases is omnipresent in the calculation of set relations. Thus, I update formulas for both the consistency and coverage of sufficiency and necessity. Finally, fourth, I discuss the implications of the updated formulas on uncovering skewed memberships (Buche 2016).

The article has not yet been published.

1.4.3 Applying QCA of High Quality

The final article can be understood as the attempt to fulfill the promise indicated by the cumulative dissertation's title "Assessing the Quality of Qualitative Comparative Analysis". If assessed comprehensively, the quality of QCA cannot only be conceptualized, evaluated, and improved. Rather, the quality of QCA should also be applicable to social science research in a high state. This final step in the assessment of the quality of QCA is conducted in the following piece:

Europeanization of Legislative-executive Relations at the Micro Level: Under Which Conditions Do Swedish MPs Interact with Ministerial Officials?

The single-authored article "Europeanization of Legislative-executive Relations at the Micro Level: Under Which Conditions Do Swedish MPs Interact with Ministerial Officials?" offers an application of QCA that takes into account the multifaceted demands of a QCA as spelled out in section 1.3. With regard to the content, the article applies fuzzy set QCA to examine the Europeanization of legislative-executive relations at the micro-level in Sweden. It asks under which conditions members of national parliaments use civil servants from the national ministries as direct information sources both for domestic and Europeanized policies. Using role

theory, I state that roles, understood as the combination of position and preference roles, can be used as reliable predictors for individual behavior, i.e. interaction. I examine four conditions of executive-legislative relations presumed to be relevant for the frequency of direct interaction between MPs and bureaucrats: membership in a governing party, strong political expertise, strong role orientation as a policy expert, and strong external role perception of civil servants as policy experts. The theoretical section offers a proper example for hypothesis-formulation in configurational set-theoretic research. Methodologically, the article draws on survey data and semi-structured interviews with 22 Swedish members of parliament. The section on data calibration discusses the common and different transparency requirements of QCAs using quantitative and qualitative data (see also Wagemann & Schneider 2015). The analysis of necessity is linked to a debate on incoherent counterfactuals (Schneider & Wagemann 2012: 198ff.) and throughout the analysis of sufficiency a systematic counterfactual thought experiment is conducted. However, the measures of relevant consistency and relevant coverage developed in are not applied in the paper, yet (Buche 2017).

The article has been published in the "COMPASSS Working Paper Series".

1.5 Conclusion

To conclude, the cumulative thesis "Assessing the Quality of Qualitative Comparative Analysis (QCA) – Evaluation, Improvement, Application" asks for the concept of quality of QCA and its implementation in QCA application. It consists of five articles and an introductory paper that jointly answer this question by providing an in-depth quality assessment of QCA and QCA applications. The introduction defines and conceptualizes the quality of QCA. Three of the five research articles deliver a broad overview on the quality of empirical applications of QCA in three scientific fields (*evaluation*), the fourth aims at enhancing the quality of QCA with a narrow discussion and an update of the major quality assessment in fsQCA (*improvement*), and the fifth applies the concept of quality of QCA scrutinized before in a specific multi-method contribution itself (*application*). Together, the cumulative dissertation delivers the most comprehensive assessment of the quality of Qualitative Comparative Analysis so far. It covers and discusses the quality of 139 articles published in peer reviewed journals in sociology, political science, and business research.

2.

Qualitative Comparative Analysis (QCA) in der Soziologie – Perspektiven, Potentiale und Anwendungsbereiche

Qualitative Comparative Analysis (QCA) and Sociology – Perspectives, Potential, and Areas of Application

(Jonas Buche, Markus B. Siewert)[23]

[23] Die Autoren danken den Teilnehmer*innen der Ad-Hoc-Gruppe „Qualitative Comparative Analysis (QCA). Perspektiven für die soziologische Forschung“ auf dem 37. Kongress der DGS in Trier für die hilfreiche Diskussion. Ein besonderer Dank gebührt Daniela Grunow, Carsten G. Ullrich, Antje Buche und Claudius Wagemann sowie zwei anonymen Gutachter*innen und dem Herausgeberkreis der ZfS für zahlreiche wertvolle und anregende Kommentare. Zudem bedanken wir uns bei Yulia Aleshchenkova und Christoph Klement für die Zuarbeit.

Abstract/Zusammenfassung

The sociologist Charles C. Ragin originally introduced Qualitative Comparative Analysis (QCA), which combines ideas of case-oriented research, configurational thinking, and set-theoretical logic. Since then – as further developed by Ragin and others – QCA has become an established tool within social science methodology which uses a set-theoretical approach to analyze social phenomena. This paper aims at highlighting perspectives and potentials which QCA as a (relatively) new method is able to offer to address sociological research questions. In order to do so, seventy-seven publications in the field of sociology are reviewed. On this basis, on the one hand, a broad overview is given regarding applications, trends, and developments of QCA. On the other hand, crucial steps within QCA are examined and potential pitfalls discussed by exemplifying practiced practices and best practices.

Qualitative Comparative Analysis (QCA) wurde von dem Soziologen Charles C. Ragin als Verbindung von fallorientierten, konfigurativen Ansätzen und mengentheoretischem Denken präsentiert. Mittlerweile hat sich QCA – von Ragin und anderen weiterentwickelt – als mengentheoretischer Ansatz zur Untersuchung sozialer Phänomene im sozialwissenschaftlichen Methodenkanon etabliert. Der vorliegende Beitrag zielt darauf ab, Forschungsperspektiven und -potentiale von QCA als (relativ) junge Methode für soziologische Fragestellungen aufzuzeigen. Auf der Grundlage einer Rundschau von 77 publizierten, soziologischen Zeitschriftenartikeln wird einerseits ein breiter Überblick über Anwendungsbereiche, aktuelle Trends und Entwicklungen von QCA in der Soziologie gegeben. Andererseits werden am Beispiel der publizierten Studien die einzelnen Analyseschritte einer QCA besprochen und dabei gängige Fallstricke aufgezeigt, wobei sowohl practiced practices als auch best practices in ihrer Anwendung herausgearbeitet werden.

Keywords/Schlagworte

Qualitative Comparative Analysis (QCA); Fuzzy-set Analysis; Set-theoretical Thinking; QCA-application; Evaluation; Standards of "Good" Practice.

Qualitative Comparative Analysis (QCA); Fuzzy-Set Analysis; mengentheoretisches Denken; Evaluation von QCA-Anwendungen; Standards „guter" Praxis.

2.1 Einleitung

Die Publikation „The Comparative Method“ (1987) des amerikanischen Soziologen Charles C. Ragin ist in vielerlei Hinsicht bahnbrechend; in erster Linie jedoch legte sie den Grundstein für eine neue methodologische Perspektive in den Sozialwissenschaften: der Verbindung von fallorientierten, konfigurativen Ansätzen und mengentheoretischem Denken. Qualitative Comparative Analysis (QCA) ist hierfür gleichsam zum Synonym geworden. In den darauffolgenden Jahren von Ragin (2000, 2008) und anderen (siehe u.a. Rihoux & Ragin 2009; Schneider & Wagemann 2012) weiterentwickelt, hat sich QCA als mengentheoretischer Ansatz zur Untersuchung sozialer Phänomene zunehmend im sozialwissenschaftlichen Methodenkanon etabliert.[24]

Insbesondere seit der Einführung sogenannter *fuzzy sets* – was etwas sperrig mit „unscharfen Mengen“ (Wagemann & Schneider 2003) übersetzt werden kann – ist ein gewisser Boom um QCA zu beobachten, der sich neben Lehrbüchern und zahlreichen methodologischen Artikeln zuvorderst in der steigenden Zahl von QCA-Anwendungen ausdrückt (vgl. Rihoux et al. 2013). Zieht man etwa Publikationen in Zeitschriften mit Begutachtungsverfahren als ein Indikator für die Popularität von QCA heran, so wurden zwischen 1987 und 2011 die zweitmeisten aller QCA-Studien zu soziologischen Themen veröffentlicht.[25] Wirft man zudem einen Blick auf die jährlichen Publikationsraten rangiert die Soziologie gemeinsam mit Politikwissenschaften und Business und Management stets unter den ersten drei Anwendungsfeldern (Rihoux et al. 2013: 177).

An der deutschsprachigen Soziologie scheint die „Ragin Revolution“ (Vaisey 2009) allerdings weitgehend vorübergegangen zu sein. So sind bis dato lediglich zwei soziologische Studien in deutscher Sprache in Zeitschriften mit Begutachtungsverfahren erschienen, die auf einer QCA beruhen (Hörisch 2012; Laux 2015). Diese Tatsache bildet den Ausgangspunkt des vorliegenden Beitrags, der dreierlei Zielsetzungen verfolgt: Erstens wird eine knappe Einführung zum Grundverständnis von QCA gegeben, welche die originäre Perspektive auf Mengentheorie beruhender Ansätze erläutert. Zweitens wird auf der Grundlage einer Rundschau von 77 soziologischen Journalartikeln ein breiter Überblick über die Anwendung von QCA in der Soziologie gegeben. Drittens werden am Beispiel zahlreicher publizierter Studien

[24] Für einen sehr guten historischen Überblick über die Entwicklung von QCA in den letzten 25 Jahren siehe Marx et al. 2014.

[25] Der Anteil der Soziologie in diesem Zeitraum lag bei etwa 34% (gemeinsam mit anthropologischen Fragestellungen). Die meisten Anwendungen finden sich mit 51% im Bereich der Politikwissenschaft. Siehe Rihoux et al. 2013: 177.

die einzelnen Analyseschritte einer QCA besprochen und dabei diverse Tücken, aber auch *best practices* in ihrer Anwendung herausgearbeitet. In der Zusammenschau zielt der vorliegende Beitrag darauf ab, Forschungspotentiale von QCA als (relativ) junge Methode für soziologische Fragestellungen aufzuzeigen, aktuelle Trends und Entwicklungen soziologischer QCA-Anwendungen zu präsentieren und grundlegende forschungspraktische Probleme und Fallstricke am Beispiel publizierter QCA-Studien aus der Soziologie zu diskutieren.

Hierzu werden in einem ersten Schritt zentrale Elemente von QCA als einem mengentheoretischen Forschungsansatz vorgestellt und erläutert. Im dritten Abschnitt wird eine Rundschau soziologischer QCA-Studien vorgelegt, welche Fragen und Analyseebenen erläutern. Im vierten Abschnitt wird die Vorgehensweise bei einer QCA skizziert, wobei anhand publizierter Studien Probleme und Fallstricke in der Analyse aufgezeigt und diskutiert werden. In Abschnitt fünf werden schließlich drei QCA-Anwendungen – je eine auf Makro-, Meso- und Mikroebene – im Sinne von *best practices* präsentiert. Abschließend wird argumentiert, dass mengentheoretische Methoden wie QCA einen deutlichen Mehrwert gerade für soziologische Fragestellungen aufweisen und den soziologischen Methodenkanon sinnvoll zu ergänzen vermögen.

2.2 Was es heißt, soziale Phänomene mit QCA zu betrachten

Charles Ragin präsentierte QCA ursprünglich als *die* vergleichende Methode für Forschungsdesigns mit geringer bis mittlerer Fallzahl (Ragin 1987, 2000). Auch wenn bis heute zahlreiche Studien die Anwendung einer QCA mit dem Argument der mittleren Fallzahl rechtfertigen, so ist dieses Kriterium keineswegs das entscheidende. Wesentlich zentraler ist vielmehr, dass mit QCA eine mengentheoretische Forschungsperspektive eingenommen wird (Schneider & Wagemann 2012: 8-12). Was dies im Einzelnen bedeutet, wird in den folgenden Abschnitten erläutert.

So steht die Suche nach *notwendigen* und *hinreichenden* Bedingungen im Zentrum einer jeden QCA (siehe zum folgenden Ragin 2008: 29-68; Schneider & Wagemann 2012: 56-90). Nehmen wir als hypothetisches Beispiel einmal an, wir sind an Begründungen für eine ‚hohe Lebenszufriedenheit' interessiert, wobei ein möglicher Erklärungsfaktor hierfür eine ‚hervorragende Gesundheit' sein könnte. Um nun zu prüfen, ob ein sehr guter Gesundheitszustand (X) eine notwendige Bedingung für eine hohe Lebenszufriedenheit (Y) darstellt, muss wann immer Y vorliegt, auch X gegeben sein. Eine hohe Lebenszufriedenheit könnte in diesem Fall nicht ohne eine hervorragende Gesundheit auftreten. Aus mengentheoretischer Perspektive ließe sich diese Beziehung so beschreiben, dass das Outcome eine Teilmenge (*subset*) der

Bedingungsmenge bzw. die Bedingung eine Übermenge (*superset*) der Outcomemenge ist (siehe Abbildungen 2.1a-c).

Abbildungen 2.1 und 2.2 Notwendige und hinreichende Bedingungen

Abbildung 2.1a:
2x2 Tabelle für notwendige Bedingungen

Y \ X	abwesend	anwesend
anwesend	Fälle nicht erlaubt	Fälle erlaubt
abwesend	Fälle erlaubt; aber irrelevant	Fälle erlaubt; aber irrelevant

Abbildung 2.2a:
2x2 Tabelle für hinreichende Bedingungen

Y \ X	abwesend	anwesend
anwesend	Fälle erlaubt; aber irrelevant	Fälle erlaubt
abwesend	Fälle erlaubt; aber irrelevant	Fälle nicht erlaubt

Abbildung 2.1b:
Venn-Diagramm für notwendige Bedingungen

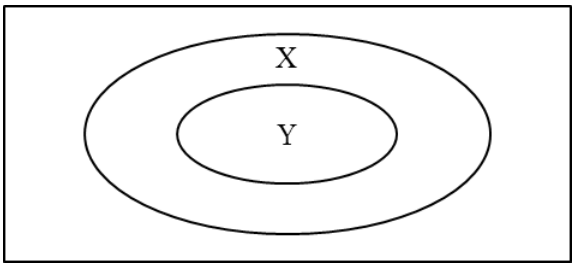


Abbildung 2.2b:
Venn-Diagramm für hinreichende Bedingungen

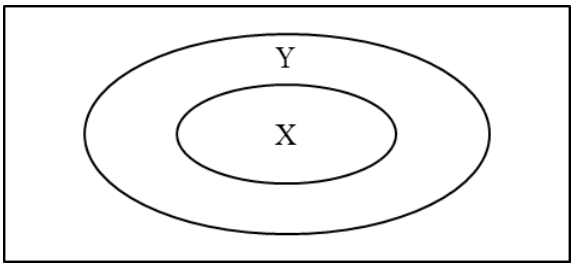


Abbildung 2.1c:
XY-Plot für notwendige Bedingungen

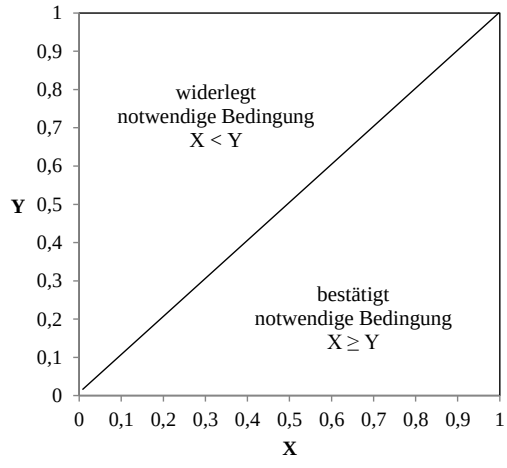
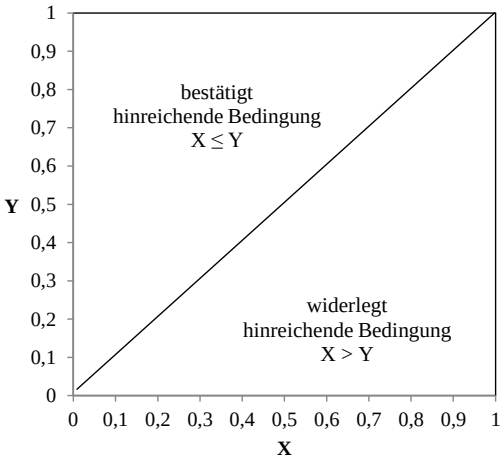


Abbildung 2.2c:
XY-Plot für hinreichende Bedingungen



Eine hinreichende Bedingung hingegen verhält sich quasi spiegelbildlich zu einer notwendigen. So könnte eine ‚hervorragende Gesundheit' als hinreichend bezeichnet werden für den Fall, dass wann immer jemand sehr gesund ist, auch die Lebenszufriedenheit hoch ist. Es dürften also keine Personen existieren, die Mitglied in der Menge ‚hervorragende Gesundheit'

sind, zugleich aber Nicht-Mitglied in der Menge ‚hohe Lebenszufriedenheit'.[26] Eine Bedingung ist also dann hinreichend, wenn wann immer X auftritt, auch Y vorliegt; oder mengentheoretisch ausgedrückt: Die Menge aller sehr gesunden Personen ist eine Teilmenge (*subset*) aller Menschen mit hoher Lebenszufriedenheit (siehe Abbildungen 2.2a-c).

Mit QCA als mengentheoretischem Ansatz geht somit eine ganz spezifische Sichtweise auf soziale Phänomene einher. Ragin (1987: 16-17; 34-68, 2004) spricht hier auch von einer fallorientierten (*case-oriented*) Perspektive, die sich grundlegend von variablen-orientierten Ansätzen wie etwa statistischen Verfahren unterscheidet. Der Unterschied liegt bereits in den zentralen Untersuchungseinheiten, nämlich Mengen (*sets*), begründet. Mengen können als „zones of inclusion and exclusion" definiert werden, in welche Fälle „according to their fit within the boundaries of a set" zugeordnet werden (Mahoney 2010 zitiert nach Schneider & Wagemann 2012: 24; siehe auch Verkuilen 2005). Die zu untersuchenden Fälle werden demnach als Konfigurationen unterschiedlicher Mengen verstanden, in denen Fälle einen bestimmten Grad an Mitgliedschaft aufweisen können. Beispielsweise kann ein Individuum volles Mitglied, volles Nicht-Mitglied oder partielles Mitglied – mehr innerhalb oder mehr außerhalb einer Menge, ggf. mit verschiedenen Abstufungen – in der Menge ‚Menschen mit hoher Lebenszufriedenheit' sein. Mengen zeichnen sich folglich dadurch aus, dass sie sowohl quantitative Abstufungen (*differences in degree*) partieller Mitgliedschaften als auch qualitative Unterschiede (*differences in kind*) zwischen Nicht-Mitglied und Mitglied in einer Menge erfassen können (Schneider & Wagemann 2012: 24-31; Ragin 2008: 29-34).[27]

Bei der Zuweisung der Mengenmitgliedschaften (*Kalibrierung*) kommt das qualitative, fallorientierte Moment gleich mehrfach zum Tragen: So erfolgt die Kalibrierung idealiter in einem iterativen Prozess des „back and forth between theory and empirical evidence" (Ragin 1987: 78) auf Grundlage theoretischen und fallspezifischen Wissens in Kombination mit empirischen Informationen des Datenkorpus wie z.B. Mittelwerten oder der Verteilung der Daten. Im Gegensatz zu Variablen, welche die Varianz eines Indikators eins zu eins abbilden, werden Mengen in Bezug auf das zugrundeliegende Konzept kalibriert. Nehmen wir beispielsweise das Konzept ‚hoher sozio-ökonomischer Status', welches wir über den Indikator

---

[26] Dies gilt natürlich nur insofern, als dass bei allen Personen die gleichen Hintergrundfaktoren (*scope conditions*) gegeben sind. Hier spielt die Fallauswahl hinsichtlich der Untersuchungsfrage eine zentrale Rolle für die kausale Homogenität und Generalisierbarkeit über die Fälle hinweg (siehe etwa Rohlfing 2012: 46f.).

[27] Bei sogenannten *fuzzy sets* kann die Mengenmitgliedschaft zwischen voller Mitgliedschaft (Mitgliedschaftswert 1) und voller Nicht-Mitgliedschaft (Mitgliedschaftswert 0) in beliebig vielen Stufen partieller Mitgliedschaft variieren (z.B. Mitgliedschaftswert 0.8 für eine teilweise Mitgliedschaft oder 0.4 für eine teilweise Nicht-Mitgliedschaft). Eine spezielle Variante sind dabei sogenannte *crisp sets*, die lediglich eine Unterscheidung zwischen Mitgliedschaft und Nicht-Mitgliedschaft bieten und somit dichotom sind.

monatliches Einkommen erheben. Für eine volle Mitgliedschaft könnte nun ein Wert von über 5.579 € (aktuelles W3-Einstiegsgehalt in Hessen) festgelegt werden; damit würde allen Personen mit einem höheren monatlichen Einkommen, egal ob 5.580 € oder 10.000 €, eine volle Mitgliedschaft in der Menge von Personen mit hohem sozio-ökonomischen Status zugewiesen. Damit wird durch die Kalibrierung von Mengen Varianz in den Indikatoren trunkiert – Ragin spricht hier von irrelevanter Varianz (2008: 74). Hierbei obliegt es der Forscherin, wie sie das Konzept in Abhängigkeit ihrer Theorie, Fallauswahl und empirischen Informationen definiert.

Mengentheoretische Ansätze wie QCA besitzen darüber hinaus ein spezifisches Verständnis von sozialer Realität (siehe zum Folgenden grundlegend Ragin 2008: 176-187; Schneider & Wagemann 2012: 76-89; Rohlfing 2012: 40-60; Goertz & Mahoney 2012: 16-74). So sind Mengenbeziehungen erstens durch *Asymmetrie* geprägt, welche sich in verschiedenen Kontexten zeigt. So werden etwa bei notwendigen und hinreichenden Bedingungen unterschiedliche Annahmen geprüft. Bei notwendigen Bedingungen kann Y nicht ohne X vorliegen. Hier interessieren uns folglich nur solche Fälle, die auch tatsächlich das Outcome aufweisen: Ob Menschen ohne hohe Lebenszufriedenheit gesund sind oder nicht, beeinflusst die Aussage über eine notwendige Bedingung dagegen in keiner Weise. Umgekehrt sind bei hinreichenden Bedingungen lediglich solche Fälle für den Wahrheitsgehalt der Aussage ausschlaggebend, die X aufweisen. Dies bedeutet, dass auf der einen Seite solche Fälle, die sowohl X als auch Y aufzeigen, eine hinreichende Aussage bestätigen. Auf der anderen Seite widersprechen aber all diejenigen Fälle einer hinreichenden Bedingung, in denen Bedingung X vorliegt, aber nicht das Outcome Y. Alle Fälle ohne Bedingung X sind hingegen irrelevant für die Aussage, dass X hinreichend für Y ist (siehe auch Abb. 1a und 2a).

Mengentheoretische Annahmen sind demnach nicht-linear. Bezogen auf das Beispiel Gesundheit und Lebenszufriedenheit würde ein linearer Zusammenhang etwa nicht nur beschreiben, dass *mehr* Gesundheit zu *mehr* Lebenszufriedenheit führt, sondern zugleich auch behaupten, dass *weniger* Gesundheit zu *weniger* Lebenszufriedenheit führt. Aus einer mengentheoretischen Perspektive hingegen würde man formulieren, *wenn* eine hervorragende Gesundheit, *dann* eine hohe Lebenszufriedenheit. Damit wird lediglich eine Aussage über den Zustand gemacht, wenn eine hervorragende Gesundheit vorliegt, allerdings nichts ausgesagt, was in ihrer Abwesenheit geschieht.

Ein weiterer Aspekt von Asymmetrie liegt in der Kalibrierung von Mengen begründet. Demnach muss der positive Pol eines Konzepts sich nicht zwangsläufig symmetrisch zum negati-

ven Pol verhalten. Nehmen wir etwa wieder das Konzept ‚sozio-ökonomischer Status'. Dieses Konzept kann einerseits symmetrisch verstanden werden, so dass das konzeptuelle Kontinuum beispielsweise zwischen ‚arm' und ‚reich' liegt. Es kann aber auch so konzeptuell erfasst werden, dass der negative Pol ‚nicht-reich' ist, was eine grundsätzlich andere Kalibrierung erforderlich macht und auch in der Analyse zu unterschiedlichen Schlussfolgerungen führen kann. Dies wiederum sollte bei der Kalibrierung eines Konzepts als Menge beachtet werden (Goertz & Mahoney 2012: 64-74, 161-165).

Zweitens ist die Perspektive einer QCA Y-zentriert, d.h. es wird – quasi in Rückschau – nach den Ursachen für ein bestimmtes Outcome (*causes-of-effects*) gesucht. Damit eng verbunden ist das *konfigurative Verständnis*, wonach Fälle als Konfigurationen von Bedingungen analysiert werden. Im Gegensatz zu X-zentrierten Forschungsdesigns zielt QCA somit nicht auf die Untersuchung durchschnittlicher Kausaleffekte (*average causal effects*) einzelner Erklärungsfaktoren; vielmehr wird nach Bedingungskombinationen gesucht, welche in ihrem Zusammenspiel ein bestimmtes Outcome erklären können. Zur Veranschaulichung können wir wieder auf das bekannte hypothetische Beispiel zurückgreifen: So können Individuen nicht nur (Nicht-)Mitglieder in der Menge ‚gesunder Menschen' (A), sondern zugleich auch in der Menge ‚Menschen in Partnerschaft' (B) sein. Aus diesen zwei Bedingungen ergeben sich die vier Idealtypen ‚gesunder Mensch in Partnerschaft' (A*B)[28], ‚gesunder Mensch nicht in Partnerschaft' (A*~B) sowie ‚nicht-gesunder Mensch in Partnerschaft' (~A*B) und ‚nicht-gesunder Mensch nicht in Partnerschaft' (~A*~B). Erweitert um eine dritte Bedingung, etwa der Menge ‚hoher sozio-ökonomischer Status' (C), ergäben sich bereits acht idealtypische Merkmalskombinationen.

In der Analyse von notwendigen und hinreichenden Bedingungen treten diese konfigurativen Elemente in Form von SUIN- und INUS-Bedingungen zu Tage. Das Akronym SUIN steht dabei für "sufficient but unnecessary part of a factor that is insufficient but necessary for the result" (Mahoney et al. 2009). Mit unserem obigen Beispiel könnte ein Ergebnis einer QCA etwa sein, dass die Kombination aus in einer Partnerschaft lebend ODER hohem sozio-ökonomischem Status eine notwendige Bedingung für eine hohe Lebenszufriedenheit darstellen. Der entsprechende Lösungsterm einer QCA sähe demnach wie folgt aus: B+C ← Y. Die Bedingungen B und C wären in diesem Fall wechselseitig substituierend und damit SUIN-Bedingungen. Liegt beispielsweise keine Partnerschaft vor, muss der Single demnach einen

[28] QCA basiert auf Boolescher Algebra. Das Multiplikationszeichen (*) steht für das logische UND, das Additionszeichen (+) für das logische ODER. Das Tildezeichen (~) steht für die Abwesenheit der Menge. Ein Pfeil impliziert eine notwendige (←) oder hinreichende (→) Bedingung.

hohen sozio-ökonomischem Status aufweisen, um überhaupt Lebenszufriedenheit erreichen zu können. INUS-Bedingungen sind hingegen solche, die "insufficient but necessary part of a condition which is itself unnecessary but sufficient for the result" (Mackie 1974: 62) sind. Wiederum ausgehend von unserem Beispiel könnte eine Lösung einer QCA für eine hinreichende Bedingung sein: A*B → Y. Dies würde bedeuten, dass eine hohe Lebenszufriedenheit durch die Kombination von gesunden, in einer Partnerschaft lebenden Menschen beschrieben werden kann. A UND B sind somit INUS-Bedingungen, da sie notwendiger Teil einer Bedingungskombination sind, die als solche hinreichend für das Outcome ist.

Drittens zielen Mengenbeziehungen darauf ab, *äquifinale Erklärungspfade* für ein Outcome aufzudecken. So könnte das obige Beispiel um einen weiteren Term erweitert werden und etwa (A*B) + (~B*C) → Y lauten. Hier wären sowohl die Kombinationen aus hervorragender Gesundheit UND in einer Partnerschaft lebend (A*B) als auch Single UND hoher sozio-ökonomischen Status (~B*C) hinreichende Bedingungen für das Outcome ‚hohe Lebenszufriedenheit'. Dabei zeigt sich wiederum die fallorientierte Perspektive von QCA, denn die äquifinalen Lösungsterme beschreiben oftmals unterschiedliche Fälle. Zudem verdeutlicht das Beispiel einen weiteren Aspekt der zuvor beschriebenen Asymmetrie, denn die Bedingung ‚in einer Partnerschaft lebend' (B) ist in der ersten Konfiguration in ihrer Anwesenheit hinreichend für das Outcome, hingegen in der zweiten Kombination in ihrer Abwesenheit.

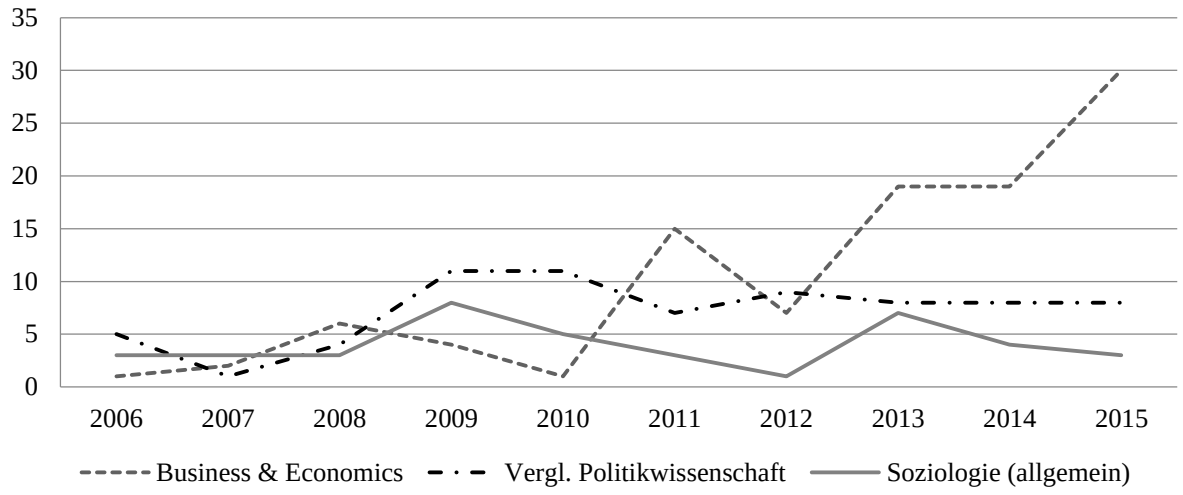
Die Entscheidung für ein Forschungsdesign mit einer QCA ist demnach eine Entscheidung für eine mengentheoretische Untersuchungsperspektive. Zusammenfassend bedeutet dies, dass erstens Konzepte als Mengen begriffen werden, in denen Fälle einen bestimmten Grad an Mitgliedschaft aufweisen können; dass zweitens ein Verständnis von komplexer sozialer Realität als konfigurativ, äquifinal und asymmetrisch vorliegt; und dass drittens Beziehungen zwischen Mengen als notwendige und hinreichende Bedingungen verstanden und untersucht werden können. Aufgrund dieser drei Merkmale eröffnen mengentheoretische Ansätze wie QCA also einen vollkommen anderen Blick auf soziale Phänomene als etwa korrelative Verfahren.

2.3 QCA in der Soziologie: Eine Rundschau

In der vergangenen Dekade hat sich QCA zweifelsohne als Forschungsdesign und -methode in den Sozialwissenschaften etabliert. Wirft man etwa einen Blick in die Datenbank publizierter Zeitschriftenartikel auf der Webseite www.compasss.org, so zeigt sich, dass allein im Jahr 2015 (Stand: 19. November 2015) bisher 91 Publikationen erschienen sind, die QCA als Methode einsetzen und sich der bisherige Trend der 2010er Jahre fortsetzt. So sind über die Hälf-

te aller in der Datenbank erfassten Publikationen, die seit „The Comparative Method“ (Ragin 1987) veröffentlicht wurden, in den letzten vier Jahren erschienen. Im Vergleich der Fachrichtungen fällt die Soziologie allerdings als Anwendungsgebiet von QCA leicht ab: Zwar rangiert sie gemeinsam mit der Politikwissenschaft und der Mikro-Ökonomie stets unter den ersten drei, doch ist gerade in den beiden anderen Disziplinen in jüngster Zeit ein größerer Zuwachs an Publikationen zu verzeichnen (siehe Abbildung 2.3).

Abbildung 2.3 Anzahl publizierter Journalartikel nach Disziplin, 2006-2015[29]



Der folgende Überblick basiert auf insgesamt 77 Studien, die über die COMPASSS-Datenbank erhoben wurden.[30] Zwar handelt es sich hierbei um eine nicht-repräsentative Auswahl, die nichtsdestotrotz ermöglicht, zentrale Trends zu QCA innerhalb der Soziologie herauszuarbeiten und zu illustrieren.

Aus der Rundschau der Zeitschriftenbeiträge ist zunächst die breite Spanne an soziologischen Fragestellungen und Gegenstandsbereichen herauszulesen, die mit Hilfe von QCA untersucht wird. Diese zeigt sich, erstens, anhand der verschiedenen Forschungsobjekte, welcher sich die untersuchten Studien annehmen. So behandelt knapp die Hälfte der Publikationen (35 von 77 Studien) makro-soziologische Themen zu Staaten (Laux 2015; Schneider & Makszin 2014a; Svevo-Cianci et al. 2010) und subnationalen (Epple et al. 2014) oder lokalen Einheiten (Blackman 2013; Wollebæk, 2010; McVeigh et al. 2006). Darüber hinaus basiert rund ein

[29] Quelle: www.compasss.org , Stand: 19. November 2015

[30] Bei der Fallauswahl beschränken wir uns auf Publikationen, die seit 2006 erschienen sind. In diesem Jahr wurden die Parameter Konsistenz und Abdeckung eingeführt (Ragin 2006), die eine wesentliche Neuerung für QCA darstellen. Neben den 40 Studien in der Rubrik ‚Soziologie (allgemein)‘ wurden 37 weitere Artikel mittels der Datenbank des Compasss-Netzwerks in den Überblick aufgenommen, die ebenfalls soziologische Fragestellungen behandeln. Eine Übersicht findet sich im Online-Appendix. Stand der erfassten Zeitschriftenbeiträge ist der 19. November 2015.

Viertel der Studien (21 von 77) auf Daten, die auf einer mittleren Analyseebene angesiedelt sind; diese haben beispielsweise soziale Bewegungen (Giugni & Nai 2013; Wright & Schaffer Boudet 2012; Kröger 2011; Giugni & Yamasaki 2009), Organisationen (Hodson et al. 2006; Grant et al. 2010) oder andere Gruppen (Vaisey 2007; Moritz et al. 2011) zum Gegenstand der Untersuchung. Ebenso viele Studien (21 von 77) betrachten schließlich die Mikro-Ebene individueller Akteure wie Schüler und Studierende (Glaesser & Cooper 2011, 2012a, 2012b; Roberts et al. 2010; Shanahan et al. 2007, 2008), Beamte und Angestellte (Baltzer et al. 2011; Braun 2013) oder Basketball-Trainer (Savage & Seebruck 2015).

Eng mit den Analysenebenen zusammenhängend, drückt sich die Vielfalt der Themen, zweitens, in den Daten aus, auf denen die einzelnen QCAs beruhen. Die allermeisten Studien greifen auf aggregierte Indikatoren oder Indizes (McLevey 2014; Hafner-Fink et al. 2013; Grant et al. 2010; Ishida et al. 2006) zurück und oftmals werden auch Daten aus Surveys (Glaesser & Cooper 2011, 2012a, 2012b; Longest & Thoits 2012; Vaisey 2007) oder Interviews (Giugni & Nai 2013; Marr 2012; Baltzer et al. 2011) zur Kalibrierung der Mengen herangezogen. Darüber hinaus finden sich aber auch qualitative Methoden der Datenerhebungen: So nutzen einige Anwendungen Informationen aus qualitativen Inhaltsanalysen (Laux 2015; Amenta et al. 2009), ethnographische Studien (Moritz et al. 2011; Hodson et al. 2006) oder teilnehmenden Beobachtungen (Vaisey 2007; Britt & Evans 2007) als Grundlage ihrer Untersuchungen.

Drittens zeigt sich die Bandbreite soziologischer Forschung, die sich QCA bedient, in den Forschungsfragen und -themen der publizierten Anwendungen. Von der Resilienz einzelner Betroffener von Naturkatastrophen (Perry & Schafer 2014) zur Entwicklung von Wohlfahrtsstaatsregimen (Ahn & Seung-yoon Lee 2012; Jang 2009), vom Ausmaß der Risiken bei der Internetnutzung durch Kinder (Bauwens et al. 2009) zu Ungleichheiten am Arbeitsmarkt (Bentele 2013; Crowley 2012, 2013, 2014; Da Roit & Weicht 2013), von sozialen Netzwerken älterer Personen (Haynes et al. 2010, 2013; Haynes 2011) zu gesundheitssoziologischen Ungleichheitsfragen (Blackman et al. 2011; Blackman & Dunstan 2010, Blackman 2008) – stets wird in vergleichender Perspektive nach dem Zusammenspiel von Bedingungen gefragt, unter denen ein bestimmtes Outcome auftritt oder nicht auftritt.

Mit Blick auf die unterschiedlichen Varianten von QCA ist bemerkenswert, dass in der Soziologie bis heute häufiger die auf die dichotome Unterscheidung zwischen Mitgliedschaft und Nicht-Mitgliedschaft begrenzte *crisp set* QCA angewendet wird als *fuzzy sets* (in 43 von 77 Studien; siehe hierzu auch Rihoux et al. 2013: 176). Dies ist insofern erstaunlich, als dass sich

soziale Phänomene in der Regel nicht (eindeutig) dichotomisieren lassen und die Entwicklung von *fuzzy sets* eine Antwort hierauf darstellt.

In Bezug auf die Anzahl der einbezogenen Erklärungsfaktoren kann konstatiert werden, dass drei Viertel der Studien (58 von 77) auf einer mittleren Zahl an Bedingungen – d.h. zwischen drei und sechs – beruhen. Allerdings werden in ca. jeder zehnten Studie (8 von 77) neun (Harris-White et al. 2013; Hussain & Howard 2013; Melinder 2007) oder gar bis zu elf (Eng & Woodside 2012; McAdam et al. 2010) Bedingungen in die Untersuchung eingespeist. Eine einzige Studie beruht hingegen nur auf dem Zusammenspiel zweier Erklärungsfaktoren (Stevenson 2013).

Die Anzahl der Bedingungen muss allerdings – wie später noch erläutert wird – immer auch im Verhältnis zur Fallzahl betrachtet werden. Hier zeigt sich, dass QCA in der Tat als Methode zur Analyse mittlerer Fallzahlen angewendet wird, denn eine große Mehrheit der Studien (48 von 77) beruht auf zehn bis ca. 50 Fällen. Allerdings ist dies nicht ausschließlich so: Während gerade einmal fünf Prozent der Studien (4 von 77) weniger als zehn Fälle untersuchen (Bleijenbergh & Roggeband 2007; Ignatow 2011; Achilov & Shaykhutdinov 2013; Da Roit & Weicht 2013), finden sich zahlreiche Publikationen, die mehrere hundert (Perry & Schafer 2014; Longest & Thoits 2012; Grant et al. 2009a, 2009b; Hodson et al. 2006) oder gar tausende Fälle (Amenta et al. 2009; Shanahan et al. 2007, 2008; Glaesser & Cooper 2011, 2012a, 2012b) einer QCA unterziehen.

2.4 Beispielstudien: Analyseschritte und potentielle Fallstricke einer QCA

Wie in Abschnitt 2 gezeigt wurde, ist QCA als Forschungsansatz tief in der qualitativ-empirischen, fallorientierten Wissenschaftstradition verwurzelt (Ragin 1994; Goertz & Mahoney 2012; Blatter & Haverland 2012). Auch wenn QCA als Methode auf Boolescher Algebra basiert und diverse Formeln, Algorithmen und Softwares zum Einsatz kommen, so muss die Forscherin in jeder Analysephase eine Vielzahl (qualitativer) Entscheidungen treffen, die sich aus ihrem Theoriewissen sowie ihrer Fallexpertise speisen sollten. So wird im Folgenden zwar eine Art „rezeptartiger“ Ablauf einer QCA skizziert (Wagemann 2015: 436), was allerdings nicht zu einer rein mechanischen Anwendung weder der Software noch existierender Maßzahlen führen darf. Vielmehr sollte jede QCA-Anwendung zwar minimale Standards beachten (vgl. Schneider & Wagemann 2010, 2012: 275-304); die zahlreichen Entscheidungen sollten aber transparent dargestellt werden (siehe Wagemann & Schneider 2015) und stets an eine fallorientierte Perspektive rückgebunden werden.

2.4.1 Schritt 1 – Kalibrierung von Mengen

Empirische Daten, die in eine QCA einfließen sollen, müssen in einem ersten Schritt zu Mengen geformt werden. Dieser Schritt wird Kalibrierung genannt. Diese Phase beinhaltet bereits ein starkes qualitatives Element, da die einer QCA zugrundeliegenden Konzepte in einem iterativen Prozess zwischen theoretischem Wissen und Empirie entwickelt, spezifiziert und zu Mengen kalibriert werden. Während die spätere Analyse mit Hilfe geeigneter Software in der Regel eher weniger Zeit in Anspruch nimmt, sind es die ‚Vorarbeiten' wie Konzeptformierung und Kalibrierung der Mengen, die den Großteil der analytischen Arbeit ausmachen.

Um Mengen sinnvoll kalibrieren zu können, d. h. den untersuchten Fällen begründet Mengenmitgliedschaften zuweisen zu können, bedarf es einer klaren Konzeptformierung (siehe etwa Goertz 2006a; Adcock & Collier 2001). Dies lässt sich in einem ersten Schritt dadurch erreichen, dass Mengen mit einem aussagekräftigen Label versehen werden, welches Mitgliedschaftsaussagen möglich macht. Während zum Beispiel ein Mitgliedschaftswert von 0.8 in Mengen wie ‚Mobilisierung' (Wright & Schaffer Boudet 2012: 743), ‚Teilzeitarbeit' (Lee 2013: 338) oder ‚Sozialbeziehung' (Marr 2012: 991) wenig aussagekräftig ist, sähe dies für alternativ bezeichnete Mengen wie ‚Soziale Bewegungen mit *hohem* Mobilisierungsgrad', ‚Staaten mit *hohem* Anteil an Teilzeitarbeit' oder ‚Personen mit *vielen* Sozialbeziehungen' bereits anders aus. Adjektive bieten sich als Labels auch deshalb an, weil sie asymmetrische Konzepte zu beschreiben vermögen. So verbleibt beispielsweise die Mengenmitgliedschaft einer Person mit einer vergleichsweise durchschnittlichen Anzahl von Sozialbeziehungen in der Menge ‚Sozialbeziehungen' unklar; sie kann sowohl Mitglied als auch Nicht-Mitglied sein. In der alternativen Menge ‚Personen mit *vielen* Sozialbeziehungen' wird dieser Fall kein Mitglied sein, da er eben nicht viele, sondern nur eine durchschnittliche Anzahl von Sozialbeziehungen aufweist. Eine genaue Bezeichnung der einer Analyse zugrundeliegenden Konzepte bietet somit nicht nur der Leserin, sondern auch der Forscherin eine Möglichkeit, die Bedeutung des Konzepts in einer Menge zu verankern.

Im Weiteren werden durch die Kalibrierung von Mengen einzelnen Fällen Mitgliedschaftswerte zugeordnet, die eine Aussage über das zugrundeliegende Konzept in sich tragen. Auch hierbei gilt es, einige Fallstricke zu vermeiden, z.B. dass eine Mengenmitgliedschaft mit dem Wert 0.5 nicht vergeben werden sollte (siehe aber z. B. Braun 2013: 831; Korczynski & Evans 2013: 775, Lee 2013: 338). Dieser Transitionspunkt bildet einen Wert, welcher genau zwischen Mitgliedschaft und Nicht-Mitgliedschaft liegt. Entsprechend wird über einen Fall mit dieser Mengenzugehörigkeit überhaupt keine Aussage getroffen, er repräsentiert das zu-

grundeliegende Konzept in keiner Weise – weder positiv noch negativ.[31] Stattdessen sollte für jeden Fall eine Entscheidung getroffen werden, ob er nun eher Mitglied oder Nicht-Mitglied ist. „[T]o avoid the ambiguity of 0.50“ vergeben etwa Harriss-White et al. (2013: 417) in ihrer symmetrischen Kalibrierung an all diejenigen Fälle, welche in ihre Mittelkategorie gehören, den Wert 0.51. Wenngleich dies eine eindeutige *numerische* Entscheidung darstellt, da nun alle diese Fälle eher Mitglied der jeweiligen Menge sind, bleibt zu beachten, dass die *konzeptuelle* Aussagekraft dieses Wertes nur sehr begrenzt ist und gegebenenfalls die Interpretation der Ergebnisse erschwert.

Ähnlich problematisch verhält es sich, wenn Daten, welche von vornherein in einer Ausprägung zwischen 0 und 1 vorliegen, nicht in Mengenmitgliedschaften übersetzt werden, sondern schlicht übernommen werden. So behalten z.B. Achilov und Shaykhutdinov (2013: 28) für ihre Menge ‚Human Development Index (HDI)‘ den Indexwert ihrer Fälle als Mengenmitgliedschaften bei. Die geringste Mengenmitgliedschaft in ‚HDI‘ kommt demnach mit einem Wert von 0.197 Kirgisistan, der höchste Mitgliedschaftswert Albanien mit 0.319 zu. Unabhängig davon, dass somit kein Fall Mitglied in der Menge ‚HDI‘ ist, da der Transitionspunkt (0.5) nicht überschritten wird, verfehlt eine solche Kalibrierung das Grundverständnis von Mengen. Demnach würden lediglich Fälle mit einem HDI von 1.00 eine volle Mitgliedschaft in der Menge ‚HDI‘ erhalten und nur solche eine volle Nicht-Mitgliedschaft, die einen HDI von 0.00 aufweisen. Dies ist weder konzeptuell noch empirisch sinnvoll. Darüber hinaus geht bei solch kleinteiligen Mitgliedschaftswerten bis zur dritten Nachkommastelle die qualitative Unterscheidbarkeit verloren. Mengenmitgliedschaften sollten stattdessen immer so vergeben werden, dass diese konzeptuell und empirisch begründbar sind.

In jedem Fall muss die Zuschreibung von Mitgliedschaftswerten für die Leserschaft nachvollziehbar und transparent gestaltet sein. Die einzelnen Entscheidungen über Vollmitgliedschaft, volle Nicht-Mitgliedschaft und Transitionspunkt sollten unter keinen Umständen unerwähnt bleiben (siehe etwa Crowley 2013, 2014) oder gar einem Computer überlassen werden, um „user-imposed biases in the analyses“ (Longest & Thoits 2012: 204) zu vermeiden. Die Kalibrierung von Mengen ist ein qualitativer Prozess, der begründeter, transparenter Entscheidungen auf der Grundlage theoretisch-konzeptuellen Wissens und der Rückbindung an die untersuchten Fälle bedarf. Um die Entscheidungen nachvollziehen zu können, sollten sowohl die Rohdaten, d. h. die nicht kalibrierten Ausgangsdaten, als auch die kalibrierten Daten jeweils

[31] Alle Fälle, die mit 0.5 kalibriert werden, können keiner Wahrheitstafelzeile eindeutig zugeordnet werden.

in einer Matrix im Text oder (Online-) Anhang präsentiert werden (siehe aber für viele McLevey 2014, Park 2013, Ahn & Seung-yoon Lee 2012).

2.4.2 Schritt 2 – Analyse notwendiger (Kombinationen von) Bedingungen

Bei der Analyse notwendiger Bedingungen werden jede einzelne Bedingung sowie ihr jeweiliges Komplement[32] betrachtet und geprüft, ob ein Outcome Y immer nur dann vorliegt, wenn auch Bedingung X gegeben ist. Sollte diese Prüfung keine notwendige Bedingung hervorbringen, so besteht noch die Möglichkeit, nach verschiedenen ODER-Kombinationen (Vereinigungsmengen / *unions*) von Bedingungen zu suchen. Technisch ist dies auf Grundlage aller Kombinationen möglich (Thiem 2014a; Bol & Luppi 2013). Konzeptuell ist es allerdings nur dann angeraten, wenn es sich bei den Bedingungen um funktionale Äquivalente handelt (Schneider & Wagemann 2012: 74). So können z. B. ‚hohe Bildung' ODER ‚hohes Erwerbseinkommen' als funktionale Äquivalente für das übergeordnete Konzepte ‚hoher sozialer Status' verstanden werden. Liegen also in sämtlichen untersuchten Fälle, die das Outcome zeigen, entweder die Bedingung ‚hohe Bildung' oder ‚hohes Erwerbseinkommen' oder beides vor, kann diese Oder-Kombination nicht nur technisch einwandfrei als notwendig identifiziert, sondern auch inhaltlich interpretiert werden. Die inhaltliche Aussagekraft von notwendigen Oder-Bedingungen ist hingegen begrenzt, wenn beispielsweise in jedem Fall, der das Outcome aufweist, ‚hohe Bildung' oder ‚in Partnerschaft' oder beides vorliegt, da diese beiden Bedingungen eben nicht in einem gemeinsamen, übergeordneten Konzept gefasst werden können.

In jedem Fall sollte die Analyse von notwendigen Bedingungen separat zur Prüfung auf hinreichende Bedingungen erfolgen (Schneider & Wagemann 2010: 8; siehe allerdings für viele Blackman 2013, Ghoshal 2013, Haynes et al. 2013). Wird sie zudem der Prüfung auf hinreichende Bedingungen vorangestellt, bringt dies einige Vorteile mit sich. Erstens kann eine etwaige notwendige Bedingung in die nachfolgende Untersuchung von hinreichenden Bedingungen einbezogen werden, etwa wenn es um den Umgang mit logischen Rudimenten geht (siehe für die sogenannte *enhanced standard analysis* Schneider & Wagemann 2012: 200-211). Zweitens unterliegt man durch die vorangestellte Analyse notwendiger Bedingungen nicht der Gefahr, „falsche notwendige Bedingungen" aus der Analyse hinreichender Bedingungen abzuleiten (Schneider & Wagemann 2012: 221-226). So identifiziert bspw. Blackman

[32] Beim Komplement einer Menge X handelt es sich um seine Negation (~X). Da Fälle in *fuzzy sets* partielle Mitgliedschaften aufweisen können, haben sie demnach auch eine partielle Mitgliedschaft im jeweiligen Komplement. Die Mitgliedschaft eines Falles im Komplement ~X berechnet sich mittels Subtraktion des Mengenmitgliedschaftswerts X von 1.

