EUI WORKING PAPERS

SPS No. 2006/06

Contradictions in
Qualitative Comparative Analysis (QCA):
Ways Out of the Dilemma

ANDREA HERRMANN AND
LASSE CRONQVIST

EUROPEAN UNIVERSITY INSTITUTE
Political and Social Sciences Department

EUROPEAN UNIVERSITY INSTITUTE
DEPARTMENT OF POLITICAL AND SOCIAL SCIENCES

Contradictions in Qualitative Comparative Analysis (QCA): Ways Out of the Dilemma

ANDREA HERRMANN AND LASSE CRONQVIST

EUI Working Paper **SPS** No. 2006/06

ISSN 1725-6755

Printed in Italy
European University Institute
Badia Fiesolana
I – 50016 San Domenico di Fiesole (FI)
Italy

http://www.iue.it/
http://cadmus.iue.it/dspace/index.jsp

Abstract

This paper aims at illustrating the circumstances in which QCA and its ramifications, fs/QCA and MVQCA, become particularly useful tools of analysis. To this end, we discuss the most pertinent problem which researchers encounter when using QCA, namely the problem of contradicting observations. In QCA analysis, contradictions arise from the sheer number of cases, as well as from the problem of dichotomisation. Therefore, we argue that, in order to handle contradictions, the method for analysing middle-sized-N situations should be chosen according to two parameters; the size of a dataset on the one hand, and the need to preserve raw-data information on the other. While QCA is an apt tool for analysing comparatively small middle-sized datasets with a correspondingly reduced necessity to preserve cluster information, the opposite holds true for fs/QCA. MVQCA strikes a balance between these two methods as it is most suitable for analysing genuinely middle-sized case sets for which some cluster information needs to be preserved

Authors:

Andrea Herrmann (European University Institute, Florence)
Lasse Cronqvist (University of Marburg)

We wish to thank Dirk Berg-Schlosser, Jaap Dronkers, Charles Ragin, Benoît Rihoux, Simon Toubeau, and Sakura Yamasaki for stimulating discussions and their comments on earlier versions of this paper.

1. Introduction

Research in the social sciences seems to be, at least partly, guided by a continuing *Methodenstreit* about the superiority of either quantitative or qualitative methods. On the one hand, scholars using qualitative or case-oriented methods argue that an in-depth understanding of a small number of cases is vital when attempting to understand causal complexity (see inter alia Muno 2003; Munck 2004). On the other hand, researchers preferring quantitative or variable-oriented methods claim that only the study of a large number of cases allows one to make reliable statements about (causal) relationships (see for example King, Keohane, and Verba 1994). From the authors' point of view, it is deplorable that social scientists seem to have divided into these two camps, because the strict adherence to one or the other group entails the risk that the preferred type of methods determines how a research question is posed. Ideally, however, the question to be explored determines the choice of method: Each question directs the researcher to a population of cases out of which s/he chooses the most representative sample. Depending on data availability and sample size, the researcher then chooses the most adequate method for analysing her dataset. In sum, the research question should determine the choice of method – not the other way round.

Several attempts have been made to bridge the methodological divide between qualitative and quantitative analyses (see e.g. Campbell 1975; Eckstein 1975; Lijphart 1971; Lijphart 1975; Smelser 1976). Probably the most renowned proposal has been advanced by King, Keohane and Verba (1994) seeking to apply a large-N logic to the analysis of a small number of cases. It must however be questioned whether it is both possible and fruitful to merge various aspects of different methods in the attempt to obtain 'superior' analytical tools. Both quantitative and qualitative methods have different features which make them more or less suitable for the analysis of a certain number of cases. The logic of a qualitative method such as process-tracing, particularly well-suited for the fine-grained analysis of one or very few cases, can hardly be transferred to a large number of cases as an in-depth analysis would be inherently difficult. In a similar vein, the effort of King, Keohane and Verba to apply the statistics-based approach of large-N methods to the analysis of a small number of cases has been seriously questioned (see McKeown 1999; Munck 1998; Ragin 2000: 14; Collier, Seawright, and Munck 2004). In sum, each method has specific characteristics which are advantageous for the analysis of one research scenario, while being disadvantageous for the analysis of another scenario.

Therefore, instead of trying to merge existing methods, it seems more promising to design new ones. A wide range of methods for the analysis of small-N[1] and large-N situations[2] exists and is under constant development (see for example Katz and Beck 2004). But only a few tools have been developed for the analysis of

[1] Following the suggestion of Charles Ragin (see Ragin 2003: 13), we use the notion of 'small-N' for samples which include 1 to 4 cases. Examples of small-N methods are hermeneutics, in-depth interviews or long-term observations (process-tracing).

[2] Following the suggestion of Charles Ragin (see Ragin 2003: 13), we use the notion of 'large-N' for samples which include more than 50 cases. Examples of large-N methods are regression analysis and its various ramifications.

middle-sized-N situations[3]. Today, the most prominent of these tools is Charles Ragin's Qualitative Comparative Analysis (henceforth 'QCA'). It was in 1987 that Charles Ragin introduced this method to the public (see Ragin 1987). Extensions to QCA have recently been proposed, leading to the naissance of *fuzzy-set QCA* (henceforth 'fs/QCA') on the one hand (Ragin 2000), and of *Multi-Value QCA* (henceforth 'MVQCA') on the other (Cronqvist 2004; Cronqvist 2005a). It was, however, not later than 1990 that QCA started to encounter harsh, and often unfounded, criticism (see for example Markoff 1990: 179). In line with the '*Methodenstreit* paradigm', various scholars pointed to the weaknesses of QCA, seeking to portray the latter as inferior to the more traditional methods (see De Meur and Rihoux 2002: 119-144).

This paper aims at illustrating under which conditions QCA and its ramifications, fs/QCA and MVQCA, are particularly useful tools of analysis. This is done by discussing the problem of 'contradictions' which constitutes the most persistent difficulty a researcher faces when using QCA. However, in contrast to the *Methodenstreit* paradigm, it is by no means our aim to portray (a ramification of) QCA as superior to any other qualitative or quantitative method. As argued above, such discussions seem inherently fruitless to us. Instead, by illustrating their different features, we aim at presenting QCA, fs/QCA and MVQCA as genuine alternatives to the more traditional qualitative (small-N) and quantitative (large-N) methods.

Overall, we argue that the explanatory power of a QCA, fs/QCA, and MVQCA analysis is a function of two parameters; The size of a case set on the one hand, and the necessity to preserve the richness of raw-data information on the other. Accordingly, a researcher should use QCA for analysing rather small middle-sized case sets whose values can be converted into dichotomous scores without a loss of important cluster information. Fs/QCA, instead, is most useful whenever a researcher wishes to analyse a comparatively large middle-sized case set which requires to preserve rich raw-data information. MVQCA, in turn, strikes a balance between QCA and fs/QCA as it constitutes the most suitable method for analysing genuinely middle-sized case sets which necessitate the conservation of some raw-data information.[4]

To illustrate these arguments, the paper is organised as follows: Section 2 briefly introduces the logic of QCA and the problem of contradicting observations which can notably limit the explanatory power of QCA. Section 3 illustrates how fs/QCA addresses the problem of contradictions, and points to the limits of this method. Section 4 shows in which circumstances MVQCA succeeds in striking a balance between QCA and fs/QCA. Section 5 concludes the analysis by summarising our argument.

2. QCA – a powerful tool for analysing middle-sized datasets?

Various scholars (see Ragin 2000: 25; see also Bollen, Entwisle, and Alderson 1993; Ragin 1989; Sigelman and Gadbois 1983) have illustrated that research in the social sciences is dominated by the analysis of either small-N or large-N situations,

[3] Consequent to the remarks of footnote 1 and 2, the notion of 'middle-sized-N' refers to samples which include 5 to 50 cases (see Ragin 2003: 13).

[4] We wish to stress that we do not want to define *small-size*, *genuinely middle-sized*, and *large-size middle-sized datasets* by suggesting precise numbers. The reason is that these definitions depend on the individual research design, i.e. the number and the conceptual richness of causal and outcome variables.

whereas very few research is carried out on the basis of a middle-sized number of cases. Until 1987, when Charles Ragin introduced QCA to the public (Ragin 1987), this bias towards the use of qualitative, and respectively quantitative techniques was surely aggravated by the lack of a method that was capable of assessing middle-sized case sets adequately. Probably the most important benefit of QCA therefore resides in the fact that it constitutes a powerful tool for the analysis of middle-sized-N situations.

Yet, we argue that, in two respects, QCA is a more limited tool of analysis for middle-sized-N situations than Ragin's text-book (Ragin 1987) might suggest. Firstly, QCA is particularly well-suited only for comparatively small middle-sized-N situations. Since QCA ignores the frequency with which a causal combination occurs, the likelihood of contradictions increases with the number of cases. Secondly, QCA is an inadequate tool whenever raw data cannot be recoded into dichotomous variables without a loss of important information. This so-called problem of dichotomisation also entails the risk of contradictions and, hence, of a situation in which a parsimonious solution only covers a small number of studied cases. Therefore, we argue that QCA is a particularly useful tool of analysis for a small middle-sized case set with a reduced necessity to preserve the raw data's richness of information.