(2013: 66) aus seiner Analyse hinreichender Bedingungen eine nur scheinbar notwendige Bedingung, weil diese in allen fünf hinreichenden Lösungspfaden enthalten ist (ebenso Kim 2011, Castellano 2010). Drittens wird durch die vorgelagerte Analyse notwendiger Bedingungen auch das Problem vermieden, „versteckte notwendige Bedingungen" zu übersehen (Schneider & Wagemann 2012: 227-232).

2.4.3 Schritt 3 – Analyse hinreichender (Kombinationen von) Bedingungen

Bei der Analyse von hinreichenden Bedingungen bildet die Wahrheitstafel (*truth table*) „the core of QCA, both in the understanding of it as an approach and as a technique" (Schneider & Wagemann 2012: 91). In der Wahrheitstafel werden alle logisch möglichen Kombinationen von Bedingungen abgetragen.[33] Jeder Fall wird anschließend auf Grundlage seiner Mitgliedschaftswerte in den Bedingungen derjenigen Wahrheitstafelzeile zugewiesen, welche seine Bedingungskonfiguration idealtypisch wiedergibt. In einem weiteren Schritt wird über alle Fälle hinweg geprüft, ob eine Wahrheitstafelzeile, also eine idealtypische Kombination von Bedingungen, hinreichend für das Outcome ist oder nicht. Am Ende dieses Prozesses steht eine vollständige Wahrheitstafel, die anzeigt, welche Bedingungskonfigurationen als hinreichend identifiziert wurden und welche nicht. Da diese Ausdrücke jedoch recht komplex sein können, werden die als hinreichend identifizierten Bedingungskombinationen in einem abschließenden Analyseschritt noch vereinfacht, d. h. mit Hilfe eines auf Boolescher Algebra beruhenden Algorithmus minimiert (Schneider & Wagemann 2012: 92-104; Ragin 2008: 124-144).

Eine Wahrheitstafel aus den kalibrierten Daten zu erstellen, stellt dank der gängigen Softwarepakete wie fsQCA oder R keine Hürde dar. Dennoch wird dieses wichtige Analyseinstrument häufig weder im eigentlichen Aufsatz noch in etwaigen (Online-)Appendizes zur Verfügung gestellt (siehe für viele Wright & Schaffer Boudet 2012, Longest & Thoits 2012, Crowley 2013). Dies ist nicht nur hinsichtlich mangelnder Transparenz nachteilig; es erschwert zudem, den Umgang mit begrenzter empirischer Vielfalt (*limited empirical diversity*) nachzuvollziehen. Diese tritt auf, wenn für eine Wahrheitstafelzeile – also für eine idealtypische Bedingungskonfiguration – kein empirisch beobachtbarer Fall vorliegt (logische Rudimente, *logical remainders*). Dies wird am Beispiel von Wright und Schaffer Boudet (2012) besonders deutlich, die 20 Fälle untersuchen und 7 Bedingungen in ihre Analyse integrieren:

[33] Sowohl bei *crisp sets* als auch bei *fuzzy sets* besteht eine Wahrheitstafel aus 2^k Zeilen, wobei k für die Anzahl der Bedingungen steht und 2 für den möglichen Wahrheitswert der Bedingung (vorhanden oder nicht-vorhanden). Werden also drei Bedingungen betrachtet, besteht die Wahrheitstafel aus acht Zeilen, bei vier Bedingungen aus 16 usw.

7 Bedingungen ermöglichen 2^7 Bedingungskombinationen und damit 128 Wahrheitstafelzeilen. Selbst wenn also jeder Fall einen anderen Idealtypus repräsentieren würde, blieben 108 logische Rudimente die Folge. Blackman (2013) sowie Hussain und Howard (2013) untersuchen sogar jeweils 9 Bedingungen und betrachten lediglich 27 (Blackman) bzw. 20 (Hussain & Howard) Fälle, woraus sich mindestens 485 bzw. 492 logische Rudimente ergeben. Das bedeutet, dass 94,7% der Wahrheitstafelzeilen der ersten Studie und 96,1% der zweiten Studie ohne empirische Fälle sind.

Begrenzte empirische Vielfalt stellt deshalb ein Problem dar, weil für jede leere Wahrheitstafelzeile eine Entscheidung getroffen werden muss, ob diese potentiell auch als hinreichende Bedingung interpretiert und damit in den Minimierungsprozess einbezogen werden kann oder grundsätzlich aus der Lösung ausgeschlossen werden muss. Hierzu stehen eine Reihe unterschiedlicher Strategien zur Verfügung wie etwa die Einbeziehung etwaiger notwendiger Bedingungen, kontrafaktisches Denken (*counterfactual reasoning*) oder auch der Rückgriff auf theoretische Annahmen und empirische Ergebnisse anderer Studien (siehe für eine ausführliche Diskussion Schneider & Wagemann 2012: 151-176, 197-219; Ragin 2008: 147-175). In jedem Fall behauptet jedes logische Rudiment, das in die Minimierung eingebracht wird, dass *wenn* Fälle in dieser Merkmalskombination vorhanden wären, sie *dann* auch das Outcome zeigen würden. Die Entscheidung über logische Rudimente beeinflusst in ihrer Konsequenz natürlich auch die Lösungsterme einer QCA.[34] Es liegt also nahe zu fordern, dass diese Entscheidungen auf einer soliden theoretischen und konzeptuellen Grundlage getroffen werden sollten. Während die mittlere (*intermediate*) Lösung (z. B. Lee 2013: 340) genau diese Grundlage bietet, bezieht die sparsamste (*most parsimonious*) Lösung (z.B. Achilov & Shaykhutdinov 2013) automatisch alle logischen Rudimente mit ein, die das Gesamtergebnis weniger komplex machen. Die weitreichenden Implikationen dieser Entscheidung werden häufig nicht thematisiert. So argumentiert z. B. Blackman (2013: 66) "[f]or reasons of space, only the parsimonious solution for the twenty-seven cases is discussed here; it is the most interpretable and interesting." Der Software wurde in diesem Fall somit für mindestens 485(!) Wahrheitstafelzeilen überlassen zu entscheiden, ob diese hinreichend für das Outcome sind oder nicht;

[34] Es können grundsätzlich drei Lösungen in einer QCA unterschieden werden: die konservative Lösung (*conservative solution*) bezieht keine logischen Rudimente ein, während die sparsamste Lösung (*most parsimonious solution*) alle diejenigen logischen Rudimente in den Minimierungsprozess aufnimmt, die das Ergebnis schlanker machen. In der mittleren Lösung (*intermediate solution*) werden nur diejenigen logischen Rudimente in den Minimierungsprozess einbezogen, die sowohl die Lösung sparsamer machen als auch den theoretischen Erwartungen der einzelnen Bedingungen entsprechen.

d.h. sie konnte aus 2^{485} möglichen Kombinationen[35] diejenige wählen, welche die technisch sparsamste Lösung hervorbringt – ohne dass dies theoretisch-konzeptuell rückgebunden werden konnte und ohne dass die genutzten logischen Rudimente transparent dargestellt wurden. Dass diese Lösung dann am einfachsten zu interpretieren ist, überrascht kaum. Grundsätzlich muss allerdings festgehalten werden, dass allen Ergebnissen, egal ob sie logische Rudimente einbeziehen oder nicht, immer die gleiche empirische Basis zugrunde liegt und sie daher alle sowohl logisch richtig sind als auch der empirischen Realität nicht widersprechen. Sie unterscheiden sich allein in ihren Annahmen bezüglich der begrenzten empirischen Vielfalt. Es ist deshalb unerlässlich, dass eine QCA den Umgang mit logischen Rudimenten so transparent wie möglich thematisiert, mögliche Strategien und Annahmen offen diskutiert und am Ende der Analyse die unterschiedlichen Lösungsterme im Spiegel des Theorie- und Fallwissens der Forscherin erörtert. Es kann nicht stark genug betont werden, dass der Umgang mit logischen Rudimenten eines der wesentlichen Gütekriterien für jede QCA darstellt.

2.4.4 Schritt 4 – Konsistenz und Abdeckung als Gütemaße einer QCA

In der sozialwissenschaftlichen Forschungspraxis lassen sich häufig (Kombinationen von) Bedingungen identifizieren, die nicht perfekt, d. h. nicht für sämtliche Fälle, hinreichend bzw. notwendig sind, sondern leichte bis starke Abweichungen aufweisen. Mit Konsistenz (*consistency*) und Abdeckung (*coverage*) liegen aus diesem Grund zwei Gütemaße vor, die zur Evaluierung nicht-perfekter Teilmengenbeziehungen sowohl bei notwendigen als auch hinreichenden Bedingungen herangezogen werden können (siehe zum Folgenden Schneider & Wagemann 2012: 119-150; Ragin 2008: 44-68). Das Konsistenzmaß gibt den Gütegrad der Teilmengenbeziehungen wieder und zeigt an, wie stark sie von einer perfekten *subset/superset*-Beziehung abweicht. Der Abdeckungsparameter hingegen muss je nachdem, ob es sich um eine notwendige oder hinreichende Bedingung handelt, auf unterschiedliche Art und Weise interpretiert werden. Bei hinreichenden Bedingungen auf der einen Seite ist die Abdeckung ein Maßstab dafür, wie „breit“ eine Erklärung ist – d. h. inwieweit ein Outcome durch eine hinreichende Bedingung erklärt werden kann. Bei notwendigen Bedingungen auf der anderen Seite zeigt die Abdeckung an, wie trivial eine Bedingung ist: Ist eine Bedingung nahezu immer vorhanden, egal ob ein Outcome vorliegt oder nicht, so kann von einer trivialen notwendigen Bedingung gesprochen werden (siehe Schneider & Wagemann 2012: 233f).

[35] Die sind 99.895.953.610.111.751.404.211.111.353.381.321.783.955.140.565.279.076.827.493.022.708.011.895.642.232.499.843.849.795.298.031.743.077.114.461.795.885.011.932.654.335.221.737.225.129.801.285.632 unterschiedliche Kombinationen und schlimmstenfalls auch ebenso viele Lösungen.

Die beiden Parameter Konsistenz und Abdeckung leisten zu unterschiedlichen Phasen der Analyse wertvolle Hilfestellungen für die Bewertung der Teilmengenbeziehungen. So ist das Konsistenzmaß bei der Analyse von hinreichenden Bedingungen ein wichtiger Indikator, wenn eine Wahrheitstafelzeile nicht perfekt konsistent ist, d. h. in manchen Fällen das Outcome vorliegt und in anderen nicht. Und auch bei der Darstellung und Interpretation der Untersuchungsergebnisse von notwendigen und hinreichenden Bedingungen sind sie zwingender Bestandteil – wie z. B. bei der Diskussion äquifinaler Lösungsterme. Es ist allerdings ausdrücklich vor einer automatisierten Anwendung der Parameter zu warnen. Zwar existieren hinsichtlich akzeptabler Konsistenz- und Abdeckungsmaße bestimmte Richtwerte (z. B. durch Ragin 2009: 121 oder Schneider & Wagemann 2012: 278-79), die jedoch nur Empfehlungen oder Anhaltspunkte darstellen. Die letztendliche Evaluierung muss durch die Forscherin auf Basis ihres Theoriewissens erfolgen sowie an multiplen Punkten der Analyse an die Fälle rückgekoppelt werden.

Jede Abweichung von einer perfekten hinreichenden Bedingung, d. h. jede Entscheidung, eine (leicht) inkonsistente Kombination dennoch als hinreichend für den Outcome zu bewerten, sollte daher sorgfältig und transparent begründet werden. Häufig lässt sich jedoch beobachten, dass eine solche Begründung fehlt und stattdessen auf empfohlene Grenzwerte zurückgegriffen wird (für viele Wright & Schaffer Boudet 2012: 740), oder gar die Existenz von Grenzwerten unbelegt behauptet wird: „A threshold of 0.8 consistency (80 percent or higher) is generally accepted for supporting claims on necessary conditions." (Achilov & Shaykhutdinov 2013: 30). Eine automatisierte Anwendung der empfohlenen Grenzwerte verschleiert sowohl den Sinn der Gütemaße als auch der Empfehlungen selbst. Die oft zitierte Aussage, „consistency levels (well) above 0.75 are advisable" (Schneider & Wagemann 2012: 279) bedeutet gerade nicht, dass 0.75 der Standardwert für die Konsistenz hinreichender Bedingungen ist. Das wäre auch dahingehend fatal, als dass jede Kombination von Bedingungen automatisch als hinreichend anzusehen wäre, sobald sie drei Viertel der abgedeckten Outcomemenge richtig beschreibt. Der Grenzwert zeigt vielmehr, dass diese Grenze keinesfalls zu unterschreiten ist. Grundsätzlich gilt, dass die Gütemaße niemals diejenigen Fälle verdecken sollten, die bspw. für eine geringe Konsistenz verantwortlich zeichnen. Stattdessen sollten diese Fälle immer identifiziert werden – bspw. mittels eines XY-Plots –, um bewerten zu können, wie stark eine Kombination von Bedingungen von einer perfekten Teilmengenbeziehung abweicht.

Ein weiteres breit auftretendes Problem der untersuchten Studien stellt die geringe Gesamtabdeckung aller äquifinalen Lösungspfade dar. QCA zielt darauf ab, ein Outcome für alle bzw.

zumindest eine größtmögliche Anzahl von Fällen, in denen es auftritt, mittels verschiedener Kombinationen von Bedingungen zu erklären. Verbleibt dennoch eine geringe Abdeckung selbst einer sehr guten, d. h. konsistenten, Gesamtlösung, weist dies darauf hin, dass die gewählten Bedingungen einen Großteil des Auftretens des Outcomes nicht zu erklären vermögen. Dieses Problem wird indes häufig nicht thematisiert (z. B. Eng & Woodside 2012), sondern es wird etwa argumentiert, eine Abdeckung von 0.43 sei „in line with many other published studies using fsQCA“ (Korczynski & Evans 2013: 777). Auch wenn Abdeckungsmaße von vielerlei Faktoren abhängig sind, sollte eine höhere Abdeckung angestrebt werden und die mit einer niedrigen Abdeckung einhergehenden Probleme wiederum offen diskutiert werden.

2.5 Aktuelle Anwendungsbeispiele von QCA in der Soziologie

Im Folgenden werden drei QCA-Anwendungen vorgestellt, die die oben beschriebenen Analyseschritte nahezu umfassend durchgeführt und transparent dargestellt haben und anhand derer das Potential von QCA für soziologische Fragestellungen beispielhaft aufgezeigt werden kann. Die Auswahl der Untersuchungen versucht einerseits, eine möglichst große Bandbreite soziologischer Themen abzudecken und andererseits eine tiefergehende, detaillierte Darstellung der Studien zu gewährleisten. Aus diesem Grund wurden drei Studien ausgewählt: je eine makrosoziologische, die Nationalstaaten zum Gegenstand der Untersuchung hat, eine Studie auf der Meso-Ebene, die sich mit sozialen Gruppen auseinandersetzt, sowie eine mikrosoziologische Studie, die sich mit Individuen beschäftigt.

2.5.1 Beispiel einer soziologischen QCA auf Makro-Ebene

Schneider und Makszin (2014a) untersuchen in ihrer Studie “Forms of Welfare Capitalism and Education-Based Participatory Inequality”, unter welchen Bedingungen Wohlfahrtsstaaten geringe bildungsbedingte Ungleichheiten in der politischen Partizipation aufweisen. Der Frage liegt die Erwartung zugrunde, dass soziale und sozio-ökonomische Ungleichheiten sich in verschiedenen Wohlfahrtsstaatstypen unterschiedlich stark in ungleiche politische Partizipation übersetzen. Allein auf der Grundlage mikrosoziologischer Annahmen, dass „low-educated, low-income earners or other socially disadvantaged individuals tend to participate less in politics“ (Schneider & Makszin 2014a: 438) lassen sich Länderunterschiede in der politischen Partizipation von Bürgern nämlich nicht erklären. Daher verknüpfen die Autoren theoretisch mikrofundierte Annahmen zum Einfluss von Ressourcen und Engagement auf politische Partizipation mit makrosoziologischen Theorien, nach welchen individuelle sozio-ökonomische Unterschiede durch staatliche Maßnahmen wie Arbeitsschutz sowie monetäre

und nicht-monetäre Unterstützungsleistungen teilweise auszugleichen vermocht werden. Sie verzichten dabei explizit auf bestehende Typologien von Wohlfahrtsstaaten und fragen, „which (combinations of) welfare capitalist traits jointly influence our dependent variable, participatory inequality.“ (Schneider & Makszin 2014a: 440).

Empirisch verknüpfen Schneider und Makszin eine zweistufige Regression (auf der Individualebene und auf der Länderebene) mit einer fsQCA. Sie betrachten insgesamt 37 Länder zu je drei Zeitpunkten (1995, 2000, 2005). Für die QCA wählen die Autoren insgesamt 77 dieser Beobachtungen als Fälle aus. Sie untersuchen dabei die Kombinationen von vier Bedingungen, weshalb die Wahrheitstafel aus 16 idealtypischen Kombinationen von Bedingungen besteht. Da sich die 77 Fälle auf 15 Wahrheitstafelzeilen verteilen, stellt begrenzte empirische Vielfalt kaum ein Problem dar; es existiert lediglich ein einziges logisches Rudiment.

Das Outcome ‚*low* participatory inequality‘ ist, wie auch die vier Bedingungen ‚*high* labor market expenditure‘, ‚*high* wage coordination‘, ‚*high* union density‘ und ‚*high* employment protection‘ mit einem aussagekräftigen Label versehen. Schneider und Makszin (2014a: 449) verweisen darauf, dass der qualitative Unterschied zwischen Mitgliedschaft und Nicht-Mitgliedschaft gerade nicht in den intervall-skalierten Variablen enthalten ist und diese mengentheoretische Verankerung durch eine eigene Entscheidung erfolgen muss. Sie nutzen für sämtliche Bedingungen die direkte Kalibrierung, welche eine logarithmische Funktion über metrische Variablen legt und anhand der drei Ankerpunkte die einzelnen Mitgliedschaftswerte der Fälle berechnet. Die Autoren stellen in einem umfassenden Online-Appendix (Schneider & Makszin 2014b: 11) die Schwellenwerte für Vollmitgliedschaft, Nicht-Mitgliedschaft und Transitionspunkt transparent dar, allerdings ohne diese qualitative Wahl eingehend zu begründen. Stattdessen führen sie einen Robustheitstest (siehe hierzu Skaaning 2011) der Kalibrierung des Outcomes durch, was als innovatives Element innerhalb einer QCA bezeichnet werden kann (Schneider & Makszin 2014b: 19-22).

Schneider und Makszin (2014a: 451) identifizieren keine notwendige Bedingung für eine niedrige Ungleichheit in der Partizipation und verweisen auf die Berechnung, die ebenfalls im Anhang zu finden ist (Schneider & Makszin 2014b: 14). In der Analyse hinreichender Bedingungen minimieren sie die im Text untergebrachte Wahrheitstafel zu drei hinreichenden Kombinationen von Bedingungen, welche dann zur Typologisierung von Wohlfahrtsstaatstypen genutzt werden. So identifizieren die Autoren eine erste Gruppe von Staaten, in denen die Kombination von starkem Arbeitsschutz für Personen mit niedrigem Bildungsstand und starker Unterstützung für Arbeitslose hinreichend für eine niedrige Ungleichheit der politischen

Partizipation ist. Sie bezeichnen diesen Wohlfahrtstypus, dessen idealtypischen Vertreter Dänemark (1995) und Slowenien (1995) sind, als ‚protective support'. Darüber hinaus zeigt sich ein zweiter Wohlfahrtsstaatstypus, in welchem die Kombination aus hohen Arbeitsmarktausgaben und starken Lohnabsprachen (*wage coordination*) hinreichend für eine niedrige partizipatorische Ungleichheit ist. Dänemark (2000) und Irland (2000) stehen für diesen Typus des ‚coordinated support' Pate. Ein dritter Wohlfahrtsstaatstypus kennzeichnet sich schließlich durch die Kombination von starkem Arbeitsschutz und schwachen Gewerkschaften. Diese Form der ‚unorganized protection', wie sie Portugal (1995-2005), Mexiko (2000, 2005), Estland (2005) und Korea (2005) idealtypisch zeigen, ist ebenfalls hinreichend für niedrige politische Ungleichheiten.

Die Darstellung der Ergebnisse ist umfangreich und transparent. Im Text wird neben der Analyse mit Gütemaßen (nach Ragin & Fiss 2009) auch eine Tabelle mit typischen Fällen der Wohlfahrtstypen präsentiert; der Anhang enthält zudem die XY-Plots der drei Lösungspfade und der Gesamtlösung (Schneider & Makszin 2014b: 15-18). Während die drei Lösungspfade und die Gesamtlösung sehr hohe Konsistenzwerte aufweisen (zwischen 0.85 und 0.89), also sehr gute Erklärungen für das Outcome bieten, vermag die Lösung das Auftreten des Outcomes bei weitem nicht in Gänze zu erklären (Gesamtabdeckung der Lösung 0.58). Schneider und Makszin (2014a: 459) diskutieren diesen Punkt offen und verweisen in ihrem Fazit darauf, dass weitere Wohlfahrtstypen durch ihre Auswahl von – immerhin nur vier – Bedingungen unberücksichtigt geblieben sind.

2.5.2 Beispiel einer soziologischen QCA auf der Meso-Ebene

Cebotari und Vink (2013) untersuchen in ihrem Artikel "A Configurational Analysis of Ethnic Protest in Europe" 29 ethnische Minderheiten in Europa hinsichtlich der Frage, unter welchen Bedingungen diese starkes bzw. schwaches Protestverhalten aufweisen. Sie kritisieren bestehende, vornehmlich quantitative Studien zum einen dahingehend, dass deren Befunde auf eine große Fallzahl auch stark unterschiedlicher ethnischer Minderheiten generalisiert werden würden. Zum anderen würden bestehende Studien aufgrund ihrer additiven Regressionsmodelle behaupten, dass die gefundenen „causal conditions can independently explain whether groups are more or less mobilized" (Cebotari & Vink 2013: 298). Die beiden Autoren verweisen daher auf die Vorteile der kombinatorischen Herangehensweise von QCA und interessieren sich für verschiedene Erklärungspfade, die das Protestverhalten unterschiedlicher ethnischer Minderheiten fallorientiert erklären. Theoretisch verknüpfen Cebotari und Vink ethnopolitisch orientierte Konfliktliteratur mit den Grundlagen politischer Partizipationsfor-

schung und definieren die Mobilisierung zu ethnisch motivierten Protesten als „non-violent form of contentious politics“ (Cebotari & Vink 2013: 301). Um die unterschiedlichen Ausformungen von schwachem und starkem gewaltlosem Protest erklären zu können, integrieren sie drei gruppenbezogene und zwei institutionelle Bedingungen in den Analyserahmen. Einer hohen ‚Demokratiestärke‘ wird zugeschrieben, einen friedlichen Protest zu ermöglichen, einer starken ‚ethnischen Zersplitterung‘ als zweiter institutioneller Bedingung hingegen, hinreichend, aber nicht notwendig für starken Protest zu sein. Bei den gruppenbezogenen Bedingungen handelt es sich (1) um hohe ‚geographische Konzentration‘, welche die Kapazität beschreibt, einen wirkungsvollen Protest überhaupt zu organisieren, (2) um hohe ‚politische Diskriminierung‘, welche der jeweiligen ethnischen Minderheit widerfährt, und als Impuls für Protest dienen kann sowie (3) um starken ‚Nationalstolz‘ innerhalb der Minderheit als Proxy für eine nationale Identität, die starke Proteste unwahrscheinlich macht.

Die Autoren stellen die Kalibrierung ihrer Mengen recht ausführlich dar. Im Anhang I des Artikels beschreiben sie neben der jeweiligen Datenquelle jeweils eine „brief elaboration of calibration“ und geben die qualitativen Ankerpunkte für Voll- und Nichtmitgliedschaft sowohl Transitionspunkt an (Cebotari & Vink 2013: 319-321). Es fällt auf, dass sie die Daten jeweils symmetrisch kalibrieren, um im Weiteren die Analyse von ‚starkem Protest‘ und ‚schwachem Protest‘ mit den gleichen Mengenmitgliedschaften durchführen zu können. Aussagekräftige Labels vergeben Cebotari und Vink indes nicht; die Bedingungen heißen ‚democracy‘, ‚ethnic fractionalization‘, ‚geographical concentration‘, ‚political discrimination‘ und ‚national pride‘.

Aus den fünf Bedingungen ergeben sich 32 logisch mögliche Kombinationen. Neben den drei arithmetischen Rudimenten (32 Wahrheitstafelzeilen – 29 Fälle) stellen mehrere Fälle den gleichen Idealtypus dar, weshalb nur 17 der 32 Idealtypen mit Fällen abgedeckt werden. Obschon die Autoren die sparsamste Lösung anwenden, gehen sie transparent mit der Nutzung der 15 logischen Rudimente um (Cebotari & Vink 2013: 315).

Die Analyse der Mengenbeziehungen führen Cebotari und Vink sowohl ausführlich als auch in der oben beschriebenen Reihenfolge aus. Für ‚starken Protest‘ identifizieren sie ‚geographische Konzentration‘ als notwendige Bedingung, für ‚schwachen Protest‘ hingegen ‚Nationalstolz‘. Da es sich nicht um perfekt konsistente notwendige Bedingungen handelt, verweisen sie einerseits auf den „standard consistency threshold of 0.9“ (Cebotari & Vink 2013: 307) von Ragin (2009) und Schneider & Wagemann (2012). Andererseits erinnern sie an die

inhaltliche Bedeutung nicht-perfekter Konsistenz und sprechen daher im Weiteren nur von „quasi-necessary conditions“ (Ragin 2000).

Die (sparsamste) Analyse hinreichender Bedingungen für das Outcome ‚starker Protest‘ bringt vier Kombinationen von Bedingungen hervor, die das Auftreten des Outcomes insgesamt sehr gut (Konsistenz 0.88) und sehr breit (Abdeckung 0.81) erklären. Interessant ist hierbei, dass nur drei der vier Pfade die quasi-notwendige Bedingung ‚geographische Konzentration‘ aufweisen. Es handelt sich dabei um die Kombination von ‚geographischer Konzentration‘ mit entweder den beiden institutionellen Bedingungen ‚Demokratiestärke‘ UND ‚ethnischer Zersplitterung‘ ODER mit ‚ethnischer Zersplitterung‘, UND ‚politischer Diskriminierung‘ ODER mit ‚politischer Diskriminierung‘ UND ‚Demokratiestärke‘. Der vierte Pfad besteht nur aus der Abwesenheit der Bedingung ‚Nationalstolz‘ (Cebotari & Vink 2013: 309).

Auch die (sparsamste) Analyse ‚schwachen Protests‘ weist sehr hohe Werte auf (Konsistenz 0.82, Abdeckung 0.89). Die ebenfalls vier Kombinationen enthalten die notwendige Bedingung ‚Nationalstolz‘. Hinreichend wird diese Bedingung in Kombination mit entweder der Abwesenheit von ‚politischer Diskriminierung‘ UND der Abwesenheit von ‚ethnischer Zersplitterung‘ ODER mit der Abwesenheit von ‚politischer Diskriminierung‘ UND der Abwesenheit von ‚Demokratiestärke‘ ODER mit der Abwesenheit von ‚Demokratiestärke‘ UND der Abwesenheit von ‚ethnischer Zersplitterung‘ ODER der Abwesenheit von ‚geographischer Konzentration‘. Die Darstellung dieser sprachlich nur recht komplex formulierbaren Ergebnisse erfolgt in einer Tabelle anhand aller Fälle, die das Outcome aufweisen (Cebotari & Vink 2013: 309, 311), jedoch ohne die graphische Unterstützung von XY-Plots. Das Papier endet mit der Einordnung von QCA und der Ergebnisse in das Forschungsfeld und mögliche methodisch und inhaltlich weiterführender Designs. Eine Rückbindung der Ergebnisse an einzelne Fälle bleibt indes leider aus.

2.5.3 Beispiel einer soziologischen QCA auf der Mikro-Ebene

Marr (2012) untersucht in seinem Artikel „Pathways out of Homelessness in Los Angeles and Tokyo: Multilevel Contexts of Limited Mobility amid Advanced Urban Marginality“ das Phänomen der Massenobdachlosigkeit in Großstädten der Welt. Er fragt, unter welchen Kombinationen von ökonomischen, organisationalen und individuellen Bedingungen Wege aus der Obdachlosigkeit gefunden werden können. Marr (2012) untersucht dazu 34 Fälle – jeweils 17 Individuen, welche in Übergangsunterkünften in Tokyo und Los Angeles mehrfach interviewt wurden. Wie auch das Outcome ‚exited‘, welches den Wohnzustand der Fälle 30 Tage nach

dem Verlassen der Übergangsunterkünfte in eine *fuzzy* Menge transformiert, sind auch die vier Bedingungen sehr transparent und mit je vier qualitativ klar unterscheidbaren *fuzzy*-Werten (Vollmitgliedschaft 1, eher Mitglied als Nicht-Mitglied 0.67, eher Nicht-Mitglied als Mitglied 0.33 und Nicht-Mitglied 0) kalibriert (Marr 2012: 991). Die Menge ‚minimum wage' gibt als ökonomische Rahmenbedingung den finanziellen Spielraum der untersuchten Personen an. Die Hypothese dazu lautet, dass ein Zugang zum Niedriglohn-Arbeitsmarkt in Tokyo den Ausstieg aus der Obdachlosigkeit weitaus stärker ermöglicht als eine Arbeitsstelle im Niedriglohnsektor in Los Angeles. Die Bedingung ‚staff ally' deckt den organisationalen Einfluss ab und betrachtet, wie intensiv und konfliktbehaftet die Beziehung der untersuchten Personen mit dem Personal der Übergangsunterkunft und den örtlichen Behörden verlief. Die individuelle Bedingung ‚social tie', beschreibt, wie intensiv persönliche Kontakte mit nicht-obdachlosen Freunden und Verwandten gehalten werden. Folgende Hypothese wird zu dieser Bedingung formuliert. Demnach erwartet der Autor, dass Personen mit starken sozialen Beziehungen in Los Angeles aufgrund der kulturellen Gegebenheiten wesentlich leichter der Obdachlosigkeit entrinnen als ebensolche Personen in Tokyo. Der Bedingung ‚recently dislocated', welche persönliche Erfahrungen wie physische oder psychische Erkrankungen, frühere Phasen von Obdachlosigkeit und ähnliches umfasst, wird schließlich in beiden Städten ein geringerer Einfluss auf den Ausstieg aus der Obdachlosigkeit zugeschrieben als den ökonomischen und sozialen Bedingungen.

In der weiterführenden Analyse *clustern* die Fälle in Los Angeles stark und bilden nur fünf der 16 Idealtypen empirisch ab (11 logische Rudimente); in Tokyo hingegen werden neun der 16 Bedingungskombinationen abgedeckt (7 logische Rudimente). Während eine Analyse notwendiger Bedingungen ausbleibt, beruht die Analyse hinreichender Bedingungen auf der mittleren (*intermediate*) Lösung. Marr (2012: 993) legt transparent dar, nach welchen theoretischen Erwartungen er die logischen Rudimente auswählt, die seine Lösungspfade schmaler machen. Die Ergebnisse beider Analysen, Los Angeles und Tokyo, präsentiert er in einer Tabelle, die auch die Gütemaße umfasst. In Los Angeles vermag von den zwei hinreichenden Kombinationen (Gesamtkonsistenz 0.77, -abdeckung 0.84) insbesondere die Verbindung von ‚staff ally' UND ‚social tie' einen großen Anteil des Outcomes zu erklären (Abdeckung dieses Pfades 0.5). Erwartungsgemäß sind es somit auch empirisch vor allem die sozialen Erklärungsfaktoren, welche den Ausstieg aus der Obdachlosigkeit in Los Angeles ermöglichen. Der zweite Lösungspfad umfasst ebenfalls die Bedingung ‚social tie', allerdings in Kombination mit ‚minimum wage', also einer Arbeit im Niedriglohnbereich UND ‚recently dislocated', also früheren Erfahrungen mit Obdachlosigkeit. Auch für die befragten Personen in Tokyo

identifiziert Marr (2012: 994) zwei gut und breit erklärende hinreichende Pfade (Gesamtkonsistenz 0.93, -abdeckung 0.93), von denen einer allein aus der Bedingung ‚minimum wage' besteht (Abdeckung dieses Pfades 0.6). Erwartungsgemäß kommt der ökonomischen Bedingung somit eine starke Erklärungskraft zu. Der zweite Pfad besteht aus der Kombination aus früheren Erfahrungen mit Obdachlosigkeit (‚recently dislocated') UND der Abwesenheit sozialer Bindungen (‚~social ties'). Abschließend bindet Marr (2012) die Ergebnisse an die theoretischen und konzeptuellen Überlegungen rück.

2.5.4 Zusammenfassend-vergleichende Betrachtung der drei Anwendungen

In der vergleichenden Analyse der drei Studien auf Makro-, Meso- und Mikroebene lässt sich im Wesentlichen dreierlei erkennen. Zum einen zeigt sich die Vielfalt der möglichen Anwendungen und Datenquellen, aus welchen Mengen kalibriert werden können, als eine zentrale Stärke von QCA. Während die ersten beiden Studien vornehmlich quantitative Indikatoren nutzen und durch qualitative Entscheidungen in aussagekräftige Mengen transformieren, die das zugrundeliegende (teilweise sogar asymmetrische) Konzept repräsentieren, nutzt Marr (2012) auch Interviewdaten, Krankenakten und Informationen aus Lebensläufen, um seine Mengen zu kalibrieren. Zum zweiten nutzen alle drei Studien die Methode QCA mit großer Sorgfalt und im Wissen um ihre Probleme und Fallstricke. Mit beträchtlichem Aufwand werden qualitative Entscheidungen transparent und umfassend begründet sowie Gütemaße und der häufig schwierige Umgang mit logischen Rudimenten besprochen. Dennoch zeigt sich zum dritten, dass selbst die hier als *best practices* vorgestellten Studien nicht perfekt sind. Schneider und Makszin (2014a) legen die Ankerpunkte ihrer Kalibrierung offen, begründen jedoch nicht, warum sie diese gewählt haben. Cebotari und Vink (2013) identifizieren aus ihren Daten je eine (quasi-)notwendige Bedingung für das Auftreten und Nicht-Auftreten des Outcomes, nutzen diese Information jedoch nicht, um ihre logischen Rudimente für die sparsamste Lösung auszuwählen (siehe Schneider & Wagemann 2012: 200-211). Marr (2012) analysiert schließlich notwendige Bedingungen überhaupt nicht. Auch dass drei Studien bei gleicher methodischer Herangehensweise drei unterschiedliche Formen der Darstellung ihrer Ergebnisse wählen, ist auf den ersten Blick suboptimal. Positiv gewendet, bieten die Studien somit eine Übersicht über die gängigsten Formen der Ergebnissicherung von QCAs: (1) eine einfache Tabelle der minimierten Lösungsterme (Marr 2012: 994), (2) eine Tabelle mit den Mitgliedschaften jedes das Outcome aufweisenden Falles in jeder hinreichenden Kombination von Bedingungen (Cebotari & Vink 2013: 209) oder (3) die von Ragin und Fiss (2009) empfohlene Darstellung von Kern- und peripheren Bedingungen, welche um XY-Plots erweitert wird (Schneider & Makszin 2014a: 452, 2014b: 15-18). Zudem weisen alle drei Darstellungs-

formen die wichtigsten Informationen auf – es werden sowohl die einzelnen Lösungspfade und die komplette Lösung mit ihren Gütemaßen korrekt wiedergegeben als auch eine Rückbindung an die untersuchten Fälle ermöglicht.

2.6 Abschließende Bemerkungen zum Mehrwert von QCA für die Soziologie

Zielsetzung dieses Beitrags war es, den originären Mehrwert sowie Forschungsperspektiven und -potentiale von QCA für soziologische Frage- und Themenstellungen am Beispiel von publizierten Journalartikeln darzulegen. Dabei sind drei Aspekte besonders hervorzuheben:

Erstens unterscheiden sich mengentheoretische Ansätze wie QCA grundlegendend von korrelativen Analyseverfahren – sei es in ihren zentralen Bausteinen der Analyse wie Mengen vs. Variablen, oder in ihren Funktionslogiken. So eröffnet die Suche nach notwendigen und hinreichenden Bedingungen einen gänzlich anderen Blick auf soziale Phänomene als beispielsweise etablierte statistische Methoden. Goertz und Mahoney (2013) sprechen in diesem Zusammenhang auch von einem methodologischen Rorschach-Test. Dabei kann die Einnahme einer mengentheoretischen Perspektive gegebenenfalls Erklärungszusammenhänge aufdecken, welche nicht durch Korrelationen erfasst werden (siehe Ragin 2008; Goertz & Mahoney 2012). Darüber hinaus ist QCA mit seinem grundlegenden Verständnis besonders sensitiv gegenüber Diversität wie beispielsweise äquifinaler Erklärungsfaktoren, INUS- und SUIN-Bedingungen oder asymmetrischer Wirkungszusammenhängen. Insofern sollte QCA eine mehr als willkommene Erweiterung des sozialwissenschaftlichen Methodenkanons darstellen.

Methodologisch bietet QCA, zweitens, einen besonderen Zugang zu *mixed- und multi-methods* Designs (vgl. u.a. Ebbinghaus 2014: 365; Hollstein 2014). Zum einen kann – und sollte – QCA als mengentheoretischer Ansatz durchaus mit anderen Methoden kombiniert werden. Studien wie beispielsweise die Arbeiten von Schneider und Makszin (2014a) oder Hollstein und Wagemann (2014) verdeutlichen auf innovative Art und Weise, wie unterschiedliche methodische Ansätze wie etwa Regressionen oder Netzwerkanalysen mit QCA verknüpft werden können. Weitere Möglichkeiten zur Verbindung von QCA mit anderen methodischen Herangehensweisen stellen detaillierte Einzelfallstudien oder *process tracing* zur Aufdeckung von Kausalmechanismen dar; hierzu existieren einige Vorschläge für eine systematische Fallauswahl im Anschluss an eine QCA (grundlegend Rohlfing & Schneider 2013; Schneider & Rohlfing 2013, 2014). Zum anderen kann QCA selbst als *mixed-methods* Ansatz verstanden werden (siehe insbesondere Hollstein 2014: 17). Während Elemente wie Kalibrierung und Konzeptformierung, iteratives Hin-und-Her zwischen Theorie und Empirie, konfigurratives Denken oder die Fallorientierung das *Qualitative* in QCA betonen, unterstreichen

Merkmale wie der systematische Fallvergleich selbst mittlerer und größerer Fallzahlen, die Transformation von Informationen in Zahlen (sogenannte *data-set observations*; siehe Brady & Collier 2004), oder der Einsatz von mathematischen Regeln, Formeln und Parametern das *Quantitative* in QCA.[36] Ragin (1987) selbst hat QCA deshalb schon früh als „dritten Weg" zwischen quantitativer und qualitativer Forschung bezeichnet (siehe auch Wagemann 2015).[37]

In diesem Sinne könnte eine stärkere Berücksichtigung von QCA, drittens, auch die allgemeine methodologische Diskussion in den Sozialwissenschaften bereichern. So werden etwa in der quantitativ ausgerichteten Soziologie aktuell Diskussionen geführt, die zahlreiche konzeptuelle Schnittmengen mit QCA aufzeigen, wie zum Beispiel die Argumentationslogiken kontrafaktischer Kausalinferenz von Morgan und Winship (2007). Schließlich würde umgekehrt auch QCA von einem konstruktiven methodologischen Dialog profitieren, da grundlegende Probleme wie Transparenz, Reliabilität und Validität, Visualisierung, Robustheit und Interpretation von Ergebnissen einen gemeinsamen Problemhorizont aller Methoden darstellen. Bestes Beispiel ist hier sicherlich die gegenwärtige (scharf geführte) Debatte um den (Un-)Sinn von Simulationen als Robustheitschecks bei QCAs (einführend siehe Rohlfing 2015a).[38]

Aus einer inhaltlichen Perspektive kann festgehalten werden, dass die Rundschau der publizierten Studien das Potential von QCA für die Soziologie deutlich offengelegt hat. So kann mittels QCA eine Vielfalt unterschiedlicher Themen- und Fragestellungen untersucht werden, gleich ob auf Mikro-, Meso- oder Makro-Ebene. Insbesondere zeigt sich, dass sowohl qualitative Informationen als auch quantitative Daten in die Analyse integriert werden können. Darüber hinaus konnte herausgestellt werden, dass QCA keineswegs auf mittlere Fallzahlen beschränkt ist, sondern durchaus auch bei large-N Designs angewendet werden kann. Entscheidendes Kriterium für eine QCA ist vielmehr das oben beschriebene Grundverständnis von QCA als mengentheoretischem und fallorientiertem Ansatz zur Analyse von notwendigen und hinreichenden Bedingungskonfigurationen.

---

[36] Gerade den deutschsprachigen Sozialwissenschaften, in denen qualitatives Arbeiten in der Regel mit interpretativ-hermeneutischen bzw. Sinn-verstehenden Methoden gleichgesetzt wird (vgl. etwa Hollstein & Ullrich 2003; Lamnek 2010), dürfte es folglich nicht unbedingt leicht fallen, das Q in QCA als qualitativ anzuerkennen.

[37] Im Französischen findet sich auch das Akronym *AQQC*, das für *Analyse Quali-Quantitative Comparée* steht (u. a. bei DeMeur und Rihoux 2002). An anderer Stelle wird sogar ganz auf das Q verzichtet und stärker das konfigurative und fallorientierte Grundverständnis in den Vordergrund der Methode gestellt (vgl. Rihoux & Ragin 2009; Hall 2003; Blatter & Haverland 2012).

[38] Für einen weiterführenden Einstieg in die Diskussion siehe Lucas & Szatrowski 2014; Collier 2014; Ragin 2014; Seawright 2014; Vaisey 2014; Krogslund & Michel 2014a; Krogslund et al. 2015; Hug 2013, 2014; Thiem 2014b; Schneider & Rohlfing 2014.

Anhang: Analysierte Zeitschriftenartikel[39]

Achilov, Dilshod, and Renat Shaykhutdinov. 2013. "State Regulation of Religion and Radicalism in the Post-Communist Muslim Republics." *Problems of Post-Communism* 60 (5):17-33.

Ahn, Sang-Hoon, and Sophia Seung-yoon Lee. 2012. "Explaining Korean Welfare State Development with New Empirical Data and Methods." *Asian Social Work and Policy Review* 6 (2):67-85.

Amenta, Edwin, Neal Caren, Sheera J. Olasky, and James E. Stobaugh. 2009. "All the Movements Fit to Print: Who, What, When, Where, and Why SMO Families Appeared in the New York Times in the Twentieth Century." *American Sociological Review* 74 (4):636-56.

Baltzer, Maria, Hugo Westerlund, Mona Backhans, and Karin Melinder. 2011. "Involvement and Structure: A Qualitative Study of Organizational Change and Sickness Absence among Women in the Public Sector in Sweden." *BMC Public Health* 11 (1):318-34.

Bauwens, Joke, Bojana Lobe, Katia Segers, and Liza Tsaliki. 2009. "A Shared Responsibility." *Journal of Children and Media* 3 (4):316-30.

Bentele, Keith Gunnar. 2013. "Distinct paths to higher inequality? A Qualitative Comparative Analysis of rising earnings inequality among U.S. States, 1980-2010." *Research in Social Stratification and Mobility* 34 (0):30-57.

Blackman, Tim. 2008. "Can Smoking Cessation Services be Better Targeted to Tackle Health Inequalities? Evidence from a Cross-Sectional Study." *Health Education Journal* 67 (2):91-101.

Blackman, Tim. 2013. "Exploring Explanations for Local Reductions in Teenage Pregnancy Rates in England: An Approach Using Qualitative Comparative Analysis." *Social Policy and Society* 12 (1):61-72.

Blackman, Tim, and Katie Dunstan. 2010. "Qualitative Comparative Analysis and Health Inequalities: Investigating Reasons for Differential Progress with Narrowing Local Gaps in Mortality." *Journal of Social Policy* 39 (03):359-73.

Blackman, Tim, Jonathan Wistow, and David Byrne. 2011. "A Qualitative Comparative Analysis of Factors Associated with Trends in Narrowing Health Inequalities in England." *Social Science & Medicine* 72 (12):1965-74.

Bleijenbergh, Inge, and Conny Roggeband. 2007. "Equality Machineries Matter: The Impact of Women's Political Pressure on European Social-Care Policies." *Social Politics: International Studies in Gender, State and Society* 14 (4):437-59.

Braun, Caelesta. 2013. "The Driving Forces of Stability: Exploring the Nature of Long-Term Bureaucracy-Interest Group Interactions." *Administration & Society* 45 (7):809-36.

Britt, David W., and Mark I. Evans. 2007. "Sometimes Doing the Right Thing Sucks: Frame Combinations and Multi-Fetal Pregnancy Reduction Decision Difficulty." *Social Science & Medicine* 65 (11):2342-56.

Cárdenas, Julián. 2012. "Varieties of Corporate Networks: Network Analysis and fsQCA." *International Journal of Comparative Sociology* 53 (4):298-322.

[39] Quelle: www.compasss.org , Stand: 19. November 2015.

Castellano, José María. 2010. "Secondary Education and Broadband Diffusion: A Qualitative Comparative Analysis." *Info* 12 (6):121-38.

Cebotari, Victor, and Maarten P. Vink. 2013. "A Configurational Analysis of Ethnic Protest in Europe." *International Journal of Comparative Sociology* 54 (4):298-324.

Crowley, Martha. 2012. "Control and Dignity in Professional, Manual and Service-Sector Employment." *Organization Studies* 33 (10):1383-406.

Crowley, Martha. 2013. "Gender, the Labor Process and Dignity at Work." *Social Forces* 91 (4):1209-38.

Crowley, Martha. 2014. "Class, Control, and Relational Indignity: Labor Process Foundations for Workplace Humiliation, Conflict, and Shame." *American Behavioral Scientist* 58 (3):416-34.

Da Roit, Barbara, and Bernhard Weicht. 2013. "Migrant Care Work and Care, Migration and Employment Regimes: A Fuzzy-Set Analysis." *Journal of European Social Policy* 23 (5):469-86.

Eng, Sereikhuoch, and Arch G. Woodside. 2012. "Configural Analysis of the Drinking Man: Fuzzy-Set Qualitative Comparative Analyses." *Addictive Behaviors* 37 (4):541-3.

Epple, Ruedi, Martin Gasser, Sarah Kersten, Michael Nollert, and Sebastian Schief. 2014. "Institutions and Gender Time Inequality: A Fuzzy-Set QCA of Swiss Cantons." *Swiss Journal of Sociology* 40 (2):259-78.

Flagg, Julia A. 2015. "Aiming for Zero: What Makes Nations Adopt Carbon Neutral Pledges?" *Environmental Sociology* 1 (3):202-12.

Ghoshal, Raj Andrew. 2013. "Transforming Collective Memory: Mnemonic Opportunity Structures and the Outcomes of Racial Violence Memory Movements." *Theory and Society* 42 (4):329-50.

Giugni, Marco und Alessandro Nai, 2013. Paths Towards Consensus: Explaining Decision Making within the Swiss Global Justice Movement. In *Swiss Political Science Review*, Vol. 19(1), 26-40.

Giugni, Marco und Sakura Yamasaki, 2009. The Policy Impact of Social Movements: A Replication Through Qualitative Comparative Analysis. In *Mobilization: An International Quarterly*, Vol. 14 (4), 467-484.

Gjølberg, Maria. 2009. "The Origin of Corporate Social Responsibility: Global Forces or National Legacies?" *Socio-Economic Review* 7 (4):605-37.

Glaesser, Judith. 2008. "Just how Flexible is the German Selective Secondary School System? A Configurational Analysis." *International Journal of Research & Method in Education* 31 (2):193-209.

Glaesser, Judith, and Barry Cooper. 2011. "Selectivity and Flexibility in the German Secondary School System: A Configurational Analysis of Recent Data from the German Socio-Economic Panel." *European Sociological Review* 27 (5):570-85.

Glaesser, Judith, and Barry Cooper. 2012a. "Educational Achievement in Selective and Comprehensive Local Education Authorities: A Configurational Analysis." *British Journal of Sociology of Education* 33 (2):223-44.

Glaesser, Judith, and Barry Cooper. 2012b. "Gender, Parental Education, and Ability: Their Interacting Roles in Predicting GCSE Success." *Cambridge Journal of Education* 42 (4):463-80.

Grant, Don, Alfonso Morales, and Jeffrey J. Sallaz. 2009. "Pathways to Meaning: A New Approach to Studying Emotions at Work." *American Journal of Sociology* 115 (2):327-64.

Grant, Don, Kathleen O'Neil, and Laura Stephens. 2009. "Pluralistic Ignorance Among Assembled Peers." *Sociological Perspectives* 52 (1):59-79.

Grant, Don, Mary N. Trautner, Liam Downey, and Lisa Thiebaud. 2010. "Bringing the Polluters back in: Environmental Inequality and the Organization of Chemical Production." *American Sociological Review* 75 (4):479-504.

Hafner-Fink, Mitja, Brina Malnar, and Samo Uhan. 2013. "The National Contexts of Post-National Citizenship." *Czech Sociological Review* 49 (6):867-901.

Harriss-White, Barbara, Wendy Olsen, Penny Vera-Sanso, and V. Suresh. 2013. "Multiple Shocks and Slum Household Economies in South India." *Economy and Society* 42 (3):398-429.

Haynes, Philip. 2011. "Are Scandinavian Countries Different? A Comparison of Relative Incomes for Older People in OECD Nations." *Social Policy & Administration* 45 (2):114-30.

Haynes, Philip, Laura Banks, and Michael Hill. 2013. "Social networks amongst older people in OECD countries: a qualitatative [*sic*] comparative analysis." *Journal of International and Comparative Social Policy* 29 (1):15-27.

Haynes, Philip, Michael Hill, and Laura Banks. 2010. "Older People's Family Contacts and Long-term Care Expenditure in OECD Countries: A Comparative Approach Using Qualitative Comparative Analysis." *Social Policy & Administration* 44 (1):67-84.

Hodson, Randy, Vincent J. Roscigno, and Steven H. Lopez. 2006. "Chaos and the Abuse of Power." *Work and Occupations* 33 (4):382-416.

Hörisch, Felix. 2012. "Die Bestimmungsfaktoren der Einführung der Unternehmensmitbestimmung in den OECD-Staaten - Eine fuzzy-set Qualitative Comparative Analysis" *Zeitschrift für Sozialreform* 58 (1):33-57.

Hussain, Muzammil M., and Philip N. Howard. 2013. "What Best Explains Successful Protest Cascades? ICTs and the Fuzzy Causes of the Arab Spring." *International Studies Review* 15 (1):48-66.

Ignatow, Gabe. 2011. "What has Globalization Done to Developing Countries' Public Libraries?" *International Sociology* 26 (6):746-68.

Ishida, Atsushi, Miya Yonetani, and Kenji Kosaka. 2006. "Determinants of Linguistics Human Rights Movements: An Analysis of Multiple Causation of LHRs Movements Using a Boolean Approach." *Social Forces* 84 (4):1938-55.

Jang, Dong-Ho. 2009. "Significance of Variations between Income Transfers and Social Care Services Development." *Journal of Comparative Social Welfare* 25 (1):37-48.

Kim, Seung Hyun. 2011. "On the Historical Determinants of Third Sector Strength: A Qualitative Comparative Analysis." *Social Science Journal* 48 (4):641-50.

Kim, Kyo-Seong, and Yeonjung Lee. 2008. "A Qualitative Comparative Analysis of Strategies for an ageing Society, with special Reference to Pension and Employment Policies." *International Journal of Social Welfare* 17 (3):225-35.

Korczynski, Marek, and Claire Evans. 2013. "Customer abuse to service workers: An analysis of its social creation within the service economy." *Work, Employment & Society* 27 (5):768-84.

Kröger, Markus. 2011. "Promotion of contentious agency as a rewarding movement strategy: Evidence from the MST-paper industry conflicts in Brazil." *Journal of Peasant Studies* 38 (2):435-58.

Laux, Thomas. 2015. "Nationalstaatliche Prozesse oder globale Strukturen? Eine Analyse der Mechanismen zur Gleichstellung von Frauen im Recht." *Berliner Journal für Soziologie* 24 (4):531-58.

Lee, Sophia Seung-Yoon. 2013. "High Non-Standard Employment Rates in the Republic of Korea and Japan: Analyzing Policy Configurations with Fuzzy-Set/QCA." *Policy and Society* 32 (4):333-44.

Longest, Kyle C., and Peggy A. Thoits. 2012. "Gender, the Stress Process, and Health: A Configurational Approach." *Society and Mental Health* 2 (3):187-206.

Marr, Matthew D. 2012. "Pathways out of Homelessness in Los Angeles and Tokyo: Multilevel Contexts of Limited Mobility amid Advanced Urban Marginality." *International Journal of Urban and Regional Research* 36 (5):980-1006.

McAdam, Doug, Hilary S. Boudet, Jennifer Davis, Ryan J. Orr, W. Richard Scott, and Raymond E. Levitt. 2010. "Site Fights: Explaining Opposition to Pipeline Projects in the Developing World." *Sociological Forum* 25 (3):401-27.

McCammon, Holly J., Soma Chaudhuri, Lyndi Hewitt, Courtney Sanders Muse, Harmony D. Newman, Carrie Lee Smith, and Teresa M. Terrell. 2008. "Becoming Full Citizens: The U.S. Women's Jury Rights Campaigns, the Pace of Reform, and Strategic Adaptation." *American Journal of Sociology* 113 (4):1104-47.

McLevey, John. 2014. "Think Tanks, Funding, and the Politics of Policy Knowledge in Canada." *Canadian Review of Sociology* 51 (1):54-75.

McVeigh, Rory, Carl Neblett, and Sarah Shafiq. 2006. "Explaining Social Movement Outcomes: Multiorganizational Fields and Hate Crime Reporting." *Mobilization: An International Journal* 11 (1):23-49.

Melinder, Karin A. 2007. "Socio-Cultural Characteristics of High versus Low Risk Societies Regarding Road Traffic Safety." *Safety Science* 45 (3):397-414.

Mitar, Miran. 2010. "Bailey's Social Entropy Theory as an Explicit Theoretical Approach for an Empirical Assessment of Security of Contemporary Societies." *Quality & Quantity* 44 (5):941-55.

Moritz, Mark, Julia Giblin, Miranda Ciccone, Andréa Davis, Jesse Fuhrman, Masoumeh Kimiaie, Stefanie Madzsar, Kyle Olson, and Matthew Senn. 2011. "Social Risk-Management Strategies in Pastoral Systems: A Qualitative Comparative Analysis." *Cross-Cultural Research* 45 (3):286-317.

Park, Sung Ho. 2013. "Capital Openness, Monetary Integration, and Wage-Setting Coordination in Developed European Countries." *Economic and Industrial Democracy* 34 (4):637-66.

Perry, Ashley M., and Mark J. Schafer. 2014. "Resilience in Louisiana FEMA Parks: A Person-Centered, Fuzzy-Set Analysis." *Sociological Spectrum* 34 (1):39-60.

Põder, Kaire, Kaie Kerem, and Triin Lauri. 2013. "Efficiency and Equity Within European Education Systems and School Choice Policy: Bridging Qualitative and Quantitative Approaches." *Journal of School Choice* 7 (1):1-36.

Portes, Alejandro, and Lori D. Smith. 2010. "Institutions and National Development in Latin America: A Comparative Study." *Socio-Economic Review* 8 (4):585-621.

Roberts, Ros, Richard Gott, and Judith Glaesser. 2010. "Students' Approaches to Open-Ended Science Investigation: The Importance of Substantive and Procedural Understanding." *Research Papers in Education* 25 (4):377-407.

Savage, Scott V., and Ryan Seebruck. 2015. "The Road to the Top: A Qualitative Comparative Analysis of Mobility in the Elite Labor Market of College Basketball Coaching." *Sociological Focus* 48 (3):208-27.

Schneider, Carsten Q., and Kristin Makszin. 2014. "Forms of Welfare Capitalism and Education-Based Participatory Inequality." *Socio-Economic Review* 12 (2):437-62.

Seeleib-Kaiser, Martin, and Timo Fleckenstein. 2009. "The Political Economy of Occupational Family Policies: Comparing Workplaces in Britain and Germany." *British Journal of Industrial Relations* 47 (4):741-64.

Shanahan, Michael J., Lance D. Erickson, Stephen Vaisey, and Andrew Smolen. 2007. "Helping Relationships and Genetic Propensities: A Combinatoric Study of DRD2, Mentoring, and Educational Continuation." *Twin Research and Human Genetics* 10 (2):285-98.

Shanahan, Michael J., Stephen Vaisey, Lance D. Erickson, and Andrew Smolen. 2008. "Environmental Contingencies and Genetic Propensities: Social Capital, Educational Continuation, and Dopamine Receptor Gene DRD2." *American Journal of Sociology* 114 (S1):S260-S86.

Stevenson, Ian. 2013. "Does Technology have an Impact on Learning? A Fuzzy Set Analysis of Historical Data on the Role of Digital Repertoires in Shaping the Outcomes of Classroom Pedagogy." *Computers & Education* 69 (0):148-58.

Svevo-Cianci, Kimberly A., Stuart N. Hart, and Claude Rubinson. 2010. "Protecting Children from Violence and Maltreatment: A Qualitative Comparative Analysis Assessing the Implementation of U.N. CRC Article 19." *Child Abuse & Neglect* 34 (1):45-56.

Vaisey, Stephen. 2007. "Structure, Culture, and Community: The Search for Belonging in 50 Urban Communes." *American Sociological Review* 72 (6):851-73.

Valtonen, Kathleen, Jacqueline Cameron Padmore, Michele Sogren, and Letnie Rock. 2009. "Lived Experiences of Vulnerability in the Childhood of Persons Recovering from Substance Abuse." *Journal of Social Work* 9 (1):39-60.

Wollebæk, Dag. 2010. "Volatility and Growth in Populations of Rural Associations." *Rural Sociology* 75 (1):144-66.

Wright, Rachel A., and Hilary Schaffer Boudet. 2012. "To Act or Not to Act: Context, Capability, and Community Response to Environmental Risk." *American Journal of Sociology* 118 (3):728-77.

Yamasaki, Sakura. 2009. "A Boolean Analysis of Movement Impact on Nuclear Energy Policy." *Mobilization: An International Journal* 14 (4):485-504.

3.

QCA and Business Research: Work in Progress or a Consolidated Agenda?

(Claudius Wagemann, Jonas Buche, Markus B. Siewert)[40]

[40] The authors are grateful to the student assistants Yulia Aleshchenkova and Christoph Klement for their help in collecting data.

Abstract

Qualitative Comparative Analysis (QCA) in its different variants has become increasingly prominent over the last years in business and management research. While this notable trend is, in general, more than welcome, the applications in the field have to hold against some minimal standards of the QCA community. This study is based on an assessment of 61 published, peer-reviewed journal articles applying QCA within the field of business and management studies. It demonstrates several major and minor flaws and presents several suggestions how to overcome these shortcomings or avoid them in the first place.

Keywords

Qualitative Comparative Analysis, Fuzzy Set Analysis, Standards & Good Practices, Assessment of Applications.

3.1 Introduction: QCA meets Business Research

When speaking about science a recurrence to scientific methods is inescapable. Hardly any historian can present her findings without indicating the mode of analyzing the sources, and hardly any research result in the natural sciences can be shared with a scientific audience without clearly describing the (experimental) proceedings. Indeed, over the centuries all sciences have developed a methodological toolkit that allows them to communicate their findings to the scientific community, but also to a wider audience. The social sciences do not differ in this regard. Nevertheless, the high differentiation of the social sciences themselves has also led to a vast variety of methodological approaches. Psychology, for instance, seems to consider itself close to the logic of the natural sciences since experiments and their derivate, random-based statistical analysis dominate the discipline. Some parts of sociology employ sophisticated mathematical models in order to achieve distinct results while others are inspired from once again divergent hermeneutic methods that are common in human sciences. The political science tradition adds comparative case studies to this portfolio, which go back to ancient understandings of reasoning derived from classic philosophy.

Studies looking at the economy (broadly understood, no matter if in the field of economics, management or business research) are often identified as being close to a quantitative template—the mathematical formalism in economics has even been characterized as "monist methodology" (Dow 2007: 457). Taking note of some exceptions such as case studies, interviews or interpretative approaches (for an overview, see Doz 2011; Priore 2006; Starr 2014), "qualitative research in economics has traditionally been relatively unimportant compared to quantitative work" (Starr 2014: 238). This presumed predominance of quantification is not surprising since quantities are also of special *substantial* importance—it seems natural to quantify central analytical categories such as earnings, losses, success, markets, and internationalization.

As the present study shows a notable amount of research is nevertheless based on an alternative methodological approach that has also become prominent in other social science disciplines such as sociology and political science: Qualitative Comparative Analysis (QCA). QCA departs from a different epistemological foundation than mainstream statistical techniques such as set theory and models causal patterns as set relations. QCA has gained early prominence in political science and sociology right after its first publication (Ragin 1987) and is increasingly applied in other disciplines such as linguistics (Ishida et al. 2006), organization

studies (Fiss et al. 2013a), or—as this study will discuss—business and management research (for publication rates across several fields, see Rihoux et al. 2013).

This article proceeds as follows. Section 2 summarizes the current state of QCA. Section 3 gives an overview of QCA applications in business research. Section 4 assesses these applications with regard to central features a properly done QCA should include by adapting a list of criteria that Schneider and Wagemann (2010) present. Section 5 presents the discussion in a broad methodological context.

3.2 QCA: Approaching an Acronym

Nearly thirty years after the start of the "Ragin Revolution" (Vaisey 2009), QCA has become a kind of bone of contention in the social sciences. A lively discussion about real pitfalls and perceived shortcomings of QCA has been started, for example, in the symposium in Sociological Methodology (Collier 2014; Lucas & Szatrowski 2014; Ragin 2014; Seawright 2014; Vaisey 2014) and in the latest newsletters by the Qualitative and Mixed-Methods Section of the American Political Science Association (Vol. 12(1) and 12(2); see also Hug 2013; Krogslund et al. 2015; Krogslund & Michel 2014a; Paine 2016; Schneider & Rohlfing 2014). These debates clearly indicate that QCA represents an important addition to the available methods in the social sciences.

QCA became known to a greater audience through the book *The Comparative Method* by Ragin (1987). Its aim was to move beyond qualitative and quantitative approaches and thus to propose a "third way". The use of Boolean algebra formalizes QCA results although it has certainly not only been Ragin who developed this logic for the social sciences (for a non-QCA centered description of applications of Boolean algebra to comparative social science research, see Caramani 2009). Nevertheless, a debate is important—which is not very intensively conducted, though—whether QCA is something different from the distinction of qualitative and quantitative methods or if it is rather a qualitative method: note that the "Q" stands for "qualitative".

QCA underwent a major boost with Ragin's (2000) second important book on QCA which focuses more on sets than on Boolean algebra. Indeed, Ragin went beyond dichotomous set membership—being a member in a set versus not being a member—and allowed for partial set memberships. With this, he opened QCA for the use of fuzzy sets that had already been around in mathematics and computer science for quite some time (Klir et al. 1997; Zadeh 1965, 1969). This had two effects on QCA. First, QCA was more applicable to social science

questions because, living in a non-dichotomous world, researchers had always had the difficulty to squeeze their analytical categories into dichotomous sets. A business had to be either successful or not, a country had to be either a democracy or a non-democracy, an economy had either to be growing or not, etc. Social science research in all disciplines becomes more realistic and thus more appropriate for a world that is not considered a mere agglomeration of dichotomies. Second, the introduction of fuzzy sets "quantified" QCA in the sense that the way in which its results were presented (graphs, formula, use of parameters of fit, in a first version including also tests of significance, Ragin 2000: 112) reminded users and observers of practices and modes in quantitative social science research. This had at least two consequences. On the one hand, the generally qualitative nature of QCA could be doubted. Qualitative researchers often feel alienated by procedures for which some technical and mathematical knowledge is required. On the other hand, looking like regression but not being it (Grofman & Schneider 2009; Seawright 2005), this quantification of QCA opened the door to a huge array of critiques from the quantitative faction which—correctly—identified QCA as different from the logic of statistics.

In the following years QCA was further consolidated (for an overview of different phases, see Marx et al. 2014). Some developments originating in the 2000 book (such as fuzzy adjustments and significance tests) were abandoned while others were newly introduced—such as parameters of fit like consistency and coverage—until all-encompassing presentations of QCA (Ragin 2008) were—also in textbook format (Schneider & Wagemann 2007, 2012) —produced. Certainly, this process of developing QCA has not ended yet. Other variants, such as multi-value QCA (mvQCA; Cronqvist 2005; Cronqvist & Berg-Schlosser 2009; for comments, see Schneider & Wagemann 2012: 255 ff.; Thiem 2013; Vink & Van Vliet 2013) and temporal QCA (tQCA; Caren & Panofsky 2005; Ragin & Strand 2008) have been proposed. A new concentration on large-N analyses (Fiss et al. 2013b; Greckhamer et al. 2013) is encouraged. New facets and problems, such as robustness checks, measurement error and simulations (Emmenegger et al. 2013; Emmenegger et al. 2014; Hug 2013; Maggetti & Levi-Faur 2013; Skaaning 2011; Schneider & Wagemann 2012) are discussed and even new procedures such as Coincidence analysis (CNA; Baumgartner 2009) or decision tree models (Krogslund & Michel 2014b) were introduced. While QCA is sufficiently consolidated—in the sense of being available for a broader range of users over the last years—the advancements in various fields of QCA are thriving and rapidly evolving.

QCA is frequently considered to be a typical "mid-sized N" method, in reference to those analyses that work with intermediate numbers of cases and are thus neither appropriate for

small-N in-depth analysis nor for those statistical analyses for which large numbers are needed. It is not wrong to see this as a central characteristic of QCA. However, an approach from an epistemological perspective is more appropriate (Schneider & Wagemann 2012: 12). QCA is clearly rooted in set theory. This has (at least) three implications. First, instead of assigning values to variables, the researcher attributes set membership scores to cases. If "internationalization" of a company is considered, then the study assesses every company under analysis in how far it belongs to the set of internationalized companies. So-called fuzzy values indicate set membership. This is not precisely identical to measuring the degree of internationalization of a company. Rather, a set is defined that comprises all "internationalized companies" and criteria are introduced that define when and in how far cases are (partial) members of the set (Ragin 2008: 71 ff.; Schneider & Wagemann 2012: 23 ff.). Secondly, working with sets enables researchers to conceptualize intersections and unions of sets. Researchers might be interested in those companies that are not only internationalized but are also firmly rooted in a long business tradition. Thus, they create the intersection between the set of internationalized companies and the set of long-standing companies. Technically speaking, this aspect of QCA is important for executing the formal analysis (Schneider & Wagemann 2012: 42 ff.). Epistemologically speaking, this alludes to the configurational reasoning that is typical for QCA (see below). Furthermore, this connects set theory and the elaboration of typologies (Cacciatore et al. 2015; Elman 2005; Kvist 2006, 2007). Thirdly, set relations can also have a causal notion (for more details, see Schneider & Wagemann 2012: 56 ff.). Subset and superset relations indicate the presence of necessary and sufficient conditions, including their more complex derivates INUS (Mackie 1965; Schneider & Wagemann 2012: 79) and SUIN (Mahoney et al. 2009: 126; Schneider & Wagemann 2012: 79 f.) conditions. Necessary conditions are those conditions that have to be present in order for an outcome to occur. There cannot be any outcome without the necessary conditions. Sufficient conditions imply the outcome. The outcome is present whenever the sufficient condition is present. INUS conditions are parts of a combination of conditions that are jointly sufficient (but not necessary) for the outcome. SUIN conditions are, to sum it up very briefly, conditions that are alternatively necessary—which means that at least one of the SUIN conditions has to be present in order that the outcome can occur.

Already this short overview of set theory makes clear that the main goal of QCA is to work out—through means of set relation analysis—the conditions for a given outcome. While sufficiency and necessity are central categories of social science analysis (Goertz & Mahoney 2012; Schneider & Wagemann 2012: 53), their formal notation is a major advancement of

QCA. Thus, in other words, QCA is a new way to conduct analyses as they had already been around before the introduction of QCA. A clear asset of QCA is the development of an algorithm and a formal system for the analysis of sufficiency and necessity of conditions for given outcomes (for the role of necessity and sufficiency in in other case oriented approaches, see Mahoney 2003; for Comparative Historical Analysis, see Mahoney & Terrie 2009; and for process tracing, see Beach & Pedersen 2012; Collier 2011; Mahoney 2012).

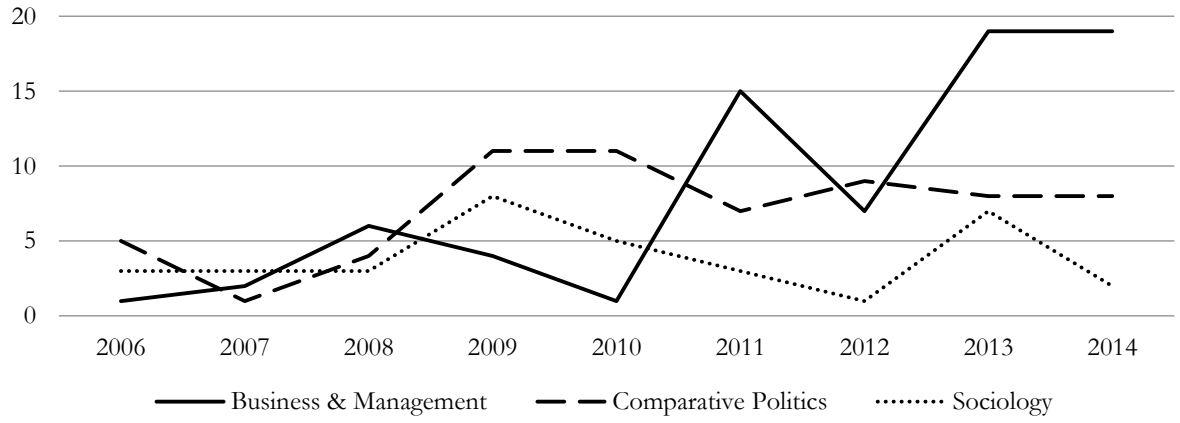
As a new way to approach social science questions QCA has important consequences. QCA literature underlines three basic features of causal complexity that are well mirrored in QCA (for more on this, see Schneider & Wagemann 2012: 76 ff.). The first of these notions lies in configural causation. This means that causes rarely occur isolated from one another; they rather work in combination. A measure of business reform might not deploy its effectiveness unless the researcher combines the measure with another measure. Secondly, QCA enables equifinal causal statements. If, for instance, a condition (or a combination of conditions) is sufficient but not necessary then this means that there has to be at least one more (usually many more) condition (or a combination of conditions) that is also sufficient but not necessary. Thus, a few to several explanations for the outcome occur, and they are all equally valid. The outcome can be explained in different modes. These modes can also represent different cases or case clusters. For example, if the success of businesses is analyzed, there might be more than one (partially overlapping) combination of conditions that implies the outcome of successful businesses. Success is explained in different ways for different firms. This stands in sharp contrast to statistical analysis, which is usually uni-finally oriented. A third feature is asymmetry: the explanation of the absence of the outcome (i.e. the negation of the phenomenon under analysis) can—apart from very specific situations—not be directly derived from the explanation of the presence of the outcome. This is also a fundamental difference from quantitative methods where the explanation for the dependent variable is the same and it does not matter whether the dependent variable takes low or high values, or if it has a value of zero.

In short, QCA is different from other methods in that it allows for an analysis of causal diversity which is characterized by equifinal, configural, and asymmetric causal relations. Causal relations are assessed through set-theoretical operations. This makes it possible to work out sufficient and necessary conditions. Such a kind of causality might correspond better to the epistemological bases of social science research than other ways of modeling and understanding complex social phenomena (Hall 2003: 374).

3.3 QCA in Business Research: A Birds-Eye View

In recent years, the publication rates of QCA applications accelerated remarkably. According to the bibliographical COMPASSS database at www.compasss.org, 39 journal articles that apply QCA were published both in 2011 and 2012, 60 in 2013, and 60 in 2014. Given the overall number of 397 empirical applications covered by the database in the period from Ragin's first book in 1987 until 2014, it is astonishing that half of all journal articles using some variant of QCA have been published during the last four years. A look at the research disciplines to which these publications belong shows that, first, most articles using QCA still originate from political science and its different subfields. Secondly, (macro-)sociology, which for a long time has produced the second largest amount of QCA studies (see Rihoux et al. 2013: 177), is now falling behind. Thirdly, business and economics as well as management and organization studies currently constitute the most active environment regarding empirical QCA publications in scientific journals in the most recent past (see figure 3.1).

Figure 3.1 Number of articles by discipline, 2006-2014



The following overview on QCA applications in business and economics as well as management and organization research focuses on 61 journal articles (see appendix) published since the implementation of the parameters of fit by Ragin (2006) until December 2014. These articles are selected from the COMPASSS database. Thus, the list does not represent an exhaustive collection of QCA publications allowing for generalizable arguments but an illustrative sample. However, the sample is both large in number and without any bias in selection. This allows to detect trends in the application of QCA in this field. The compilation shows a very large variety of topics, levels and units of analysis, number of cases, number of conditions, and even research designs QCA can be used for.

The examined issues are highly diverse. Many applications of QCA focus on classical macro level data like the relationships between developed and developing economies towards their delivery of wealth and equality (Judge et al. 2014). Others deal with new EU member states' flat tax systems (Freitas et al. 2011) or business regulation policies (Allen & Aldred 2013). And yet other studies address micro level questions like individual costumers' assessments of service facets in beauty salon and spa treatments (Wu et al. 2014), or the tourism behavior of Australia visitors from different countries (Woodside et al. 2011). Finally, some studies are classifiable as meso-level QCA applications dealing with, for example, energy policy planning networks and the key positions within those networks held by some of the leading energy firms (Crawford 2012), U.S. industry sectors and their (non-)innovativeness as well as (non-)profitableness (Stanko & Olleros 2013), or multinational enterprises and environmental performance standardization (Aguilera-Caracuel et al. 2014).

Moreover, the variety of case numbers is striking. The studies with the largest numbers of cases (*N*) in the sample analyze the different forms of family involvement in 6.611 firms (García-Castro & Casasole 2011) and the impact on family involvement on the business and financial performance of 6.592 firms, respectively (García-Castro & Aguilera 2014). The smallest *n* in the sample can be found in a study on the adoption of information and communication technology by five Australian and five Croatian enterprises (Skoko et al. 2006). Yet another article on the evolution of large firms in paper industry investigates the particular "flagship" firm of six European countries (Järvinen et al. 2009) while another conducts a QCA on nine semi-structured interviews in the American, British and Japanese offices of a major Japanese pharmaceutical company (Magnier-Watanabe & Senoo 2008). However, most of the studies provide a QCA on a medium sized number of cases as, for instance, the empirical assessment of the national business systems typology that is based on 30 OECD countries (Hotho 2014), the analysis of inter-organizational technology transfer success of 68 firms (Leischnig et al. 2014), or the investigation on strategic management of customer satisfaction which relies on 51 detailed telephone interviews with customers undertaking live housing projects (Forsythe 2012).

The number of explanatory factors taken into consideration differs. While Weissenberger-Eibl and Teufel (2013) combine three conditions (as do Chang et al. 2013 and Park 2013), most studies rely on a number between four (e.g. Cárdenas 2012; Grandori & Furnari 2008; Marx 2008) and eight (e.g. Fiss 2011; Judge et al. 2014; Woodside & Zhang 2013) conditions. However, some applications include even nine (e.g. Forsythe 2012; Hotho 2014), ten (Freitas et al. 2013; Woodside & Zhang 2012) or eleven (Boudet et al. 2011) conditions.

The list of studies exposes another key feature of QCA: sets can be calibrated from a wide range of data sources, for example from book length ethnographies (Crowley 2012; Korczynski & Evans 2013) and interview data (Bakker et al. 2011; Crilly et al. 2012; Winand et al. 2011) to survey data (Cheng et al. 2013; Ordanini et al. 2014) and purely quantitative indicators like unemployment rates (Park 2013), annual growth rates in the consumption of paper and board products (Järvinen et al. 2012) or market capitalization of indigenous listed companies and high-technology exports (Allen & Aldred 2011). Indeed, many authors triangulate qualitative and quantitative information sources. For instance, Winand et al. (2011) who study performance scores of sport governing bodies from the French-speaking community in Belgium rely on quantitative performance scores for the calibration of their outcome while 36 qualitative interviews are the sources for the calibration of the conditions. Ordanini and Maglio (2009) base their outcome set on in-depth telephone interviews with 39 hotel managers but, "for the sake of robustness" (Ordanini & Maglio 2009: 611), add observed measures of performance to the calibration.

Interestingly, the kinds of data sources do not tell much about the calibration technique applied by the authors. This regards, for instance, the question whether to use a (dichotomous) crisp-set QCA or a fuzzy-set QCA. While some studies relying on statistical measures opt for the fuzzy-set version (Boudet et al. 2011; Schneider et al. 2010), others do not (Seeleib-Kaiser & Fleckenstein 2009; Winand et al. 2013). In general, however, and in contrast to other fields (see Rihoux et al. 2013: 176), in business and management research, the ratio of the variants is clearly in favor of fuzzy sets. More than two thirds of the sample studies calibrate their sets not just as dichotomies by defining differences in kind, but additionally define differences in degree (see Schneider & Wagemann 2012: 14 f.). By using fuzzy set calibration they avoid the loss of information that crisp sets imply by default.

Within the universe of fuzzy set calibration, the number and (a)symmetry of different degrees of memberships varies remarkably. It ranges from different qualitative differentiations between four (e.g. Chandra & Prabhu 2014; García-Castro et al. 2013; Meuer 2014) and seven fuzzy membership values (e.g. Harriss-White et al. 2013; Maggetti 2007, 2014: Woodside & Zhang 2012) to the so-called "continuous" calibration (e.g. Bell et al. 2014; Chang & Cheng 2014; Fiss 2011; Kim 2013) which is based on the direct calibration option implemented in fsQCA software (Ragin & Davey 2014). For instance, Crilly (2011) considers the outcome set "scope of stakeholder orientation" that varies between the conceptual poles "broad" and "narrow" to fit in a four-scale fuzzy set with the values 0 and 0.33 indicating a (partial) non-membership and 0.67 and 1 indicating a (partial) membership. In contrast, García-Castro and

Casasole (2011) opt for the direct calibration of their outcome set "family ownership" in order to grasp fine-grained differences in the underlying variable.

This overview shows the potential of QCA in at least five points. First, QCA is not limited to a specific field of research, but constitutes an appropriate method for a broad variety of topics. Secondly, the units of analysis can be chosen from different levels, for example, from the micro level of firms or costumers to the macro level of countries. Thirdly, QCA encompasses both qualitative and quantitative data. Fourthly, QCA is not limited to mid-sized *n* but can deal with very small and (very) large numbers of cases. Fifthly, a major feature of QCA is present in all studies of the sample—for instance in the understanding of social reality as an array of complex phenomena that calls for an approach that is able to capture configurations of conditions, asymmetric patterns, and equifinal explanations.

3.4 Does QCA in Business Research Avoid Common Pitfalls?

Methods do not just exist in textbooks: the very kind of their application convinces readers of their usefulness. Concerning QCA Schneider and Wagemann (2010) elaborate on "standards of good practice in QCA" (updated in Schneider & Wagemann 2012: 275 ff.; applied to a sample of 77 peer-reviewed articles in sociology in Buche & Siewert 2015). While selected voices see this as too authoritative (e.g. Cooper et al. 2014; Thiem 2014a), the necessity of such a "code of good conduct" is demonstrated by its various breaches that can be observed.

As mentioned above, QCA is applied to a wide array of case samples ranging from a minimum of five up to more than 6.500 cases. QCA was initially introduced on the premise that it provides a useful tool for the analysis of small- to medium-N samples overcoming what is often referred to as the problem of "too many variables, too few cases" (e.g. Lijphart 1971, 1975). However, this is only partially true because the ratio between the number of cases and the number of conditions has several consequences: first, the introduction of every additional condition doubles the number of truth table rows—and thus, amplifies the possibility of rows without empirical reference (*limited diversity*, see Schneider & Wagemann 2012: 151 ff.). The risk connected to this phenomenon is to base findings on too little variation of available cases. Additionally, the results of a QCA can easily encompass too many conditions within a single solution term so that a meaningful interpretation is difficult. However, increasing the number of cases does not guarantee that these problems are avoidable. For instance, even when case numbers are large, cases can cluster in only a few truth table rows. If the sample is rather heterogeneous, the probability of inconsistencies increases which then has an influence on the parameters of fit such as consistency and coverage.

Usually, this becomes especially problematic when very few cases are analyzed: Allen and Aldred (2011, 2013), for instance, analyze ten cases with seven respectively eight conditions, Järvinen et al. (2009) six cases with seven conditions, and Järvinen et al. (2012) eight cases with six conditions, while Skoko et al. (2006) examine just five cases with seven conditions. As the previous section outlines, several studies in the sample even use more than eight conditions which is problematic, even if notable numbers of cases are investigated. For example, Valliere et al. (2008) include ten conditions and 40 cases in their analyses resulting in at least 984 logical remainders. Harriss-White et al. (2013) base their QCA on 91 cases and twelve conditions which means that 4.096 truth table rows exist out of which at least 4.005 are not covered by empirically observable cases. Boudet et al. (2011) use eleven conditions to explain the outcome in 15 cases which results in at least 2.033 logical remainders. In all these publications, conclusions follow from very low portions of empirical information in comparison to the theoretically imaginable diversity.

Another aspect is the question of calibration, meaning the assignment of (fuzzy) set membership values to single cases. This is a core activity of every QCA. It is directly linked to concept formation and thus has to meet shared standards of validity, reliability, and replicability. The thresholds defining set memberships need to explicitly discussed and justified on the grounds of the insights gained from theory, case knowledge or the underlying empirical data. Just around one third of all studies in the database either present a table displaying the calibration thresholds or the set data matrix (or both)—something that would be a minimal standard of transparency to allow the reader to track calibration decisions (Wagemann & Schneider 2015). Unfortunately, several studies in the database do not live up to these criteria. Some (e.g. Aguilera-Caracuel et al. 2014; Cardenas 2012; Leischnig et al. 2014; Ordanini et al. 2014) indicate the anchor points for the calibration of the outcome and the conditions (Schneider & Wagemann 2012: 27 ff.). However, they do not provide sound discussions on how they arrive at these thresholds. Other studies (e.g. Chang & Cheng 2014; Cheng et al. 2013; Crowley 2012; Freitas et al. 2013; Ganter & Hecker 2014; Seeleib-Kaiser & Fleckenstein 2009) give virtually no information about the calibration decisions.

Besides describing and justifying the assignment of membership thresholds, calibration should go beyond a mere transformation of an ordinal or metric variable into a set with membership values ranging between 0 and 1. Rather, membership scores should be able to capture the underlying concept as closely as possible. Ragin (2008: 79) defines this as the interpretation of data related to specific concepts under study. Calibrations such as those, for example by Allen and Allen (2015) or Leischnig et al. (2014), which assign the full membership score

to the case with the observed maximum, the full non-membership to the observed minimum, and the crossover point to the mean are problematic for several reasons. First, the conceptual meaning of these anchor points is completely obscure. Secondly, only one single case receives a full membership and another case a full non-membership score. This contradicts the notion of "irrelevant variance" that is attributed to sets (seminally Ragin 2008: 71 ff.). Thirdly, the anchor points should be, at best, external to the distribution of cases within the data. While this is often not feasible because strong theoretical arguments regarding the thresholds are missing from the literature, with calibrations like the ones mentioned above, set membership values change when the research adds or removes cases. This is particularly questionable with regard to the 0.5 anchor definition using the arithmetic mean, because adding or removing cases necessarily changes the mean so that cases are "moved" across the threshold. Ultimately, the difference in kind is altered. For example, cases that conceptually belonged to the set might not belong to it any more without the set membership criteria having conceptually changed.

If possible, avoid the empirical allocation of the 0.5 crossover threshold to cases in fuzzy sets. This has, on the one hand, technical reasons because cases with a set value of 0.5 in one of the conditions cannot be univocally assigned to a truth table row which results at best in just software difficulties and interpretational dilemmas, and at worst, in excluding cases from the analysis by default. From a conceptual point of view, the 0.5 set value means that cases are neither in nor out of the set. Thus, no statements are being made about these cases with regard to the given concept. In the consequence, they are not useful for any conclusion based on that concept. Within the database 17 out of 42 fsQCA studies assigned the 0.5 anchor to at least one condition or the outcome set (e.g. Chang et al. 2013; Cheng et al. 2013; Korczynski & Evans 2013; Magetti 2014; Meuer 2014; Young & Poon 2013). Other studies (starting with Fiss 2011, and from there spreading to Crilly 2011, Greckhamer 2011 and Woodside & Zhang 2013) add a constant of 0.001 to every set value below 1.0 in order to avoid the allocation of the 0.5 anchor. This is arbitrary and should not become common practice. All previously undecided cases are automatically rather considered members of the set than not, something that does not have much to do with a decision about set membership. Following a similar logic, the value of 0.001 could also be subtracted from the actual values (e.g. Crilly et al. 2012).

Turning to the "analytical moment" (Schneider & Wagemann 2012: 11) of QCA, it is recommended to analyze necessity and sufficiency separately—another aspect of asymmetry in set theory. While some authors—without further justification—claim that the analysis of necessary conditions should be performed after checking for sufficient conditions (Marx et al.

2013; see for an example Kim 2013) the reversed order is advisable for two reasons. First, results from the analysis of necessary conditions can be useful for the analysis of sufficiency and the treatment of logical remainders. In case a necessary condition can be identified, truth table rows (no matter if existing ones or logical remainders) that do not show this condition can be automatically excluded from the minimization process (Schneider & Wagemann 2012: 201 f.). Secondly, by analyzing necessity first, the pitfall of deducing a "false" necessary condition from the analysis of sufficiency is avoidable. This refers to the possibility that a condition is part of every sufficiency solution term and risks to be mistaken for a necessary condition without being it. Only analyzing necessary conditions first can draw the attention to such a situation (Schneider & Wagemann 2012: 220 ff.).

Strikingly, approximately three fourth of the studies in the database do not check for necessary conditions at all, thus demonstrating the "sufficiency bias" in QCA (Schneider & Wagemann 2012: 220). Skoko et al. (2006), Bakker et al. (2011), Ordanini and Maglio (2009), Ordanini et al. (2014), or Judge et al. (2014) take the bait and derive their necessary conditions from sufficiency tests, thus risking the pitfall of "false" necessary conditions.

With regard to the analysis of sufficiency, nearly all studies nicely describe (sometimes even in textbook style) the process of constructing a truth table. The actual truth table, however, appears only in 17 out of 61 publications. Again, this limits the transparency and replicability of these studies because the truth table gives important information about the consistency of each configuration as well as the question of limited empirical diversity. Its publication is thus essential to track the decisions made by the researcher and should be included either directly in the publication or some form of (online) appendix. Virtually all fsQCA publications on which this article is based actually do make their decision about consistency cut offs transparent. In the dataset, the accepted level of inconsistencies varies between thresholds of 0.65 (Pajunen 2008) and 0.96 (Järvinen et al. 2012). In the justification of their decisions the vast majority of publications refers to the benchmarks between 0.75 and 0.85 once established by Ragin (2006, 2008) or Schneider and Wagemann (2012). Another often applied strategy is looking at gaps between consistency scores (e.g. Crilly 2011; Crilly et al. 2012; Schneider et al. 2010). While these practices are commonly applied, they should not be implemented mechanically. The suggested benchmarks are rather guidelines, above all regarding lower bounds. Still, researchers should critically review their decisions, for instance by looking for those cases to which lower consistency scores can be traced.

Another striking feature of many sufficiency analyses is the low coverage of the solution formula. For instance, an overall coverage of 0.039 (Garcia-Castro & Aguilera 2014) means that all equifinal solution paths together only cover 3.9% of the outcome set—different from what the authors think, it does not mean that 3.9% of the cases are covered. This low value is not even an exception among the articles under scrutiny (0.13 and 0.15 in Woodside & Zhang 2012; 0.21 and 0.27 in Stanko & Olleros 2013; 0.22 in Ganter & Hecker 2014; 0.27 in one of the analyses presented in Järvinen et al. 2009; 0.36 in Fiss 2011; 0.38 in Garcia-Castro et al. 2013; 0.43 in Korczynski & Evans 2013; and 0.43 and 0.47 in Judge et al. 2014). Thus, the empirical explanatory power is only marginal because a large portion of the outcome is not explained by the solution term. Such low coverage scores can be the result of different aspects such as calibration, neglected explanatory factors, or heterogeneity and skewedness of cases' set memberships. It can also be the case that the conditions just do not explain the outcome. While establishing specific benchmarks on coverage is clearly unrewarding, the researcher should problematize and discuss such low scores and their possible sources (see for example Ganter & Hecker 2014).

Another pitfall the researcher has to deal within a truth table analysis is the question of limited diversity, i.e. single truth table rows that are "empty" because they do not refer to empirical references. While it was impossible for 11 out of 61 studies (Bell at al. 2014; Chang et al. 2013; Crowley 2012; Grandori & Furnari 2008; Kim 2013; Skoko et al. 2006; Stanko & Olleros 2013; Vailliere et al. 2008; Weissenberger-Eibl & Teufel 2013; Wooodside et al. 2011; Young & Poon 2013) to reconstruct whether logical remainders existed at all—due to the missing truth table and no discussion of the issue of limited diversity—basically all studies in the dataset encounter limited diversity. In general, the awareness of logical remainders is rather high. Most authors refer to the problem when outlining the different options for solutions (conservative, intermediate and most parsimonious). Nevertheless, the treatment of logical remainders applies nearly always mechanically via the software and decisions about the inclusion and exclusion of logical remainders are not discussed (but see Greckhamer 2011; Winand et al. 2011, 2013). This is particularly problematic when the subsequent interpretation of the results is based on the most parsimonious solution—for which the software includes all logical remainders that make the solution term more parsimonious—but also for the intermediate solution—where the software decides on logical remainders based on theoretical assumptions processed by the researcher. For instance, none of these standardized procedures provides any safeguards against making contradictory simplifying assumptions or including impossible logical remainders (Schneider & Wagemann 2012: 200 ff.). In these cases, it is the research-

er's duty to deal actively with empty truth table rows, for example by employing counterfactual reasoning, using insights from the analysis of necessity or of external information derived from previous studies. Of course the more logical remainders exist the more problematic becomes the non-discussion of their handling, and the more questionable are any interpretations based on intermediate or most parsimonious results. Allen and Aldred (2013), for instance, base their analysis on the intermediate solutions by examining ten cases with eight conditions leading to a minimum of 246 logical remainders. In their two QCAs, Harriss-White et al. (2013) have to deal with at least 1.957 respectively 4.005 logical remainders without discussing their treatment or even naming which solution they present. Thus, for the most parsimonious solution, the researchers let the software pick the most parsimonious solution out of $2^{1.957}$ and $2^{4.005}$ possible solutions.

3.5 QCA and Business Research: How Big Is the Intersection?

There is an impressive application of QCA in all social science disciplines, and the recent surge in business and management studies is a most notable trend. The range of application of QCA in this field is very broad and by no means limited to the niches of the discipline. While there are certainly contributions that correctly observe the recommendations made by the "standards of good practice" (Schneider & Wagemann 2010) and the broader QCA-community, important aspects of a good-quality QCA are overseen, neglected, or even violated. Observing the proposals for such a standard is no luxury option. Rather, deviations from it either render the analysis nontransparent and thus violate the standard of replicability (for example the non-provision of important information about the decisions taken), or they can actually lead to wrong interpretations (such as basing the analysis on too much limited diversity paired with a bad relation between the number of cases and the number of conditions, or arriving at conclusions that only cover very small percentages of the outcome). A first systematic analysis of articles in sociology (Buche & Siewert 2015) or insights in political science sub-fields (Wagemann 2014, 2015) show that these phenomena are not limited to business studies. Nevertheless, the impression is that the latter deviations are rather frequent in business research and even extreme in the sense that transparency issues and recommendations about coverages or limited diversity are not only marginally but substantially violated. In some studies, the violations are even so strong that the research results are questionable. Note that all the studies went through a peer-review process.

The reasons behind these observations remain unclear. However, there is a general tendency in business research to render QCA analyses similar to research modes in standard statistical

analyses. Limited diversity is not discussed frequently in statistics (although being present, Schneider & Wagemann 2012: 157 ff.), low R^2 values (as equivalents to coverage) are usually accepted, as long as they are significant, parsimony in the results is to be aimed at, case knowledge is less important than net effects of variables, etc. Moreover, new software is in development. A group of scholars has started to develop packages for R (Schneider et al. 2012; Thiem & Dusa 2013) that find an ever-larger diffusion, thus replacing the freeware fsQCA, which many researchers did not evaluate as particularly user-friendly. Eventually, R will be the default option. While this underlines the very lively interest to improve the practicalities of QCA and to make the method more attractive for potential users, a mere concentration on software development also enhances the view of QCA as an automatized algorithm.

Another development contributing to the quantification of QCA is the revival of a discussion that challenges QCA from a mainly quantitatively oriented point of view (Collier 2014; Hug 2013, 2014; Krogslund & Michel 2014a; Krogslund et al. 2015; Lucas & Szatrowski 2014; Paine 2016; Ragin 2014; Schneider & Rohlfing 2014; Seawright 2014; Thiem 2014b; Vaisey 2014). Apart from more fundamental and epistemologically inspired doubts about the usefulness of QCA, specific analytical aspects need discussion, such as the robustness of QCA analyses (Emmenegger et al. 2013; Emmenegger et al. 2014; Hug 2013; Maggetti & Levi-Faur 2013; Skaaning 2011; Thiem 2014c). In most of these critiques, the mission and foundations of QCA are either not well understood or wrongly interpreted as wanting to be the better regression, no matter whether the attacks are more fundamental or more problem-centered. Nevertheless, this stands in contrast to a development since the discussion about QCA relates back to a debate about case study methods in general. Over the last years, case studies were presented with regard to theory development (George & Bennett 2005; Rueschemeyer 2003). In addition, process tracing literature has recently been rendered more systematic (Beach & Pedersen 2012; Bennett & Checkel 2014; Collier 2011; Mahoney 2010, 2012) and important textbooks offer easy-to-read approaches to case study research (Blatter & Haverland 2012; George & Bennett 2005; Rohlfing 2012).

However, approaching QCA as Qualitative Comparative Analysis reflects at least three fundamental and partially overlapping principles for proper applications, namely transparency, the need for justification, and awareness. Transparency means that both the quality and the replicability of QCA rely to a large extent on an open discussion about which qualitative decisions are made (Wagemann & Schneider 2015). This refers to case selection, the choice of explanatory factors and their operationalization, calibration, the presentation of analytical steps, and the thresholds for consistency and coverage values. The need for justification leads

back to the question why these qualitative decisions were made. While theoretical and case knowledge is the gold standard but not always at hand, empirical knowledge provides another source of suitable motivations. Finally, awareness implies the knowledge on the consequences of qualitative decisions. For instance, what consequences a certain ratio of cases and conditions has on the analysis. What happens if the most parsimonious solution is applied or which cases are responsible for the low consistency score. The combination of transparency, justification, and awareness reduces the risk of an automatized, mechanistic application of QCA and might even be necessary for a proper QCA, although—unfortunately—not sufficient.

Appendix: Analyzed Journal Articles[41]

Aguilera-Caracuel, Javier, Eugenio M. Fedriani, and Blanca L. Delgado-Márquez. 2014. "Institutional Distance among Country Influences and Environmental Performance Standardization in Multinational Enterprises." *Journal of Business Research* 67 (11):2385-92.

Allen, Matthew M. C., and Maria L. Aldred. 2011. "Varieties of Capitalism, Governance, and High-Tech Export Performance: A Fuzzy-Set Analysis of the new EU Member States." *Employee Relations* 33 (4):334-55.

Allen, Matthew, and Maria L. Aldred. 2013. "Business Regulation, Inward Foreign Direct Investment, and Economic Growth in the New European Union Member States." *Critical Perspectives on International Business* 9 (3):301-21.

Allen, Matthew M. C., and Maria L. Allen. 2015. "Companies' Access to Finance, Cooperative Industrial Relations, and Economic Growth: A Comparative Analysis of the States of South Eastern Europe." *Research in International Business and Finance* 33 (1):167-77.

Bakker, René M., Bart Cambré, Leonique Korlaar, and Joerg Raab. 2011. "Managing the Project Learning Paradox: A Set-Theoretic Approach Toward Project Knowledge Transfer." *International Journal of Project Management* 29 (5):494-503.

Bell, R. Greg, Igor Filatotchev, and Ruth V. Aguilera. 2014. "Corporate Governance and Investors' Perceptions of Foreign IPO Value: An Institutional Perspective." *Academy of Management Journal* 57 (1):301-20.

Boudet, Hillary S., Dilanka C. Jayasundera, and Jennifer Davis. 2011. "Drivers of Conflict in Developing Country Infrastructure Projects: Experience from the Water and Pipeline Sectors." *Journal of Construction Engineering and Management* 137 (7):498-511.

Cárdenas, Julián. 2012. "Varieties of Corporate Networks: Network Analysis and fsQCA." *International Journal of Comparative Sociology* 53 (4):298-322.

Chandra Balodi, Krishna, and Jaideep Prabhu. 2014. "Causal Recipes for High Performance." *International Journal of Entrepreneurial Behaviour & Research* 20 (6):542-61.

Chang, Chia-Wen, Ting-Hsiang Tseng, and Arch G. Woodside. 2013. "Configural Algorithms of Patient Satisfaction, Participation in Diagnostics, and Treatment Decisions' Influences on Hospital Loyalty." *Journal of Services Marketing* 27 (2):91-103.

Chang, Man-Ling, and Cheng-Feng Cheng. 2014. "How Balance Theory Explains High-Tech Professionals' Solutions of Enhancing Job Satisfaction." *Journal of Business Research* 67 (9):2008-18.

Cheng, Cheng-Feng, Man-Ling Chang, and Chu-Shiu Li. 2013. "Configural paths to successful product innovation." *Journal of Business Research* 66 (12):2561-73.

Crawford, Seth. 2012. "What is the Energy Policy-Planning Network and Who Dominates it? A Network and QCA Analysis of Leading Energy Firms and Organizations." *Energy Policy* 45 (0):430-9.

Crilly, Donal. 2011. "Predicting Stakeholder Orientation in the Multinational Enterprise: A Mid-Range Theory." *Journal of International Business Studies* 42 (5):694-717.

[41] Source: www.compasss.org as of December 2014.

Crilly, Donal, Maurizio Zollo, and Morten T. Hansen. 2012. "Faking It or Muddling Through? Understanding Decoupling in Response to Stakeholder Pressures." *Academy of Management Journal* 55 (6):1429-48.

Crowley, Martha. 2012. "Control and Dignity in Professional, Manual and Service-Sector Employment." *Organization Studies* 33 (10):1383-406.

Evans, Anthony John, and Paul Dragos Aligica. 2008. "The Spread of the Flat Tax in Eastern Europe." *Eastern European Economics* 46 (3):49-67.

Fiss, Peer C. 2011. "Building Better Causal Theories: A Fuzzy Set Approach to Typologies in Organizational Research." *Academy of Management Journal* 54 (2):393-420.

Forsythe, Perry John. 2012. "Profiling Customer perceived Service Quality Expectations in made-to-order Housing Construction in Australia." *Engineering, Construction and Architectural Management* 19 (6):587-609.

Freitas, J. S., Gonçalves, C. A., Cheng, L. C., & Muniz, R. M. (2011). Parsimonious determinants of pre-incubated academic spin-offs initial performance: A configurational perspective. *Journal of Technology Management & Innovation*, 6(2), 50–65.

Freitas, Jonathan Simões, Carlos Alberto Gonçalves, Lin Chih Cheng, and Reynaldo Maia Muniz. 2013. "Structuration Aspects in Academic Spin-Off Emergence: A Roadmap-Based Analysis." *Technological Forecasting and Social Change* 80 (6):1162-78.

Ganter, Alois, and Achim Hecker. 2014. "Configurational Paths to Organizational Innovation: Qualitative Comparative Analyses of Antecedents and Contingencies." *Journal of Business Research* 67 (6):1285-92.

Garcia-Castro, Roberto, and Ruth V. Aguilera. 2014. "Family Involvement in Business and Financial Performance: A Set-Theoretic Cross-National Inquiry." *Journal of Family Business Strategy* 5 (1):85-96.

García-Castro, Roberto, Ruth V. Aguilera, and Miguel A. Ariño. 2013. "Bundles of Firm Corporate Governance Practices: A Fuzzy Set Analysis." *Corporate Governance: An International Review* 21 (4):390-407.

García-Castro, Roberto, and José Casasola. 2011. "A Set-Theoretic Analysis of the Components of Family Involvement in Publicly Listed and Major Unlisted Firms." *Journal of Family Business Strategy* 2 (1):15-25.

Grandori, Anna, and Santi Furnari. 2008. "A Chemistry of Organization: Combinatory Analysis and Design." *Organization Studies* 29 (3):459-85.

Greckhamer, Thomas. 2011. "Cross-Cultural Differences in Compensation Level and Inequality across Occupations: A Set-Theoretic Analysis." *Organization Studies* 32 (1):85-115.

Harriss-White, Barbara, Wendy Olsen, Penny Vera-Sanso, and V. Suresh. 2013. "Multiple Shocks and Slum Household Economies in South India." *Economy and Society* 42 (3):398-429.