To illustrate our argument, we use a dataset derived from the studies of Tatu Vanhanen (see Vanhanen 1984) which Berg-Schlosser and De Meur analysed in one of the first published applications of QCA (Berg-Schlosser and De Meur 1994). Studying the causes of breakdown of democratic regimes in the interwar period, Vanhanen constructed three socio-economic indices which he identified as pillars of democratisation (see table 1). While the first index (*Index of Occupational Diversification – IOD*) reports the arithmetic mean of urban population and non-agricultural population in a country, the second index (*Index of Knowledge Distribution – IKD*) combines measures of literacy and university education. The third measure (*Family Farms – FF*) indicates the percentage of family-sized landholding as a percentage of the total area of holdings (Vanhanen 1984: 38). In line with Berg-Schlosser and De Meur (1994), we focus our analysis on sixteen cases which comprise 'all (…) major 'breakdown'-cases (…) [as well as] the major 'survivors', including some of the smaller countries which often tend to be overlooked' (Berg-Schlosser and De Meur 1994: 254)[5]. To keep explanations simple, we only perform QCA, fs/QCA, and MVQCA analyses for those cases in which democracy collapsed during the interwar period. Accordingly, we assign a score of 1 to all 'democratic breakdown countries', while we assign a score of 0 to those countries in which democracy endured the interwar-period.

[5] Accordingly, our analysis includes Austria (AUS), Belgium (BEL), Czechoslovakia (CZE), Finland (FIN), France (FRA), Germany (GER), Greece (GRE), Hungary (HUN), Italy (ITA), the Netherlands (NET), Poland (POL), Portugal (POR), Romania (ROM), Spain (SPA), Sweden (SWE), and the United Kingdom (UK).

Table 1: Raw dataset on causes of democracy breakdown in the interwar period

Case	IOD (Index of Occupational Diversification)	IKD (Index of Knowledge Distribution)	FF (Family Farms)	Outcome (Breakdown of Democracy)
AUS	51,5	55	45	1
BEL	64	51,5	30	0
CZE	38,5	49	40	0
FIN	21,5	46,5	47	0
FRA	48	50,5	35	0
GER	53	54	54	1
GRE	34	28	28	1
HUN	37	47	40	1
ITA	38	39,5	22	1
NET	61	51,5	40	0
POL	17,5	37,5	53	1
POR	30,5	18,5	20	1
ROM	16,5	25	41	1
SPA	35	33	20	1
SWE	39,5	52,5	50	0
UK	78,5	50	25	0

Source: Vanhanen (1984)

It is not the aim of this paper to review how a QCA, fs/QCA, and MVQCA analysis are carried out in detail. Instead, we limit our illustrations to those analytical steps which are central to the understanding of our argument. Thus, suffice it to say here that a QCA analysis is carried out in four steps which we will exemplify on the basis of the Vanhanen dataset.

In a nutshell, QCA consists in applying the logic of Mill's method of difference so as to reduce causal complexity (see Mill 1872: 451-452). Once a researcher has determined which cases s/he wants to study, the **first step** consists in drawing up a summary table that recapitulates – for each case – whether the respective causal conditions and the outcome are present or absent. Importantly, a QCA analysis can *only* be carried out on the basis of dichotomous variables. Therefore, any ordinal or scale variables of the raw dataset must be recoded into dichotomous scores. Turning back to the Vanhanen dataset presented in table 1, we see that all three independent variables need to be recoded. In so doing, and contrary to Berg-Schlosser and De Meur (1994), we use a cut-off value of 45 for both IOD and IKD, and a cut-off value of 38 for FF because an in-depth cluster analysis shows that these thresholds are most representative. Table 2 reports the results obtained from recoding Vanhanen's raw dataset into dichotomous scores.