Hotho, Jasper J. 2014. "From Typology to Taxonomy: A Configurational Analysis of National Business Systems and their Explanatory Power." *Organization Studies* 35 (5):671-702.

Järvinen, Joonas, Juha-Antti Lamberg, Johan-Peter Murmann, and Jari Ojala. 2009. "Alternative Paths to Competitive Advantage: A Fuzzy-Set Analysis of the Origins of Large Firms." *Industry and Innovation* 16 (6):545-74.

Järvinen, Joonas, Juha-Antti Lamberg, and Lauri Pietinalho. 2012. "The Fall and the Fragmentation of National Clusters: Cluster Evolution in the Paper and Pulp Industry." *Journal of Forest Economics* 18 (3):218-41.

Judge, William Q., Stav Fainshmidt, and J. Lee Brown III. 2014. "Which Model of Capitalism best delivers both Wealth and Equality?" *Journal of International Business Studies* 45 (4):363-86.

Kim, Minyoung. 2013. "Many Roads lead to Rome: Implications of Geographic Scope as a Source of Isolating Mechanisms." *Journal of International Business Studies* 44 (9):898-921.

Korczynski, Marek, and Claire Evans. 2013. "Customer abuse to service workers: An analysis of its social creation within the service economy." *Work, Employment & Society* 27 (5):768-84.

Krasnozhon, Leonid A. 2013. "Institutional Stickiness of Democracy in post-communist States: Can prevailing Culture explain it?" *Review of Austrian Economics* 26 (2):221-37.

Leischnig, Alexander, Anja Geigenmueller, and Stefanie Lohmann. 2014. "On the Role of Alliance Management Capability, Organizational Compatibility, and Interaction Quality in Interorganizational Technology Transfer." *Journal of Business Research* 67 (6):1049-57.

Maggetti, Martino. 2007. "De Facto Independence after Delegation: A Fuzzy-Set Analysis." *Regulation & Governance* 1 (4):271-94.

Maggetti, Martino. 2014. "Promoting Corporate Responsibility in Private Banking: Necessary and Sufficient Conditions for Joining the Wolfsberg Initiative Against Money Laundering." *Business & Society* 53 (6):787-819.

Magnier-Watanabe, Rémy, and Dai Senoo. 2008. "Organizational Characteristics as Prescriptive Factors of Knowledge Management Initiatives." *Journal of Knowledge Management* 12 (1):21-36.

Martin, Drew, and Arch G. Woodside. 2011. "Tourists' Dual-Processing Accounts of Reasoning, Judgment, and Actions." *International Journal of Culture, Tourism and Hospitality Research* 5 (2):195-212.

Marx, Axel. 2008. "Limits to non-State Market Regulation: A Qualitative Comparative Analysis of the International Sport Footwear Industry and the Fair Labor Association." *Regulation & Governance* 2 (2):253-73.

Meuer, Johannes. 2014. "Archetypes of Inter-Firm Relations in the Implementation of Management Innovation: A Set-theoretic Study in China's Biopharmaceutical Industry." *Organization Studies* 35 (1):121-45.

Ordanini, Andrea, and Paul P. Maglio. 2009. "Market Orientation, Internal Process, and External Network: A Qualitative Comparative Analysis of Key Decisional Alternatives in the New Service Development." *Decision Sciences* 40 (3):601-25.

Ordanini, Andrea, A. Parasuraman, and Gaia Rubera. 2014. "When the Recipe Is More Important Than the Ingredients: A Qualitative Comparative Analysis (QCA) of Service Innovation Configurations." *Journal of Service Research* 17 (2):134-49.

Pajunen, Kalle. 2008. "Institutions and Inflows of Foreign Direct Investment: A Fuzzy-Set Analysis." *Journal of International Business Studies* 39 (4):652-69.

Park, Sung Ho. 2013. "Capital Openness, Monetary Integration, and Wage-Setting Coordination in Developed European Countries." *Economic and Industrial Democracy* 34 (4):637-66.

Sawyer, Steve, Robert Schrier, Jane Fedorowicz, Martin Dias, Christine Williams, and Mike Tyworth. 2013. "U.S. Public Safety Networks: Architectural Patterns and Performance." *Information Polity* 18 (2):139-56.

Schneider, Martin R., Conrad Schulze-Bentrop, and Mihai Paunescu. 2010. "Mapping the Institutional Capital of High-Tech Firms: A Fuzzy-Set Analysis of Capitalist Variety and Export Performance." *Journal of International Business Studies* 41 (2):246-66.

Seeleib-Kaiser, Martin, and Timo Fleckenstein. 2009. "The Political Economy of Occupational Family Policies: Comparing Workplaces in Britain and Germany." *British Journal of Industrial Relations* 47 (4):741-64.

Skaaning, Svend-Erik. 2007. "Explaining Post-Communist Respect for Civil Liberty: A Multi-Methods Test." *Journal of Business Research* 60 (5):493-500.

Skoko, Hazbo, Branka Krivokapic-Skoko, Marinko Skare, and Arnela Ceric. 2006. "ICT Adoption Policy of Australian and Croatian SMEs." *Managing Global Transitions* 4 (1):25-40.

Stanko, Michael A., and Xavier Olleros. 2013. "Industry Growth and the Knowledge Spillover Regime: Does Outsourcing harm Innovativeness but help Profit?" *Journal of Business Research* 66 (10):2007-16.

Valliere, Dave, Ni Na, and Sean Wise. 2008. "Prior Relationships and M&A Exit Valuations: A Set-Theoretic Approach." *Journal of Private Equity* 11 (2):60-72.

Weissenberger-Eibl, Marion A., and Benjamin Teufel. 2013. "The Innovation Process as a Political Process - How Organisational Politics can Influence the Technological Newness of Product Innovations." *International Journal of Innovation Management* 17 (4):1-34.

Winand, Mathieu, Rihoux Benoît, David Qualizza, and Thierry Zintz. 2011. "Combinations of Key Determinants of Performance in Sport Governing Bodies." *Sport, Business and Management: An International Journal* 1 (3):234-51.

Winand, Mathieu, Benoît Rihoux, Leigh Robinson, and Thierry Zintz. 2013. "Pathways to High Performance: A Qualitative Comparative Analysis of Sport Governing Bodies." *Nonprofit and Voluntary Sector Quarterly* 42 (4):739-62.

Woodside, Arch G., Shih-Yun Hsu, and Roger Marshall. 2011. "General Theory of Cultures' Consequences on International Tourism Behavior." *Journal of Business Research* 64 (8):785-99.

Woodside, Arch, and Mann Zhang. 2012. "Identifying X-Consumers Using Causal Recipes: "Whales" and "Jumbo Shrimps" Casino Gamblers." *Journal of Gambling Studies* 28 (1):13-26.

Woodside, Arch G., and Mann Zhang. 2013. "Cultural Diversity and Marketing Transactions: Are Market Integration, Large Community Size, and World Religions Necessary for Fairness in Ephemeral Exchanges?" *Psychology & Marketing* 30 (3):263-76.

Wu, Pei-Ling, Shih-Shuo Yeh, Tzung-Cheng Huan, and Arch G. Woodside. 2014. "Applying Complexity Theory to Deepen Service Dominant Logic: Configural Analysis of Customer Experience-and-Outcome Assessments of Professional Services for Personal Transformations." *Journal of Business Research* 67 (8):1647-70.

Young, Raymond, and Simon Poon. 2013. "Top Management Support - Almost always Necessary and sometimes Sufficient for Success: Findings from a Fuzzy Set Analysis." *International Journal of Project Management* 31 (7):943-57.

4.

Fuzzy Logic or Fuzzy Application?
A Response to Stockemer's 'Fuzzy Set or Fuzzy Logic?'

(Antje Buche, Jonas Buche, Markus B. Siewert)

Abstract

In a recent article in this journal, Daniel Stockemer (2013) characterizes fuzzy-set Qualitative Comparative Analysis (fsQCA) in comparison to Ordinary Least Squares regression as a 'poor methodological choice' because of its 'suboptimal nature' for the study of descriptive female representation in national assemblies across the globe. This article seeks to demonstrate that his judgments are based on two misconceptions: first, a misunderstanding of set-theoretical thinking in general, and specifically Qualitative Comparative Analysis (QCA); and second, a misinformed application throughout various steps of the fsQCA, e.g. the calibration process, the analysis of necessary and sufficient conditions, and the interpretation of the results. In pointing out the weaknesses of Stockemer's application of OLS, we argue – in contrast to Stockemer – that fsQCA can be a valuable tool for the comparative study of social phenomena which offers a fundamentally different analytical perspective from standard quantitative techniques.

Keywords

Women's Representation; Qualitative Comparative Analysis (QCA); Regression Analysis; Mixed Methods; Evaluation; Standards of "Good" Practice.

4.1 Introduction

The descriptive representation of women in legislatures is a central topic of today's comparative political science research (for an overview see e.g. Wängnerud 2009; Krook & Childs 2010; Paxton & Hughes 2014). While women constitute more than half of the world's population, as of 1 April 2015 the global average of female parliamentarians is only around 22.1 per cent (IPU 2015). A closer look reveals two further aspects. First, variation across the globe is immense. While some countries have close to equal gender representation in their national legislatures, others have only a few, or even no, female representatives. Second, parliaments with a high share of female representation are not restricted to countries which are classically associated with gender equality, such as the Nordic countries. Indeed, contrary to popular perceptions, countries as diverse as Bolivia, Cuba, Ecuador, Rwanda, Seychelles and South Africa have joined Finland and Sweden at the top of the global gender equality rankings. At the same time, there are several prominent European Union (EU) and Organisation for Economic Co-operation and Development (OECD) countries, such as Chile, Japan, Ireland, Malta and the United States of America (US), that exhibit a low share of women in their national assemblies (IPU 2015).

Existing scholarship identifies three main sets of factors that affect the share of women in legislatures. First, institutional approaches focus on the design of the electoral system. According to some studies, proportional systems and multi-member districts further the equal representation of women while majoritarian electoral systems and single-member districts seem to have a negative effect (e.g. McAllister & Studlar 2002; Rosen 2012; Roberts et al. 2013). Additionally, statutory or party-based gender quotas have been identified as effective drivers of higher women's representation (e.g. Dahlerup 2006; Tripp & Kang 2008; Krook 2009). A second strand of research emphasizes socio-economic conditions. From these perspectives, higher proportions of women in parliaments are strongly interlinked with equal access to educational systems and emancipation at the workplace (e.g. Matland 1998; Rosenbluth et al. 2006; Viterna et al. 2008). Third, cultural approaches refer to the importance of femininity models, gender roles and religious norms within societies as explanations for descriptive political representation (e.g. Inglehart & Norris 2001; Paxton & Kunovich 2003). While these examples of explanatory factors that affect the share of women in national parliaments are well defined, their impact is still largely contested. Therefore, newer studies have concentrated on the interplay of these factors and varying patterns across different contexts

such as status of development or geographical location (e.g. Krook 2010; Rosen 2012; Ruedin 2012; Stockemer 2013, 2015).

This newer line of research has also led to methodical innovations. For instance, Mona Lena Krook's study (2010) of women's representation in sub-Saharan Africa and Western countries introduced Qualitative Comparative Analysis (QCA) to the field of gender politics in general, and women's representation specifically (e.g. Ragin 2008; Rihoux & Ragin 2009; Schneider & Wagemann 2012). At first sight, there are several reasons why QCA seems to fit nicely with the above-mentioned trends in research on descriptive women's representation: it strives for a more case-oriented, and therefore, context-specific assessment; it looks for combinations of conditions and not so much on the average effects of single variables; it focuses on equifinal configurations that explain various sets of cases; and it introduces the notion of asymmetry which opens the possibility for different explanations for high and low shares of women's representation.

In a recent article, Daniel Stockemer (2013) offered a stark contrast to these assertions by concluding 'that QCA as a method is a poor methodological choice for explaining the factors, or combinations of variables, that lead to high or low women's representation in national assemblies' (Stockemer 2013: 95). For Stockemer, QCA is unable 'to provide any additional insights' (Stockemer 2013: 96). He arrived at this conclusion after comparing a fuzzy-set QCA (fsQCA) and a standard OLS regression model to the same dataset of 54 Latin American and Asian-Pacific countries. In his study, Stockemer applied the two methods separately and sequentially in order to reveal 'whether they yield different results and, if so, which method is more appropriate for the study of women's parliamentary representation?' (Stockemer 2013: 87). In doing so, a range of seven institutional and socio-economic explanatory factors – democratic or communist status, electoral systems, gender quotas and women's suffrage, female economic activity rate and the country's level of education – are included in the analysis of women's representation in national legislatures.

Somewhat puzzled by Stockemer's harsh and absolute conclusions, we argue that his outright dismissal of fsQCA as a valuable tool for the study of women's representation is largely based on two misconceptions: a misunderstanding of the Boolean and set-theoretic roots of QCA and a misinformed application of fsQCA. In addressing both, we build our argument on two pillars. First, we engage in a methodological discussion that emphasizes QCA's origins in a qualitative-empirical, case-oriented paradigm, while standard regression analysis stems from a quantitative, variable-oriented research tradition (Ragin 1987, 2000, 2008; Goertz & Mahoney

2012; Schneider & Wagemann 2012; Thiem et al. 2016). Second, we highlight three major shortcomings in Stockemer's argument: 1) by debating specific elements of his research design (like case selection and operationalization); 2) by replicating his fsQCA and OLS regressions; and 3) by assessing his interpretation of the research findings.

4.2 Misunderstanding QCA

The central aim of Stockemer's study is to assess the value of fsQCA and standard OLS regression with regard to descriptive women's representation. Acknowledging that these two approaches are based on completely different logics, he asks 'whether they yield different results and, if so, which method is more appropriate for the study of women's parliamentary representation?' (Stockemer 2013: 87). Stockemer's comparative yardstick is clear: '[i]f fsQCA helps identify parsimonious combinations of factors that explain high women's representation, then this method is a powerful tool. In contrast, if we find few unique combinations of indicators and little empirical coverage of these factor combinations, then this method has some serious limitations when applied to the study of women's representation' (Stockemer 2013: 87f.). From our perspective, both the overall goal and the benchmark of Stockemer's study are based on two major misunderstandings and therefore are to no avail. On the one hand, the rationales and technicalities of QCA and standard regressions are so fundamentally different that it is neither valuable nor feasible to compare results produced by the two approaches against each other (Goertz & Mahoney 2013). On the other hand, parsimony does not have to be a central aim of QCA. Quite the contrary, researchers applying QCA are usually interested in complex and multiple configurations of explanatory factors.

The logic behind QCA is deeply linked to Boolean algebra and set-theoretical thinking (Ragin 1987, 2000, 2008; Goertz & Mahoney 2012; Schneider & Wagemann 2012). From this basic logic follow some major ramifications that distinguish QCA from standard quantitative techniques like regression.

First of all, the central units of QCA and regression analysis are different. While the later makes use of variables, the former utilizes sets as building blocks for the analysis. This difference is far from superficial. A variable stands for a latent concept that is measured directly by one or several indicators. Here the underlying assumption typically is that the value of the indicator reflects the latent concept – the variable – in a linear way. The linkage between sets and concepts is more about assessing the meaning of a concept. This regularly goes hand-in-hand with an asymmetric transformation of the data in order to capture the semantic meaning of the concept – a process called set calibration. Put differently, in constructing sets, the re-

searcher interprets the data, truncates unnecessary variation across an indicator, and (through the process of assigning set memberships) attaches conceptual meaning to the set value. In the case of women's representation, a variable captures the relative share of female parliamentarians while sets are calibrated in a way to assess high, low, or medium levels of female representation (Ragin 2008: 71-84; Goertz & Mahoney 2012: 139-149; Schneider & Wagemann 2012: 24-31; Thiem et al. 2016: 749).

Secondly, once sets are calibrated, QCA enables the researcher to examine subset-superset-relations, or in other words, to check for necessity and sufficiency. Looking for necessity involves the analysis of shared antecedents of a given outcome. In order to be necessary a condition has to be present if the outcome is present; thus, the outcome cannot occur without the antecedent. Sufficiency, on the other side, aims at the analysis of shared outcomes. So the focus is on the question as to whether the outcome is present if a certain (combination) of condition(s) is present? From this short depiction it should become clear that hypotheses based on necessary and sufficient conditions are different from those based on standard regressions. While the latter aim at correlational relationships which are symmetric (e.g. the higher or lower the share of female workforce, the higher or lower the proportion of women in legislatures), set relations are non-linear by nature (Goertz & Mahoney 2012: 16-38; Schneider & Wagemann 2012: 83ff; Thiem et al. 2016: 752ff).

Thirdly, researchers trained in set-theory tend to think in terms of configurations, equifinality, and asymmetry. Again, this offers a completely different perspective of social science phenomena from standard regression analysis. In contrast to quantitative techniques, QCA does not focus on the strength and significance of the net-effects of single variables. Rather, cases are perceived as configurations of conditions. For instance, a high share of women's representation is explained by the simultaneous presence or absence of certain conditions. Thus, one is interested in the concurrence of several explanantes that are associated with a given explanandum. In contrast to standard regressions, which usually focus on the one model that fits best, QCA aims to detect equifinal patterns, uncovering multiple configurations of conditions that explain an outcome (like a high share of female parliamentarians). Asymmetry refers to the fact that configurations of conditions that contribute to a high proportion of female representatives can be different from the configurations that explain a low share of women in legislatures. Again, this is dissimilar to standard regression approaches which assume a symmetrical relationship in the sense that if: 1) a higher share of women with a high educational degree leads to a higher share of women in legislatures, then 2) a lower share of women with high

education should lead to a lower proportion of female parliamentarians (Schneider & Wagemann 2012: 5ff, 76-90).

This abbreviated juxtaposition shows that set-theoretic approaches (like QCA) and standard regressions (like OLS) start out from dissimilar assumptions, are embedded in opposite logics, and apply different procedures. If one accepts these fundamental differences, how is it then possible to compare the results of these two approaches, and how should we assess the value of QCA and regressions for the study of women's representation? The short answer to these questions is that we cannot compare them and we cannot assess them against each other. If two methods use different building blocks in their analysis, aim at answering different hypotheses, and rivet on different aspects of the social world, it is not at all surprising that their results also differ. Rather, it is the task of the researcher to make use of the different insights gained through different analytical perspectives and methodological choices in a constructive way to shed light on social science problems.

4.3 Misinformed Application

4.3.1 Debating Case Selection and Operationalization

We start our discussion of Stockemer's (2013) application with his case selection. In his study, Stockemer analyses fifty-four Latin American and Asian-Pacific countries. Although his own literature review points to the fact that 'indicators driving women's representation are often embedded in a specific regional, social, and/or economic context' (Stockemer 2013: 87), for three reasons he rates the selection of two different world regions as a 'a propitious laboratory to help determine the usefulness of fsQCA and/or regression analysis' (Stockemer 2013: 90). This assertion is based, first, on the region's roughly similar level of development, measured by the region-wide average gross domestic product (GDP). Second, 'other key explanatory variables in the data set' (Stockemer 2013: 90) are similar. Third, also the dependent variable of the study, the number of female parliamentarians is comparable on average. However, cultural differences of the regions are left out completely. Moreover, the difficulty of comparing micro-states, such as Bhutan or Papua-New Guinea, with countries like Australia, Brazil or China are not discussed at all. While this might not necessarily be problematic, this decision influences the interpretation of the results. One way to avoid problems that stem from case selection is to run separate analyses with more context-sensitive measures, with different calibration thresholds, and probably even partially different conditions (e.g. Krook 2010).

The decision to put more emphasis on region-specific differences comes at a price. On the one hand, if two QCAs and two regressions are conducted, the number of cases decreases to thirty-two Latin American or twenty-two Asian countries, respectively. Thus, to apply an OLS regression for each sub-sample would have limited statistical power (Cohen 1988), while for a QCA these numbers of cases are not problematic per se (Schneider & Wagemann 2012: 317). On the other hand, one could include a region-specific dummy – or crisp set – to control for the basic cultural differences. While keeping the number of cases constant, this would lead to a higher number of explanatory factors. Hence, both approaches might not offer a convenient solution for the problem. This is especially true as the original analysis with fifty-four cases and seven explanatory factors is already tangent to the 'many variables, small N problem' (Lijphart 1971: 686).

According to this problem, a certain number of cases in an analysis might not be able to display the diversity that is conceptually included by the number of explanatory factors. With seven conditions included to the analysis, 128 combinations of these conditions are logically possible and by far outnumber the number of cases. Every combination of conditions that lacks empirical evidence by a case is a so-called logical remainder (Ragin 2008; Schneider & Wagemann 2012: 151 ff.). For OLS regressions, the problem of limited empirical diversity also holds because sufficient variance on the explanatory factors and the dependent variable cannot be guaranteed. Additionally, the problem of influential observations (outliers) appears in this context. This means that some observations have a 'demonstrably larger impact on the calculated values of the various estimates (coefficients, standard errors, t-values, etc.) than is the case for most of the other observations' (Belsley et al. 1980: 11). Given the relatively small number of fifty-four cases, this problem does occur and weakens the statistical power of the model.

Furthermore, a crucial factor in both QCA and OLS regression analysis is not just the number of cases and explanatory factors, but also the operationalization of the latter. This is despite differences between the operationalization of variables and sets (see tables 4.A1 and 4.A2 for the original data used by Stockemer). To exemplify a problem associated with operationalization and calibration, this article focuses on the two ordinal variables 'quotas' and 'electoral systems', with three categories each.

Using these variables in an OLS regression contradicts the logic of linear regression techniques. This is because dependent and independent variables have to be continuous or, as in the case of the independent variable, be dummy variables. To this end, using other types of

variables for linear regression is possible only if some important conditions are fulfilled (Urban & Mayerl 2011: 275). In specific terms:

(1) The variables must at least have five categories;

(2) The categories of variables must be in a ranked order;

(3) The distances between the categories have to be equal (both in semantic meaning and numeric value); and

(4) The categories can be interpreted as intervals for the continuous latent variable.

Stockemer neglects these conditions in his operationalization of the two variables. First, 'quotas' (quota clauses in parties), is coded into zero, 0.5 and one, where zero means there are no quota clauses in all parties, and 0.5 stands for at least one party that uses quotas. If all parties have quotas (constitutionally-implemented quota clauses), the variable becomes one. Second, the 'electoral system' is coded one for proportional systems, zero for majoritarian systems, and 0.5 for all others. These two variables violate the first and third conditions outlined above: they are limited to just three categories each and the distance between the categories is neither semantically nor numerically the same. Here, Stockemer should have used categorical dummies or, for a more parsimonious way, recoded the data to get at least two dummy variables. However, as the second option would lead to a loss of information, we use categorical dummies for the replication.

A similar problem with the same variables occurs under the calibration procedure in QCA. In fsQCA, set memberships vary between full membership (set membership score equals 1) and non-membership (membership score equals 0). Stockemer states that as the variables 'electoral system' and 'quotas' already have 'values between 0 and 1 [… they] consequently do not need to be transformed any more (Stockemer 2013: 93). This means that every country with a mixed party system is calibrated as a 0.5 member in the set of 'electoral system type'. The same is true for every country with at least one party that uses a quota system (but not all). But in QCA the 0.5 anchor defines the transition point between membership and non-membership, i.e. the 'point of maximum ambiguity' (Ragin 2008: 30). Conceptually, assigning this value to cases means that we cannot say anything about their set membership: they are neither members nor non-members of the respective set. Technically, these cases cannot be assigned to a single truth table row as they belong in at least two. These cases would display multiple combinations of conditions ideal typically.

This produces two additional problems. First, the computer software will 'hide' these cases from the truth table, and this may cause false logical remainders. Second, if these cases are covered by an explanatory conjunction that encompasses the condition with the 0.5 calibration, the software cannot display this since the cases' membership score in the conjunction is also 0.5. Thus, the cases can only be explained if the problematically-calibrated condition is dropped via the minimization process. To be clear, none of the cases with a 0.5 membership in at least one of the conditions is excluded from the analysis by default. But due to this easily-avoidable calibration issue, Stockemer risks losing the explanatory power of his model for those seventeen cases which he assigned a 0.5 set membership to (in either 'electoral system type', 'quota' or both). As this covers one third of the cases, the analysis runs the risk of explaining basically nothing at all. The reason for this is not down to the choice of QCA as a method, but rather to its application.

Taking these points seriously, we change the calibration for the two most problematic conditions in the replication (, i.e. 'electoral system type' and 'quota' without the allocation of the 0.5 anchor). In order to avoid this membership score we have to decide whether the conceptual meaning of a '0.5' corresponds more to a (partial) membership or a (partial) non-membership in the particular set. For the set 'quota' we proceed in the following way. As a case was coded 0.5 'in which at least one party represented in the national assembly […] has a party quota' (Stockemer 2013: 91), the underlying concept of the set is (partially) represented by those cases. Thus, we recode all 0.5 set memberships in the set 'quota' to 0.66. Alternatively, the coding of 0.5 for the variable 'electoral system type' includes any 'semi-proportional or mixed system' (Stockemer 2013: 91). Here, without case knowledge we cannot allocate the same set membership score for all of the countries under scrutiny. Thus, we allocate a 0.66 membership in the set 'proportional electoral system type' to all countries with a 'Mixed Member Proportional System (MMP)' (IDEA 2015) as in this type of mixed system the 'List PR system compensates for the disproportionality in the results from the plurality/majority system' (IDEA 2015). In our understanding, this represents the concept of a proportional electoral system, i.e. such a case is more in the set than out of the set. By contrast, a 'Parallel System is a mixed system […] where no account is taken of the seats allocated under the first system in calculating the results in the second system' (IDEA 2015). Consequently, cases with a parallel system are treated as more out of the set (than in the set); i.e. they are allocated a 0.33 membership (for our calibration see table 4.A3).

4.3.2 Replicating the Analysis

In replicating Stockemer's analysis with categorical dummies, the results look very similar to the original analysis (for a comparison see Stockemer's results in table 4.A4). There are no significant effects for the electoral system and, ceteris paribus, a significant, approximately ten percentage points higher women's representation for countries with implemented quotas (D_Quotas3). But there is one elementary difference between the results. Stockemer treats the variables as continuous. Due to linearity assumption, Stockemer (2013) has to interpret the coefficient as follows: every change in the independent variable 'quotas' (0; 0.5; 1) causes a change in the dependent variable 'women's representation'. In his model, ceteris paribus, women's representation is approximately five percentage points higher in countries with quotas for at least one party but not for all (quotas has the value 0.5), than in countries without quotas (D_Quotas2). But as one can see in table 4.1, there is no significant difference between these countries and countries without quotas . Thus, the effect only occurs between 'no quota' and 'quotas in all parties', proving the weakness of the OLS's ordinal variable.

Table 4.1 Results of OLS regression with categorical dummy variables

	B	SE	Sig
D_Quotas2	3.98	3.01	0.202
D_Quotas3	10.35	2.38	0.000
D_PR2	-2.15	2.90	0.46
D_PR3	-0.38	2.43	0.876
Education	14.79	9.57	0.129
Year of Women Suffrage	0.06	0.09	0.535
Female Activity Rate	-0.04	0.08	0.605
Degree of Democracy	0.09	0.05	0.060
D_Communist State	15.51	3.41	0.000
Constant	-1.91	7.00	0.787

Notes: $R^2 = 0.51$, adj. $R^2 = 0.41$, N = 54.

Furthermore, in his analysis Stockemer argues about the impact of the degree of democracy. This impact is not statistically significant at the 0.05 level, but it is at the 0.1 level. The problem is caused because of the absence of a diagnostic test to confirm the results of the analysis. As mentioned above, regression analyses with a small number of observations are very sensitive and one has to account for influential observations (Jann 2009). A look at the data reveals that there are some countries with a great deal of influence. This holds especially true for the variable 'degrees of democracy'. Here, the effect is mainly driven by New Zealand and Australia, both of which have a long democratic tradition (at the time of writing 129 and 106 years of democracy respectively) and a rather high percentage of women's representation

(33.1 per cent and 26.7 per cent respectively). Figure 4.1 shows that without these two countries there is no significant effect for 'degrees of democracy' at all. Even if one does not eliminate the two (or at least New Zealand) countries for the analysis, this figure shows the sensitivity of the regression results. At this point, we might question whether OLS regression is an adequate method for use on these data.

Figure 4.1 Added variable plot

Regarding the replication of the QCA, one should start with the analysis of necessary conditions completely missing so far. As can be seen in table 4.2, no single condition is necessary due to the non-perfect subset relations indicated by the low consistency values.

Table 4.2 Analysis of necessity of single conditions

Condition	Consistency necessity	Coverage necessity	Condition	Consistency necessity	Coverage necessity
edu	0.86	0.49	~[42] edu	0.75	0.63
dem	0.27	0.54	~ dem	0.21	0.62
eco	0.28	0.48	~ eco	0.34	0.86
ws	0.56	0.54	~ ws	0.51	0.73
pr	0.52	0.45	~ pr	0.49	0.63
com	0.21	0.78	~ com	0.04	0.22
q	0.58	0.57	~ q	0.35	0.51

Notes: edu = education, dem = consolidation/longevity of democracy, eco = female economic activity rate, ws = women's suffrage, pr = proportional electoral system, com = communist regime type, q = quota provisions

The analysis of sufficient combinations of conditions with the newly calibrated data reveals a solution term with three equifinal combinations consisting of six conditions each (see table 4.3; the truth table is provided in table 4.A5). Although this is somewhat more parsimonious

[42] The tilde (~) indicates the absence of the respective condition.

than Stockemer's (2013: 101) result, which consists of four complex solution paths, the problems of low coverage values and rather complex conjunctions remain.

Table 4.3 Analysis of sufficient (combinations of) conditions, conservative solution

Parameters of Fit	edu*~dem*~eco* ~pr*com*~q	edu*dem*~eco* ~pr*~com*q	~edu*~dem*~eco ~ws*~pr*com
Consistency	1.00	0.98	1.00
Raw Coverage	0.10	0.08	0.10
Unique Coverage	0.07	0.08	0.06
Cases	Cuba, Vietnam	Costa Rica, Australia	Nepal, Laos
Solution consistency	*0.99*		
Solution coverage	*0.24*		

Notes: edu = education, dem = consolidation/longevity of democracy, eco = female economic activity rate, pr = proportional electoral system, q = quota provisions, com = communist regime type, ws = women's suffrage

In order to deal with these problems, the logical remainders can be assessed. The high number of conditions combined with a rather medium number of cases result in 128 logically possible combinations of the seven conditions, i.e. truth table rows (27). Only twenty-nine of these are covered by empirical cases. Thus, another ninety-nine combinations cannot be included in the analysis in order to reduce the model's complexity. However, in addition to the so-called conservative strategy of excluding all logical remainders two more solution strategies can be applied. First, the most parsimonious solution term includes all those logical remainders into the minimization process that reduce the model complexity (Schneider & Wagemann 2012: 169 ff.). With ninety-nine possible logical remainders to include, the most parsimonious solution is way less complex, relative to the conservative one (see table 4.4).

Table 4.4 Analysis of sufficient (combinations of) conditions, most parsimonious solution

Parameters of Fit	~pr*com	dem*q
Consistency	0.92	0.94
Raw Coverage	0.18	0.12
Unique Coverage	0.18	0.12
Cases	Cuba, Vietnam, Nepal, Laos	Costa Rica, Australia
Solution consistency	*0.93*	
Solution coverage	*0.30*	

Notes: dem = consolidation/longevity of democracy, pr = proportional electoral system, q = quota provisions, com = communist regime type

While the results are clearly more parsimonious, they might not be as trustworthy as before due to the 'computerized strategy' of selecting logical remainders for the minimization process. However, there is a third solution term in QCA: the so-called intermediate solution term (Schneider & Wagemann 2012: 169 f.). This solution term offers the opportunity to select logical remainders on a theoretical basis. In contrast to the conservative strategy, the interme-

diate solution term allows for the inclusion of logical remainders. But, contrary to the most parsimonious solution only those remainders that are in line with the directional expectations implied by the researcher on a theoretical basis are included. Applying this accepted strategy to the data at hand, the complexity of the initial solution by Stockemer is resolved into two combinations of conditions (see table 4.5).

Table 4.5 Analysis of sufficient (combinations of) conditions, intermediate solution

Parameters of Fit	~pr*com	edu*dem*q
Consistency	0.92	0.94
Raw Coverage	0.18	0.12
Unique Coverage	0.18	0.12
Cases	Cuba, Vietnam, Nepal, Laos	Costa Rica, Australia
Solution consistency	*0.93*	
Solution coverage	*0.30*	

Notes: edu = education, dem = consolidation/longevity of democracy, pr = proportional electoral system, q = quota provisions, com = communist regime type

Our strategy of the QCA replication can be summarized as follows. We changed the calibration for two out of seven conditions in order to avoid allocating the 0.5 anchor. We restrict ourselves to these technically most problematic ones in order to keep the replication close to the original analysis. However, regarding the content Stockemer's calibration decisions on qualitative anchors of the other five conditions are without any doubt questionable as well (see notes table 4.A2). The results for the conservative strategy are quite similar to the one done by Stockemer. Although we are to some extent able to reduce the complexity, the single solution paths still comprise rather complex configurations. However, applying the intermediate strategy produces quite parsimonious and interpretable results. The remaining problem of rather low coverage is discussed in the following section.

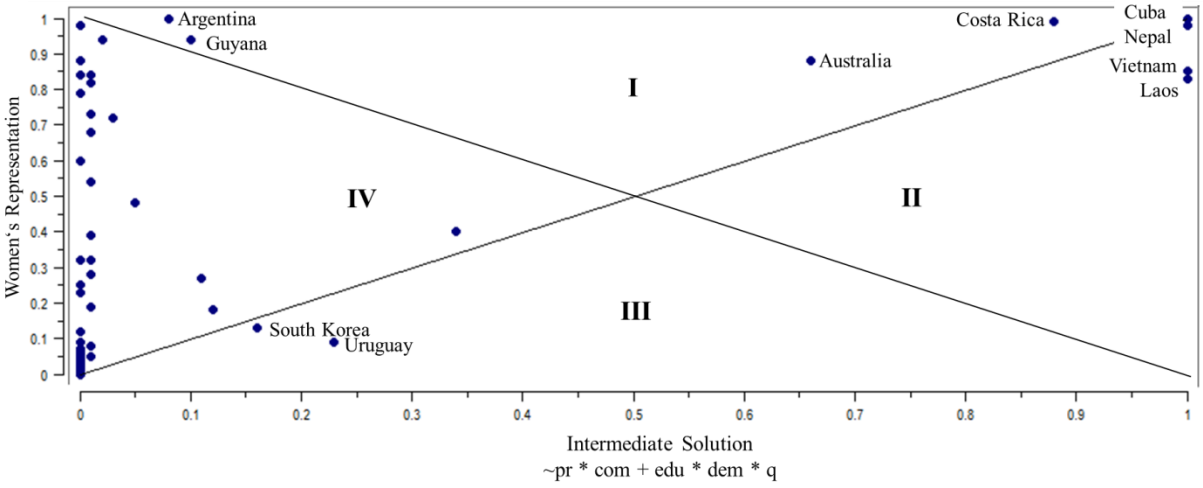
4.3.3 Assessing the Interpretation

Stockemer's interpretation of the effects resulting from OLS is correct and clearly formulated – except for the question of using quotas and electoral system as continuous variables. One further issue is the use of the R2 in order to interpret the quality of the analysis. Based on the R2, Stockemer concludes 'the current model explains 50 per cent of the variance in the dependent variable' (2013: 96). However, due to 'Occam's Razor' it is compulsory to interpret the adjusted R2 which accounts for the number of independent variables. Even though the interpretation of the adjusted R2 is not as intuitive as the R2, it works better by evaluating the accuracy of fit between data and model. In Stockemer's analysis the adjusted R2 is 0.42, which is still considerably high.

Shifting our focus towards the fsQCA, the interpretation of the conservative solution paths is a complex task. As stated above, this can be explained by the combination of few cases and many conditions, which subsequently result in a high number of logical remainders. The opportunity for minimization is thereby potentially reduced. Thus, either we reduce the number of explanatory factors in order to match it with the number of cases, or we make use of logical remainders. Applying the intermediate solution on a solid theoretical foundation is a widely-accepted strategy for dealing with the almost inevitable problem of limited empirical diversity in the social science (Schneider & Wagemann 2012: 152f, 2013). Moreover, it reduces the complexity of the explanatory model and makes the results more interpretable without the problem of untenable assumptions one has to face using the most parsimonious solution. According to these results, either the combination of a communist state with no proportional electoral system or the combination of quota provisions, a consolidated democracy and a high level of education is sufficient for a high share of female deputies.

However, as we have already outlined above, the coverage of the intermediate solution is still rather low. Although being highly consistent (with a consistency score of 0.95), roughly one third (coverage of 0.3) of the outcome set 'high women's representation' can be explained. This is nicely visualized by a XY-plot from which we can deduce three things (see figure 4.2).

Figure 4.2 XY-plot of sufficient (combinations of) conditions, intermediate solution



First, we can see that we do not have cases that strongly contradict the statement of sufficiency showing the combination of conditions but not the outcome. Nevertheless, the consistency is not perfect because five countries (Laos, Nepal, South Korea, Uruguay and Vietnam) have a slightly greater membership in the configuration of the intermediate solution than in the outcome set.

Second, in the end only six countries out of twenty (in the set 'high levels of women's representation') are covered by the overall solution: Australia, Costa Rica, Cuba, Laos, Nepal, and Vietnam. While Stockemer is right that this is not a satisfying result for a case-oriented method, which is interested in explaining as many cases as possible, he draws the wrong conclusions. The low coverage does not mean that QCA is a poor methodological tool for the analysis of women's representation in national assemblies. The clustering of cases in the upper left quadrant of the XY-plot – what we call the equifinality-corner – indicates that the most relevant factors (or indicators) have not been selected for the analysis and that others – equifinal ones – have to be found to sufficiently explain high women's representation in the Latin American and Asian-Pacific countries.

Third, and closely related to this, the data allocation is highly skewed (see Schneider & Wagemann 2012: 232-48). Just ten out of fifty-four cases are located outside the left triangle of the XY plot. Thus, all other cases do not contradict the sufficiency of the intermediate solution for the outcome, and for the non-outcome (alike). This indication of simultaneous subset relations raises the questions as to whether the skewedness stems from single variables (e.g. communism) and subsequently, whether the explanatory factors have been selected and calibrated in a proper way.

4.4 Conclusion

The starting point of our discussion was the devastating critique by Daniel Stockemer, who dismisses QCA due to its 'suboptimal nature' (Stockemer 2013: 88) and calls it a 'poor methodological choice' (Stockemer 2013: 94) for studying female representation in national assemblies that does not 'provide any additional insight' (Stockemer 2013: 96). However, from our perspective these judgments stem from a misunderstanding of central ideas and features of QCA, and a misinformed application. In our opinion, there is nothing 'inherently problematic' (Stockemer 2013: 97) with either fsQCA or OLS regressions. Both approaches are rooted in fundamentally different logics, apply completely different techniques, and have their distinct pitfalls.

If we take these points seriously for the study of descriptive women's representation in national legislatures across the globe we cannot subscribe to Stockemer's observations. Set-theoretical thinking differs from standard quantitative approaches in so many ways that it is not possible to contrast their findings against each other. Thus, questions like 'which method is more appropriate' (Stockemer 2013: 87) or which one has 'greater leverage' over the other (Krook 2010: 887) are not very expedient and fruitful. It is not about more or less leverage but

about different leverages. In this sense, the comparison is like the simile of the blind men touching and feeling different parts of an elephant and coming to completely different conclusions about what kind of animal they have in front of them. Goertz and Mahoney (2012, 2013) have called this the 'Rorschach principle': social scientists perceive social phenomena differently if they look at them through set-theoretic or quantitative lenses. Therefore, the real challenge is to reconcile the different perspectives and to strive for complementary conclusions.

Appendix

Table 4.A1 "The original data set" by Stockemer (2013: 100), i.e. raw data matrix

country	wr	edu	dem	eco	ws	pr	com	q
Argentina	40	0,95	14	68	61	1	0	1
Australia	26,7	0,99	106	56	106	0	0	0,5
Bahamas	12,2	0,88	35	91	47	0	0	0
Bangladesh	15,1	0,5	0	34	36	0	0	1
Barbados	10	0,96	40	83	68	0	0	0
Belize	6,7	0,77	21	52	54	0	0	0
Bhutan	8,5	0,49	0	45	55	0	0	0
Bolivia	16,9	0,87	0	74	70	0,5	0	0,5
Brazil	9	0,88	12	70	76	1	0	0
Cambodia	16,3	0,69	0	74	53	1	0	0
Chile	15	0,96	17	51	67	1	0	0,5
China	21,3	0,84	0	52	59	1	1	0
Colombia	8,4	0,87	0	75	54	1	0	0
Costa Rica	36,8	0,88	61	54	59	0	0	1
Cuba	43,2	0,95	0	59	74	0	1	0
Dominican R.	19,7	0,83	11	55	66	1	0	1
Ecuador	25	0,86	0	72	79	1	0	1
El Salvador	16,7	0,77	16	61	69	1	0	0,5
Guatemala	12	0,69	0	41	62	1	0	0,5
Guyana	29	0,94	15	53	63	1	0	1
Haiti	4,1	0,54	0	67	58	0	0	0
Honduras	23,4	0,77	0	59	53	1	0	1
India	9,1	0,62	57	31	58	0	0	0
Indonesia	11,6	0,83	8	50	63	1	0	0
Jamaica	13,3	0,79	38	73	64	0	0	0
Japan	9,4	0,95	58	31	61	0,5	0	0
Kazakhstan	15,9	0,97	0	63	84	0,5	0	0
Korea. South	13,7	0,98	19	39	60	0,5	0	1
Kyrgyzstan	25,6	0,92	0	57	90	0	0	1
Laos	25,2	0,66	0	51	50	0	1	0
Malaysia	10,8	0,84	0	40	51	0	0	0
Mexico	23,2	0,86	7	49	61	0,5	0	1
Mongolia	4,2	0,91	15	54	84	0	0	0
Nepal	33,2	0,52	0	19	57	0	1	1
New Zealand	33,1	0,99	129	53	115	0,5	0	0
Nicaragua	18,5	0,75	0	41	53	1	0	0,5
Pakistan	22,5	0,47	0	26	61	0,5	0	1
Panama	16,7	0,88	16	63	67	1	0	1
Papua New Guinea	0,9	0,52	0	72	45	0	0	0
Paraguay	10	0,85	0	76	47	1	0	1
Peru	29,2	0,87	5	71	53	1	0	1
Philippines	20,5	0,89	0	61	71	0	0	0,5
Saint Lucia	11,1	0,88	28	67	84	0	0	0
Singapore	24,4	0,91	0	44	61	0	0	0
Sri Lanka	5,8	0,81	0	46	77	1	0	0
Suriname	25,5	0,85	33	52	60	1	0	0
Tajikistan	17,5	0,9	0	57	84	0,5	0	0
Thailand	11,7	0,86	0	54	76	0,5	0	0,5
Trinidad	26,8	0,87	36	61	62	0	0	0
Turkmenistan	16	0,9	0	64	81	0	0	0
Uruguay	12,1	0,94	22	71	76	1	0	1
Uzbekistan	17,5	0,91	0	60	70	0	0	1
Venezuela	18,6	0,87	0	67	62	0,5	1	0
Vietnam	25,8	0,82	0	41	62	0	1	0

Notes: wr = share of female deputies in national parliament, edu = education index, dem = years democracy, eco = the female activity rate in per cent, ws = number of years women have had the right to vote, pr = electoral system type, q = quota provisions

Table 4.A2 Anchor points for the direct calibration of metric variables as done by Stockemer (2013: 92-94)

Anchor	Women's representation	UNDP Education Index	Women's suffrage	Consolidation of Democracy	Female activity rate
1 anchor	30%	0.95	Before WWI	50 years	90%
0.5 anchor	20%	0.75	End of WWII	30 years	70%
0 anchor	10%	0.55	1962	10 years	50%

Notes: The calibration procedure by Stockemer (2013: 92-94) is justified weakly, e.g. the outcome set 'high share of female deputies':

The threshold for a full set membership is defined at 30 percent female parliamentarians, since "approximate one-third women in parliament is frequently perceived as the critical mass enabling women to exert meaningful influence on politics" (Stockemer 2013: 97).

The benchmark for the 0.5 anchor is defined at 20 percent for the 54 Asian and Latin American countries, since it "approximates the global mean women's representation rate". For the two countries Dominican Republic and Philippines, this means that the former case is not a member of the set 'high share of female deputies' (due to 19.7% female parliamentarians) while the latter is a member (due to 20.5%). If this corresponds to the cases and/or the underlying concept of the set might to be questioned.

Finally, the full non-membership is positioned at 10 percent without any justification.

Table 4.A3 Fuzzy data matrix with cases' set memberships in condition and outcome sets

country	wr	edu	dem	eco	ws	pr	com	q
Argentina	1	0,95	0,08	0,43	0,29	1	0	1
Australia	0,88	0,97	1	0,11	0,98	0	0	0,66
Bahamas	0,09	0,88	0,68	0,96	0,03	0	0	0
Bangladesh	0,19	0,02	0,01	0	0	0	0	1
Barbados	0,05	0,96	0,82	0,88	0,55	0	0	0
Belize	0,02	0,57	0,21	0,06	0,11	0	0	0
Bhutan	0,03	0,02	0,01	0,02	0,13	0	0	0
Bolivia	0,28	0,86	0,01	0,65	0,6	0,66	0	0,66
Brazil	0,04	0,88	0,06	0,5	0,72	1	0	0
Cambodia	0,25	0,29	0,01	0,65	0,09	1	0	0
Chile	0,18	0,96	0,12	0,05	0,52	1	0	0,66
China	0,6	0,79	0,01	0,06	0,23	1	1	0
Colombia	0,03	0,86	0,01	0,68	0,11	1	0	0
Costa Rica	0,99	0,88	0,99	0,08	0,23	0	0	1
Cuba	1	0,95	0,01	0,16	0,68	0	1	0
Dominican R.	0,48	0,77	0,05	0,1	0,5	1	0	1
Ecuador	0,82	0,84	0,01	0,57	0,78	1	0	1
El Salvador	0,27	0,57	0,11	0,21	0,57	1	0	0,66
Guatemala	0,08	0,29	0,01	0,01	0,33	1	0	0,66
Guyana	0,94	0,95	0,1	0,07	0,37	1	0	1
Haiti	0,01	0,04	0,01	0,39	0,2	0	0	0
Honduras	0,73	0,57	0,01	0,16	0,09	1	0	1
India	0,04	0,12	0,98	0	0,2	0	0	0
Indonesia	0,07	0,77	0,04	0,05	0,37	1	0	0
Jamaica	0,12	0,65	0,77	0,61	0,41	0	0	0
Japan	0,04	0,95	0,99	0	0,29	0,33	0	0
Kazakhstan	0,23	0,96	0,01	0,26	0,85	0,66	0	0
Korea. South	0,13	0,97	0,16	0,01	0,26	0,33	0	1
Kyrgyzstan	0,84	0,93	0,01	0,12	0,91	0	0	1
Laos	0,83	0,21	0,01	0,05	0,06	0	1	0
Malaysia	0,06	0,79	0,01	0,01	0,07	0	0	0
Mexico	0,72	0,84	0,03	0,04	0,29	0,66	0	1

Mongolia	0,01	0,92	0,1	0,08	0,85	0	0	0
Nepal	0,98	0,03	0,01	0	0,17	0	1	1
New Zealand	0,98	0,97	1	0,07	0,99	0,66	0	0
Nicaragua	0,39	0,5	0,01	0,01	0,09	1	0	0,66
Pakistan	0,68	0,01	0,01	0	0,29	0,33	0	1
Panama	0,27	0,88	0,11	0,26	0,52	1	0	1
Papua New Guinea	0	0,03	0,01	0,57	0,02	0	0	0
Paraguay	0,05	0,82	0,01	0,71	0,03	1	0	1
Peru	0,94	0,86	0,02	0,54	0,09	1	0	1
Philippines	0,54	0,89	0,01	0,21	0,62	0	0	0,66
Saint Lucia	0,06	0,88	0,43	0,39	0,85	0	0	0
Singapore	0,79	0,92	0,01	0,02	0,29	0	0	0
Sri Lanka	0,01	0,71	0,01	0,03	0,74	1	0	0
Suriname	0,84	0,82	0,61	0,06	0,26	1	0	0
Tajikistan	0,32	0,9	0,01	0,12	0,85	0,33	0	0
Thailand	0,08	0,84	0,01	0,08	0,72	0,33	0	0,66
Trinidad	0,88	0,86	0,71	0,21	0,33	0	0	0
Turkmenistan	0,23	0,9	0,01	0,29	0,81	0	0	0
Uruguay	0,09	0,95	0,23	0,54	0,72	1	0	1
Uzbekistan	0,32	0,92	0,01	0,18	0,6	0	0	1
Venezuela	0,4	0,86	0,01	0,39	0,33	0,66	1	0
Vietnam	0,85	0,74	0,01	0,01	0,33	0	1	0

Notes: wr = high women's representation, edu = education, dem = consolidation/longevity of democracy, eco = female economic activity rate, ws = women's suffrage, pr = proportional electoral system, com = communist regime type, q = quota provisions

Table 4.A4 "Results of the regression equation" by Stockemer (2013: 95)

	B	SE	Sig
D_Quotas	10.34	2.35	0.000
D_PR	-0.61	2.38	0.799
Education	14.24	9.41	0.137
Year of Women Suffrage	0.04	0.09	0.665
Female Activity Rate	-0.03	0.08	0.729
Degree of Democracy	0.09	0.05	0.051
D_Communist State	15.76	3.32	0.000
Constant	-1.51	6.89	0.828

R^2 = 0.498, N = 54.

Table 4.A5: Truth table for the analysis of 'high women's representation' (short version without logical remainder rows)

edu	dem	eco	es	pr	com	q	nr. of cases	sufficiency statement	raw cons.	PRI
0	0	0	0	0	1	0	1	1	1,00	1,00
0	0	0	0	0	1	1	1	1	1,00	1,00
1	0	0	1	0	1	0	1	1	1,00	1,00
1	0	0	0	0	1	0	1	1	1,00	1,00
1	1	0	1	0	0	1	1	1	0,97	0,96
1	1	0	0	0	0	1	1	1	0,97	0,93
1	1	0	1	1	0	0	1	0	0,83	0,73
1	1	0	0	1	0	0	1	0	0,77	0,59
1	0	1	1	1	0	1	3	0	0,75	0,48
1	0	0	0	1	0	1	4	0	0,74	0,54
1	0	1	0	1	0	1	2	0	0,73	0,50
0	0	0	0	1	0	1	1	0	0,73	0,22
1	0	0	0	1	1	0	2	0	0,72	0,34
1	0	0	1	0	0	1	4	0	0,68	0,40
1	0	0	0	0	0	1	1	0	0,65	0,06
1	0	0	1	1	0	1	3	0	0,65	0,31
0	0	0	0	0	0	1	2	0	0,64	0,29
0	0	1	0	1	0	0	1	0	0,64	0,00
1	0	1	0	1	0	0	1	0	0,57	0,00
1	0	0	0	0	0	0	3	0	0,56	0,25
1	0	0	0	1	0	0	1	0	0,53	0,10
0	0	1	0	0	0	0	1	0	0,47	0,02
1	0	0	1	0	0	0	4	0	0,46	0,08
1	1	0	0	0	0	0	2	0	0,45	0,28
1	0	0	1	1	0	0	2	0	0,43	0,04
1	1	1	1	0	0	0	1	0	0,39	0,10
0	1	0	0	0	0	0	1	0	0,32	0,01
0	0	0	0	0	0	0	2	0	0,31	0,01
1	1	1	0	0	0	0	2	0	0,29	0,05

Notes: edu = education, dem = consolidation/longevity of democracy, eco = female economic activity rate, ws = women's suffrage, pr = proportional electoral system, com = communist regime type, q = quota provisions, nr. of cases = number of cases displayed by this ideal typical configuration of conditions, sufficiency statement = truth value on whether the truth table row is sufficient for the outcome (value 1) or not (value 0), raw cons. = consistency of sufficiency for outcome high women's representation, PRI = proportional reduction in inconsistency

5.

Relevant Consistency and Relevant Coverage in Fuzzy Set Qualitative Comparative Analysis (fsQCA)

(Jonas Buche)[43]

[43] The author is deeply grateful for the intellectual support by and lively debates on asymmetry, skewedness, diversity, and especially relevance with Antje Buche.

Abstract

The development of the 'parameters of fit' (Ragin 2006) can be seen as one if not the major invention in the rather short history of fuzzy set QCA. However, the formulas have one major shortcoming. For consistency of sufficiency, cases with the condition set X absent might strongly influence the statement that the presence of X is sufficient for Y or not. This contradicts the notion of asymmetry, which can be seen as one major argument in favor of QCA (Ragin 2008).

In this paper I, first, identify different types of relevant and irrelevant cases and their effects on the consistency of sufficiency. Second, I propose a formula for 'relevant consistency' and apply it to artificial data inspired by Anscombe's quartet (Anscombe 1973). While the Ragin formula – due to the influence of irrelevant cases – produces the same consistency scores, the 'relevant consistency' displays major differences. Third, I argue that the effect of irrelevant cases is omnipresent in the calculation of set relations and update formulas for both the consistency and coverage of sufficiency and necessity. Finally, fourth, I discuss the implications of the updated formulas on uncovering skewed membership distributions and offer a formula for 'non-skewedness' in QCA.

Keywords

Asymmetric Information; Causal Inference; Consistency; Coverage; Qualitative Comparative Analysis (QCA).

5.1 Introduction

A decade ago, Charles Ragin (2006) in this journal[44] introduced the measures of consistency and coverage for the evaluation of set relations in social research. The aim was to deliver "simple descriptive measures for evaluating the strength of the empirical support for theoretical arguments describing set relations" (Ragin 2006: 292). Until today, on the one hand, the development of these parameters can be seen as one if not the major improvement of the quality assessment in the (short) history of fuzzy set QCA. On the other hand, the formulas for consistency and coverage values have one major shortcoming. For instance, Schneider and Wagemann (2012: 126) state for the calculation of the consistency of sufficiency that the formula "does not take into account whether an inconsistent case is above or below the qualitative anchor of 0.5 in X and/or Y." Put differently, cases where a potentially sufficient condition is absent might strongly influence the statement that the presence of this condition is sufficient for the outcome or not. The same applies to the consistency of necessity. Here, cases where the outcome is absent influence the strength of the empirical support for the claim that the outcome cannot be present without the condition being present. In short, all the formulas presented by Ragin (2006) rely on cases that are irrelevant for the respective set relation (see for the discussion on "good cases" Cooper & Glaesser 2011).

In the following I argue that both the formulas for consistency and coverage for both sufficiency and necessity contradict the notion of asymmetry which can be seen as one major argument in favor of QCA (Ragin 2008, Schneider & Wagemann 2012). For the statement of sufficiency, asymmetry in QCA refers to the fact that from the sufficiency of the presence of a certain condition X for a certain outcome Y one cannot draw any conclusion on whether the outcome is present or not if the condition is absent (i.e. ~X is present).[45] Simply the other way around, whether the outcome is present or absent (or both) if the condition is absent cannot influence the evaluation of the strength of the empirical support for the sufficiency claim on the presence of that condition for the presence of the outcome. But by including all cases into the calculation, this is the underlying assumption implemented in the formula for the consistency of sufficiency (Ragin 2006). Necessary set relations are asymmetric, too. The claim that a certain outcome Y cannot occur without the necessary condition X being present is independent from any situation where the outcome is absent (i.e. ~Y is present). Put differently,

---

[44] "this journal" means Political Analysis.

[45] Using the principles of Boolean algebra one can only draw from the statement X → Y (reads X is sufficient for Y) the conclusion that ~X ← ~Y (reads not-X is necessary for not-Y).

whether the necessary condition is present or absent (or both) if the outcome is absent cannot influence the evaluation of the strength of the empirical support for the necessity claim on the presence of that condition for the presence of the outcome. Again, by including all cases – regardless of their relevance for the respective set relation – into the calculation of consistency and coverage the notion of asymmetry is violated.

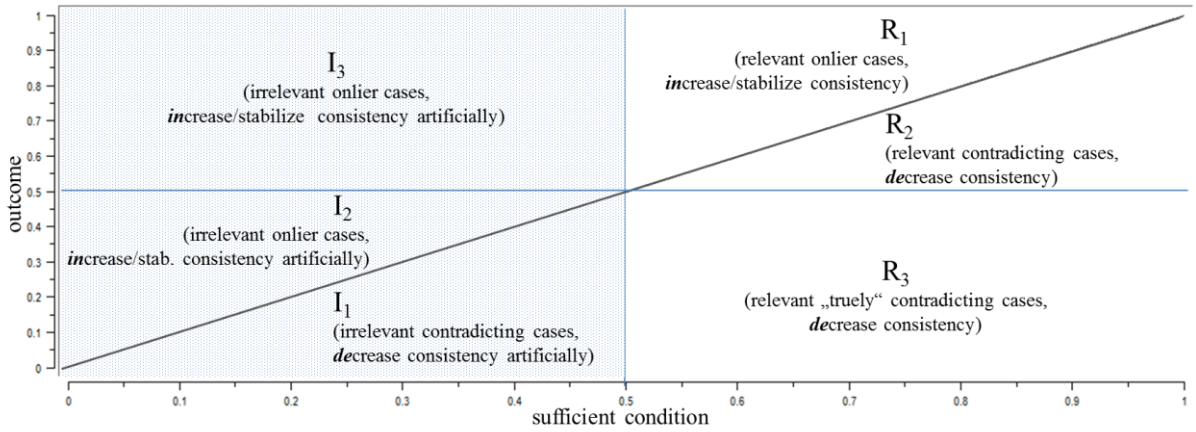
The aim of this paper is twofold. On the one hand, it identifies relevant and irrelevant cases for set relations. On the other hand, it offers remedy by a modification of the Ragin formulas excluding irrelevant cases. In the next section I discuss relevant and irrelevant cases for the consistency of sufficiency and trace their differing influence on the consistency score, i.e. the detection of false positives (type I error, see also Braumoeller 2015; Schwellnus 2013) but also false negatives (type II error). Based on the asymmetry argument I modify the formula by Ragin (2006) by excluding irrelevant cases from the calculation. Moreover, I apply the formula on 'relevant consistency of sufficiency' to artificial data. Inspired by Anscombe's quartet (Anscombe 1973) I use four datasets that have nearly identical simple statistical properties, but appear very different when graphed in XY-plots. While the Ragin formula – due to the influence of irrelevant cases – produces the same consistency scores for all four data sets regardless of the presence or absence of 'true logical contradictory cases' (Schneider & Wagemann 2012: 127), the 'relevant consistency' displays major differences. Section three argues that the effect of irrelevant cases is not limited to the formula of consistency of sufficiency. Instead, the influence of irrelevant cases is omnipresent in the calculation of set relations. Thus, I update the formula for coverage of sufficiency and apply the 'relevant coverage' formula to the artificial data as well. In section four, I discuss the mirror-image relation of sufficiency and necessity and draw conclusions for irrelevant cases and modified formulas for consistency and coverage of necessity. While the fifth section discusses implications of the updated formulas for the uncovering of skewed memberships, section six introduces a formula for the calculation of non-skewedness in QCA. The final section concludes.

5. 2 Relevant Consistency of Sufficiency

A sufficient condition X is defined as whenever X is present, the outcome Y will be present as well. However, the analysis of social reality more often than not reveals that perfect subset relations are hardly detectable. Especially when using the fuzzy set version of QCA, imperfect subset relations are almost inevitable. The reason for that is either an (comparatively small) inconsistency produced by cases being both members of the outcome and the condition with the latter membership being higher than the former. Or there might even be cases that contra-

dict the sufficiency claim by being a member in X but not in Y. According to the notion of asymmetry of set theory in general and QCA in particular, all other cases are irrelevant for the consistency of sufficiency (see figure 5.1).

Figure 5.1: Identifying relevant and irrelevant cases for the consistency of sufficiency



I argue that only cases that are members of the explanatory (combination of) condition(s) set are relevant for the calculation of consistency, i.e. $R_{1\text{-}3}$. Cases with a (more-than-not-) membership in both the condition and the outcome are relevant as they support the sufficiency claim. While R_1 includes all relevant onlier cases contributing to a perfect consistent sufficiency statement ($0.5<X\leq Y$), cases located in the area R_2 contradict the sufficiency claim slightly ($X>Y>0.5$). Furthermore, cases that are located in the area R_3 are 'true logical contradictory cases' for the statement of sufficiency ($X>0.5>Y$, see Schneider & Wagemann 2012: 123-129) as they are members of the condition set but not of the outcome set.

In contrast, all cases with the condition being absent ($I_{1\text{-}3}$) are irrelevant for the consistency of sufficiency. To include them into the calculation of consistency of sufficiency, on the one hand, contradicts the foundation of asymmetry: The statement that X is sufficient for Y is not at all influenced by cases that do not show X. This is even true for fuzzy sets. Given a sound calibration the allocation of a fuzzy membership value of $X<0.5$ is a clear decision that a certain case is not a member of the set X but of its complement ~X (difference in kind, see Schneider & Wagemann 2012), regardless of its partial (non-)membership in X (difference in degree, ibid.). On the other hand, including irrelevant cases in the calculation of consistency of sufficiency yields artificially increased or decreased values. Depending on research design aspects like the number of cases or the calibration procedure the influence of irrelevant cases will differ in extent:

To start with, both areas I_1 and I_2 are irrelevant for the sufficiency statement since all cases located in these areas are neither members in X nor Y. But they differ in their effect on the consistency of sufficiency. Cases in area I_1 (0.5>X>Y) decrease the consistency to artificially low values. For the minimization process in QCA this might lead to a situation that a (combination of) condition(s) is wrongly interpreted as inconsistent for the outcome (false negative, type II error). In contrast, cases in area I_2 increase – or stabilize a perfect – consistency value artificially since they do not contradict the statement of sufficiency (X≤Y≤0.5). The same is true for cases located in area I_3. Being non-members of the condition set but having a (more-than-not-) membership in the outcome set (X<0.5<Y) they do not contradict the consistency of sufficiency either. Together, all cases in area I_2 and I_3 are irrelevant for the statement of sufficiency but increase or stabilize the consistency value artificially. This can also lead to problematic decisions in the minimization process, because conditions might be judged as sufficient for the outcome although this judgment is (mainly) driven by irrelevant cases (false positives, type I error; see Braumoeller 2015, Schwellnus 2013).

Based on the asymmetry argument and the identification of irrelevant cases and their influence on consistency, I modify the formula for consistency of sufficiency as developed by Ragin 2006 in the following way:

Equation 5.1a: Relevant consistency of sufficiency

$$\text{Relevant Consistency}_{\text{sufficient conditions } (x_i \leq y_i)} = \frac{\sum_{i=1}^{I} min(X_i > 0.5, Y_i)}{\sum_{i=1}^{I} (X_i > 0.5)}$$

The updated formula avoids the use of irrelevant cases by calculating the consistency of sufficiency only for cases that are members in the condition set, i.e. with X>0.5.

In order to test the formula I use four artificial datasets with a small number of cases (n=10). While the x-values are the same in all four data sets ranging from 0 to 1 in increments of 0.1,[46] some of the y-values differ between the data sets in order to identify the varying influence of irrelevant cases. According to the Ragin formula the values for consistency (and coverage) of sufficiency are exactly the same in all four data sets (see table 5.1). I chose exactly 0.8 as consistency score as it is the value that Ragin (2008) states as the lower consistency threshold for the assessment of sufficiency. Moreover, it is in line with the statement by Schneider and Wagemann that only „consistency levels (well) above 0.75 are advisable" (Schneider & Wagemann 2012: 279).

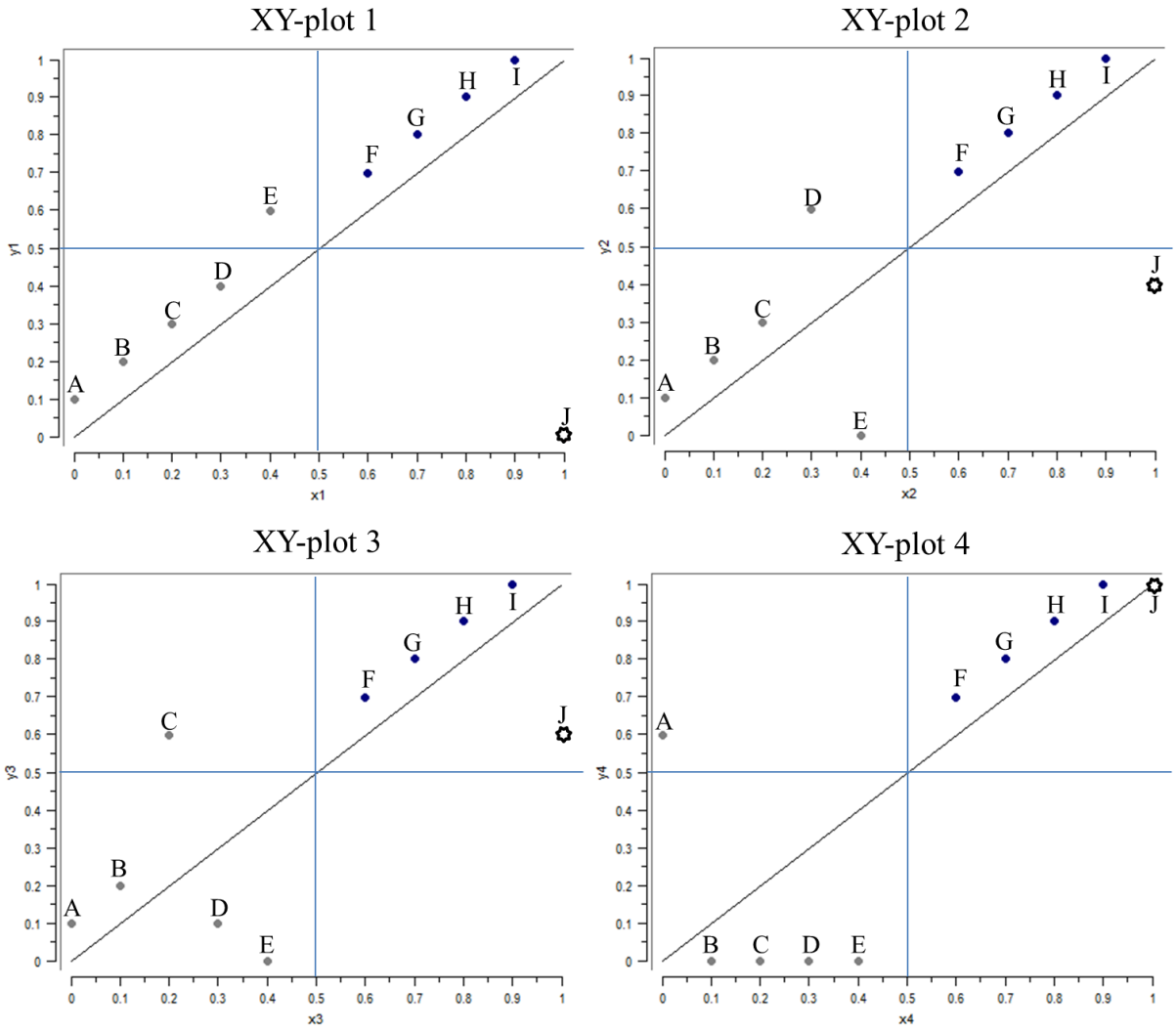
[46] I avoid the allocation of the 0.5 cross-over point to any of the cases since as the "point of maximum ambiguity" (Ragin 2008: 30) it (n)either displays a member (n)or a non-membership in the respective set.

Table 5.1: Data sets and respective consistency/coverage of sufficiency ($X_{1-4} \rightarrow Y_{1-4}$)

Cases	X_1	Y_1	X_2	Y_2	X_3	Y_3	X_4	Y_4
A	0	0,1	0	0,1	0	0,1	0	0,6
B	0,1	0,2	0,1	0,2	0,1	0,2	0,1	0
C	0,2	0,3	0,2	0,3	0,2	0,6	0,2	0
D	0,3	0,4	0,3	0,6	0,3	0,1	0,3	0
E	0,4	0,6	0,4	0	0,4	0	0,4	0
F	0,6	0,7	0,6	0,7	0,6	0,7	0,6	0,7
G	0,7	0,8	0,7	0,8	0,7	0,8	0,7	0,8
H	0,8	0,9	0,8	0,9	0,8	0,9	0,8	0,9
I	0,9	1	0,9	1	0,9	1	0,9	1
J	1	0	1	0,4	1	0,6	1	1
Ragin_Consistency of sufficiency ($X_{1-4} \rightarrow Y_{1-4}$)	0.80		0.80		0.80		0.80	
Ragin_Coverage of sufficiency ($X_{1-4} \rightarrow Y_{1-4}$)	0.80		0.80		0.80		0.80	

Although these data sets share the same consistency and coverage values, not all conditions X_{1-4} might be (correctly) classified as sufficient for the outcomes Y_{1-4}. Inspired by Anscombe (1973) who used a similar outline with four datasets with nearly identical statistical properties that appear very different when graphed, I suggest the use of XY-plots in order to, on the one hand, identify true logical contradictory cases (as proposed by Schneider & Wagemann 2010, 2012). On the other hand, the XY-plot helps to identify irrelevant cases and their influence on consistency values. In figure 5.2, the graphs vary considerably, but are identical when examined using Ragin (2006) consistency and coverage.

Figure 5.2: Visualizing the influence of 'irrelevant' cases on the consistency of sufficiency



As mentioned above, all XY-plots depict five members and five non-members in each explanatory condition $X_{1\text{-}4}$. Moreover, the four relevant onlier cases in the upper right corner of each plot do not change their position at all (cases F, G, H, and I). Regarding relevant cases (i.e. x>0.5) only the one case accentuated by a star symbol (case J) changes its position remarkably from data set to data set. Both in XY-plot 1 and 2 it truly contradicts the statement of sufficiency by being a full member in the condition set but a (full) non-member in the outcome set. Taking fuzzy logic and the differences in degree into account the logical contradiction in plot 2 is less of a problem compared to plot 1 – but still it is a true logical contradiction. In contrast, in plot 3 case J only slightly contradicts the statement of sufficiency, because it is both a member in X_3 and Y_3. Finally, in XY-plot 4 there is no inconsistent relevant case at all.

Turning the attention toward irrelevant cases (A, B, C, D, and E), XY-plot 1 shows five irrelevant cases that because of their location stabilize or even increase the consistency value according to Ragin's formula (compare figure 5.1). In other words, these cases partly compensate the inconsistency produced by the true logical contradictory case and lead to an artificially high consistency score. XY-plot 2 looks quite similar to the first one. In comparison, irrelevant case E changed its position below the diagonal (visualizing X=Y). While the remaining irrelevant cases still increase the consistency of sufficiency, this case equalizes the otherwise increased consistency produced by the star-symbol case. In XY-plot 3, another irrelevant case (D) crosses the diagonal downward and by that keeps the consistency low – although there is no true logical contradiction among the relevant cases. Finally, plot 4 displays four irrelevant cases (B-E) below the diagonal. All of them decrease the consistency artificially. Given only relevant onlier cases (F-J) there is no relevant inconsistency at all.

The differences in the calculation of consistency of sufficiency by the Ragin formula and the formula presented above are displayed in table 5.2. While the Ragin consistency is constant over all four data sets due to the influence of irrelevant cases, the relevant consistency differs remarkably. As there are no relevant inconsistent cases in the fourth data set, the sufficiency statement is perfectly consistent. The other values indicate minor or intermediate inconsistencies and clearly help to identify differences in the set relation between condition and outcome.

Table 5.2: Comparison of values using Ragin consistency and relevant consistency

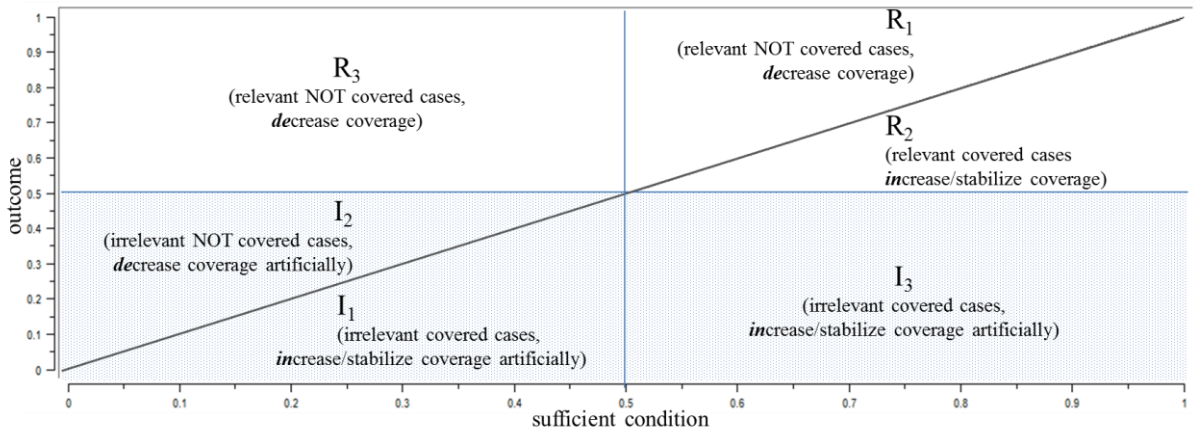
Sufficiency statement	Ragin Consistency	Relevant Consistency
$X_1 \rightarrow Y_1$	0.80	0.75
$X_2 \rightarrow Y_2$	0.80	0.85
$X_3 \rightarrow Y_3$	0.80	0.90
$X_4 \rightarrow Y_4$	0.80	1.00

5.3 Relevant Coverage of Sufficiency

The problematic influence of irrelevant cases is not limited to the consistency of sufficient conditions. Instead, it appears to be relevant for the calculation of both the consistency and coverage of both sufficiency and necessity. To start off with, the coverage of sufficiency numeralizes how much a sufficient condition set X covers the outcome set Y. Put differently, it displays the empirical importance of the sufficiency claim (X→Y). I argue that only cases that are members of the outcome set (Y>0.5) are relevant for the statement how much of the out-

come is covered by cases that are members of the explanatory (combination of) condition(s) set (see figure 5.3).

Figure 5.3: Identifying relevant and irrelevant cases for coverage of sufficiency



While cases that are located in areas $R_{1\text{-}3}$ are relevant for the assessment of the empirical importance of the sufficiency claim, cases in the areas $I_{1\text{-}3}$ just artificially influence the coverage value. To start off with, cases in area R_1 are not perfectly covered by the sufficiency statement. They are (more-than-not-) members in both the condition and the outcome set but less a member in the condition set than in the outcome set ($0.5<X<Y$). Thus, they decrease the coverage value because, expressed in a simplified manner, they do not explain the entire (fuzzy) presence of the outcome but just the (largest) part of it. Cases being located in R_2 are also both members of condition and outcome but they do not decrease coverage because they – simplified – explain the entire (fuzzy) presence of the outcome ($X \geq Y > 0.5$). Finally, cases in area R_3 are also relevant for the assessment of empirical importance of sufficiency because are not members in the condition set but in the outcome set ($X<0.5<Y$). Thus, all of those cases are not covered by sufficiency claim. Instead, they call for an equifinal explanatory condition that covers this part of the outcome.

In contrast, all cases that are not members of the outcome set ($Y \leq 0.5$) are irrelevant for the assessment of the coverage of sufficiency. Again, cases in areas I_1 and I_2 are irrelevant for the sufficiency statement since they are neither members in X nor Y. But they differ in their influence on the coverage value as well. Cases in area I_2 decrease the coverage value artificially as they are less a member in the condition compared to the outcome set ($X<Y<0.5$). This might lead to a situation that the empirical importance of an explanatory (combination of) condition(s) is undervalued although it covers all members of the outcome. In contrast, cases in I_1 increase or (stabilize a perfect) coverage value since they – simplified – cover the case's

entire fuzzy outcome by having a greater or equal membership in the condition ($0.5>X\geq Y$). Likewise, cases located in area I_3 increase or stabilize the coverage of the sufficiency statement artificially, i.e. without even being members of the outcome ($X>0.5\geq Y$). This might lead to a situation that the empirical importance of the sufficiency claim is overestimated, because cases that logically contradict this statement are involved in the calculation.

Based on this argument I modify the formula for coverage of sufficiency as developed by Ragin (2006) in the following way:

Equation 5.1b: Relevant coverage of sufficiency

$$Relevant\ Coverage_{sufficient\ conditions\ (x_i \leq y_i)} = \frac{\sum_{i=1}^{I} min(X_i, Y_i > 0.5)}{\sum_{i=1}^{I}(Y_i > 0.5)}$$

With the exclusion of all irrelevant cases from the calculation the formula bases only on cases with $Y>0.5$ and, thus, expresses the relevant coverage of sufficiency.

Testing the formula with the four data sets presented above (see table 5.1) the influence of irrelevant cases on the calculation of coverage of consistency becomes visible in comparison (see table 5.3).

Table 5.3: Comparison of values using Ragin coverage and relevant coverage

Sufficiency statement	Ragin Coverage	Relevant Coverage
$X_1 \rightarrow Y_1$	0.80	0.85
$X_2 \rightarrow Y_2$	0.80	0.83
$X_3 \rightarrow Y_3$	0.80	0.83
$X_4 \rightarrow Y_4$	0.80	0.80

The differences between the coverage values are not as remarkable as between the consistencies of sufficiency (see table 5.2). This is due to the rather small changes implemented in the data sets for cases that are relevant for coverage (see figure 5.4).

Figure 5.4: Visualizing the influence of 'irrelevant' cases on the coverage of sufficiency

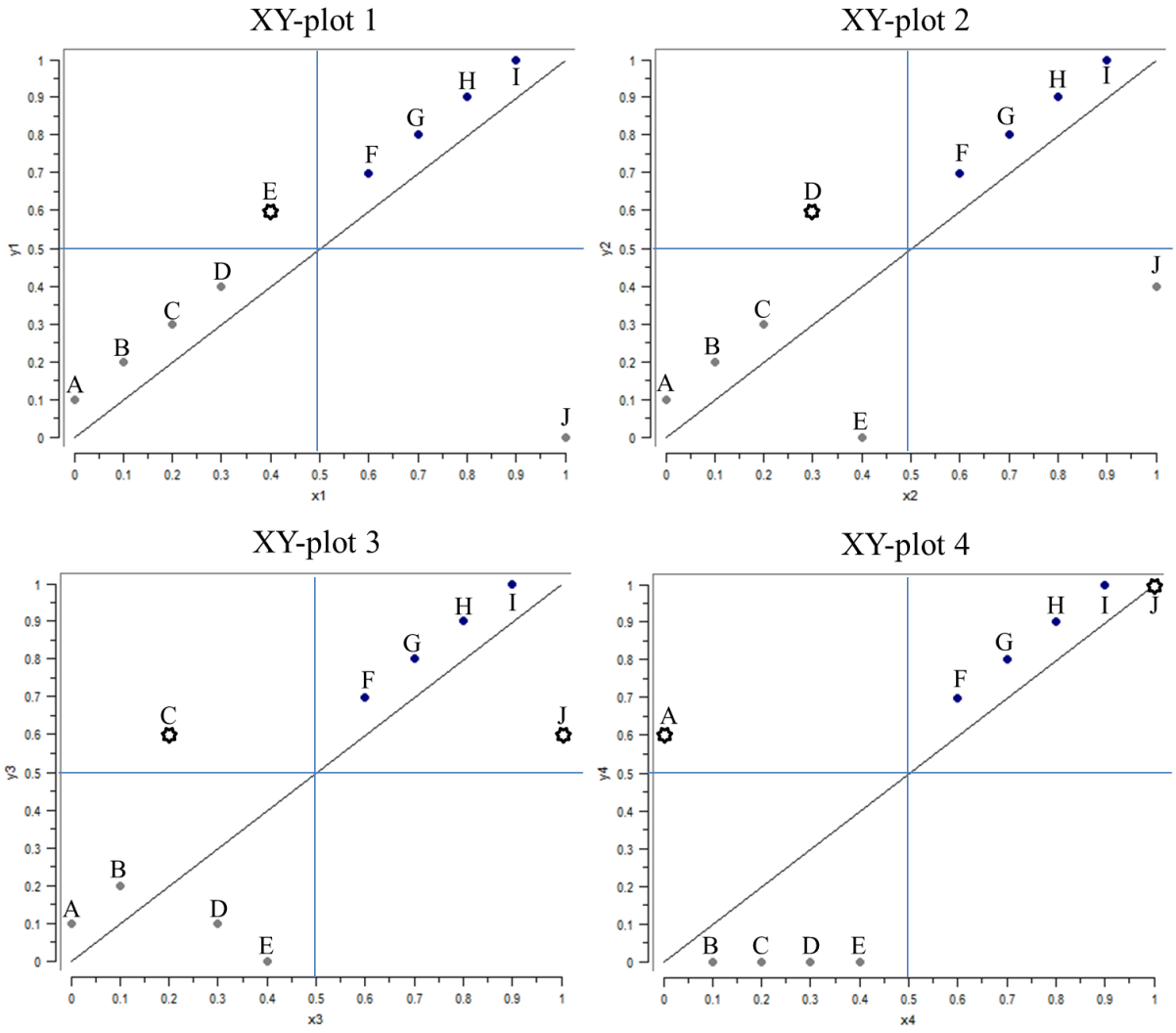


Figure 5.4 displays the same XY-plots as figure 5.2 but highlights those relevant cases with a star symbol that are responsible for the changing coverage values. In all four plots there are four slightly non-covered cases that are members in both condition and outcome in the upper right quadrant (cases F, G, H, and I). Moreover, there is only one relevant case that is clearly not covered by the respective sufficiency statement because of being a member in the outcome but not in the condition set. In XY-plot 1 the difference between case E's membership in the condition and outcome set is rather small. From plot 2 to 4 the non-covered part of the outcome raises because the X-values of the respective cases D, C, and A get smaller (from E_1=0.4 in the first plot to A_4=0 in the fourth plot) with the $Y_{1\text{-}4}$-values staying constant ($Y_{1\text{-}4}$=0.6). Thus, the relevant coverage decreases from plot 1 to plot 4 due to the respective non-covered case in the upper left quadrant. But, the additional relevant (and perfectly covered) case J that appears in the upper right quadrant (in XY-plot 3 and 4) partly equalizes the de-

crease of relevant coverage. Again, while the Ragin coverage is constant over all four data sets due to the influence of irrelevant cases, the relevant coverage differs (slightly).

5.4 Relevant Consistency and Coverage of Necessity

As stated above, the problematic influence of irrelevant cases on the parameters of fit in QCA is not limited to sufficiency. Given the mirror-image relation of necessity and sufficiency, one easily can draw conclusions from the irrelevance of cases for 'consistency of sufficiency' to 'coverage of necessity' and from the irrelevance of cases for 'coverage of sufficiency' to 'consistency of necessity'.

A necessary condition X is defined as whenever the outcome Y is present, X has to be present as well. Put differently, the outcome Y cannot occur without the presence of the necessary condition X. Stressing the asymmetry argument again, whether condition X is present or absent in cases that are not members of the outcome set Y is irrelevant for the necessity claim. However, as all cases are included in the calculation of consistency of necessity as developed by Ragin (2006) the notion of asymmetry is violated here, too. Taking the asymmetry argument seriously, all cases that are not members of the outcome set (i.e. Y≤0.5) have to be excluded from the calculation of the consistency of necessity (compare figure 5.3). Thus, the formula of relevant consistency of necessity is the same as the calculation of coverage of sufficiency:

Equation 5.2a: Relevant consistency of necessity

$$Relevant\ Consistency_{necessary\ conditions\ (x_i \geq y_i)} = \frac{\sum_{i=1}^{I} min(X_i, Y_i > 0.5)}{\sum_{i=1}^{I} (Y_i > 0.5)}$$

An equivalent mirror-image relation appears with the coverage of necessity and the consistency of sufficiency. Those cases that are not members in condition set X (i.e.X≤0.5) are irrelevant for the calculation of coverage of necessity as well since they cannot contribute to the empirical importance of the statement that the occurrence of X is necessary for the occurrence of Y (compare figure 5.1). In other words, only cases that have a (more-than-not) membership in X are relevant for the empirical assessment of how much the outcome set Y covers the necessary superset X. Thus, the formula of relevant coverage of necessity equals the calculation of relevant consistency of sufficiency:

Equation 5.2b: Relevant coverage of sufficiency

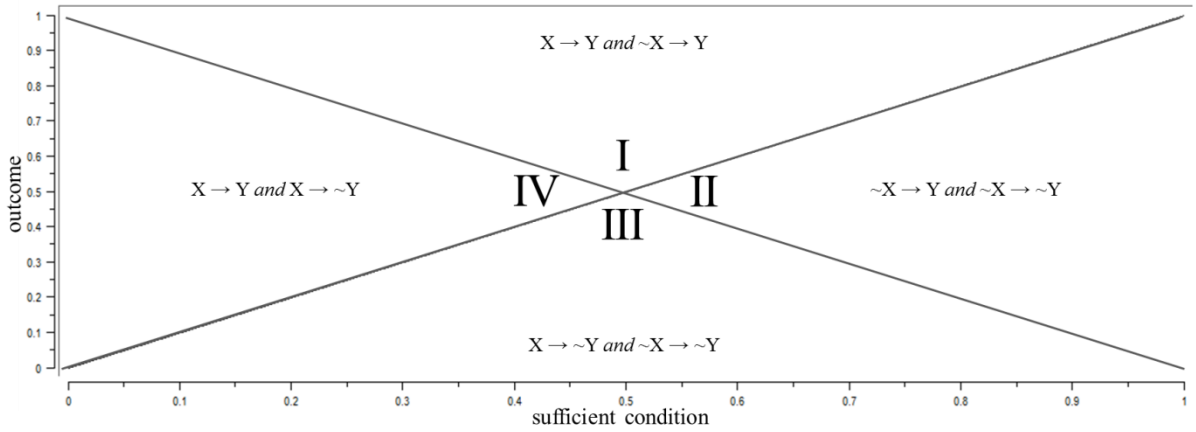
$$Relevant\ Coverage_{necessary\ conditions\ (x_i \geq y_i)} = \frac{\sum_{i=1}^{I} min(X_i > 0.5, Y_i)}{\sum_{i=1}^{I} (X_i > 0.5)}$$

The suggested formula for relevant coverage should not be confused with the discussion of the 'relevance of necessity' by Schneider and Wagemann (2012: 235) or the 'trivialness of necessity' by Goertz (2006b). In contrast to the asymmetry argument put forward here, these two formulas aim at detecting the analytic consequences of skewed set-membership scores for necessity set relations. However, in the next part I argue that the calculation of relevant consistency offers a solution for the problem of skewedness in QCA at least for sufficiency set relations.

5.5 Skewedness and Relevant Consistency of Sufficiency

Skewed distributions of set memberships might lead to inaccurate inferences about set relations. To be more precise, skewedness may result in simultaneous subset (and superset) relations of the presence and/or absence of a certain condition X on the presence and/or absence of the outcome Y (Schneider & Wagemann 2012: 232). Using a XY-plot, Schneider and Wagemann (2012: 245) graphically identify four areas of skewedness (see figure 5.5). Whenever all cases are located in only one of the four areas, simultaneous subset relations occur.[47]

Figure 5.5: Skewedness and simultaneous sufficiency relations



Source: Schneider & Wagemann 2012: 245

To begin, if all cases are located in area I, both the presence and absence of X is sufficient for the outcome (X→Y and ~X→Y). Likewise, if all cases are falling in area III, the presence and absence of X is sufficient for the non-outcome (X→~Y and ~X→~Y). Both problems occur only if the outcome is (strongly) skewed toward membership (area I) or non-membership (area III). Schneider and Wagemann (2012: 248) argue that skewedness in the outcome set is not

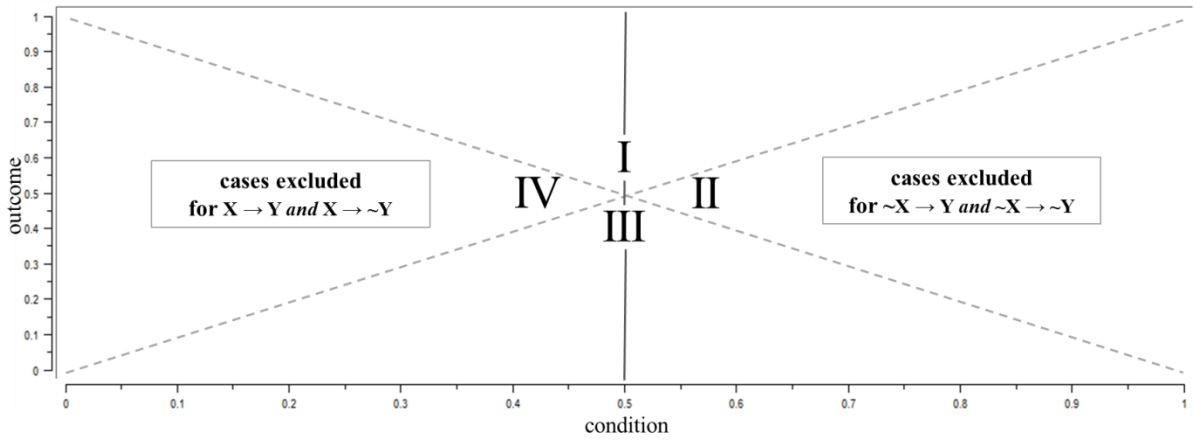
[47] Of course, the same is true for superset relations, i.e. simultaneous necessity claims for the presence and/or absence of a certain condition X on the presence and/or absence of the outcome Y (see table 5.A1 in the appendix). However, for higher clarity of the argument I restrict the discussion at this point on simultaneous sufficiency relations.

so much of a problem due to the Y-orientation in applied QCA. Since outcome sets usually consist of a single set only, skewed membership scores are very likely to be noticed.

In contrast, conditions X might be complex combinations of sets (e.g. truth table rows). This hinders to detect skewedness in X even if the single condition sets are not skewed. Given the Boolean logic of AND-combinations using the minimum rule for the intersection of sets, skewedness in X is more likely to tend to non-membership. Consequently, cases might cluster in area IV (see figure 5.5) which would lead to simultaneous subset relations of X in Y and ~Y. In other words, condition X would be sufficient for both the outcome and the non-outcome (X→Y and X→~Y). The very same might occur for the absence of the condition, i.e. low membership score in ~X. Then all cases might be located in area II and the absence of the condition would be sufficient for both the outcome and the non-outcome (~X→Y and ~X→~Y).

Based on the asymmetry discussion above, I argue that the calculation of relevant consistency of sufficiency offers remedy for consequences of skewedness in X. For instance, all cases that cluster in the left triangle (i.e. area IV) would be excluded from the calculation of relevant consistency of sufficiency for the presence of condition X because they only artificially contribute to sufficiency statements X→Y and X→~Y. Likewise, all cases that are located in area II would be excluded from the calculation of relevant consistency of sufficiency for the absence of condition X because they only artificially contribute to sufficiency statements ~X→Y and ~X→~Y. To conclude, the discussion on skewedness reveals another benefit of relevant consistency: the calculation addresses and solves skewedness in X toward membership or non-membership (see figure 5.6).

Figure 5.6: The effect of 'relevant consistency of sufficiency' on simultaneous sufficiency relations



For necessary set relations, however, the effect of 'relevant consistency' is reduced to solve problems of skewedness in Y (see figure 5.A2 in the appendix). But, as examined by both Schneider and Wagemann (2012: 235) and Goertz (2006b) the problem of skewedness for necessity is especially relevant if condition X is a (quasi-)constant; i.e. if cases cluster in areas II or IV, again. For instance, if all cases are located in area II not only a simultaneous sufficiency relation of ~X for both Y and ~Y occur (see figure 5.5). Moreover, the presence of X is simultaneously necessary for the presence and the absence of the outcome (X←Y and X←~Y; see figure 5.A1). This holds true for the calculation of 'relevant consistency of necessity', too although it takes into account different cases. For the necessity of X for Y the relevant consistency relies on only those cases that are in the upper half of area II, i.e. members of Y. In contrast, for the calculation of relevant consistency for X as necessary condition for ~Y only cases that are not members in Y are considered. Nevertheless, in such an empirical situation the necessity claim is perfectly trivial since condition X is present regardless of the presence or absence of the outcome. Consequently, Schneider and Wagemann (2012: 235) and Goertz (2006b) start off from Ragin's coverage measure to deal with consequences of skewed memberships.

However, if I apply the 'relevant coverage of necessity' skewedness cannot be detected since the value displays the *relevant* empirical importance of the necessity statement by including only cases that have a (more-than-not) membership in X. But skewedness can only be identified if information on the relation of membership and non-membership in X is at hand. Whenever (almost) all cases or (almost) no case is included in the calculation of relevant consistency or sufficiency of necessity or sufficiency, skewed distributions of set memberships are the reason. Hence, the 'share of cases' included in the calculation might be an important information.[48]

Equation 5.3a: Share of cases for relevant consistency of sufficiency

$$Share\ of\ Cases\ Relevant\ Consistency_{sufficient\ conditions\ (x_i \leq y_i)} = \frac{N_{X_i>0.5}}{N_{X_i}}$$

Equation 5.3b: Share of cases for relevant coverage of sufficiency

$$Share\ of\ Cases\ Relevant\ Coverage_{sufficient\ conditions\ (x_i \leq y_i)} = \frac{N_{Y_i>0.5}}{N_{Y_i}}$$

[48] Equations 3a and 3b take the example of sufficiency again. But due to the mirror-image relation of sufficiency and necessity discussed above, the calculation of the 'share of cases' for 'relevant consistency of sufficiency' is the same as for 'relevant coverage of necessity' and vice versa (see section 4).

Applied to the artificial data from table 5.1, the share of relevant cases for consistency of sufficiency is constant at 0.5 since membership and non-membership in X is equally distributed (see table 5.4). Thus, the relevant consistency of sufficiency bases on 50 per cent of all cases. Likewise, the share of relevant cases for coverage of sufficiency is almost constant due to controlled balance of membership in Y. Hence, the relevant coverage of sufficiency is based on 50 per cent of all cases for $X_{1\text{-}2}$ and on 60 per cent of all cases for $X_{3\text{-}4}$.

Table 5.4: Relevant consistency/coverage and share of cases

Sufficiency statement	Relevant Consistency	Share of Cases Consistency	Relevant Coverage	Share of Cases Coverage
$X_1 \rightarrow Y_1$	0.75	0.5 (5/10)	0.85	0.5 (5/10)
$X_2 \rightarrow Y_2$	0.85	0.5 (5/10)	0.83	0.5 (5/10)
$X_3 \rightarrow Y_3$	0.90	0.5 (5/10)	0.83	0.6 (6/10)
$X_4 \rightarrow Y_4$	1.00	0.5 (5/10)	0.80	0.6 (6/10)

But although the value 'share of cases' is an appropriate tool to measure the distribution of cases underlying the relevant consistency and coverage, the interpretation is complicated for at least two reasons. First, as pointed out for the artificial data a share of 5/10 means that the cases are equally distributed among membership and non-membership. Put differently, for this example the value of 0.5 depicts that skewedness is not a problem at all. By contrast, values of 1 and 0 would indicate high levels of skewedness since either all or no cases are included to the calculation. Thus, in contrast to every other parameter in QCA the value 1 is not the optimal result for non-skewedness. More important, second, the 'perfect share' of 0.5 is valid only for single conditions, not for combinations of conditions. When the complexity of configuration increases the interpretability gets even more puzzling. Hence, in the final section I argue that the calculation of skewedness with the simple 'share of cases' is appropriate only if single conditions are analyzed. Moreover, I suggest the calculation of 'non-skewedness' by adapting the Blau-Index of heterogeneity (Blau 1977: 78).

5.6 Non-Skewedness in QCA

So far, there is no formula for the calculation of (non-)skewedness of the entire explanatory model in QCA. Skewedness is caused by deviances from perfectly equal allocations of set memberships and non-memberships over all conditions included into the explanatory model. As pointed out above, the hypothetical example deals with sufficiency statements that base on one single condition $X_{1\text{-}4}$, i.e. with the two simple categories of membership and non-membership. Thus, a 'share of cases' of 0.5 displays perfect non-skewedness. But for suffi-

ciency, the logic of Boolean AND-combinations (mathematical minima) results in low membership scores, i.e. a tendency toward non-membership. Hence, the question is not whether but how strongly complex configurations of INUS conditions (Mackie 1974) tend to non-membership. Put differently, with more complex configurations there are different categories of non-membership that need to be considered for the calculation of skewedness. For instance, a truth table with three conditions consists of 2^3=8 logically possible configurations, i.e. truth table rows. For statements of skewedness for every single truth table row it does not make much sense to base the calculation on the two categories membership and non-membership only. If one did so, every truth table row would need to cover half of the cases to reach a 'share of cases' at 0.5 – which is logically impossible. Thus, if the 'share of cases' for relevant consistency of sufficiency is to be calculated for one truth table row, the number of categories needs to be increased. For the hypothetical truth table with three conditions the number of categories needs to be 8, i.e. one category of membership and seven categories for different non-memberships. If all cases are evenly distributed among membership and non-membership in every condition, i.e. if all three single conditions are not at all skewed, every of the 8 truth table rows will cover the same amount of cases. Thus, a perfectly non-skewed 'share of cases' in such a situation equals 1/8 per row. However, a value of 0.125 indicates a rather low empirical support for the sufficiency statement and is hardly interpretable as non-skewed. To perpetuate, for a truth table with four conditions, a non-skewed 'share of cases' would be about 1/16 (numerical 0.0625) for every row, for five conditions 1/32 (numerical 0.03125), etc. Likewise, for statements of necessity the logic of Boolean OR-combinations (mathematical maxima) results in high membership scores and a tendency toward set membership. Hence, the relation of membership and non-membership in complex configurations of SUIN-conditions (Mahoney et al. 2009) will be skewed if all cases are evenly distributed among membership and non-membership in every single condition.

In short, every configuration of conditions will not display an equal distribution of membership and non-membership unless the entire explanatory model is (highly) skewed. Thus, the interpretation of the simple 'share of cases' needs not only to consider the empirical diversity of cases, but also the complexity of configurations, because every increase of complexity displays another variety of categories for membership or non-membership of cases. These problems are mathematical issues that can be addressed by enhancing the calculation of (non-)skewedness.

Non-skewedness can be defined as the level of empirical diversity included into the explanatory model toward an equal distribution of set memberships and non-memberships. I suggest

calculating the extent of non-skewedness with the Blau-Index (Blau 1977: 78; see also "index of fractionalization" or inversed Herfindahl-Index). This index bases on the calculation of the share of one category in the total number of categories. By that it displays diversity in distributions of cases among a random number of categories:

Equation 5.4a: Non-skewedness based on Blau-Index

$$Non - Skewedness \quad = 1 - \sum_{k=1}^{K} {p_k}^2$$

Variable *p* stands for the proportional share of cases belonging to the k^{th} category. The number of possible categories ranges from 1 to *K*. The value of non-skewedness ranges between 0 and the theoretical maximum *(K-1)/K*. Thus, in order to compare the values for non-skewedness among different numbers of categories, the formula has to be standardized to values between 1 and 0 by dividing the Blau-Index by its theoretical maximum:

Equation 5.4b: Standardized Non-skewedness based on Blau-Index

$$Standardized\ Non - Skewedness \quad = \frac{1 - \sum_{k=1}^{K} {p_k}^2}{(K-1)/K}$$

In order to apply the formula as a first example I use the data of table 5.1 again. Here, there are two categories (*K*=2), membership and non-membership. As membership and non-membership is equally distributed in the sets $X_{1\text{-}4}$, non-skewedness takes the theoretical maximum value of 0.5 which implies a perfect standardized non-skewedness of 1 (see table 5.5). Regarding the relevant coverage the statements of sufficiency $X_{1/2} \rightarrow Y_{1/2}$ also bases on perfectly non-skewed memberships in $Y_{1/2}$, while memberships in $Y_{3/4}$ are slightly skewed.

Table 5.5: Relevant consistency/coverage, share of cases, and standardized non-skewedness for 10 cases with 2 categories

Sufficiency statement	Relevant consistency	Share of cases	Stand. non-skewedness	Relevant Coverage	Share of cases	Stand. non-skewedness
$X_1 \rightarrow Y_1$	0.75	0.5 (5/10)	1	0.85	0.5 (5/10)	1
$X_2 \rightarrow Y_2$	0.85	0.5 (5/10)	1	0.83	0.5 (5/10)	1
$X_3 \rightarrow Y_3$	0.90	0.5 (5/10)	1	0.83	0.6 (6/10)	0.96
$X_4 \rightarrow Y_4$	1.00	0.5 (5/10)	1	0.80	0.6 (6/10)	0.96

As a second example, I apply the formula for non-skewedness to a more complex example of a truth table with three conditions. As outlined above, the number of categories increases ($K=2^3=8$). Next to membership in one truth table row, i.e. the respective ideal typical configu-

ration of the three conditions, every case is non-member in seven other truth table rows. Thus, a perfectly non-skewed distribution of cases would mean that every configuration has the same number of members and non-members.

Table 5.6: Standardized non-skewedness for different distribution of 8 cases over 8 categories

Setting	Logically possible configurations (truth table rows)								Share of cases row ABC	Stand. non-skewedness
	ABC	AB~C	A~BC	A~B~C	~ABC	~AB~C	~A~BC	~A~B~C		
S_1	1	1	1	1	1	1	1	1	1/8	1
S_2	2	0	1	1	1	1	1	1	2/8	0.96
S_3	3	0	0	1	1	1	1	1	3/8	0.89
S_4	4	0	0	0	1	1	1	1	4/8	0.79
S_5	5	0	0	0	0	1	1	1	5/8	0.64
S_6	6	0	0	0	0	0	1	1	6/8	0.46
S_7	7	0	0	0	0	0	0	1	7/8	0.25
S_8	8	0	0	0	0	0	0	0	8/8	0

Table 5.6 displays eight different distributions of eight cases among the eight logically possible configurations of conditions A, B, and C.[49] In setting S_1 the cases are equally distributed over all configurations. From setting S_2 to S_8 the distribution of cases gets less equal, more and more cases cluster in the second column representing the configuration A*B*C. Column 10 presents the 'share of cases' for the configuration A*B*C increasing from 1/8 (setting S_1) to 8/8 (setting S_8). Taking into account membership and non-membership in row ABC only, setting S_4 with half of the cases being members is the least skewed one. But as explained above, non-skewedness in one configuration implies skewedness in (some of) the other seven configurations. The 'standardized non-skewedness' (see table 5.6, last column) takes into account skewed (non-)memberships over all eight configurations of the entire explanatory model. The more cases cluster into one (or few) configurations, the smaller gets the value for non-skewedness calculated with the standardized Blau-Index.[50]

However, deviances of non-skewedness in applied QCA are more than expectable. In the analysis of social science reality, cases will cluster in some configurations while other combi-

[49] Of course, these are just a very small share of all $\binom{8_{cases} + 8_{rows} - 1}{8_{cases}} = \frac{15!}{8! \times (15-8)!} = 6{,}435$ possible distributions of 8 cases to 8 different rows.

[50] To relate this value back to the simple share of cases included in the calculation of relevant consistency/coverage of sufficiency/necessity, one can state that the share of cases calculated by the division of 1 by the number of categories displays the least skewed distribution of cases among the truth table rows.

nations of conditions will be logical remainders. Thus, a value of 1 for the standardized non-skewedness is neither necessary nor sufficient for a theoretically proper calibration or analysis. Quite the reverse, a low value of non-skewedness indicates problems of skewedness in the overall explanatory model and can act as a warning sign that simultaneous subset relations might influence inferences about set relations.

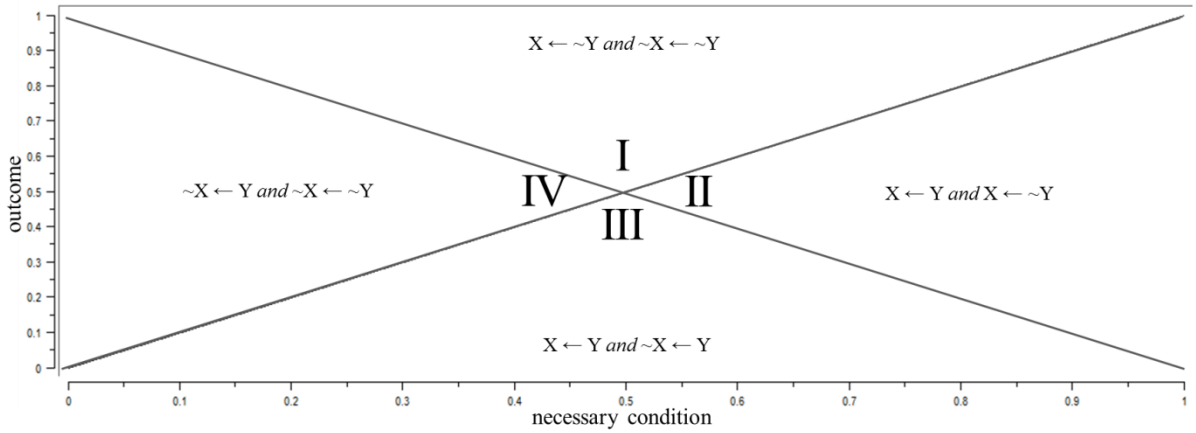
5.7 Conclusion

The aim of this contribution was to discuss the parameters consistency and coverage proposed by Charles Ragin (2006) for the assessment of empirical support for set relations. I argued that both measures contradict the notion of asymmetry as the calculation includes cases that are irrelevant for the respective set relations of sufficiency or necessity. Starting with the classification of different types of relevant and irrelevant cases and their influence on set relations, I modified the formulas for consistency and coverage for sufficiency and necessity by excluding irrelevant cases from the calculation. Following Anscombe (1973) I applied the updated formulas to artificial data that clearly illustrate the effects of irrelevant cases on the original calculation of coverage and consistency. Moreover, I related the new measures to the debate on skewed distributions of set membership scores and show for sufficient conditions that the exclusion of irrelevant cases solves the analytical problem of simultaneous subset relations, i.e. inaccurate inferences about sufficiency due to skewedness in X. As this is not simply transferable to skewedness in X for necessity claims I proposed the Blau-Index (Blau 1977) as an appropriate formula for non-skewedness in QCA. It is able to detect the extent of unequal distributions of set memberships and non-memberships across all logically possible configurations of conditions. Again, the measurement of non-skewedness is just a tool to visualize the level of (un-)equal distributions of set (non-)memberships that base on the – theoretically driven – calibration of raw data.[51] In contrast, I expect the updated formulas for relevant consistency and coverage to circumvent the influence of irrelevant cases, to avoid the analytical problems for the analysis of sufficiency that occur due to skewedness in X, and, finally, to enhance the analysis of asymmetric set relations.

[51] A value of 1 only displays perfectly equal distributions of cases. It does not imply a sound (justification of) calibration.

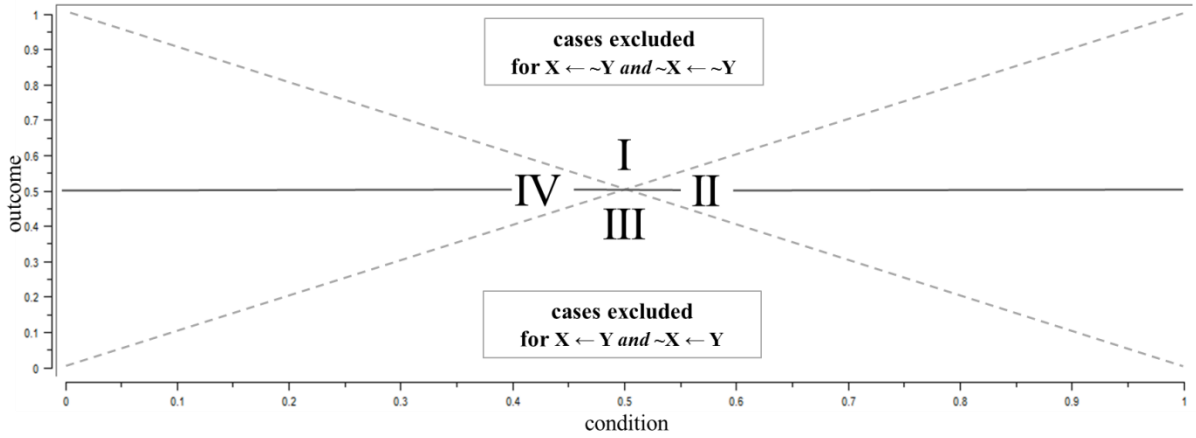
Appendix

Figure 5.A1: Skewedness and simultaneous necessity relations



Source: Schneider & Wagemann 2012: 245

Figure 5.A2: The effect of 'relevant consistency of necessity' on simultaneous necessity relations



6.

Europeanization of Legislative-executive Relations at the Micro Level – Under Which Conditions Do Swedish MPs Interact with Ministerial Officials?

(Jonas Buche)[52]

[52] The author is grateful for the support by and fruitful discussions with Antje Buche, Aron Buzogany, Julia Fleischer, Dani Nogens, Markus Siewert, Andrej Stuchlik, Claudius Wagemann, and Alexandra Wenhuda. My special thanks go to the members and committee staff of the Swedish Riksdag and the ministries of the environment and of social affairs. They provided me with detailed information about their information gathering and processing strategies in domestic and EU affairs. All interviews and surveys have been conducted in the context of an international comparative research project on 'role perception and interaction patterns of MPs and civil servants in Europeanized policies' led by Sabine Kropp at the German Research Institute for Public Administration.

Abstract

This article applies Fuzzy Set Qualitative Comparative Analysis to examine the Europeanization of legislative-executive relations at the micro-level. It asks under which conditions members of national parliaments use civil servants from the national ministries as direct information sources both for domestic and Europeanized policies. Using role theory, I state that roles, understood as the combination of position and preference roles, can be used as reliable predictors for individual behavior, i.e. interaction. I examine four conditions of executive-legislative relations presumed to be relevant for the frequency of direct interaction between MPs and bureaucrats: membership in a governing party, strong political expertise, strong role orientation as a policy expert, and strong external role perception of civil servants as policy experts. Empirically, the paper draws on survey data and semi-structured interviews with 22 members of the Swedish Riksdag conducted between 2010 and 2012. Sweden is chosen as the most likely case for frequent direct interaction both in domestic and Europeanized policies. The results indicate differences between interactions on domestic and EU-related issues. While in Swedish policies the membership in a government party is necessary for frequent interaction with civil servants, in European affairs MPs' political and policy expertise account for frequent interaction.

Keywords

Fuzzy Set Qualitative Comparative Analysis (fsQCA); Legislative-Executive Relations; Parliamentary Europeanization; Role Theory.

6.1 Introduction

Research on the Europeanization of executive-legislative relations has often focused on institutional adaptation (for an overview, see Kropp et al. 2012). Within-case studies and cross-national analyses especially shed light on formal arrangements of parliaments and executives in European affairs. For national parliaments these formal rights in scrutiny regard time and scope of information rights and the bindingness of a mandate prior to council meetings (Bergman 1997, 2000; Raunio 2005, 2009, 2011; Karlas 2012). Informal relations, though acknowledged to be important, are still largely neglected or remain a merely speculative reference (for exceptions, see Auel 2006; Kropp 2010). Holzhacker (2008: 151) notes that "there needs to be more focus on how the key actors involved in the process of scrutiny, the parliamentary party groups, individual MPs and the ministers and ministries actually use these legal instruments and institutions in practice". However, still most studies focus on formal institutions and the functionality of the European Affairs Committees (EAC) and often lack microfoundation. I share the view that "formal rules do not suffice" (Hegeland & Neuhold 2002: 12), that for instance the formal privilege to receive all EU documents does not imply a proper scrutiny. Instead, it is important to discover how single MPs collect, filter and prioritize information in order to differentiate between more and less important EU documents. In the following, I focus on individual MPs' information strategies. In order to gain knowledge in the policy process, members of national parliaments in the EU have a variety of information sources, such as party staff, NGOs, and the ministerial bureaucracy. Aberbach et al. (1981: 210) state that „[t]he intricacy of contemporary policy agendas may compel politicians to direct their attention toward the bureaucratic apparatus in order to absorb specialized knowledge relevant to policymaking" (Aberbach et al. 1981). To study how elected members of parliaments and appointed bureaucrats interact is a crucial factor for the understanding of policy-making at the micro level – both in domestic and the European policies. However, we lack studies about the interactions between these two groups (Goetz & Meyer-Sahling 2008; Kropp & Ruschke 2010; Kropp et al. 2012). I assess the interaction patterns between MPs and ministerial officials in domestic and Europeanized policies in a comparative perspective. The aim is to discover differences in information gathering posed by the Europeanization of policy making in national parliaments.

According to Aberbach et al. (1981), both for parliamentarians from parties in government and in opposition, it seems to be more and more profitable to maintain close contact with officials in order to stay informed. But individual parliamentarians have different opportunity

structures to rely on ministries as information sources. While it may be considered easy for parliamentarians belonging to one of the governing parties to contact line officials, opposition MPs face difficulties. Hence, one could argue that the "new dualism" (Eberbach-Born & Kropp 2013: 15) between government and opposition explains variation in MPs' interaction patterns. Moreover, since time is a scarce resource MPs' ideal information is already both filtered and 'politically weighted' – which does not fit a public administration's focus on detailed and neutral expertise. Thus, on the one hand one could argue that MPs generally rely on other information resources than officials. On the other hand, especially in Europeanized policies national public administrations have become a dominating actor through their extensive involvement in drafting and implementing European law. Following Aberbach et al. (1981), I expect the intricacy of Europeanized policies to strengthen the importance of ministerial bureaucrats as information source for individual MPs. Thus, the question is raised whether, and if so, under which conditions MPs interact directly with civil servants from the national ministries.

One possibility to shed light on these interactions is to look at roles that allow for predicting behavior (Derlien & Mayntz 1988; Kropp et al. 2012). In terms of neo-institutionalism, roles can be seen as a way to link micro-level behavior to macro-level structures (Peters 2005; Scharpf 1997). I argue that next to parliamentary position and preference roles (Searing 1991, 1994) role perceptions of bureaucrats can be used as reliable predictors for individual behavior, i.e. individual interaction patterns.

Empirically, the paper bases on semi-structured interviews with 22 members of the Swedish *Riksdag* on their interaction patterns and a structured questionnaire with the same persons on their self-perception as policy experts and their external perception of civil servants as policy experts. Due to high formal parliamentary power in domestic and Europeanized policies (Bergman 1997, 2000, 2003; Karlas 2012; Raunio 2005) and the low functional politicization of the ministerial bureaucracy (Pierre 2004), Sweden is argued to be the *most likely case* for frequent and intense interactions between parliamentary and administrative actors.

Methodologically, a fuzzy set Qualitative Comparative Analyses (fsQCA) (Ragin 1987, 2000, 2008; Schneider & Wagemann 2012) is applied to the original micro-level data on the interaction between Swedish parliamentarians and bureaucrats. The aim of QCA is to detect complex set relations among purposefully selected cases. Instead of the net effect of single variables on a dependent variable, QCA is interested in conjunctions of explanatory factors jointly being necessary or sufficient for an outcome. Here, I expect the outcome set 'frequent direct interac-

tion with national civil servants' on both I) domestic and II) EU-policies to be a result of complex conjunctions of the four conditions a) 'membership in a governing party' (*insider*), b) 'strong political expertise' (*experience*), c) the 'strong policy expertise' (role *orientation*), and d) the 'strong policy expertise of civil servants' (role *perception*).

The paper is structured as follows: in the next section, I discuss role theory as the theoretical framework of the paper. The third section outlines the research design, data and methods. Section four introduces the concept and fuzzy set calibration of the outcomes, i.e. MPs' direct interaction with bureaucrats in domestic and Europeanized policies, as well as the four conditions expected to explain the outcome. In section five, the fsQCA is conducted and the final section concludes.

6.2 Theoretical foundation – role theory

As outlined above, many studies so far reduce the role of national parliaments in the EU on formal rights. The explanatory factors for the strength of parliamentary scrutiny can be broadly assigned to historical, rational-choice, and sociological institutionalist approaches (for an overview, see Kropp et al. 2012). Rational-choice and sociological institutionalist logics are not mutually exclusive. Political action is neither generally explicable as based exclusively on the logic of consequences nor as based exclusively on the logic of appropriateness. Also individual behavior such as interaction patterns involves elements of each. Political actors are constituted both by their interests, by which they evaluate their anticipations of consequences, and by the rules embedded in their identities and political institutions. They calculate consequences and follow rules, and the relation between the two is often subtle (Peters 2005: 31). Thus, role theory might be a potential framework to bring together micro-foundation and institutionalist approaches. Roles combine and provide a vital link between the two logics and can be regarded as a possibility to conceptualize the multiple expectations of political and administrative actors in Europeanized, but also in domestic policies. Until now, it remains largely unclear whether MPs taking different parliamentary roles in the domestic arena also respond differently to the challenges posed by Europeanization (but see Kropp & Ruschke 2010; Kropp et al. 2011, 2012). Our knowledge on the Europeanization of administrative roles is slightly better (e.g. Trondal 2007; Trondal & Veggeland 2003), but still confined to officials' role orientations. Finally, it is widely acknowledged that it is not only the self-understanding of actors but also their mutual perceptions that shape interactions (Scharpf 1997). But we lack information on which roles MPs and bureaucrats assign to each other, whether or not these roles are compatible, and whether and under which conditions mutual

role perceptions help to establish and stabilize trust and interaction routines between the political and administrative elite.

Thus theoretically, this paper argues that role theory can enhance the neo-institutionalist research agenda by combining sociological and rational choice institutionalism. While early role theory in parliamentary research has its roots in structural-functionalism of the Parsonian tradition and the interactional approach of Mead's symbolic interactionism (Biddle 1986), the second wave of role theory was marked by the seminal contribution on Westminster MPs by Searing (1994; for an overview, see Saalfeld & Müller 1997). Criticizing the functional and interactionist usages of role theory, Searing developed the 'motivational' approach. This approach incorporates insights from both sociological (structural and interactional) and rational choice traditions. It integrates these two traditions by recognizing that the roles of politicians are embedded in institutional contexts, while at the same time treating the role players as actors with independent preferences (Searing 1994). According to this distinction, Searing separates position roles from preference roles, the first being "closely tied to, and highly defined by, prominent positions in the institutional structure" (Searing 1991: 1255) while preference roles "allow considerable scope for individual preferences to shape role interpretations" (Searing 1991: 1253).

6.2.1 Position role: membership in government or opposition party

In the following I argue that in the context of legislative-administrative interaction, membership in a government or opposition party can be considered as a position role limiting MPs' strategies of information gathering, i.e. constraining their ability to simply (inter)act on the bases of their preferences. This argument is based on bureaucratic role theory, more precisely the representation role attributed to the Swedish civil servants in the literature (e.g. Premfors & Sundström 2007: 131). According to this, ministerial officials are regarded as acting in line with the classical normative ideal of the Weberian type bureaucrat. In other words, they are seen as being loyal representatives of the government. No matter which parties the current government consists of, bureaucrats will perform their duties, prepare information for the political leadership, and critically supervise the policy process from a legalistic perspective. Taking this perception of bureaucratic roles seriously, parliamentarians with the position role 'opposition MP' – theoretically – will refrain from getting in touch with ministerial officials. In the sense of sociological institutionalism, bureaucrats will simply not be considered as opposition MPs' appropriate interaction partner in policy making. Thus, membership in a governmental party can be expected to be necessary for frequent direct interaction. On the contra-

ry, Sweden is selected as the most likely case for frequent legislative-executive interaction due to its issue-minded, open, and egalitarian bureaucracy. Taking the argument of low functional politicization (Pierre 2004) seriously, membership in a governmental party will not be necessary for frequent interaction. Consequently, I expect this condition to be an INUS condition, i.e. a necessary part of at least one (but not necessarily all) sufficient configuration(s).[53]

6.2.2 Preference roles: strong political and policy expertise

Moreover, I consider MPs' political and policy expertise as preference roles. This is based on the seminal study on the British House of Commons by Searing (1994: 40) who classifies 60 percent of the British backbenchers to adopt the preference role 'specialists'. Specialists "assume that knowledge provides the springboard for influence" (Searing 1994: 39). Looking at the committee structures in working parliaments, I expect this role to be of even greater importance for virtually all MPs in Sweden. Conceptually, one can distinguish between different types of knowledge and expertise that are important for both domestic and EU-related policy making and parliamentary control: "technical knowledge that defines the context of a policy issue and political knowledge of the relative strength of the competing claims and of the consequences of alternative decisions on policy issues" (Truman 1951: 333f.).

Political knowledge means MPs' political expertise. As parliamentarians do not suffer from a lack of information, but from a time-consuming information overload, political expertise is defined as the capacity to filter and prioritize information (Workman et al. 2009; see also Buzogany & Kropp 2013; Webber 1992). I expect political expertise to be an important factor in sufficient combinations of conditions (i.e. an INUS condition) especially for frequent interactions with civil servants in Europeanized policies. This is based on the assumption that MPs with a strong political expertise will try to keep close contact with the dominating actors in European policy making, the public administration. On the contrary, the expectation on the influence of MPs' political expertise in the national arena is rather unclear. As ministerial bureaucrats are not as dominating in domestic policies, strong political expertise might also enable MPs to identify other actors than line officials as most important information supplier.

In contrast, technical knowledge represents policy expertise[54] that is needed in order to influence policies in certain sectors (Blomgren & Rozenberg 2012: 23). Hence, I argue that policy expertise will be an important part in any sufficient combination of conditions explaining

---

[53] An INUS condition is *I*nsufficient itself but *N*ecessary for a configuration that is *U*nnecessary but *S*ufficient for the outcome (see Mackie 1974: 62).

[54] For the specification of "Europeanized policy experts" see Kropp 2010.

MPs' interaction frequency – either in its presence or in its absence. On the one hand, one could argue that a strong role orientation as a policy expert will lead to frequent interactions with civil servants since based on common expertise political-administrative "networks of cooperation" evolve (Benz 2008: 208). On the other hand, a strong self-ascription of policy expertise could make frequent interactions with ministerial officials superfluous since MPs hold the expertise themselves. The same is true for a non-strong, i.e. moderate or weak role orientation as a policy expert. Whether these MPs will interact frequently with officials in order to compensate for the missing policy expertise or simply emphasize their role as generalists who do not need specialized policy expertise is an empirical question. Theoretically, assumptions on the direction of influence on the outcome are unclear.

Next to MPs' perception of their own expert role, i.e. their role orientations, I include MPs' role perception on the policy expertise of ministerial bureaucrats in domestic and Europeanized policies as another explanatory factor into the analysis (see for an overview on the role concept in administrative research, see Kropp & Ruschke 2010; for an approach for Swedish civil servants, see Ehn 1998). Here, the theoretical direction is rather easy to formulate. Only if MPs consider the bureaucrats as strong policy experts, they will frequently interact with them on those issues – both on domestic and on Europeanized issues. Hence, the perception of officials as strong policy experts is assumed to be necessary for the outcome.

To conclude, in this paper role theory is used as a way to assess parliamentary oversight in domestic and Europeanized policies at the micro level. I adopt Searing's (1991, 1994) distinction of institutionally shaped position roles that define MPs' appropriate behavior and preference roles enabling MPs to act on the basis of individual preferences. Next to role orientations I include MPs' role perception of the expertise of their counterparts in the ministries. A summary of my theoretical expectations is given in table 6.1.

Table 6.1 Theoretical expectations about the influence of single conditions on the presence of the outcome

Role	Expected influence on the outcome in DOMESTIC policies	Expected influence on the outcome in EUROPEANIZED policies
Position role: membership in government party (insider)	**Presence** will be necessary or INUS condition	**Presence** will be INUS condition
Preference role: strong political expertise (experience)	**Unclear** presence *or* absence might be INUS conditions	**Presence** will be INUS condition
Preference role: strong policy expertise (role orientation)	**Unclear** presence *or* absence might be INUS conditions	**Unclear** presence *or* absence might be INUS conditions
Expected Preference role: strong bureaucratic policy expertise (role perception)	**Presence** will be INUS condition	**Presence** will be INUS condition

6.3 Design, data and methods

From the research design, this study is a *most likely* single case study on the Swedish Riksdag. As a working parliament with a powerful committee structure, the Swedish parliament is rated as a strong scrutinizer both in domestic and EU affairs (Bergman 1997, 2000, 2003; Karlas 2012; Raunio 2005). At the same time, Swedish ministries are regarded as egalitarian, issue-minded, and open in terms of freedom of information, i.e. the functional politicization is very low (Pierre 2004). In the language of Lijphart (1999) Sweden can be ranked as a consensus democracy, typically offering cooperative, open, and consensual policy making. Moreover, Swedish civil servants are ranked as very loyal to the government (Premfors & Sundström 2007: 129). These scope conditions make it *most likely* that one will detect intense and frequent legislative-administrative interaction patterns in domestic and Europeanized policies in Sweden.

Empirically, this paper bases on a twofold data generating process with Swedish MPs. On the one hand, the outcome data on MPs' interaction frequency with ministerial officials are assessed by semi-structured interviews with 22 Swedish MPs conducted between 2010 and 2012 (for information on the interviewees, see appendix, table 6.A1). As I am interested in expertise in domestic and Europeanized policies, most of the MPs are not members of Swedish EAC but of standing committees on social insurance and on environment and agricul-

ture.[55] The committees selected differ in their 'degree' of Europeanization. While the committee on environment and agriculture is strongly Europeanized and decisions are largely taken at the European level, the committee on social insurance deals mainly – but not exclusively – with domestic issues. On the other hand, MPs' role orientations and role perceptions are assessed by a survey conducted with the same 22 parliamentarians. In accordance to the motivational role approach (Searing 1991, 1994), the survey seeks to reconstruct roles as they are seen by the actors themselves.

Methodologically, I apply fuzzy set Qualitative Comparative Analysis (fsQCA) to identify the relation between MPs' role perceptions and the frequency of interaction with ministerial officials (Ragin 1987, 2000, 2008; Schneider & Wagemann 2012; software used: Duşa & Thiem 2014; Ragin & Davey 2014). QCA is rooted in set theory and its central aim is to identify necessary and sufficient conditions within a population of cases. Moreover, QCA is an appropriate way to examine social reality's complexity as it assumes set relations to be conjunctural, equifinal, and asymmetric. Conjunctural means that cases are seen as configurations of conditions which are expected to affect the outcome – in contrast to single conditions' net effects. I expect conjunctions of the afore-mentioned four conditions to explain frequent interaction. Moreover, equifinality implies that several mutually non-exclusive configurations of conditions may explain to the same outcome; different cases can be covered by several sufficient conjunctions. I expect different combinations of the four conditions to explain interaction patterns of different MPs. Finally, asymmetry is an important concept for QCA in many ways. One important facet is the following: if a condition is sufficient for the occurrence of an outcome this does not tell us anything about the (non-)occurrence of the outcome if the condition is absent. If I, for instance, identify the 'membership in a government party' as a sufficient condition for frequent interaction in domestic policies I cannot deduce from this information whether opposition MPs do or do not interact with civil servants. The example reveals another aspect of asymmetry: a condition might not just be sufficient for an outcome in its presence in one configuration. At the same time, the absence of the very same condition might be sufficient as well – in combination with other conditions.

QCA is based on fuzzy logic (Ragin 2000) and assumes cases to have a (partial) set membership in a set of conditions. Membership score may vary between full membership (score 1) and full non-membership (score 0). Of special importance is the crossover point (score 0.5)

[55] See on the de/centralization of EU scrutiny, i.e. the importance of sectoral committees in day-to-day Europeanization Kropp et al. 2012.

that marks the border between membership and non-membership (Schneider & Wagemann 2012: 32). As the allocation of set memberships is determined solely by the researcher, the chosen qualitative anchors have to be both transparently discussed and properly substantiated by theoretical considerations (Schneider & Wagemann 2010, Wagemann & Schneider 2015). But except for the condition 'insider' which considers MPs' membership in a governmental party, there is no universal criterion defining the qualitative anchors, i.e. full membership, crossover point and full non-membership. Consequently, the theoretically driven choices have to be motivated, which makes data calibration the centerpiece of QCA. In the following, the outcome and condition sets are presented in more detail and the data calibration is substantiated.

6.4 Data calibration of outcome(s) and conditions

6.4.1 MPs' interactions with ministerial officials – OUTCOME

Legislative-administrative interaction patterns in Sweden have mainly been discussed as a part of the Michigan Comparative Elites Project (see for the first project phase Mellbourn 1979; Anton 1980; Linde 1982; for the second phase Ehn 1998; Wallin et al. 1999; Ehn et al. 2003). While these scholars capture role orientations of civil servants and politicians with semi-structured interviews, all data on interaction patterns are collected via structured questionnaires (Ehn 1998, appendix 1-3). The standardized answers for different interaction partners cover the frequency of contacts ranging from daily, weekly, monthly, and yearly to no contact.[56]

In contrast, in my semi-structured interviews the data on frequency is collected qualitatively, i.e. without pre-existing categories. In addition, I distinguish different types of legislative-administrative interaction according to it's in/formality, immediacy, and initiative. In/formality aims at the important distinction between *ordinary* interaction MPs hardly can avoid (e.g. under committee hearings or in parliamentary commissions) and *direct* interaction, which are defined as informal contacts via telephone, email or in person. Immediacy targets the distinction of *direct* and *indirect* interaction with the latter being defined as the informal collection of ministerial information via other actors. For example, Anton (1980: 97, 132-133, 182) detects a rather low frequency of direct parliamentary-bureaucratic interaction. If there is some legislative-executive interaction, Swedish civil servants primarily are in contact with the

[56] The same approach has been applied by Premfors and Sundström (2007) and, for administrative-political interactions within the Swedish executive, by Larsson (1986, 327-33) and Niemann (2013, 34-35).

Riksdag administration (see also Premfors & Sundström 2007: 169-171). Likewise, many MPs keep close contact to the parliamentary administration in order to gain the ministries' information indirectly (Buche & Fleischer 2013, 2015). Finally, initiative meets the problem that the origin of interaction patterns is not grasped by frequencies; i.e. the question on who initiates information exchange (see for this discussion Wallin et al. 1999: 113). To conclude, the measurement of frequency alone is a rather weak indicator for the intensity of interaction processes (but see Mellbourn 1979: 14). With frequency, in/formality, immediacy, and initiative the complex concept of interaction intensity is grasped more in detail. This allows me to differentiate interaction intensity types. In the following I will focus on the frequency of the highest intensity type, i.e. *direct* MPs' interactions with ministerial bureaucrats. In other words, I do not aim at explaining every legislative-administrative contact but focus on informal, immediate interaction initiated by MPs. The outcome sets are calibrated according to the data on 'frequent direct interaction with civil servants' in domestic and Europeanized policies.

However, the interview data reveal that MPs' direct interaction frequency differs remarkably. It ranges from no contact at all, which means that policy information is never obtained by direct interaction with ministerial officials, to frequent and standardized direct interactions. This category of direct interaction indicates that MPs join a minister's working meetings with his administrative staff. The calibration of these two extreme types of direct (non-)interaction is rather easy. All MPs having no direct contact with the ministries are full non-members in the set of 'frequent direct interaction with civil servants in domestic (or Europeanized) policies'. Similarly, MPs with frequent and standardized direct interaction are determined full members of this set. Empirically, opposition MPs do never join ministerial meetings and thus can never achieve full set membership by default. This holds true for both domestic and EU-related policy making. For opposition MPs, the most intense type of direct interaction is frequent, but non-standardized. This means that MPs frequently interact directly with line officials via email or telephone in order to clarify ongoing issues. Consequently, this interaction pattern is regarded as being more 'in' than 'out' of the outcome sets, which is indicated by the membership score of 0.75.[57] The remaining category is non-frequent, spontaneous interaction when more profound information is needed. Although this is not frequent interaction it differs from no interaction as those MPs contact civil servants directly if an important or hardly understandable issue has been transferred to the Riksdag. These non-frequent interaction pat-

[57] As robustness checks reveal, the membership score of 0.75 could be replaced by any other value indicating a partial membership (such as 0.66) without changes in the results (see tables 6.A4 and 6.A5). The same is true for 0.25 as a value indicating partial non-membership. The qualitative calibration stresses the conceptual meaning of the value rather than the number itself.

terns are regarded to be more 'out' than 'in' the outcome set, i.e. are assigned a value of 0.25. The assignment of outcome set memberships to individual cases is displayed and exemplified in table 6.A2.

In domestic policies, the data suggest a rather clear difference between MPs in government and opposition. While nine out of ten MPs in the two non-frequent interaction categories are members of opposition parties, eight out of twelve MPs having more frequent contact with civil servants on domestic policies are MPs from government parties. In this regard, interaction patterns in EU-related policies differ from the domestic ones. Of the thirteen MPs that affirm to stay in frequent contact with civil servants frequently, less than the half are members of governmental parties. However, in the following I examine explanations considering the membership in government or opposition parties, political experience, and individual role orientations and role perceptions on policy expertise.

6.4.2 Membership in a governmental party – INSIDER

As outlined above, the membership in either a government or opposition party can be regarded as a position role limiting MPs' strategies of information gathering. The data on membership in a governmental party is calibrated into the set 'insider'. A high membership score in this set implies strong links to the political executive branch; i.e. ministers. Empirically, three groups of MPs can be differentiated. First, there are nine members of those four bourgeois parties that formed the government coalition (so called Alliance for Sweden, *allians för Sverige*) under the time of data collection, i.e. Moderates, Liberals, Christian Democrats, and members of the Center Party. These MPs are regarded as full members in the conditions set *insider*.[58] Second, there are seven members of opposition parties that have never been part of the Swedish government, i.e. members from the Left Party and the Green Party. They are regarded as full non-members of the set 'insider'.[59] In between the Alliance and the pure outsiders there are six members from the Social Democrats. Being in opposition at the time of data collection, this party governed Sweden for decades, at last from 1994 to 2006. All of the six social democratic MPs in the sample have been in parliament at least since 2002. Thus, as

[58] Of course, different set memberships could be assigned to members of the several governmental parties too, since one could argue that the size of the party group has an influence on their 'weight' in government. But since I am especially interested in differences between opposition and government MPs, I refrain from further categorization.

[59] Again, as especially the Green Party has had an enormous influence on policy making under the social democrats' minority government until 2006, one could argue for a higher value for them in contrast to the Left Party. But as their position role has been members of an opposition party, I keep the clear separation of government and opposition.

they are members of a former government party, I expect them to be clearly more insider than non-insider. This means for the calibration that I allocated the fuzzy value 0.75 toward them displaying a partial set membership.

6.4.3 Strong political expertise – (EU_)EXPERIENCE

For the calibration of the set 'strong political expertise', I consider parliamentary experience to be an appropriate indicator. If a MP has been in parliament for a long time, then he or she will very likely have developed strategies for reducing the complexity of information. For deciding the crossover point between membership and non-membership in this set, I rely on interview data. For many parliamentarians the first legislative term is considered to be a time for learning parliamentary rules and norms. In the second term, these rules are already internalized (in almost the same manner expressed by expert interviews SEGO[60] 8, 11 and SEOP 9, 13). Thus, in *domestic* policies, MPs with parliamentary experience of more than one legislative term have a partial membership in the condition set 'strong political expertise'. Full membership in this set is assigned to those MPs who have been in parliament for at least three legislative terms, because above three legislative terms I do not expect a further increase of political expertise in domestic policy making.[61]

In *EU-related* policy making, set membership indicating a 'strong political expertise' is assigned differently. Here, only those parliamentary experiences are considered that are linked to EU-affairs. Having no parliamentary experience at all in Europeanized policies means a full non-membership in the condition set. A partial non-membership is assigned to those MPs who have worked in an EU-related committee (e.g. committee on environment and agriculture, or EU-affairs committee) for up to one legislative term. Consequently, to be more in than out of the condition set, MPs need to have worked in an EU-related committee for more than one legislative term. Full membership is allocated to MPs who have worked in one of these committees for more than two legislative terms. Furthermore, former members of the European parliament are considered full members in this set as well (see table 6.A1).

---

[60] For anonymity reasons, the interviewees were coded the following way: SE: Member of the Swedish Parliament, OP: member of a current opposition party, GO: member of a current government party, 1-22 is the consecutive number according to the chronological order of the interviews.

[61] For the principle of "unnecessary variance" in set calibration see Ragin (2008, 77ff.) and Schneider and Wagemann (2012, 29-30). Another argument to assign the full membership in 'strong political expertise' to MPs with more than three legislative terms is the following: the Swedish greens' party rules do not allow MPs to stay in parliament for more than three legislative terms (expert interview SEOP 17). Thus, a higher threshold would have excluded members of the green party from full 'strong political expertise' by default.

6.4.4 Strong policy expertise – (EU_)ORIENTATION

The data for calibration of the set 'strong policy expertise' stem from a standardized questionnaire which all interviewees filled in directly after the interview. I asked for MPs' role orientations as policy experts using the item: "How much do you perceive yourself as an expert in [environmental/social] policies?" for domestic policies and "How much do you perceive yourself as an expert in Europeanized policies?" for Europeanized policies. The possible answers range on a seven-point Likert-scale from "very strong (7)" to "not at all (1)". Data calibration was conducted as follows: In both condition sets – "strong policy expertise" in domestic and Europeanized policy – full membership was only assigned to those MPs, who perceive themselves as a policy expert "very strong(ly)", i.e. value 7. Partial membership was allocated to those who rated their policy expertise with the values 5 and 6, meaning "rather strong" or "strong". Partial non-membership was allocated to MPs checking "undecided" (value 4). Finally, full non-membership was assigned to MPs not rating themselves as policy experts (smaller or equal to value 3). The data indicate that the overall role orientation as a policy expert is much stronger in domestic policies compared to Europeanized policies (see table 6.A1).

6.4.5 Strong policy expertise of civil servants – PERCEPTION

For the calibration of the set 'strong policy expertise of civil servants' I asked MPs for their role perception of officials as policy experts. I used the item: "How much do you perceive ministerial officials as policy experts in their policies?" for both domestic and Europeanized policies. Although the same seven-point Likert-scale from "very strong (7)" to "not at all (1)" is applied, the calibration rules have been changed in comparison to MPs' orientation of policy expertise. The reason for that are the expected high values for administrative expertise and the will to grasp interesting variance (Ragin 2008: 77ff.). Thus, the qualitative anchor for the crossover point has been adapted to the higher standard. Full membership in the condition set "strong policy expertise of civil servants" was only assigned to those MPs who perceive bureaucrats as policy experts "very strong(ly)", i.e. chose value 7. Partial set membership was allocated to interviewees who rated their policy expertise with the value 6, i.e. "strong". Partial non-membership was assigned to MPs rating administrations' policy expertise as "rather strong" (value 5). Finally, full non-membership was assigned to MPs who are at least "undecided" (value 4 or less) whether civil servants are experts or not (see table 6.A1).

The complete fuzzy-set data matrix covering all calibrated condition and outcome set memberships of all 22 MPs can be found in the appendix (see table 6.A3).

6.5 Analysis

According to the standards of good practice in QCA (Schneider & Wagemann 2010, 2012: 221) I check for necessity prior to sufficiency. Thus first, I analyze necessary conditions for the outcomes 'frequent direct interaction with civil servants in *domestic* policies' and 'frequent direct interaction with civil servants in *Europeanized* policies'. Subsequently, I assess sufficient conditions for interactions on both arenas. What needs to be mentioned is that although possibly interesting I do not asses the negation of the outcomes. More precisely, due to calibration asymmetry I do not analyze necessary and sufficient conditions for MPs' '*non-frequent* direct interactions with civil servants'.[62]

6.5.1 Necessary conditions for MPs' interaction with civil servants

A condition is regarded necessary if the outcome cannot occur without the condition being present. While perfect necessary conditions are almost never to find in social reality – especially when working with micro-level data –, slight inconsistencies can be allowed if justified in a proper way. According to Schneider and Wagemann (2012: 143) the consistency threshold for necessary conditions should not be lower than 0.9 to consider a condition necessary. In addition to this threshold I check for contradictory cases in a XY-Plot to evaluate whether a condition can be regarded as necessary or not.

The investigation of necessary conditions for the outcome 'frequent direct interaction in *domestic* policies' reveals that the presence of the condition 'insider' is consistently necessary for the outcome (see table 6.2 on the left and for the XY-plot appendix, figure 6.A1). As expected above, in domestic policies frequent interaction between MPs and ministerial officials requires MPs to be members of a – current or former – government party. In contrast to my expectancy, the perception of civil servants as policy experts is no necessary condition for frequent direct interaction, neither in domestic nor Europeanized policies. Moreover, the investigation for the outcome "frequent direct interaction in *Europeanized* policies" reveals no necessary condition at all (see table 6.2, right hand side). Interaction patterns in Europeanized

[62] To refrain from the study of the non-outcome needs to be mentioned since one cannot simply deduce the sufficiency statements for the non-occurrence of the outcome from the sufficient conditions for its occurrence. The reason for me to do this – next to limited space – is calibration asymmetry. For instance, the outcome set 'frequent direct interaction' is calibrated in a way that its negation covers 'no', 'seldom', and an 'average number of' interaction. Thus, the explanatory power of the analysis of necessary and sufficient configurations for 'non-frequent direct interaction' would be conceptually rather limited. A proper an analysis of the opposite conceptual pole of frequent interaction such as '*no* interaction' would require a completely different calibration of the outcome. The same is true for the different instances of 'non-strong experience' and 'non-strong political/policy expertise' that are or interest for 'no interaction'. In short, the analysis of the non-occurrence of frequent direct interaction is either conceptually unconvincing in its results or requires an entirely new analysis.

policies seem to be more diverse than in domestic policies, as they former can be achieved by parliamentarians without necessarily having a high membership in one of the conditions under scrutiny.

Table 6.2 Analysis of Necessity

Condition	Necessity for 'frequent interaction on *DOMESTIC* policies'		Necessity for 'frequent interaction on *EUROPEANIZED* policies'	
	Consistency[63]	Coverage	Consistency	Coverage
Insider	**0.93**	**0.78**	0.79	0.70
~Insider	0.18	0.24	0.31	0.44
Experience	0.73	0.69	-	-
~Experience	0.40	0.45	-	-
EU_Experience	-	-	0.63	0.79
~EU_Experience	-	-	0.50	0.48
Role Orientation	0.73	0.57	-	-
~Role Orientation	0.47	0.70	-	-
EU_Role Orientation	-	-	0.48	0.72
~EU_Role Orientation	-	-	0.65	0.55
Role Perception	0.82	0.71	0.79	0.73
~Role Perception	0.44	0.56	0.42	0.56

6.5.2 Sufficient conditions for MPs' interaction with civil servants in DOMESTIC policies

The truth table can be regarded as the most important tool of data structuring in QCA for at least three reasons: first, it helps to describe differences and commonalities among cases. For instance, the truth table for the outcome "frequent interaction in *domestic* policies" indicates that seven cases share the same configuration (table 6.3, row 1). This group consists of four opposition MPs and three government MPs all being more members than non-members in all of the four conditions. Second, the truth table enables to detect contradictions among cases in the same truth table row. Although cases share the same configuration of conditions not all of them might show the (non-)outcome. Sticking to the seven MPs in row 1, not all of them con-

[63] The consistency value measures how consistent the condition is necessary. The very slight inconsistency here is caused by cases that neither are members in the condition nor the outcome set – thus, they are irrelevant cases for the statement that 'insider' is necessary for the outcome (Schneider & Wagemann 2012, 71). The coverage value of consistent necessary conditions depicts the triviality of such a condition; the lower the value, the more trivial is the necessity of the condition (Schneider/ & Wagemann 2012, 233-234). However, 'insider' can be regarded as non-trivial for the occurrence of the outcome (see figure 6.A1).

firm the statement that the presence of the four conditions in conjunction leads to 'frequent interaction in *domestic* policies' since the consistency is smaller than 1. In contrast, in truth table rows four, nine, and ten the statement of sufficiency for the outcome is perfectly consistent. Third, the truth table uncovers limited empirical diversity, i.e. configurations of conditions that are not covered by any cases, so called logical remainders. Here, six out of the 16 rows do not contain empirical information (table 6.3, row 11-16).

Table 6.3 Truth Table for MPs' frequent interaction with civil servants in DOMESTIC policies

Row	Insider	Experience	Role orientation	Role perception	Cases in configuration	Consistency of sufficiency	PRI (proportional reduction in inconsistency)
1	1	1	1	1	SEOP 6, 9, 14, 22 SEGO 1, 11, 16	0.89	0.84
2	0	0	1	1	SEOP 13, 15, 17	0.20	0.00
3	0	0	1	0	SEOP 4, 18	0.30	0.00
4	1	0	1	1	SEGO 11, 22	1.00	1.00
5	1	1	0	1	SEOP 10, SEGO 20	0.94	0.89
6	1	1	1	0	SEOP 7, SEGO 3	0.75	0.43
7	0	0	0	0	SEOP 12	0.57	0.00
8	0	0	0	1	SEOP 19	0.50	0.00
9	1	0	0	0	SEGO 5	1.00	1.00
10	1	0	1	0	SEGO 2	1.00	1.00
11	0	1	0	0	-	1.00	-
12	0	1	0	1	-	0.86	0.00
13	0	1	1	0	-	0.83	0.00
14	0	1	1	1	-	0.80	0.00
15	1	0	0	1	-	1.00	1.00
16	1	1	0	0	-	1.00	1.00

Note: The columns in light gray boxes indicate the results are judged as consistently sufficient for the outcome. The columns in dark grey boxes indicate logical remainders that are included in the minimization.

In addition to consistent truth table rows covering empirical information, logical remainders can be taken into the minimization process as well.[64] As 'insider' is a necessary condition for the outcome, just logical remainders with this condition present possibly can be included in

[64] While this may decrease the level of complexity, it also bases the results on logical possible configurations without empirical evidence. Thus, the choice of logical remainders has to be justified properly.

the minimization, i.e. row 15 and 16.[65] Applying counterfactual reasoning, I regard just row 15 as sufficient for the outcome.[66]

Table 6.4 Solution of sufficient conditions for the outcome "frequent direct interaction in domestic policies"

Parameters of Fit	INSIDER* ROLE_PERCEPTION	**OR**	INSIDER*~EXPERIENCE
Consistency	0.92		1
Raw Coverage	0.76		0.33
Unique Coverage	0.53		0.11
Cases	SEOP 6, 9, 10, 14, 22 SEGO 1, 8, 11, 16, 20, 21		SEGO 2, 5, 8, 21
Solution consistency	0.93		
Solution coverage	0.87		

Note: For the robustness-check of the analysis' results with alternative calibration see table 6.A4.

The solution covers two equifinal paths explaining MPs' frequent direct interaction on domestic policies (see table 6.4). On the one hand, the results indicate that the conjunction of being an insider AND having a strong role perception of civil servants as policy experts, leads to a high interaction frequency in domestic policies. On the other hand, being an insider without being strongly experienced in domestic policy making leads to the presence of the outcome as well. Two MPs' interaction frequency can be explained by both paths. The solution is both highly consistent and covers all MPs interacting directly.[67]

6.5.3 Sufficient conditions for MPs' interaction with civil servants in EURO-PEANIZED policies

For the analysis of sufficient conditions for the outcome "frequent direct interaction in *Europeanized* policies" six truth table rows are regarded to be sufficient (see table 6.5). The diver-

---

[65] See for "incoherent counterfactuals" and the so called "enhanced standard analysis" Schneider and Wagemann (2012, 201-211).

[66] The counterfactual thought experiment is the following: First, I know that the conjunction in row 9 (presence of 'insider' combined with absence of all other conditions) is sufficient for 'frequent interaction'. Second, I consider the presence of a strong 'role expectation' as beneficial for 'frequent interaction'. Third, the conjunction in row 15 combines the presence of the conditions 'insider' with the presence of 'role expectation' (and the absence of the other two conditions). Thus, if already the presence of 'insider' in row 9 is sufficient for the outcome, then the conjunction with the additional presence of 'role expectation' with all other condition being constant can also be regarded as sufficient (see for the idea of easy counterfactuals Schneider & Wagemann 2012, 198-200). In contrast, the additional presence of 'experience' (row 16) is not expected to contribute to 'frequent interaction' in domestic policies (see theoretical foundation, table 6.1).

[67] As one can see from the XY plot displaying the final solution (see figure 6.A2), one MP contradicts the statement of sufficiency (SEOP 6). Thus, a short case study is conducted in the appendix (see note after figure 6.A2) in order to shed light on the deviant case. This allows me to explain this contradiction with mistrust and prejudices due to bad experience in former interaction. Thus, being able to explain the inconsistency I interpret the solution to be sufficient for the outcome.

sity of the 22 cases is remarkable since just three out of the 16 truth table rows lack empirical evidence. In contrast to the first analysis, none of these three logical remainder rows is included into the minimization process.[68]

Table 6.5 Truth Table for MPs' frequent interaction with civil servants in EUROPEANIZED policies

Row	Insider	Experience	Role orientation	Role perception	Cases in configuration	Consistency of sufficiency	PRI (proportional reduction in inconsistency)
1	1	0	0	1	SEOP 6, 9, 10, SEGO 1, 8, 11	0.82	0.75
2	1	1	1	1	SEOP 14 SEGO 16, 21	1.00	1.00
3	0	0	0	0	SEOP 4, 12	0.43	0.00
4	1	0	0	0	SEGO 2, 5	0.62	0.29
5	0	0	0	1	SEOP 13	0.70	0.40
6	0	0	1	1	SEOP 17	0.25	0.00
7	0	1	0	1	SEOP 15	0.55	0.00
8	0	1	1	0	SEOP 18	0.88	0.67
9	0	1	1	1	SEOP 19	0.86	0.67
10	1	0	1	0	SEGO 3	0.57	0.25
11	1	0	1	1	SEOP 22	0.57	0.25
12	1	1	0	1	SEGO 20	0.91	0.86
13	1	1	1	0	SEOP 7	0.88	0.80
14	0	0	1	0	-	0.67	0.00
15	0	1	0	0	-	0.78	0.00
16	1	1	0	0	-	0.89	0.75

Note: The columns in light gray boxes indicate the results are judged as consistently sufficient for the outcome.

Again, the solution consists of two equifinal sufficient configurations (see table 6.6). The results indicate that the conjunction of having a strong role orientation as a Europeanized policy expert AND being strongly experienced in European policy making, i.e. having both political and policy EU-expertise, leads to a high frequency of direct interaction in Europeanized policies. That means no matter whether these MPs are insiders or not and whether they regard civil servants as policy experts or not, they frequently interact with civil servants in Europeanized policy issues because of their own expertise. In addition, the lack of Europeanized policy expertise is sufficient as well, but just in combination with a strong role perception of civil

[68] This holds true both for own counterfactual reasoning and the intermediate solution implemented in the software (directional expectations according to theoretical foundation, see table 6.1).

servants as policy experts AND with being an insider. The solution is both highly consistent and covers all but one MP showing the outcome.[69]

Table 6.6 Solution of sufficient conditions for the outcome "frequent direct interaction in Europeanized policies"

Parameters of Fit	ROLE_ORIENTATION * EXPERIENCE	**OR**	~ROLE_ORIENTATION * INSIDER * ROLE_PERCEPTION
Consistency	0.91		0.85
Raw Coverage	0.42		0.48
Unique Coverage	0.33		0.40
Cases	SEOP 7, 14, 18, 19 SEGO 16, 21		SEOP 6, 9, 10 SEGO 1, 8, 11, 20
Solution consistency	0.87		
Solution coverage	0.81		

Note: For the robustness-check of the analysis' results with alternative calibration see table 6.A5.

6.6 Discussion and conclusion

The aim of this paper was to examine the Europeanization of legislative-executive relations at the micro-level. It asked under which conditions Swedish parliamentarians directly interact with civil servants from the ministries in Europeanized policies and contrasted it with interaction patterns in domestic policies. Theoretically, I state that parliamentary position roles and preference roles can be used as reliable predictors for individual behavior; here direct interactions. I expect that next to the classical government-opposition divide and parliamentary experience, the ascription of policy expertise – both MPs' role orientation and the role perceptions towards ministerial officials – are important explanatory factor for legislative-executive interaction patterns. The analysis conducted on the empirical basis of interview and survey data of 22 Swedish MPs reveals that interaction patterns differ remarkably between domestic and Europeanized policies.

In domestic policies, to be an insider, i.e. a member of one of the four current or the former government party groups, was identified a necessary condition for frequent interaction with bureaucrats. On the one hand, this is in line with the theoretical expectation regarding the classical Weberian bureaucratic representation role in Sweden. According to this officials are seen as loyal representatives of the government which brings opposition MPs to look for in-

[69] As one can see from the XY plot displaying the final solution (see figure 6.A3), the same MP contradicts the statement of sufficiency (SEOP 6) while one MP's direct interaction is not covered by the solution (SEOP 13).

formation elsewhere. On the other hand, this finding is surprising since Sweden was chosen as the most likely case for frequent direct interaction due to its lowly politicized bureaucracy. In this reading, officials do not care about party politics but act upon normative values such as bureaucratic credibility and integrity, which "sit deep in the Swedish administrative tradition" (Pierre 2004: 52). Interestingly, membership in a (also former) government party turns out to be necessary in domestic policies only. In Europeanized policies being an insider is no necessity for frequent direct interaction. Thus, one could argue that the "new dualism" (Eberbach-Born & Kropp 2013: 15) between government and opposition does not account so much for Europeanized policies as it does for domestic ones.

The analysis of sufficiency depicts two configurations that lead to frequent direct interaction each on domestic and Europeanized policies. In domestic policies, being an insider without strong experience, i.e. without much political expertise, is sufficient for frequent direct interaction with civil servants. This is fully in line with the theoretical assumptions and can be regarded as the standard behavior of newly elected MPs of governmental parties to keep close contact to the ministries in order to absorb or develop policy expertise. Empirically, the interaction patterns of four rather inexperienced MPs are covered by this configuration. Next to this, (former) government MPs, i.e. insiders, with a strong role perception of civil servants as policy experts also directly interact with officials on domestic policies. Again, this is in line with the theoretical expectations and to a certain extent describes the ideal type collaboration of more experienced MPs from governmental parties and ministries. Regardless of their own role orientation as policy experts, these MPs contact bureaucrats they feel associated with and they expect to be policy experts. This configuration covers eleven MPs, five of which are members of the former governmental party.

This second sufficient configuration explaining domestic interaction patterns applies almost equally for Europeanized policies as well. But not all insiders who consider bureaucrats as policy experts interact directly with them. Instead, in Europeanized policies only those MPs who additionally do not ascribe strong policy expertise to themselves seek bureaucratic assistance in policy making. As expected theoretically and in contrast to interaction patterns in domestic policies, MPs' role orientation is a necessary part of the explanatory configuration, i.e. an INUS condition.

The second sufficient configuration for frequent direct interaction in Europeanized policy also includes MPs' role orientation. But in contrast to the first path, not the absence of policy expertise but its presence is part of the sufficient configuration. If combined with strong experi-

ence in Europeanized policy making, a strong role orientation as Europeanized policy expert consistently explains direct interaction. Explaining more than 40 percent of the outcome set, this explanation is far from superficial. For these MPs, membership in a (former) government party is no necessity; political and policy expertise alone account for frequent direct interaction which both fits the theoretical assumptions and the expectations on Sweden as the most likely case.

To conclude, the empirical results suggest that some individual interaction patterns persist and others differ between domestic and Europeanized policy issues. To start with similarities, in both arenas insiders with a strong expectation of bureaucratic policy expertise interact directly. Although in Europeanized policies this is limited to MPs not perceiving themselves as policy experts, the basic pattern is the same: insiders have the opportunity to directly get in touch with the civil servants just by calling an official they expect to be helpful due to bureaucratic policy expertise. The differences of interaction patterns are twofold. While in domestic policies newly elected government MPs keep direct contact frequently, there is no comparable sufficient configuration of non-experienced insiders in Europeanized policies. Quite on the contrary, experience in combination with policy expertise is sufficient for frequent direct interaction on EU issues – regardless the position role as a member of government or opposition party. Overall, the analysis shows that taking roles into account is a valuable approach for examining legislative-executive relations at the micro level. In Sweden, the government-opposition divide does not explain MPs' interaction patterns on Europeanized policies. Instead, both role orientations and role expectations turn out to be important explanatory conditions.

Appendix

Table 6.A1 Raw data matrix on conditions (information on interviewees)

MPs	Party affiliation	Committee affiliation	Parliamentary Experience		Role orientation as policy expert		Role Perception civil servants as policy experts
			Legislative terms completed	Leg. Terms in Europeanized policies	Domestic policies	European policies	
SEGO 01	Center Party	Environment	2	1	Strong	Rather weak	Strong
SEGO 02	Christian Dem.	Environment	1	<1	Strong	Rather weak	Rather strong
SEGO 03	Liberals	Social Insurance	4	0	Very strong	Strong	Rather strong
SEOP 04	Green Party	Social Insurance	1	0	Strong	Very weak	Rather weak
SEGO 05	Center Party	Social Insurance	1	0	Undecided	Undecided	Rather strong
SEOP 06	Social Dem.	Social Insurance	4	<1	Strong	Rather weak	Very strong
SEOP 07	Social Dem.	Environment	2	2	Strong	Strong	Weak
SEGO 08	Moderates	Environment	1	<1	Rather strong	Weak	Strong
SEOP 09	Social Dem.	Social Insurance	4	0	Rather strong	Rather weak	Strong
SEOP 10	Social Dem.	Environment	4	1	Rather weak	Very weak	Strong
SEGO 11	Moderates	Social Insurance	3	0	Strong	Undecided	Very strong
SEOP 12	Left Party	Environment/ Social insurance	1	1	Weak	Rather weak	Rather strong
SEOP 13	Green Party	Environment	0	<1	Rather strong	Rather weak	Strong
SEOP 14	Social Dem.	Environment	3	>2	Rather strong	Strong	Strong
SEOP 15	Left Party	Environment	1	>1	Rather strong	Rather weak	Strong
SEGO 16	Moderates	Environment	2	>2	Rather strong	Strong	Strong
SEOP 17	Green Party	Environment	0	<1	Very strong	Rather strong	Strong
SEOP 18	Left Party	Environment	0	>1 (EP)	Very strong	Very strong	Rather strong
SEOP 19	Left Party	EU Affairs	0	>2 (EP)	Rather weak	Strong	Strong
SEGO 20	Liberals	EU Affairs	5	>2	Undecided	Undecided	Strong
SEGO 21	Moderates	Social Insurance	1	<2	Very strong	Very strong	Strong
SEOP 22	Social Dem.	Social Insurance	4	1	Strong	Strong	Strong

Table 6.A2 Raw data matrix on outcomes (information on frequent direct interaction)

MPs	Membership score in outcomes	Justification with example quote
SEGO 01	Dom: 1	"I have very much contact to the ministry on environment, both with political and non-political officials. They are very helpful if I need information or any documents."
	EU: 1	"I get most information on what happens at the EU level from the government office."
SEGO 02	Dom: 1	"I'm in permanent contact with the governmental officials, both with political and non-political. The latter are employees of whatever government in place."
	EU: 0.25	"In European environmental issues I can also contact our bureaucrats from the government office. […] I sometimes do."
SEGO 03	Dom: 0.25	"Yes, there is some informal contact. I to call them or maybe write an email. […] But most of the time I go to the state secretary or the minister."
	EU: 0.25	
SEOP 04	Dom: 0	"It is very difficult to get information from the ministries. There is just the official way with their answers to my interpellations […] I could call them but I would not get any answer."
	EU: 0	
SEGO 05	Dom: 1	I have a lot of contact to government, ministers and the bureaucrats. My best information source is the ministry and it's personnel.
	EU: 0	"Yes, I can contact them [on EU-issues too]. But let me think whether I did this already… No, I have never asked them concrete questions on this area."
SEOP 06	Dom: 0	"No, they do not have nice bureaucrats. I never call them, never. No, not the bureaucrats from the alliance, no. I would never do that, no."
	EU: 0	
SEOP 07	Dom: 0.25	"Yes, I can do so [to call officials on domestic policies]. And yes, it happens also."
	EU: 0.75	"When we speak about EU fishing policy the government established an information group. One from every party gets information directly from the ministry."
SEGO 08	Dom: 0.75	"I have rather frequently issue-driven contact with non-political officials."
	EU: 1	"We governmental MPs have some direct, more formalized meetings with officials where we discuss the upcoming issues on EU-policies."
SEOP 09	Dom: 0.75	"I have the right to request information, as long it is classified as working paper. The rules on freedom of information are very precise. [...] Thus, it is quite easy to contact and to get along with ministerial officials."
	EU: 0.75	
SEOP 10	Dom: 0.75	"You develop a network of contacts under government time. And of course now it is different from our government time. But they are non-political officials and still answer directly"
	EU: 0.75	
SEGO 11	Dom: 1	"We have consultations with the minister for social insurance and his bureaucrats every second week. That includes a very open access to information from both the ministry and the government office."
	EU: 0.75	"I mainly work on social insurance and the European dimension of this subject is reduced to comprehensive solutions that enable people to commute between the member states."

SEOP 12	Dom: 0.25	"We are very open here. If there is something I do not understand or know I just call the ministry and ask for an expert on that issue and they connect me. […] Yes, I also have done this already."
	EU: 0.25	
SEOP 13	Dom: 0.25	"First I ask the committee secretariat on that. But if need something very fast and they are not around I contact the ministerial officials myself. "
	EU: 0.75	"In EU-affairs officials are very service oriented. This is because the committee can call the minister and state secretary to explain things in parliament. I think it is the officials' obligation to give me the information and to prevent the minister from permanent invitations to our committee."
SEOP 14	Dom: 0.75	"Even if we are opposition now, one can simply call and test the ministries […] they can decide whether they answer or not. So I call and ask in the ministries. [...] And it is not the case that they say »No, we cannot answer your question« but they tell me what they know."
	EU: 0.75	
SEOP 15	Dom: 0.25	"Yes, you can do this [asking ministerial officials for information]. I'm not really used to do this that often."
	EU: 0	"Last legislative term we had informal contact on EU fishing policy with the ministries. But that is over now."
SEGO 16	Dom: 0.75	"I have a lot of contact with both [political and non-political officials]. It depends on the question I have."
	EU: 1	"We [governmental MPs from the committee on environment] meet the minister on environment and his officials every second week and talk about current issues. There we have an very interactive working style and exchange of information."
SEOP 17	Dom: 0	"No, no, no! The ministry belongs to the government. They are enemy troops."
	EU: 0	
SEOP 18	Dom: 0	"I know it is not easy for bureaucrats; they shall be loyal toward the ministry, more precisely the minister. But we have freedom of information here in Sweden. And I think that those officials lack expertise in EU-affairs. That is why I stay in close contact with them."
	EU: 0.75	
SEOP 19	Dom: 0	"We can contact them and they are mostly very open. I'm used to do so. Both when I was sitting in the European Parliament and now in the Riksdag I call(ed) an official from the ministry of environment regarding European regulation proposals and get very sound information."
	EU: 0.75	
SEGO 20	Dom: 1	"Due to my position I have much contact with the Prime Minister and the state secretaries and their staff on all issues."
	EU: 1	"My committee is the central hub for EU policy. That is why we work very close, especially prior to important EU summits."
SEGO 21	Dom: 0.75	"If I need to follow an issues – and normally several non-political bureaucrats are involved in drafting this – then I just call and ask them what is going on."
	EU: 1	"In preparation of council meetings, the minister on migration meets with officials and government MPs. […] We have a very informal information flow between parliament and government office."
SEOP 22	Dom: 0.75	Normally, we social democrats are in opposition. But, of course, there are same old links into the departments left. [..] These are the colleagues you approach.
	EU: 0.25	"I can call them if I want some information on e.g. regulation 1408/71, the free movement. […] But if I need a qualified judgement in EU-affairs, I like to use the parliamentary scientific service."

Table 6.A3 Fuzzy set data matrix

MPs	OUTCOME SETS		CONDITION SETS					
	Interaction frequency		Insider	Experience		Role orientation		Role perception
	Domestic	European		Domestic	European	Domestic	European	
SEGO 01	1	1	1	0.75	0.25	0.75	0	0.75
SEGO 02	1	0.25	1	0.25	0.25	0.75	0	0.25
SEGO 03	0.25	0.25	1	1	0.25	1	0.75	0.25
SEOP 04	0	0	0	0.25	0	1	0	0
SEGO 05	1	0	1	0.25	0	0.25	0.25	0.25
SEOP 06	0	0	0.75	1	0	0.75	0	1
SEOP 07	0.25	0.75	0.75	0.75	0.75	0.75	0.75	0
SEGO 08	0.75	1	1	0.25	0.25	0.75	0	0.75
SEOP 09	0.75	0.75	0.75	1	0	0.75	0	0.75
SEOP 10	0.25	0.75	0.75	1	0.25	0	0	0.75
SEGO 11	1	0.75	1	1	0	0.75	0.25	1
SEOP 12	0.25	0.25	0	0.25	0.25	0	0	0.25
SEOP 13	0.25	0.25	0	0	0.25	0.75	0	0.75
SEOP 14	0.75	0.75	0.75	1	1	0.75	0.75	0.75
SEOP 15	0.25	0	0	0.25	0.75	0.75	0	0.75
SEGO 16	0.75	1	1	0.75	1	0.75	0.75	0.75
SEOP 17	0	0	0	0	0	1	0.75	0.75
SEOP 18	0	0.75	0	0	1	1	1	0.25
SEOP 19	0	0.75	0	0	1	0	0.75	0.75
SEGO 20	0	0	1	1	1	0.25	0.25	0.75
SEGO 21	0.75	1	1	0.25	0.75	1	1	0.75
SEOP 22	0.75	0.25	0.75	1	0.25	0.75	0.75	0.75

Table 6.A4 Robustness-check I: solution of sufficient conditions for the outcome "frequent direct interaction in domestic policies" with alternative calibration

Parameters of Fit	INSIDER* ROLE_PERCEPTION	**OR**	INSIDER*~EXPERIENCE
Consistency	0.92		0.99
Raw Coverage	0.70		0.33
Unique Coverage	0.47		0.09
Cases	SEOP 6, 9, 10, 14, 22 SEGO 1, 8, 11, 16, 20, 21		SEGO 2, 5, 8, 21
Solution consistency	0.93		
Solution coverage	0.79		

Notes: Robustness-check was conducted by the following change of calibration: every set membership score indicating a partial membership in the respective set was changed from 0.75 to 0.66. Likewise, every score indicating non-membership in a condition was changed from 0.25 to 0.33. It turns out that the results are the same with slightly changing consistency and coverage values.

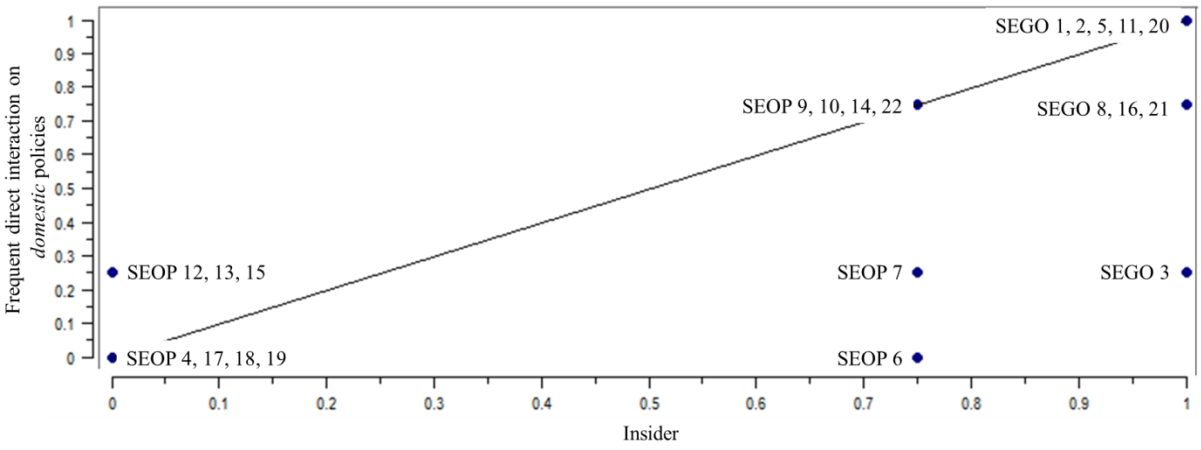
.

Table 6.A5 Robustness-check II: solution of sufficient conditions for the outcome "frequent direct interaction in Europeanized policies" with alternative calibration

Parameters of Fit	ROLE_ORIENTATION * EXPERIENCE	**OR**	~ROLE_ORIENTATION * INSIDER * ROLE_PERCEPTION
Consistency	0.87		0.85
Raw Coverage	0.40		0.49
Unique Coverage	0.28		0.37
Cases	SEOP 7, 14, 18, 19 SEGO 16, 21		SEOP 6, 9, 10 SEGO 1, 8, 11, 20
Solution consistency	0.84		
Solution coverage	0.77		

Notes: Robustness-check was conducted by the following change of calibration: every set membership score indicating a partial membership in the respective set was changed from 0.75 to 0.66. Likewise, every score indicating non-membership in a condition was changed from 0.25 to 0.33. It turns out that the results are the same with slightly changing consistency and coverage values.

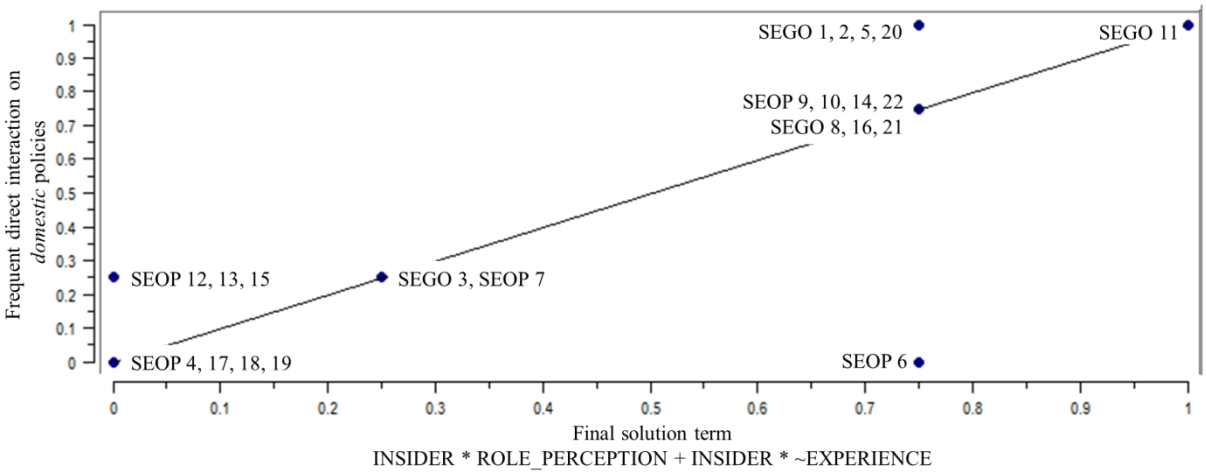
Figure 6.A1 XY plot of necessity set relation 'insider' ←'frequent direct interaction on domestic policies'



Notes: Consistency: 0.93; coverage: 0.78, see table 6.2.

If a condition is consistently necessary, the outcome cannot occur without the necessary conditions being present. Regarding the fuzzy membership scores, the membership in the condition has to be greater or equal to the membership in the outcome set. In the XY-plot this is displayed by the location of cases below the diagonal. As one can see, three cases in the lower left corner are above the diagonal (SEOP 12, 13, 15). Regarding the calculus, they decrease the consistency value. But as these cases are neither members of the outcome nor of the condition set, they can be regarded as irrelevant for the statement that the membership in 'insider' is necessary for the membership in the outcome set. Thus regarding the content, 'insider' is a necessary condition for 'frequent direct interaction on domestic policies'.

Figure 6.A.2 XY plot of final solution for the outcome 'frequent direct interaction on domestic policies'



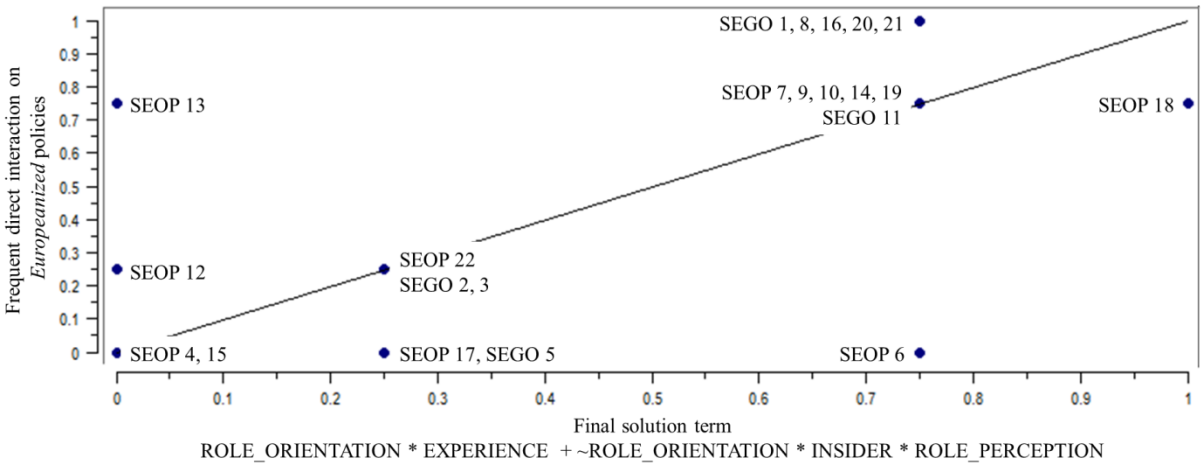
Notes: Solution consistency: 0.93, solution coverage: 0.87, see table 6.4.

If a condition is consistently sufficient, the outcome occurs whenever the sufficient (combination of) conditions are present. Regarding the fuzzy membership scores, the membership in the condition has to be smaller or equal to the membership in the outcome set. In the XY-plot this is displayed by the location of cases above the diagonal. One case is located in the lower right corner (SEOP 6). This case is problematic because it is a member of the otherwise sufficient configuration but not of the outcome. In other words, it is a logical contradiction to the statement of sufficiency. This MP is expected to be a perfect onlier for the outcome since s/he is an insider (social democrat), has both a strong external and self-perception of policy expertise, and is very experienced at the same time.

In order to explain the contradiction I returned to the respective interview. It turned out that this person has made bad experience when asking for information via direct interaction with a civil servant once. The following statement sheds light on the deviant case: "Once I got an EU regulation on environmental issues – and understood nothing. […] Finally, I called an official from the ministry and said: »You have to come here and explain me what is written in this paper since I do not understand a single word.« They came! But I was not allowed to do so. I got my explanation and was told what this text means. But then they said that I'm not allowed to do this and I should get in touch with our own political staff." This brings the MP to the following statement calibrated in non-membership in the outcome sets: "No, they do not have nice bureaucrats. I never call them, never. No, not the bureaucrats from the alliance, no. I would never do that, no" (see table 6.A2).

While from the technical perspective this case is a logical contradiction which brings down the solution consistency, a return to the context allows me to explain this contradiction with mistrust and prejudices. The statement is based on some bad experiences in the last legislative period. Since then, the MP does not use the ministerial bureaucracy as a direct information source anymore. As one can see in the XY plot, this MP is the only case violating the statement of sufficiency. Thus, being able to explain the inconsistency I interpret the solution to be sufficient for the outcome.

Figure 6.A3 XY plot of final solution for the outcome 'frequent direct interaction on Europeanized policies'



Notes: Solution consistency: 0.87, solution coverage: 0.81, see table 6.6.

The XY plot of the final solution shows eleven onlier cases on or above the diagonal in the upper right corner, one case slightly below and one strongly contradictory case. As before, the deviant case is the social democratic MP with the negative personal image of bureaucrats working for the bourgeois government parties (SEOP 6; see notes below table 6.A2). All the other cases in the plot are deemed irrelevant as they neither show the outcome nor the combination of conditions.

7.

Concluding Remarks

(Jonas Buche)

7.1 Assessing the quality of Qualitative Comparative Analysis (QCA)

The starting point of this dissertation was an empirical observation about the number of applications of Qualitative Comparative Analysis. Developed some 30 years ago, QCA in the last decade underwent a major boost of publication rates in journals from various research fields, especially in the social sciences. However, although all of these articles, which have been collected in a bibliographical database by the compasss (2016) network, passed a peer-review process, the rather low methodological quality of some applications of QCA has been used to substantiate claims that the value of QCA as a method *per se* is questionable (e.g. Seawright 2005; Hug 2013; Krogslund et al. 2015). Agreeing with John Gerring (2012: 350) that "the potential utility of a method should be differentiated from its actual employment" the idea for this cumulative dissertation was to *assess the quality of QCA* concerning both the method itself and its empirical employment. Thus, the overarching question was posed:

"What is the concept of quality of QCA and how is it implemented in QCA applications?"

Concerning the empirical applications of QCA, the systematical *evaluation* of the implementation of the quality of QCA will be addressed in the next section. It combines the findings regarding the concept of quality of QCA as presented in chapter 1 and the actual evaluation of the empirical applications of QCA in chapters 2-4. With reference to the method itself, the task of the *improvement* of the quality of QCA will be approached in section 7.3. The improvement bases on a weakness in the connection of mathematical and set-theoretic foundation of QCA as discussed in chapter 5. Finally, the extensive demands and requisites for a high quality QCA are brought together in the concluding section 7.4. In consideration of the best practice research article, the *application* of QCA is finally examined.

7.2 Evaluation

The evaluation of the quality of QCA has been carried out in three articles jointly covering 139 applications of QCA in the fields of sociology, business research, and political science. However, the proceeding of the evaluations has been developed in chapter 1 that systematically exposes the conceptual framework of the dissertation. The concept starts from five research steps that need to be considered by every QCA application. Subsequently, all five research steps are substantiated with (more or less directly measureable) indicators that, then, have been used for the actual evaluation of QCA applications:

First, with regard to the research design, the evaluation reveals that the set-theoretic basis of QCA is reflected only very sporadically in the empirical applications. Instead, oftentimes the

number of cases is argued to be the only reason to choose a QCA-based research design. Consequently, most studies under scrutiny apply QCA to a mid-sized number of cases such as 30 OECD-countries (Hotho 2014) or 68 firms (Leischnig et al. 2014) and use a medium number of conditions between four (e.g. Grandori & Furnari 2008) and six (e.g. Crowley 2013). In many but not the majority of QCA applications, the selection of cases and conditions follow clear theoretical criteria. However, many studies do not consider the ratio of cases and conditions and run into the problem of a very high number of logical remainders not because of limited diversity, but because of including an unreasonable number of conditions (e.g. Blackman 2013). To conclude, the criteria for a high quality research design in QCA are met only occasionally.

Second, the calibration of sets displays a huge problem across most of the QCA applications considered here. Just a small amount of studies transparently discusses the thresholds for the qualitative anchors (e.g. Cebotari & Vink 2013). Others at least display the anchor points without further justifying them (e.g. Ordanini et al. 2014), but many studies give virtually no information about the calibration decisions (e.g. Freitas et al. 2013). Several studies also allocate the 0.5 anchor to cases, which means that those cases are both members and non-members of the respective sets (e.g. Korczynski & Evans 2013). Moreover, the labeling of sets is oftentimes not implemented during calibration as sets like 'social tie' (Marr 2012) or 'Human Development Index' (Achilov & Shaykhutdinov 2013) reveal. Likewise, the identification of unnecessary variance is almost never to find. Instead, there are empirical applications assigning the full membership to the case with the observed maximum, the full non-membership to the observed minimum, and the crossover point to the mean (e.g. Allen & Allen 2015). Finally, some studies even leave the calibration over to the computer in order to avoid „user-imposed biases in the analyses" (Longest & Thoits 2012: 204). Again, consider that calibration of sets is the centerpiece of every QCA as it bridges conceptual knowledge and case knowledge. Based on the empirical evaluation, the empirical quality of calibration in many QCA applications needs to be judged as non-high.

Third, more than two thirds of all evaluated studies in the fields of sociology and business research do not check for necessary conditions at all! Several others do not conduct it prior to the analysis of sufficiency but deduce 'false' necessary conditions from the sufficient solution term (e.g. Bakker et al. 2011). Moreover, the justified handling of the parameters of fit is almost never at hand. If at all, Schneider and Wagemann's suggestion of a „standard consistency threshold of 0.9" (Cebotari & Vink 2013: 307) for necessity is cited; once even incorrectly: „A threshold of 0.8 consistency (80 percent or higher) is generally accepted for supporting

claims on necessary conditions.“ (Achilov & Shaykhutdinov 2013: 30). Very occasionally, necessary conjunctions of two or even more sets are presented, but if so, no discussion on theoretically useful functional equivalents has been conducted (e.g. Svevo-Cianci et al. 2010). The empirical evaluation of the quality of the analysis of necessity reveals major problems and confirms the “sufficiency bias” (Schneider & Wagemann 2012: 220) in QCA.

Consequently, fourth, the analysis of sufficiency is discussed in much more detail in all studies under scrutiny. However, with regard to the handling of consistency, a majority of studies justifies the chosen level of inconsistency mechanically by quoting standard thresholds by either Ragin (2008) or Schneider and Wagemann (2012; e.g. Wright & Schaffer Boudet 2012: 740). Others do not present any information on the consistency cutoff (e.g. Blackman et al. 2011). Likewise, the tradeoff between the level of consistency and coverage is almost never debated. One reason might be that a relatively large number of studies presents final solution formulas that cover just a very small share of the outcome set (e.g. about 4% by Garcia-Castro & Aguilera 2014; 22% Ganter & Hecker 2014). Moreover, only a very small number of studies do not face any logical remainders (e.g. Ghoshal 2013). But while especially in business research the vast majority openly approaches the problem of limited diversity, most studies in sociology do not. Here, also the influence of logical remainders on the three solution terms of the standard analysis is oftentimes not discussed. Thus, the actual treatment of logical remainders in the logical minimization process stays unclear in many studies. Finally, skewedness of set membership scores is not considered in any of the studies although some present highly skewed conditions (e.g. Stockemer 2013) or even outcomes (e.g. Braun 2013). To summarize, the quality of the analysis of sufficiency in QCA applications is – in comparison to the analysis of necessity and calibration – higher among the studies. Yet, especially the awareness and the treatment of the parameters of fit and of logical remainders need a less mechanical utilization and a stronger justification in all fields of application.

Finally, fifth, the quality of the interpretation of QCA findings differs strongly across studies. Several studies link their findings back to the cases (e.g. Bentele 2013) while other interpret only the sufficient configurations and not the covered cases (e.g. Giugni & Nai 2013). Interestingly, also the selection of the one solution term for the interpretation differs remarkably. Several applications stick to the conservative solution by not including any simplifying assumptions (e.g. Glaesser & Cooper 2011), while others use the intermediate (e.g. Korczynski & Evans 2013) or the most parsimonious solution (e.g. Wollebæk 2010). Just a few studies discuss all three solutions (e.g. Hafner-Fink et al. 2013) and a large amount of studies do not provide any information about which solution has been selected for interpretation (e.g. Jang

2009). The evaluation of typical forms of visualization in a broader sense than XY-plots, which are almost never used (but see Schneider & Makszin 2014a), reveals that also other graphical tools typically are not utilized in both fields of research. Even information that is essential for the understanding of the application like the truth table is available in about every third study only. Still, tables covering either numerical or symbolical (see Ragin & Fiss 2009) information on the quality of set relations including the final solution terms can be regarded as standard information.

In conclusion, the question on how to evaluate the quality of QCA applications can be answered with regard to the concept of quality of QCA as developed in chapter 1. The respective research steps and indicators provided the necessary tools for a proper evaluation. However, the evaluation of the quality of QCA applications revealed for the largest amount of the 139 studies under scrutiny at least smaller, but regularly larger deviations from a high quality application as conceptualized in chapter 1. A striking issue is the enormous lack of transparency both regarding the handling and processing of empirical data and the qualitative decisions that need to be taken during the analysis. By that, important features of QCA are either overlooked or ignored which in many studies lead to empirical results that are neither traceable in their formation nor understandable and replicable for the reader. In short, the overall quality of QCA in sociological and business research applications definitely cannot be evaluated as high but as problematic, and the results of several studies might be questioned.

7.3 Improvement

A major quality indicator that is applied under several research steps in the concept of quality of QCA is the set-theoretic foundation of QCA that needs to be considered constantly. Likewise, it has been applied to check the quality of the mathematical underpinning of QCA. More precisely, the parameters of fit as developed by Ragin (2006) have been criticized to contradict the notion of asymmetry. As turned out, for the Ragin consistency of sufficiency measure, cases with the condition set X being absent might strongly influence the statement that the presence of X is sufficient for Y or not. The other way around, the consistency of necessity formula might utilize cases where the outcome is absent to measure the strength of the empirical support for the claim that the outcome cannot be present without the condition being present. The improvement article, in short, discovered that all the formulas presented by Ragin (2006) rely on cases that are irrelevant for the respective set relation, which contradicts the set-theoretic foundation of QCA.

Thus, on the one hand, an update of these formulas by excluding irrelevant cases from the calculation has been developed. On the other hand, the new measures' influence on the analytical problem of simultaneous subset relations (i.e. skewedness) has been discussed. The formula for 'relevant consistency of sufficiency' helps to avoid inaccurate inferences about sufficiency due to skewedness in X. However, as this was not simply transferable to skewedness in X for necessity claims, the standardized Blau-Index (Blau 1977) has been transferred as an appropriate formula for non-skewedness in QCA.

To conclude, the improvement of the quality of QCA took place with regard to the set-theoretic notion of asymmetry that was contradicted by the Ragin formulas. In addition, a new parameter that is able to measure the extent of unequal distributions of set memberships and non-memberships across all logically possible configurations of conditions has been developed.

7.4 Application

At this final point of the cumulative dissertation on the quality of QCA two major findings need to be combined. On the one hand, the conceptualization of the quality of QCA as presented in chapter 1 reveals that the specific requisites for a proper application of QCA are numerous. On the other hand, the evaluation of the quality of QCA applications exposes that these criteria are hardly ever met by scientific articles. It still strikes the eye that all papers included into the empirical evaluation of the quality of QCA applications in chapter 2, 3, and 4 have been subject to a (more or less rigorous) peer-review process. But still, most of them have not met the criteria for a high quality of QCA application. Are the criteria set out in chapter 1 too demanding and do they have to be reformulated in a less restrictive way? The final paper of the cumulative dissertation is to be understood as an attempt to offer a best practice application of QCA that takes into account the various demands emphasized by the conceptualization:

The research design combines a reasoned ratio of 22 cases and four conditions (i.e. 16 truth table rows) in order to keep the number of logical remainders at a minimum level. Moreover, the selection of conditions and outcome set is theoretically driven. For instance, all conditions included into the analysis are different position roles and preference roles that can be used as reliable predictors for individual behavior; here direct interactions. Moreover, set calibration was conducted transparently, with clear labels and by taking into account unnecessary variance. The analysis of necessity has been conducted separately and first, and the necessity statement has been used for the subsequent counterfactual argument for the treatment of logi-

cal remainders. Finally, the interpretation went back to the cases and discussed in particular a deviant case. All findings have been visualized by a XY-plot.

By that, it strongly argues against a softening of the criteria from the concept of quality of QCA applications. Instead, I argue, that there is a serious lack of QCA expertise by authors of QCA-based articles, and maybe by reviewers, too.

Literature

Aberbach, Joel D., Robert D. Putnam, and Bert A. Rockman. 1981. *Bureaucrats and Politicians in Western Democracies*. Cambridge: Harvard University Press.

Achilov, Dilshod, and Renat Shaykhutdinov. 2013. "State Regulation of Religion and Radicalism in the Post-Communist Muslim Republics." *Problems of Post-Communism* 60 (5):17-33.

Adcock, Robert, and David Collier. 2001. "Measurement Validity: A Shared Standard for Qualitative and Quantitative Research." *American Political Science Review* 95 (3): 529-46.

Aguilera-Caracuel, Javier, Eugenio M. Fedriani, and Blanca L. Delgado-Márquez. 2014. "Institutional Distance among Country Influences and Environmental Performance Standardization in Multinational Enterprises." *Journal of Business Research* 67 (11):2385-92.

Ahn, Sang-Hoon, and Sophia Seung-yoon Lee. 2012. "Explaining Korean Welfare State Development with New Empirical Data and Methods." *Asian Social Work and Policy Review* 6 (2):67-85.

Allen, Matthew M. C., and Maria L. Aldred. 2011. "Varieties of Capitalism, Governance, and High-Tech Export Performance: A Fuzzy-Set Analysis of the new EU Member States." *Employee Relations* 33 (4):334-55.

Allen, Matthew, and Maria L. Aldred. 2013. "Business Regulation, Inward Foreign Direct Investment, and Economic Growth in the New European Union Member States." *Critical Perspectives on International Business* 9 (3):301-21.

Allen, Matthew M. C., and Maria L. Allen. 2015. "Companies' Access to Finance, Cooperative Industrial Relations, and Economic Growth: A Comparative Analysis of the States of South Eastern Europe." *Research in International Business and Finance* 33 (1):167-77.

Ambuehl, Mathias, Michael Baumgartner, Ruedi Epple, Alexis Kauffmann, and Alrik Thiem. 2015. *cna: A Package for Coincidence Analysis. R Package Version 1.0-3. http://cran.r-project.org/package=cna*.

Amenta, Edwin, Neal Caren, Sheera J. Olasky, and James E. Stobaugh. 2009. "All the Movements Fit to Print: Who, What, When, Where, and Why SMO Families Appeared in the New York Times in the Twentieth Century." *American Sociological Review* 74 (4):636-56.

Anscombe, Francis J. 1973. "Graphs in Statistical Analysis." *American Statistician* 27 (1):17-21.

Anton, Thomas J. 1980. *Administered Politics: Elite Political Culture in Sweden*. Boston: Nijhoff Publisher.

Auel, Katrin. 2006. "The Europeanisation of the German Bundestag: Institutional Change and Informal Adaptation." *German Politics* 15 (3):249-68.

Bakker, René M., Bart Cambré, Leonique Korlaar, and Joerg Raab. 2011. "Managing the Project Learning Paradox: A Set-Theoretic Approach Toward Project Knowledge Transfer." *International Journal of Project Management* 29 (5):494-503.

Baltzer, Maria, Hugo Westerlund, Mona Backhans, and Karin Melinder. 2011. "Involvement and Structure: A Qualitative Study of Organizational Change and Sickness Absence among Women in the Public Sector in Sweden." *BMC Public Health* 11 (1):318-34.

Baumgartner, Michael. 2009. "Uncovering Deterministic Causal Structures: A Boolean Approach." *Synthese* 170 (1):71-96.

Baumgartner, Michael. 2013. "Detecting Causal Chains in Small N Data." *Field Methods* 25 (1):3-24.

Baumgartner, Michael. 2015. "Parsimony and Causality." *Quality & Quantity* 49 (2):839-56.

Bauwens, Joke, Bojana Lobe, Katia Segers, and Liza Tsaliki. 2009. "A Shared Responsibility." *Journal of Children and Media* 3 (4):316-30.

Beach, Derek, and Rasmus B. Pedersen. 2012. *Process tracing methods. Foundations and guidelines*. Ann Arbor: University of Michigan Press.

Bell, R. Greg, Igor Filatotchev, and Ruth V. Aguilera. 2014. "Corporate Governance and Investors' Perceptions of Foreign IPO Value: An Institutional Perspective." *Academy of Management Journal* 57 (1):301-20.

Belsley, David A., Edwin Kuh, and Roy E. Welsch. 1980. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. Hoboken: Wiley-Interscience.

Bennett, Andrew, and Jeffrey T. Checkel (Eds.). 2014. *Process Tracing. From Metaphor to Analytic Tool*. Cambridge: Cambridge University Press.

Bentele, Keith Gunnar. 2013. "Distinct paths to higher inequality? A Qualitative Comparative Analysis of rising earnings inequality among U.S. States, 1980-2010." *Research in Social Stratification and Mobility* 34 (0):30-57.

Benz, Arthur. 2008. *Der moderne Staat: Grundlagen der politologischen Analyse*. München: Oldenbourg.

Bergman, Torbjörn. 1997. "National Parliaments and EU Affairs Committees: Notes on Empirical Variation and Competing Explanations." *Journal of European Public Policy* 4 (3):373-87.

Bergman, Torbjörn. 2000. "The European Union as the Next Step of Delegation and Accountability." *European Journal of Political Research* 37 (3):415-29.

Bergman, Torbjörn. 2003. "Sweden: From Separation of Power to Parliamentary Supremacy - and Back Again?" In Strøm, Kaare, Wolfgang C. Müller, and Torbjörn Bergman (Eds.). *Delegation and Accountability in Parliamentary Democracies*. Oxford: Oxford University Press, 594-619.

Biddle, Bruce J. 1986. "Recent Developments in Role Theory." *Annual Review of Sociology* 12 (1):67-92.

Blackman, Tim. 2008. "Can Smoking Cessation Services be Better Targeted to Tackle Health Inequalities? Evidence from a Cross-Sectional Study." *Health Education Journal* 67 (2):91-101.

Blackman, Tim. 2013. "Exploring Explanations for Local Reductions in Teenage Pregnancy Rates in England: An Approach Using Qualitative Comparative Analysis." *Social Policy and Society* 12 (1):61-72.

Blackman, Tim, and Katie Dunstan. 2010. "Qualitative Comparative Analysis and Health Inequalities: Investigating Reasons for Differential Progress with Narrowing Local Gaps in Mortality." *Journal of Social Policy* 39 (3):359-73.

Blackman, Tim, Jonathan Wistow, and David Byrne. 2011. "A Qualitative Comparative Analysis of Factors Associated with Trends in Narrowing Health Inequalities in England." *Social Science & Medicine* 72 (12):1965-74.

Blatter, Joachim, and Markus Haverland. 2012. *Designing Case Studies: Explanatory Approaches in Small-N Research*. Basingstoke: Palgrave Macmillan.

Blau, Peter M. 1977. *Inequality and Heterogeneity: A Primitive Theory of Social Structure*. New York: Free Press.

Bleijenbergh, Inge, and Conny Roggeband. 2007. "Equality Machineries Matter: The Impact of Women's Political Pressure on European Social-Care Policies." *Social Politics: International Studies in Gender, State and Society* 14 (4):437-59.

Blomgren, Magnus, and Olivier Rozenberg. 2012. "Legislative Roles and Legislative Studies – The Neo-institutionalist Turning Point?" In Blomgren, Magnus, and Olivier Rozenberg (Eds.). *Parliamentary Roles in Modern Legislatures*. London: Routledge, 8-36.

Bol, Damien, and Francesca Luppi. 2013. "Confronting Theories Based on Necessary Relations: Making the Best of QCA Possibilities." *Political Research Quarterly* 66 (1):205-10.

Boudet, Hillary S., Dilanka C. Jayasundera, and Jennifer Davis. 2011. "Drivers of Conflict in Developing Country Infrastructure Projects: Experience from the Water and Pipeline Sectors." *Journal of Construction Engineering and Management* 137 (7):498-511.

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

Braumoeller, Bear F. 2015. "Guarding Against False Positives in Qualitative Comparative Analysis." *Political Analysis* 23 (4):471-87.

Braumoeller, Bear F., and Gary Goertz. 2000. "The Methodology of Necessary Conditions." *American Journal of Political Science* 44 (4):844-58.

Braumoeller, Bear F., and Gary Goertz. 2003. "The Statistical Methodology of Necessary Conditions." In Goertz, Gary, and Harvey Starr (Eds.). *Necessary Conditions: Theory, Methodology, and Applications*. Lanham: Rowman & Littlefield, 197-223.

Braun, Caelesta. 2013. "The Driving Forces of Stability: Exploring the Nature of Long-Term Bureaucracy-Interest Group Interactions." *Administration & Society* 45 (7):809-36.

Britt, David W., and Mark I. Evans. 2007. "Sometimes Doing the Right Thing Sucks: Frame Combinations and Multi-Fetal Pregnancy Reduction Decision Difficulty." *Social Science & Medicine* 65 (11):2342-56.

Buche, Antje, Jonas Buche und Markus B. Siewert. 2016. "Fuzzy Logic or Fuzzy Application? A Response to Stockemer's "Fuzzy Set or Fuzzy Logic?"" *European Political Science* 15(2): 359-378.

Buche, Jonas. 2016. "Relevant Consistency and Relevant Coverage in Fuzzy Set Qualitative Comparative Analysis (fsQCA)." *Unpublished manuscript*.

Buche, Jonas. 2017. "Europeanization of Legislative-executive Relations at the Micro Level – Under Which Conditions Do Swedish MPs Interact with Ministerial Officials?" *COMPASSS Working Paper Series* 2017-87.

Buche, Jonas, and Julia Fleischer. 2013. *Silent Servants or Mighty Masters? The Role of National Parliamentary Committee Secretariats in EU Affairs*. Bordeaux: ECPR General Conference.

Buche, Jonas, and Julia Fleischer. 2015. *Supporting Parliamentary Behaviour in European Union Affairs : The Power of Parliamentary Bureaucracy*. Warschau: ECPR Joint Sessions 2015.

Buche, Jonas, and Markus B. Siewert. 2015. "Qualitative Comparative Analysis (QCA) in der Soziologie – Perspektiven, Potentiale und Anwendungsbereiche." *Zeitschrift für Soziologie* 44 (6):386-406.

Buzogány, Aron C., and Sabine Kropp. 2013. "Sisyphos und der Marmorblock: Parlament und Expertise in Zeiten der Euro-Krise." *Zeitschrift für Politikberatung* 6 (1):3-14.

Cacciatore, Federica, Alessandro Natalini, and Claudius Wagemann. 2015. "Clustered Europeanization and National Reform Programmes: A Qualitative Comparative Analysis." *Journal of European Public Policy* 22 (8):1186-211.

Caramani, Daniele. 2009. *Introduction to the Comparative Method with Boolean Algebra*. Los Angeles: Sage.

Cárdenas, Julián. 2012. "Varieties of Corporate Networks: Network Analysis and fsQCA." *International Journal of Comparative Sociology* 53 (4):298-322.

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

Castellano, José María. 2010. "Secondary Education and Broadband Diffusion: A Qualitative Comparative Analysis." *Info* 12 (6):121-38.

Cebotari, Victor, and Maarten P. Vink. 2013. "A Configurational Analysis of Ethnic Protest in Europe." *International Journal of Comparative Sociology* 54 (4):298-324.

Chandra Balodi, Krishna, and Jaideep Prabhu. 2014. "Causal Recipes for High Performance." *International Journal of Entrepreneurial Behaviour & Research* 20 (6):542-61.

Chang, Chia-Wen, Ting-Hsiang Tseng, and Arch G. Woodside. 2013. "Configural Algorithms of Patient Satisfaction, Participation in Diagnostics, and Treatment Decisions' Influences on Hospital Loyalty." *Journal of Services Marketing* 27 (2):91-103.

Chang, Man-Ling, and Cheng-Feng Cheng. 2014. "How Balance Theory Explains High-Tech Professionals' Solutions of Enhancing Job Satisfaction." *Journal of Business Research* 67 (9):2008-18.

Cheng, Cheng-Feng, Man-Ling Chang, and Chu-Shiu Li. 2013. "Configural paths to successful product innovation." *Journal of Business Research* 66 (12):2561-73.

Clark, William R., Michael J. Gilligan, and Matt Golder. 2006. "A Simple Multivariate Test for Asymmetric Hypotheses." *Political Analysis* 14 (3):311-31.

Cohen, Jacob .1988. *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale: Erlbaum.

Collier, David. 2011. "Understanding process tracing." *Political Science and Politics* 44 (2):823-30.

Collier, David. 2014. "Comment: QCA Should Set Aside the Algorithms." *Sociological Methodology* 44 (1):122-26.

Compasss. 2016. *Bibliography*. http://www.compasss.org/bibdata.htm, last access: 02.07.2016.

Cooper, Barry, and Judith Glaesser. 2011. "Paradoxes and Pitfalls in Using Fuzzy Set QCA: Illustrations from a Critical Review of a Study of Educational Inequality." Sociological Research Online 16 (3):8.

Cooper, Barry, Judith Glaesser, and Stephanie Thompson. 2014. "Schneider and Wagemann's Proposed Enhanced Standard Analysis for Ragin's Qualitative Comparative Analysis: Some Unresolved Problems and Some Suggestions for Addressing Them." *COMPASSS Working Paper Series* 2014-77.

Crawford, Seth. 2012. "What is the Energy Policy-Planning Network and Who Dominates it? A Network and QCA Analysis of Leading Energy Firms and Organizations." *Energy Policy* 45 (0):430-9.

Crilly, Donal. 2011. "Predicting Stakeholder Orientation in the Multinational Enterprise: A Mid-Range Theory." *Journal of International Business Studies* 42 (5):694-717.

Crilly, Donal, Maurizio Zollo, and Morten T. Hansen. 2012. "Faking It or Muddling Through? Understanding Decoupling in Response to Stakeholder Pressures." *Academy of Management Journal* 55 (6):1429-48.

Cronqvist, Lasse. 2005. "Introduction to Multi-value Qualitative Comparative Analysis (mvQCA)." *COMPASSS Didactics Paper* 2005-4.

Cronqvist, Lasse. 2016. *Tosmana: Tool for Small-N Analysis [Computer Programme], Version 1.4*. Trier: Universität Trier.

Cronqvist, Lasse, and Dirk Berg-Schlosser. 2009. "Multi-value QCA (mvQCA)." In Rihoux, Benoît, and Charles C. Ragin (Eds.). *Configurational Comparative Methods. Qualitative Comparative Analysis (QCA) and Related Techniques*. Thousand Oaks: Sage, 69-86.

Crowley, Martha. 2012. "Control and Dignity in Professional, Manual and Service-Sector Employment." *Organization Studies* 33 (10):1383-406.

Crowley, Martha. 2013. "Gender, the Labor Process and Dignity at Work." *Social Forces* 91 (4):1209-38.

Crowley, Martha. 2014. "Class, Control, and Relational Indignity: Labor Process Foundations for Workplace Humiliation, Conflict, and Shame." *American Behavioral Scientist* 58 (3):416-34.

Da Roit, Barbara, and Bernhard Weicht. 2013. "Migrant Care Work and Care, Migration and Employment Regimes: A Fuzzy-Set Analysis." *Journal of European Social Policy* 23 (5):469-86.

Dahlerup, Drude (Ed.). 2006. *Women, Quotas, and Politics*. New York: Routledge.

DeMeur, Gisèle, and Benoît Rihoux. 2002. *L'analyse Quali-quantitative Comparée (AQQC-QCA)*. Louvain-La-Neuve: Academia Bruylant.

Derlien, Hans-Ulrich, and Renate Mayntz. 1988. *Comparative Elite Study II: Einstellungen der politisch-administrativen Elite des Bundes 1987*. Bamberg: Universität Bamberg.

Dow, Sheila C. 2007. "Variety of Methodological Approach in Economics." *Journal of Economic Surveys* 21 (3):447-65.

Doz, Yves L. 2011. "Qualitative Research for International Business." *Journal of International Business Studies* 42 (5):582-90.

Dusa, Adrian. 2016. *Modern Functions for Qualitative Comparative Analysis. R Package Version 2.4. http://cran.r-project.org/package=QCAGUI*

Duşa, Adrian, and Alrik Thiem. 2014. *QCA: A Package for Qualitative Comparative Analysis. R Package Version 1.1-4. http://cran.r-project.org/package=QCA.*

Ebbinghaus, Bernhard. 2014. "Renzension Goertz, Gary, and James Mahoney. 2012. A Tale of Two Cultures. Qualitative and Quantitative Research in the Social Sciences. New Jersey: Princeton University Press." *Soziologische Revue* 37 (3):363-65.

Eberbach-Born, Birgit, and Sabine Kropp. 2013. "Parlamentarische Kontrolle und Europäische Union: Chancen und Restriktionen nach dem Lissabon-Vertrag – einleitende Überlegungen." In Eberbach-Born, Birgit, Sabine Kropp, Andrej Stuchlik, and Wolfgang Zeh (Eds.). *Parlamentarische Kontrolle und Europäische Union*. Baden-Baden: Nomos, 11-20.

Ehn, Peter. 1998. *Maktens administratörer: Ledande svenska statstjänstemäns och politikers syn på tjänstemannarollen i ett förändringsperspektiv*. Stockholm: Stockholms universitet.

Ehn, Peter, Magnus Isberg, Claes Linde, and Gunnar Wallin. 2003. "Swedish Bureaucracy in an Era of Change." *Governance* 16 (3):429-58.

Elman, Colin. 2005. "Explanatory Typologies in Qualitative Studies of International Politics." *International Organization*, 59 (2):293–326.

Emmenegger, Patrick, Jon Kvist, and Svend-Erik Skaaning. 2013. "Making the Most of Configurational Comparative Analysis: An Assessment of QCA Applications in Comparative Welfare State Research." *Political Research Quarterly*, 66 (1):185-90.

Emmenegger, Patrick, Dominik Schraff, and André Walter. 2014. "QCA, the Truth Table Analysis and Large-N Survey Data: The Benefits of Calibration and the Importance of Robustness Tests." *COMPASSS Working Paper Series* 2014-79.

Eng, Sereikhuoch, and Arch G. Woodside. 2012. "Configural Analysis of the Drinking Man: Fuzzy-Set Qualitative Comparative Analyses." *Addictive Behaviors* 37 (4):541-3.

Epple, Ruedi, Martin Gasser, Sarah Kersten, Michael Nollert, and Sebastian Schief. 2014. "Institutions and Gender Time Inequality: A Fuzzy-Set QCA of Swiss Cantons." *Swiss Journal of Sociology* 40 (2):259-78.

Fiss, Peer C. 2011. "Building Better Causal Theories: A Fuzzy Set Approach to Typologies in Organizational Research." *Academy of Management Journal* 54 (2):393-420.

Fiss, Peer C., Bart Cambré, and Axel Marx (Eds.). 2013a. *Configurational theory and methods in organizational research*. Bingley: Emerald.

Fiss, Peer C., Dmitry Sharapov, and Lasse Cronqvist. 2013b. "Opposites Attract? Opportunities and Challenges for Integrating Large-N QCA and Econometric Analysis." *Political Research Quarterly* 66 (1):191-98.

Forsythe, Perry John. 2012. "Profiling Customer perceived Service Quality Expectations in made-to-order Housing Construction in Australia." *Engineering, Construction and Architectural Management* 19 (6):587-609.

Freitas, J. S., Gonçalves, C. A., Cheng, L. C., & Muniz, R. M. (2011). Parsimonious determinants of pre-incubated academic spin-offs initial performance: A configurational perspective. *Journal of Technology Management & Innovation*, 6(2), 50–65.

Freitas, Jonathan Simões, Carlos Alberto Gonçalves, Lin Chih Cheng, and Reynaldo Maia Muniz. 2013. "Structuration Aspects in Academic Spin-Off Emergence: A Roadmap-Based Analysis." *Technological Forecasting and Social Change* 80 (6):1162-78.

Ganter, Alois, and Achim Hecker. 2014. "Configurational Paths to Organizational Innovation: Qualitative Comparative Analyses of Antecedents and Contingencies." *Journal of Business Research* 67 (6):1285-92.

Garcia-Castro, Roberto, and Ruth V. Aguilera. 2014. "Family Involvement in Business and Financial Performance: A Set-Theoretic Cross-National Inquiry." *Journal of Family Business Strategy* 5 (1):85-96.

García-Castro, Roberto, Ruth V. Aguilera, and Miguel A. Ariño. 2013. "Bundles of Firm Corporate Governance Practices: A Fuzzy Set Analysis." *Corporate Governance: An International Review* 21 (4):390-407.

García-Castro, Roberto, and José Casasola. 2011. "A Set-Theoretic Analysis of the Components of Family Involvement in Publicly Listed and Major Unlisted Firms." *Journal of Family Business Strategy* 2 (1):15-25.

George, Alexander L., and Andrew Bennett. 2005. *Case studies and theory development in the social sciences*. Cambridge: MIT Press.

Gerring, John. 2012. *Social Science Methodology. A Unified Framework.* . Cambridge: Cambridge University Press.

Ghoshal, Raj Andrew. 2013. "Transforming Collective Memory: Mnemonic Opportunity Structures and the Outcomes of Racial Violence Memory Movements." *Theory and Society* 42 (4):329-50.

Giugni, Marco und Alessandro Nai, 2013. Paths Towards Consensus: Explaining Decision Making within the Swiss Global Justice Movement. In *Swiss Political Science Review*, Vol. 19(1), 26-40.

Giugni, Marco und Sakura Yamasaki, 2009. The Policy Impact of Social Movements: A Replication Through Qualitative Comparative Analysis. In *Mobilization: An International Quarterly*, Vol. 14 (4), 467-484.

Glaesser, Judith, and Barry Cooper. 2011. "Selectivity and Flexibility in the German Secondary School System: A Configurational Analysis of Recent Data from the German Socio-Economic Panel." *European Sociological Review* 27 (5):570-85.

Glaesser, Judith, and Barry Cooper. 2012a. "Educational Achievement in Selective and Comprehensive Local Education Authorities: A Configurational Analysis." *British Journal of Sociology of Education* 33 (2):223-44.

Glaesser, Judith, and Barry Cooper. 2012b. "Gender, Parental Education, and Ability: Their Interacting Roles in Predicting GCSE Success." *Cambridge Journal of Education* 42 (4):463-80.

Goertz, Gary. 2006a. *Social Sciences Concepts. A User's Guide.* Princeton: Princeton University Press.

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

Goertz, Gary, and James Mahoney. 2012. *A Tale of Two Cultures: Qualitative and Quantitative Research in the Social Sciences.* Princeton: Princeton University Press.

Goertz, Gary, and James Mahoney. 2013. "Methodological Rorschach Tests: Contrasting Interpretations in Qualitative and Quantitative Research." *Comparative Political Studies* 46 (2): 236-51.

Goetz, Klaus H., and Jan-Hinrik Meyer-Sahling. 2008. "The Europeanisation of National Political Systems: Parliaments and Executives." *Living Reviews in European Governance* 3 (2).

Grandori, Anna, and Santi Furnari. 2008. "A Chemistry of Organization: Combinatory Analysis and Design." *Organization Studies* 29 (3):459-85.

Grant, Don, Alfonso Morales, and Jeffrey J. Sallaz. 2009. "Pathways to Meaning: A New Approach to Studying Emotions at Work." *American Journal of Sociology* 115 (2):327-64.

Grant, Don, Kathleen O'Neil, and Laura Stephens. 2009. "Pluralistic Ignorance Among Assembled Peers." *Sociological Perspectives* 52 (1):59-79.

Grant, Don, Mary N. Trautner, Liam Downey, and Lisa Thiebaud. 2010. "Bringing the Polluters back in: Environmental Inequality and the Organization of Chemical Production." *American Sociological Review* 75 (4):479-504.

Greckhamer, Thomas. 2011. "Cross-Cultural Differences in Compensation Level and Inequality across Occupations: A Set-Theoretic Analysis." *Organization Studies* 32 (1):85-115.

Greckhamer, Thomas, Vilmos F. Misangyi, and Peer C. Fiss. 2013. "The Two QCAs: From Small-N to a Large-N Set-Theoretical Approach." In Fiss, Peer C., Bart Cambré, and Axel Marx (Eds.). *Configurational Theory and Methods in Organizational Research*. Bingley: Emerald, 49-75.

Grofman, Bernard, and Carsten Q. Schneider. 2009. "An Introduction to Crisp Set QCA, with a Comparison to Binary Logistic Regression." *Political Research Quarterly* 62 (4):662-72.

Hafner-Fink, Mitja, Brina Malnar, and Samo Uhan. 2013. "The National Contexts of Post-National Citizenship." *Czech Sociological Review* 49 (6):867-901.

Hall, Peter A. 2003. "Aliging Ontology and Methodology in Comparative Research." In Mahoney, James, and Dietrich Rueschemeyer (Eds.). *Comparative Historical Analysis in the Social Sciences*. Cambridge: Cambridge University Press, 373-404.

Harriss-White, Barbara, Wendy Olsen, Penny Vera-Sanso, and V. Suresh. 2013. "Multiple Shocks and Slum Household Economies in South India." *Economy and Society* 42 (3):398-429.

Haynes, Philip. 2011. "Are Scandinavian Countries Different? A Comparison of Relative Incomes for Older People in OECD Nations." *Social Policy & Administration* 45 (2):114-30.

Haynes, Philip, Laura Banks, and Michael Hill. 2013. "Social networks amongst older people in OECD countries: a qualitatative [*sic*] comparative analysis." *Journal of International and Comparative Social Policy* 29 (1):15-27.

Haynes, Philip, Michael Hill, and Laura Banks. 2010. "Older People's Family Contacts and Long-term Care Expenditure in OECD Countries: A Comparative Approach Using Qualitative Comparative Analysis." *Social Policy & Administration* 44 (1):67-84.

Hegeland, Hans, and Christine Neuhold. 2002. "Parliamentary Participation in EU Affairs in Austria, Finland and Sweden: Newcomers with Different Approaches." *European Integration online Papers* 6 (10).

Hodson, Randy, Vincent J. Roscigno, and Steven H. Lopez. 2006. "Chaos and the Abuse of Power." *Work and Occupations* 33 (4):382-416.

Holzhacker, Ronald. 2008. "Parliamentary Scrutiny." In Graziano, Paolo, and Maarten Peter Vink (Eds.). *Europeanization: New research agendas*. New York: Palgrave Macmillan, 141-53.

Hörisch, Felix. 2012. "Die Bestimmungsfaktoren der Einführung der Unternehmensmitbestimmung in den OECD-Staaten - Eine fuzzy-set Qualitative Comparative Analysis" *Zeitschrift für Sozialreform* 58 (1):33-57.

Hollstein, Betina. 2014. "Mixed Methods Social Networks Research: An Introduction." In Domínguez, Silvia, and Betina Hollstein (Eds.). *Mixed Methods Social Networks Research: Design and Applications*. Cambridge: Cambridge University Press, 3-34.

Hollstein, Betina, and Carsten G. Ullrich. 2003. "Einheit trotz Vielfalt? Zum konstitutiven Kern qualitativer Forschung." *Soziologie* 32(4): 29-43.

Hollstein, Betina, and Claudius Wagemann. 2014. "Fuzzy-Set Analysis of Network Data as Mixed Method: Personal Networks and the Transition from School to Work." In Domínguez, Silvia, and Betina Hollstein (Eds.). *Mixed Methods Social Networks Research: Design and Applications*. Cambridge: Cambridge University Press, 237-268.

Hotho, Jasper J. 2014. "From Typology to Taxonomy: A Configurational Analysis of National Business Systems and their Explanatory Power." *Organization Studies* 35 (5):671-702.

Huang, Ronggui. 2014. *QCA3: Yet another Package for Qualitative Comparative Analysis. R Package Version 0.0-7. http://cran.r-project.org/web/packages/QCA3/index.html*.

Hug, Simon. 2013. "Qualitative Comparative Analysis: How Inductive Use and Measurement Error Lead to Problematic Inference." *Political Analysis* 21 (2):252-65.

Hug, Simon. 2014. "We Need an Open Discussion of QCA's Limitations: A Comment on Thiem. Qualitative and Multi-Method Research." *Newsletter of the APSA Section for Qualitative and Multi-Method Research* 11 (1):24-27.

Hussain, Muzammil M., and Philip N. Howard. 2013. "What Best Explains Successful Protest Cascades? ICTs and the Fuzzy Causes of the Arab Spring." *International Studies Review* 15 (1):48-66.

IDEA [International Institute for Democracy and Electoral Assistance]. 2015. *Glossary of Terms*. Available at http://www.idea.int/esd/glossary.cfm, accessed 22 May 2015.

Ignatow, Gabe. 2011. "What has Globalization Done to Developing Countries' Public Libraries?" *International Sociology* 26 (6):746-68.

Inglehart, Ronald, and Pippa Norris. 2001. *Rising Tide: Gender Equality and Cultural Change Around the World*. Cambridge: Cambridge University Press.

IPU [Inter-Parliamentary Union]. 2015. *Women in National Parliaments*. Available at http://www.ipu.org/wmn-e/classif.htm, accessed 12 May 2015.

Ishida, Atsushi, Miya Yonetani, and Kenji Kosaka. 2006. "Determinants of Linguistics Human Rights Movements: An Analysis of Multiple Causation of LHRs Movements Using a Boolean Approach." *Social Forces* 84 (4):1938-55.

Jang, Dong-Ho. 2009. "Significance of Variations between Income Transfers and Social Care Services Development." *Journal of Comparative Social Welfare* 25 (1):37-48.

Jann, Ben. 2009. "Diagnostik von Regressionsschätzungen bei kleinen Stichproben (mit einem Exkurs zu logistischer Regression)." In Kriwy, Peter, and Christiane Gross (Eds.). *Klein aber fein! Quantitative empirische Sozialforschung mit kleinen Fallzahlen*. Wiesbaden: VS Verlag, 93-125.

Järvinen, Joonas, Juha-Antti Lamberg, Johan-Peter Murmann, and Jari Ojala. 2009. "Alternative Paths to Competitive Advantage: A Fuzzy-Set Analysis of the Origins of Large Firms." *Industry and Innovation* 16 (6):545-74.

Järvinen, Joonas, Juha-Antti Lamberg, and Lauri Pietinalho. 2012. "The Fall and the Fragmentation of National Clusters: Cluster Evolution in the Paper and Pulp Industry." *Journal of Forest Economics* 18 (3):218-41.

Johnson, Paul E. 2009. *Working with Ordinal Predictors*. Paper presented at the Annual Meeting of the MPSA. Chicago, March 30, 2009.

Judge, William Q., Stav Fainshmidt, and J. Lee Brown III. 2014. "Which Model of Capitalism best delivers both Wealth and Equality?" *Journal of International Business Studies* 45 (4):363-86.

Karlas, Jan. 2012. "National Parliamentary Control of EU Affairs: Institutional Design after Enlargement." *West European Politics* 35 (5):1095-113.

Kim, Seung Hyun. 2011. "On the Historical Determinants of Third Sector Strength: A Qualitative Comparative Analysis." *Social Science Journal* 48 (4):641-50.

Kim, Minyoung. 2013. "Many Roads lead to Rome: Implications of Geographic Scope as a Source of Isolating Mechanisms." *Journal of International Business Studies* 44 (9):898-921.

Klir, George J., Ute H. St. Clair, and Bo Yuan. 1997.*Fuzzy set theory. Foundations and applications*. Upper Saddle River, NJ: Prentice Hall.

Korczynski, Marek, and Claire Evans. 2013. "Customer Abuse to Service Workers: An Analysis of its Social Creation within the Service Economy." *Work, Employment & Society* 27 (5):768-84.

Kröger, Markus. 2011. "Promotion of Contentious Agency as a Rewarding Movement Strategy: Evidence from the MST-paper Industry Conflicts in Brazil." *Journal of Peasant Studies* 38 (2):435-58.

Krogslund, Chris, Donghyun Danny Choi, and Mathias Poertner. 2015. "Fuzzy Sets on Shaky Ground: Parameter Sensitivity and Confirmation Bias in fsQCA." *Political Analysis* 23 (1):21-41.

Krogslund, Chris, and Katherine Michel. 2014a. "A Larger-N, Fewer Variables Problem? The Counterintuitive Sensitivity of QCA." *Qualitative & Multi-Method Research* 14 (1):25-33.

Krogslund, Chris, and Katherine Michel. 2014b. *Set Theoretic Comparative Methods and the Recovery of Data Generating Processes.* Washington, D.C.: APSA Annual Meeting.

Krook, Mona L. 2009. *Quotas for Women in Politics: Gender and Candidate Selection Reform Worldwide*. Oxford: Oxford University Press.

Krook, Mona L. 2010. "Women's Representation in Parliament: A Qualitative Comparative Analysis." *Political Studies* 58 (5):886-908.

Krook, Mona L., and Sarah Childs. 2010. *Women, Gender, Politics. A Reader*. Oxford: Oxford University Press.

Kropp, Sabine. 2010. "German Parliamentary Party Groups in Europeanised Policymaking: Awakening from the Sleep? Institutions and Heuristics as MPs' Resources." *German Politics* 19 (2):123-47.

Kropp, Sabine, Jonas Buche, and Aron C. Buzogány. 2011. "Parlamentarisch-exekutive Steuerung in europäisierten Fachpolitiken – ein Blick auf die Mikroebene." In Abels, Gabriele, and Annegret Eppler (Eds.). *Auf dem Weg zum Mehrebenenparlamentarismus?: Funktionen von Parlamenten im politischen System der EU*. Baden-Baden: Nomos, 227-42.

Kropp, Sabine, Aron C. Buzogány, and Jonas Buche. 2012. "Von den Schwierigkeiten, Zusammengehöriges zu vereinen - Nationale Parlamente und Exekutiven als Gegenstand der Europäisierungsforschung." *Politische Vierteljahresschrift* 53 (1):109-34.

Kropp, Sabine, and Matthias Ruschke. 2010. "Parlament und Verwaltung: in Rollenpartnerschaft vereint? Ein Plädoyer für die Verknüpfung von Parlaments- und Verwaltungsforschung." *Zeitschrift für Parlamentsfragen* 41 (3):638-60.

Kvist, Jon. 2006. "Diversity, Ideal Types and Fuzzy Sets in Comparative Welfare State Research." In Rihoux, Benoît, and Heike Grimm (Eds.). *Innovative Comparative Methods for Policy Analysis: Beyond the Quantitative-qualitative Divide*. New York: Springer, 167-84.

Kvist, Jon. 2007. "Fuzzy set ideal type analysis." *Journal of Business Research* 60 (5):474-81.

Lamnek, Siegfried. 2010. *Qualitative Sozialforschung*. Weinheim: Beltz.

Laux, Thomas. 2015. "Nationalstaatliche Prozesse oder globale Strukturen? Eine Analyse der Mechanismen zur Gleichstellung von Frauen im Recht." *Berliner Journal für Soziologie* 24 (4):531-58.

Lee, Sophia Seung-Yoon. 2013. "High Non-Standard Employment Rates in the Republic of Korea and Japan: Analyzing Policy Configurations with Fuzzy-Set/QCA." *Policy and Society* 32 (4):333-44.

Leischnig, Alexander, Anja Geigenmueller, and Stefanie Lohmann. 2014. "On the Role of Alliance Management Capability, Organizational Compatibility, and Interaction Quality in Interorganizational Technology Transfer." *Journal of Business Research* 67 (6):1049-57.

Lijphart, Arend. 1971. "Comparative Politics and the Comparative Method." *The American Political Science Review* 65 (3):682-93.

Lijphart, Arend. 1975. "The Comparable-cases Strategy in Comparative Research." *Comparative Political Studies*, 8 (2):158-77.

Lijphart, Arend. 1999. *Patterns of Democracy: Government Forms and Performance in Thirty-six Countries*. New Haven: Yale University Press.

Linde, Claes. 1982. *Departement och verk: Om synen på den centrala statsförvaltningen och dess uppdelning - i en förändrad offentlig sektor*. Stockholm: Stockholms universitet.

Longest, Kyle C., and Peggy A. Thoits. 2012. "Gender, the Stress Process, and Health: A Configurational Approach." *Society and Mental Health* 2 (3):187-206.

Lucas, Samuel R., and Alisa Szatrowski. 2014. "Qualitative Comparative Analysis in Critical Perspective." *Sociological Methodology* 44 (1):1-79.

Mackie, John L. 1965. "Causes and Conditions." *American Philosophical Quarterly* 2 (4):245-64.

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

Maggetti, Martino. 2007. "De Facto Independence after Delegation: A Fuzzy-Set Analysis." *Regulation & Governance* 1 (4):271-94.

Maggetti, Martino, and Levi-Faur, David. 2013. "Dealing with errors in QCA." *Political Research Quarterly* 66 (1): 199-204.

Maggetti, Martino. 2014. "Promoting Corporate Responsibility in Private Banking: Necessary and Sufficient Conditions for Joining the Wolfsberg Initiative Against Money Laundering." *Business & Society* 53 (6):787-819.

Magnier-Watanabe, Rémy, and Dai Senoo. 2008. "Organizational Characteristics as Prescriptive Factors of Knowledge Management Initiatives." *Journal of Knowledge Management* 12 (1):21-36.

Mahoney, James. 2003. "Strategies of Causal Assessment in Comparative Historical Analysis." In Mahoney, James, and Dietrich Rueschemeyer (Eds.). *Comparative Historical Analysis in the Social Sciences*. Cambridge: Cambridge University Press, 337-72.

Mahoney, James. 2010. "After KKV: The new methodology of qualitative research." *World Politics* 62 (1):120-47.

Mahoney, James. 2012. "The Logic of Process Tracing Tests in the Social Sciences." *Sociological Methods & Research* 41 (4): 570-97.

Mahoney, James, Erin Kimball, and Kendra L. Koivu. 2009. "The Logic of Historical Explanation in the Social Sciences." *Comparative Political Studies* 42 (1):114-46.

Mahoney, James, and P. Larkin Terrie. 2009. "Comparative-Historical Analysis in Contemporary Political Science." In Box-Steffensmeier, Janet M., Henry E. Brady, and David Collier (Eds.). *The Oxford Handbook of Political Methodology*. Oxford: Oxford University Press, 737-56.

Marr, Matthew D. 2012. "Pathways out of Homelessness in Los Angeles and Tokyo: Multilevel Contexts of Limited Mobility amid Advanced Urban Marginality." *International Journal of Urban and Regional Research* 36 (5):980-1006.

Marx, Axel. 2008. "Limits to non-State Market Regulation: A Qualitative Comparative Analysis of the International Sport Footwear Industry and the Fair Labor Association." *Regulation & Governance* 2 (2):253-73.

Marx, Axel, Bart Cambré, and Peer C. Fiss. 2013. "Crisp-Set Qualitative Comparative Analysis in Organizational Studies." In Fiss, Peer C., Bart Cambré, and Axel Marx (Eds.). *Configurational Theory and Methods in Organizational Research.* Bingley: Emerald, 23-47.

Marx, Axel, Benoît Rihoux, and Charles C. Ragin. 2014. "The Origins, Development, and Application of Qualitative Comparative Analysis: The First 25 Years." *European Political Science Review* 6 (1):115-42.

Matland, Richard E. (1998) "Women's Representation in National Legislatures: Developed and Developing Countries." *Legislative Studies Quarterly* 23 (1):109-25.

Mello, Patrick A. 2012. *A Critical Review of Applications in QCA and Fuzzy-Set Analysis and a 'Toolbox' of Proven Solutions to Frequently Encountered Problems.* New Orleans: APSA Annual Meeting 2012.

Meuer, Johannes. 2014. "Archetypes of Inter-Firm Relations in the Implementation of Management Innovation: A Set-theoretic Study in China's Biopharmaceutical Industry." *Organization Studies* 35 (1):121-45.

McAdam, Doug, Hilary S. Boudet, Jennifer Davis, Ryan J. Orr, W. Richard Scott, and Raymond E. Levitt. 2010. "Site Fights: Explaining Opposition to Pipeline Projects in the Developing World." *Sociological Forum* 25 (3):401-27.

McAllister, Ian, and Donley T. Studlar. 2002. "Electoral Systems and Women's Representation: A Long-Term Perspective." *Representation* 39 (1):227-49.

McLevey, John. 2014. "Think Tanks, Funding, and the Politics of Policy Knowledge in Canada." *Canadian Review of Sociology* 51 (1):54-75.

McVeigh, Rory, Carl Neblett, and Sarah Shafiq. 2006. "Explaining Social Movement Outcomes: Multiorganizational Fields and Hate Crime Reporting." *Mobilization: An International Journal* 11 (1):23-49.

Mellbourn, Anders. 1979. *Byråkratins ansikten: Rolluppfattningar hos svenska högre statstjänstemän*. Stockholm: LiberFörlag.

Melinder, Karin A. 2007. "Socio-Cultural Characteristics of High versus Low Risk Societies Regarding Road Traffic Safety." *Safety Science* 45 (3):397-414.

Morgan, Stephen L., and Christopher Winship. 2007. *Counterfactuals and Causal Inference. Methods and Principles for Social Research.* Cambridge: Cambridge University Press.

Moritz, Mark, Julia Giblin, Miranda Ciccone, Andréa Davis, Jesse Fuhrman, Masoumeh Kimiaie, Stefanie Madzsar, Kyle Olson, and Matthew Senn. 2011. "Social Risk-Management Strategies in Pastoral Systems: A Qualitative Comparative Analysis." *Cross-Cultural Research* 45 (3):286-317.

Ordanini, Andrea, and Paul P. Maglio. 2009. "Market Orientation, Internal Process, and External Network: A Qualitative Comparative Analysis of Key Decisional Alternatives in the New Service Development." *Decision Sciences* 40 (3):601-25.

Ordanini, Andrea, A. Parasuraman, and Gaia Rubera. 2014. "When the Recipe Is More Important Than the Ingredients: A Qualitative Comparative Analysis (QCA) of Service Innovation Configurations." *Journal of Service Research* 17 (2):134-49.

Paine, Jack. 2016. "Set-Theoretic Comparative Methods: Less Distinctive Than Claimed." *Comparative Political Studies* 49 (6):703-41.

Pajunen, Kalle. 2008. "Institutions and Inflows of Foreign Direct Investment: A Fuzzy-Set Analysis." *Journal of International Business Studies* 39 (4):652-69.

Park, Sung Ho. 2013. "Capital Openness, Monetary Integration, and Wage-Setting Coordination in Developed European Countries." *Economic and Industrial Democracy* 34 (4):637-66.

Paxton, Pamela M., and Melanie M. Hughes. 2014. *Women, Politics, and Power: A Global Perspective*. Thousand Oaks: Pine Forge Press.

Paxton, Pamela M., and Sheri Kunovich. 2003. "Women's Political Representation: The Importance of Ideology" *Social Forces* 82 (1):87-113.

Perry, Ashley M., and Mark J. Schafer. 2014. "Resilience in Louisiana FEMA Parks: A Person-Centered, Fuzzy-Set Analysis." *Sociological Spectrum* 34 (1):39-60.

Peters, Brainard G. 2005. *Institutional Theory in Political Science: The "New Institutionalism"*. London: Continuum.

Pierre, Jon. 2004. "Politicization of the Swedish Civil Service. A Necessary Evil – Or Just Evil?" In Peters, Brainard G., and Jon Pierre (Eds.). *Politicization of the Civil Service in Comparative Perspective: The Quest for Control*. London: Routledge, 41-54.

Premfors, Rune, and Göran Sundström. 2007. *Regeringskansliet*. Malmö: Liber.

Priore, Michael J. 2006. "Qualitative Research: Does it fit in Economics?" *European Management Review* 3 (1):17-21.

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

Ragin, Charles C. 1994. *Constructing Social Research. The Unity and Diversity of Method.* Thousand Oaks: Pine Forge Press.

Ragin, Charles C. 2000. *Fuzzy-set Social Science*. Chicago: University of Chicago Press.

Ragin, Charles C. 2004. "Turning the Tables: How Case-Oriented Research Challenges Variable-Oriented Research." In Brady, Henry E., and David Collier (Eds.). *Rethinking Social Inquiry. Diverse Tools, Shared Standards.* Lanham: Rowman & Littlefield, 123-38.

Ragin, Charles C. 2006. "Set Relations in Social Research: Evaluating Their Consistency and Coverage." *Political Analysis* 14 (3):291-310.

Ragin, Charles C. 2008. *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago: University of Chicago Press.

Ragin, Charles C. 2009: "Qualitative Comparative Analysis using Fuzzy Sets (fsQCA)." In Rihoux, Benoît, and Charles C. Ragin (Eds.). 2009. *Configurational Comparative Methods. Qualitative Comparative Analysis (QCA) and Related Techniques.* Thousand Oaks: Sage, 87-121.

Ragin, Charles C. 2014. "Comment: Lucas and Szatrowski in Critical Perspective." *Sociological Methodology* 44 (1): 80-94.

Ragin, Charles, and Sean Davey. 2014. *fs/QCA [Computer Programme]*. Irvine: University of California.

Ragin, Charles C., and Peer C. Fiss. 2009. "Net Effects Analysis Versus Configurational Analysis: An Empirical Demonstration." In Ragin, Charles C. (Eds.). *Redesigning Social Inquiry: Set Relations in Social Research.* Chicago: University of Chicago Press, 190-212.

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

Raunio, Tapio. 2005. "Holding Governments Accountable in European Affairs: Explaining Cross-national Variation." *Journal of Legislative Studies* 11 (3/4):319-42.

Raunio, Tapio. 2009. "National Parliaments and European Integration: What We Know and Agenda for Future Research." *Journal of Legislative Studies* 15 (4):317-34.

Raunio, Tapio. 2011. "The Gatekeepers of European Integration? The Functions of National Parliaments in the EU Political System." *Journal of European Integration* 33 (3):303-21.

Rihoux, Benoît, Priscilla Álamos-Concha, Damian Bol, Axel Marx, and Ilona Rezsöhazy. 2013. "From Niche to Main-stream? A Comprehensive Mapping of QCA Applications in Journal Articles from 1984 to 2011." *Political Research Quarterly* 66 (1):175-84.

Rihoux, Benoît, and Charles C. Ragin. (Eds.). 2009. *Configurational Comparative Methods. Qualitative Comparative Analysis (QCA) and Related Techniques.* Thousand Oaks: Sage.

Rihoux, Benoît, Ilona Rezsöhazy, and Damien Bol. 2009. "Qualitative Comparative Analysis (QCA) in Public Policy Analysis: an Extensive Review." *German Political Studies* 7 (3): 9-82.

Roberts, Andrew, Jason Seawright, and Jennifer Cyr. 2013. "Do Electoral Laws Affect Women's Representation?" *Comparative Political Studies* 46 (12):1555-81.

Roberts, Ros, Richard Gott, and Judith Glaesser. 2010. "Students' Approaches to Open-Ended Science Investigation: The Importance of Substantive and Procedural Understanding." *Research Papers in Education* 25 (4):377-407.

Rohlfing, Ingo. 2012. *Case Studies and Causal Inference. An Integrative Framework*. Basingstoke: Palgrave Macmillan.

Rohlfing, Ingo. 2015a. "Why Simulations Are Appropriate for Evaluating Qualitative Comparative Analysis." *Quality & Quantity*. DOI: 10.1007/s11135-015-0251-8.

Rohlfing, Ingo. 2015b. "Mind the Gap: A Review of Simulation Designs for Qualitative Comparative Analysis." *Research & Politics* 2 (4). DOI: 10.1177/2053168015623562

Rohlfing, Ingo, and Carsten Q. Schneider. 2013. "Formalized Case Selection for Process Tracing after QCA: The Case of Necessity." *Political Research Quarterly* 66 (1):220-35.

Rosen, Jennifer. 2012. "The Effects of Political Institutions on Women's Political Representation. A Comparative Analysis of 168 Countries from 1992-2010." *Political Research Quarterly* 66 (2):306-21.

Rosenbluth, Frances M., Rob Salmond, and Michael F. Thies. 2006. "Welfare Works: Explaining Female Legislative Representation." *Politics and Gender* 2 (2):165-92.

Ruedin, Didier. 2012. "The Representation of Women in National Parliaments. A Cross-National Comparison." *European Sociological Review* 28 (1):96-109.

Rueschemeyer, Dietrich. 2003. "Can One or a Few Cases Yield Theoretical Gains?" In Mahoney, James, and Dietrich Rueschemeyer (Eds.). *Comparative Historical Analysis in the Social Sciences*. Cambridge: Cambridge University Press, 305-36.

Saalfeld, Thomas, and Wolfgang C. Müller. 1997. "Roles in Legislative Studies: A Theoretical Introduction." *Journal of Legislative Studies* 3 (1):1-16.

Savage, Scott V., and Ryan Seebruck. 2015. "The Road to the Top: A Qualitative Comparative Analysis of Mobility in the Elite Labor Market of College Basketball Coaching." *Sociological Focus* 48 (3):208-27.

Scharpf, Fritz W. 1997. *Games Real Actors Play: Actor-centered Institutionalism In Policy Research*. New York: Westview Press.

Schneider, Carsten Q., and Kristin Makszin. 2014a. "Forms of Welfare Capitalism and Education-Based Participatory Inequality." *Socio-Economic Review* 12 (2):437-62.

Schneider, Carsten Q., and Kristin Makszin. 2014b. *Online Appendix to Forms of Welfare Capitalism and Education-Based Participatory Inequality*. Available at http://ser.oxfordjournals.org/content/suppl/2014/03/25/mwu010.DC1, accessed 12 July 2015.

Schneider, Carsten Q., and Ingo Rohlfing, 2014: "Clarifying Misunderstandings, Moving Forward: Towards Standards and Tools in Set-Theoretic Methods." Qualitative and Multi-Method Research. *Newsletter of the APSA Section for Qualitative and Multi-Method Research* 12 (1):27-34.

Schneider, Carsten Q., and Ingo Rohlfing. 2013. "Combining QCA and Process Tracing in Set-Theoretic Multi-Method Research." *Sociological Methods & Research* 42 (4):559-97.

Schneider, Carsten Q., and Claudius Wagemann. 2007. *QCA und fsQCA. Ein einführendes Lehrbuch für Anwender und jene, die es werden wollen.* Opladen: Barbara-Budrich-Verlag.

Schneider, Carsten Q., and Claudius Wagemann. 2008. "Standards guter Praxis in Qualitative Comparative Analysis (QCA) und Fuzzy-Sets." In Pickel, Susanne, Gerd Pickel, Hans-Joachim Lauth, and Detlef Jahn (Eds.). *Methoden der vergleichenden Politik- und Sozialwissenschaft: Neue Entwicklungen und Anwendungen*. Wiesbaden: Verlag für Sozialwissenschaften, 361-86.

Schneider, Carsten Q., and Claudius Wagemann. 2010. "Standards of Good Practice in Qualitative Comparative Analysis (QCA) and Fuzzy Sets." *Comparative Sociology* 9 (3):397-418.

Schneider, Carsten Q., and Claudius Wagemann. 2012. *Set-theoretic methods for the social sciences: A guide to qualitative comparative analysis*. Cambridge: Cambridge University Press.

Schneider, Carsten Q., and Claudius Wagemann. 2013. "Doing Justice to Logical Remainders in QCA. Moving Beyond the Standard Analysis." *Political Research Quarterly* 66 (1):211-20.

Schneider, Martin R., Conrad Schulze-Bentrop, and Mihai Paunescu. 2010. "Mapping the Institutional Capital of High-Tech Firms: A Fuzzy-Set Analysis of Capitalist Variety and Export Performance." *Journal of International Business Studies* 41 (2):246-66.

Schwellnus, Guido. 2013. *Eliminating the Influence of Irrelevant Cases on the Consistency and Coverage of Necessary and Sufficient Conditions in Fuzzy-Set QCA*. Bordeaux*:* ECPR General Conference 2013.

Searing, Donald D. 1991. "Roles, Rules, and Rationality in the New Institutionalism." *American Political Science Review* 85 (4):1239-60.

Searing, Donald D. 1994. *Westminster's World: Understanding Political Roles*. Cambridge: Harvard University Press.

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

Seawright, Jason. 2014. "Comment: Limited Diversity and the Unreliability of QCA." *Sociological Methodology* 44 (1):118-21.

Seeleib-Kaiser, Martin, and Timo Fleckenstein. 2009. "The Political Economy of Occupational Family Policies: Comparing Workplaces in Britain and Germany." *British Journal of Industrial Relations* 47 (4):741-64.

Shanahan, Michael J., Lance D. Erickson, Stephen Vaisey, and Andrew Smolen. 2007. "Helping Relationships and Genetic Propensities: A Combinatoric Study of DRD2, Mentoring, and Educational Continuation." *Twin Research and Human Genetics* 10 (2):285-98.

Shanahan, Michael J., Stephen Vaisey, Lance D. Erickson, and Andrew Smolen. 2008. "Environmental Contingencies and Genetic Propensities: Social Capital, Educational Continuation, and Dopamine Receptor Gene DRD2." *American Journal of Sociology* 114 (S1):S260-S86.

Skaaning, Svend-Erik. 2011. "Assessing the Robustness of Crisp-Set and Fuzzy-Set QCA Results." *Sociological Methods & Research* 40 (2):391-408.

Skoko, Hazbo, Branka Krivokapic-Skoko, Marinko Skare, and Arnela Ceric. 2006. "ICT Adoption Policy of Australian and Croatian SMEs." *Managing Global Transitions* 4 (1):25-40.

SOEP (German Socio-Economic Panel Study). 2016. *Satisfaction With Life Today*. https://data.soep.de/soep-core/data/cp/cp9601, last access: 02.10.2016.

Stanko, Michael A., and Xavier Olleros. 2013. "Industry Growth and the Knowledge Spillover Regime: Does Outsourcing harm Innovativeness but help Profit?" *Journal of Business Research* 66 (10):2007-16.

Starr, Martha A. (2014). "Qualitative and Mixed-Methods Research in Economics: Surprising Growth, Promising Future." *Journal of Economic Surveys* 28 (2):238-64.

Stevenson, Ian. 2013. "Does Technology have an Impact on Learning? A Fuzzy Set Analysis of Historical Data on the Role of Digital Repertoires in Shaping the Outcomes of Classroom Pedagogy." *Computers & Education* 69 (0):148-58.

Stockemer, Daniel. 2013. "Fuzzy Set or Fuzzy Logic? Comparing the Value of Qualitative Comparative Analysis (fsQCA) versus Regression Analysis for the Study of Women's Legislative Representation." *European Political Science* 12 (1):86-101.

Stockemer, Daniel. 2014. "Women's Descriptive Representation in Developed and Developing Countries." *International Political Science Review* 36 (4):393-408.

Svevo-Cianci, Kimberly A., Stuart N. Hart, and Claude Rubinson. 2010. "Protecting Children from Violence and Maltreatment: A Qualitative Comparative Analysis Assessing the Implementation of U.N. CRC Article 19." *Child Abuse & Neglect* 34 (1):45-56.

Thiem, Alrik. 2013. "Clearly Crisp, and Not Fuzzy: A Reassessment of the (Putative) Pitfalls of Multi-value QCA." *Field Methods* 25 (2):197-207.

Thiem, Alrik. 2014a. "Navigating the Complexities of Qualitative Comparative Analysis: Case Numbers, Necessity Relations, and Model Ambiguities." *Evaluation Review* 38 (6):487-513.

Thiem, Alrik. 2014b. "Mill's Methods, Induction and Case Sensitivity in Qualitative Comparative Analysis: A Comment on Hug (2013)." *Qualitative & Multi-Method Research* 12 (2):19-24.

Thiem, Alrik. 2014c. "Membership Function Sensitivity of Descriptive Statistics in Fuzzy-Set Relations." *International Journal of Social Research Methodology* 17 (6):625-42.

Thiem, Alrik. 2016. *QCApro: Professional Functionality for Performing and Evaluating Qualitative Comparative Analysis. R Package Version 1.1-1. http://www.alrik-thiem.net/software/.*

Thiem, Alrik, and Adrian Duşa. 2013. *Qualitative Comparative Analysis with R: A User's Guide*. New York: Springer.

Thiem, Alrik, Michael Baumgartner, and Damien Bol. 2016. "Still Lost in Translation! A Correction of Three Misunderstandings between Configurational Comparativists and Regressional Analysts." *Comparative Political Studies* 49 (6):742-74.

Thomson Reuters. 2016. *2015 Journal Citation Reports® Social Sciences Edition*.

Tripp, Aili M., and Alice Kang. 2008. "The Global Impact of Quotas: The Fast Track to Female Representation." *Comparative Political Studies* 41 (3):338-61.

Trondal, Jarle. 2007. "The Public Administration Turn in Integration Research." *Journal of European Public Policy* 14 (6):960-72.

Trondal, Jarle, and Frode Veggeland. 2003. "Access, Voice and Loyalty: The Representation of Domestic Civil Servants in EU Committees." *Journal of European Public Policy* 10 (1):59-77.

Truman, David B. 1951. *The Governmental Process: Political Interests and Public Opinion*. Westport: Greenwood Press.

Urban, Dieter, and Jochen Mayerl. 2011. *Regressionsanalyse: Theorie, Technik und Anwendung*. Wiesbaden: VS Verlag.

Vaisey, Stephen. 2007. "Structure, Culture, and Community: The Search for Belonging in 50 Urban Communes." *American Sociological Review* 72 (6):851-73.

Vaisey, Stephen. 2009. "QCA 3.0: The "Ragin Revolution" Continues." *Contemporary Sociology. A Journal of Reviews* 38 (4):308-12.

Vaisey, Stephen. 2014. "Comment: QCA Works - When Used with Care." *Sociological Methodology* 44 (1):108-12.

Valliere, Dave, Ni Na, and Sean Wise. 2008. "Prior Relationships and M&A Exit Valuations: A Set-Theoretic Approach." *Journal of Private Equity* 11 (2):60-72.

Verkuilen, Jay. 2005. "Assigning Membership in a Fuzzy Set Analysis." *Sociological Methods & Research* 33 (4):462-96.

Vink, Maarten P., and Olaf van Vliet. 2013. "Potentials and Pitfalls of Multi-value QCA: Response to Thiem." *Field Methods* 25 (2):208-13.

Viterna, Jocelyn, Kathleen M. Fallon, and Jason Beckfield. 2008. "How Development Matters: A Research Note on the Relationship Between Development, Democracy and Women's Political Representation." *International Journal of Comparative Sociology* 49 (6):455-77.

Wagemann, Claudius. 2014. "Qualitative Comparative Analysis: What it Is, What it Does, and How it Works." In Della Porta, Donatella (Ed.). *Methodological Practices in Social Movement Rresearch*. Oxford: Oxford University Press, 43-66.

Wagemann, Claudius. 2015. "Qualitative Comparative Analysis." In Wenzelburger, Georg, and Reimut Zohlnhöfer (Eds.). *Handbuch Policyforschung*. Wiesbaden: VS Springer, 429-52.

Wagemann, Claudius, Jonas Buche, and Markus B. Siewert. 2015. "QCA and Business Research: Work in Progress or a Consolidated Agenda?" *Journal of Business Research* 69 (7):2531–40.

Wagemann, Claudius, and Carsten Q. Schneider. 2003. "Fuzzy Set Qualitative Comparative Analysis (fs/QCA): Ein Zwei-Stufen-Modul." In Pickel, Susanne, Gerd Pickel, Hans-Joachim Lauth, and Detlef Jahn (Eds.). *Vergleichende politikwissenschaftliche Methoden – Neue Entwicklungen und Diskussionen*. Opladen: Westdeutscher Verlag, 105-34.

Wagemann, Claudius, and Carsten Q. Schneider. 2015. "Transparency Standards in Qualitative Comparative Analysis." *Newsletter of the APSA Section for Qualitative and Multi-Method Research* 13 (1):38-42.

Wängnerud, Lena. 2009. "Women in Parliaments: Descriptive and Substantive Representation." *Annual Review of Political Science* 12 (1):51-69.

Wallin, Gunnar, Peter Ehn, Magnus Isberg, and Claes Linde. 1999. *Makthavare i fokus: Attityder och verklighetsuppfattningar hos toppskikten inom politik och förvaltning*. Stockholm: SNS förlag.

Webber, Douglas. 1992. "The distribution and use of policy knoledge in the policy process." In Dunn, Walter N., and Richard M. Kelly (Eds.). *Advances in Policy Studies since 1950*. New Brunswick: Transaction Publishers, 383-418.

Weissenberger-Eibl, Marion A., and Benjamin Teufel. 2013. "The Innovation Process as a Political Process - How Organisational Politics can Influence the Technological Newness of Product Innovations." *International Journal of Innovation Management* 17 (4):1-34.

Winand, Mathieu, Rihoux Benoît, David Qualizza, and Thierry Zintz. 2011. "Combinations of Key Determinants of Performance in Sport Governing Bodies." *Sport, Business and Management: An International Journal* 1 (3):234-51.

Winand, Mathieu, Benoît Rihoux, Leigh Robinson, and Thierry Zintz. 2013. "Pathways to High Performance: A Qualitative Comparative Analysis of Sport Governing Bodies." *Nonprofit and Voluntary Sector Quarterly* 42 (4):739-62.

Wollebæk, Dag. 2010. "Volatility and Growth in Populations of Rural Associations." *Rural Sociology* 75 (1):144-66.

Woodside, Arch G., Shih-Yun Hsu, and Roger Marshall. 2011. "General Theory of Cultures' Consequences on International Tourism Behavior." *Journal of Business Research* 64 (8):785-99.

Woodside, Arch, and Mann Zhang. 2012. "Identifying X-Consumers Using Causal Recipes: "Whales" and "Jumbo Shrimps" Casino Gamblers." *Journal of Gambling Studies* 28 (1):13-26.

Woodside, Arch G., and Mann Zhang. 2013. "Cultural Diversity and Marketing Transactions: Are Market Integration, Large Community Size, and World Religions Necessary for Fairness in Ephemeral Exchanges?" *Psychology & Marketing* 30 (3):263-76.

Workman, Samuel, Bryan D. Jones, and Ashley E. Jochim. 2009. "Information Processing and Policy Dynamics." *Policy Studies Journal* 37 (1):75-92.

Wright, Rachel A., and Hilary Schaffer Boudet. 2012. "To Act or Not to Act: Context, Capability, and Community Response to Environmental Risk." *American Journal of Sociology* 118 (3):728-77.

Wu, Pei-Ling, Shih-Shuo Yeh, Tzung-Cheng Huan, and Arch G. Woodside. 2014. "Applying Complexity Theory to Deepen Service Dominant Logic: Configural Analysis of Customer Experience-and-Outcome Assessments of Professional Services for Personal Transformations." *Journal of Business Research* 67 (8):1647-70.

Young, Raymond, and Simon Poon. 2013. "Top Management Support - Almost always Necessary and sometimes Sufficient for Success: Findings from a Fuzzy Set Analysis." *International Journal of Project Management* 31 (7):943-57.

Zadeh, Lofti A. 1965. "Fuzzy Sets." *Information and Control* 8 (3):338-53.

Zadeh, Lofti A. 1969. "Fuzzy algorithms." *Information and Control* 12(1): 99-102.

Zadeh, Lofti A. 1978. "Fuzzy Sets as a Basis for a Theory of Possibility." *Fuzzy Sets and Systems* 1 (1):3-28.

Lebenslauf Jonas Buche

Wissenschaftlicher Werdegang

05/2017	Promotion zum Dr. phil. an der Leibniz Universität Hannover
Seit 11/2016	Wissenschaftsmanager im Präsidialstab der Leibniz Universität Hannover / Referat für Berufungsangelegenheiten
10/2016 – 11/2016	Wissenschaftlicher Mitarbeiter am Deutschen Zentrum für Hochschul- und Wissenschaftsforschung (DZHW) Hannover
Seit 10/2013	Freiberuflicher Dozent
10/2012 – 09/2016	Wissenschaftlicher Mitarbeiter an der Professur für Gesellschaftswissenschaften mit dem Schwerpunkt qualitative empirische Sozialforschung, Goethe-Universität Frankfurt am Main
08/2009 – 09/2012	Forschungs- und Sektionsreferent am Deutschen Forschungsinstitut für öffentliche Verwaltung (FÖV) Speyer
05/2009	Magister Artium an der Universität Kiel
10/2003 – 05/2009	Magisterstudium der Politischen Wissenschaft, Neuesten Geschichte und Philosophie an den Universitäten Kiel und Göteborg (S)

Publikationen

Buche, Jonas. 2017. "Europeanization of Legislative-executive Relations at the Micro Level – Under Which Conditions Do Swedish MPs Interact with Ministerial Officials?" *COMPASSS Working Paper Series* 2017-87.

Buche, Antje, Jonas Buche und Markus B. Siewert. 2016. "Fuzzy Logic or Fuzzy Application? A Response to Stockemer's "Fuzzy Set or Fuzzy Logic?"" *European Political Science* 15(2): 359-78.

Buche, Jonas, and Markus B. Siewert. 2015. "Qualitative Comparative Analysis (QCA) in der Soziologie – Perspektiven, Potentiale und Anwendungsbereiche." *Zeitschrift für Soziologie* 44 (6):386-406.

Wagemann, Claudius, Jonas Buche, and Markus B. Siewert. 2015. "QCA and Business Research: Work in Progress or a Consolidated Agenda?" *Journal of Business Research* 69 (7):2531-40.

Buche, Jonas. 2013. "Europäisierung parlamentarischer Kontrolle im Norden Europas: Dänemark, Finnland und Schweden im Vergleich" In Eberbach-Born, Birgit, Sabine Kropp, Andrej Stuchlik, and Wolfgang Zeh (Eds.). *Parlamentarische Kontrolle und Mitsteuerung unter den Bedingungen der Europäisierung*. Baden-Baden: Nomos, 365-94.

Kropp, Sabine, Aron C. Buzogány, and Jonas Buche. 2012. "Von den Schwierigkeiten, Zusammengehöriges zu vereinen - Nationale Parlamente und Exekutiven als Gegenstand der Europäisierungsforschung." *Politische Vierteljahresschrift* 53 (1):109-34.

Kropp, Sabine, Jonas Buche, and Aron C. Buzogány. 2011. "Parlamentarisch-exekutive Steuerung in europäisierten Fachpolitiken – ein Blick auf die Mikroebene." In Abels, Gabriele, and Annegret Eppler (Eds.). *Auf dem Weg zum Mehrebenenparlamentarismus?: Funktionen von Parlamenten im politischen System der EU*. Baden-Baden: Nomos, 227-42.

Buche, Jonas. 2010. "Die Europäisierung von Parteien und Parteisystemen – Eine Analyse am Beispiel Schwedens vom Beitritt zur EU 1995 bis zur Reichstagswahl 2006." *FÖV Discussion Paper* 62.