Table 2: QCA summary table on causes of democracy breakdown in the interwar period

Case	**IOD** (Index of Occupational Diversification)	**IKD** (Index of Knowledge Distribution)	**FF** (Family Farms)	**Outcome** (Breakdown of Democracy)
AUS	**1**	**1**	**1**	**1**
BEL	**1**	**1**	**0**	**0**
CZE	**0**	**1**	**1**	**0**
FIN	**0**	**1**	**1**	**0**
FRA	**1**	**1**	**0**	**0**
GER	**1**	**1**	**1**	**1**
GRE	**0**	**0**	**0**	**1**
HUN	**0**	**1**	**1**	**1**
ITA	**0**	**0**	**0**	**1**
NET	**1**	**1**	**1**	**0**
POL	**0**	**0**	**1**	**1**
POR	**0**	**0**	**0**	**1**
ROM	**0**	**0**	**1**	**1**
SPA	**0**	**0**	**0**	**1**
SWE	**0**	**1**	**1**	**0**
UK	**1**	**1**	**0**	**0**

Source: Vanhanen (1984), recoded as described in the text

The **second** step consists in converting the obtained dataset into a so-called 'truth table' which lists all logically possible combinations of causes. Accordingly, a QCA truth table contains 2^k rows of possible causal combinations, whereby k stands for the number of causal conditions (Ragin 1987: 87-89). It is important to note that the predominant concern of QCA is akin to qualitative methods in that it only recapitulates *whether* a causal combination is observed and which *result* the latter produces. However, no attention is paid to the *number of times* a certain combination occurs. Converting the dichotomous Vanhanen dataset into a truth table leads to the outcome presented in table 3.

Table 3: QCA truth table on causes of democracy breakdown in the interwar period

Case	**IOD** (Index of Occupational Diversification)	**IKD** (Index of Knowledge Distribution)	**FF** (Family Farms)	**Outcome** (Breakdown of Democracy)	**Number of observed cases**
Por, Gre, Spa, Ita	0	0	0	1	4
Rom, Pol	0	0	1	1	2
	0	1	0	?	0
Fin, Cze, Swe Hun	0	1	1	1 : 0 (Contradiction)	3 : 1
	1	0	0	?	0
	1	0	1	?	0
Fra, Bel, Uk	1	1	0	0	3
Net Aus, Ger	1	1	1	1 : 0 (Contradiction)	1 : 2

Source: Vanhanen (1984), recoded and summarised as described in the text

In a **third step**, the researcher resorts to Boolean algebra so as to derive the lowest common denominator of causal conditions which produce a certain outcome. Akin to Mill's method of difference (see Mill 1872: 453), the fundamental rule for reducing causal complexity is the so-called 'minimisation rule': 'If two Boolean expressions differ in only one causal condition yet produce the same outcome, then the causal condition that distinguishes the two expressions can be considered irrelevant and can be removed to create a simpler, combined expression' (Ragin 1987: 93). Using this rule iteratively, so-called 'prime implicants' are determined which cover only those cases that lead to the studied outcome. These prime implicants are then combined to the shortest possible solutions so as to explain all cases with the outcome in question (see Ragin 1987: 92-101). If we apply this procedure to the Vanhanen dataset – thereby including all logical remainders into, and excluding all contradictions from the minimisation procedure of our analysis – we obtain the following Boolean equation:

1 (Breakdown of Democracy) = ikd

In (other) words, breakdown of democratic regimes results from an unequal knowledge distribution.

The last step of a QCA analysis consists in interpreting the obtained result with regard to its sufficiency and/ or necessity (see Ragin 1987: 99-100). For our example, we find that the absence of an equal knowledge distribution is both a necessary and a sufficient criterion for the failure of democratic regimes. In other words, the Vanhanen dataset shows that democratic regimes came to an end during the interwar period whenever knowledge was concentrated among a small elite of people.

This highly parsimonious explanation for democracy breakdown seems to suggest that QCA is an ideal tool for analysing middle-sized datasets. However, it is important to note that this parsimony could only be obtained by excluding all contradicting cases from the QCA minimisation procedure. This, in turn, indicates that QCA is a more limited analytical tool than Ragin's textbook (1987) suggests.

The problem of contradicting observations in QCA analysis

While QCA has encountered much unfounded criticism (see De Meur and Rihoux 2002: 123-141), the difficulty of dealing with contradicting cases often keeps researchers from using QCA on a more than experimental level. We argue that this makes QCA a more limited tool for the analysis of middle-sized-N situations than one might think when reading Ragin's 1987 textbook. More precisely, we argue that QCA is a particularly useful method only in the study of comparatively small middle-sized datasets which can be transformed into dichotomous scores without a loss of important (cluster-) information.

To understand this argument, it firstly is important to note that the occurrence of contradictions does not constitute a problem *per se*. On the contrary, the fact that a researcher needs to take a decision about how to deal with contradicting observations constitutes a particular strength of QCA (see Ragin 1987: 113-118). In essence, contradictory observations indicate that the researcher's analysis is still incomplete to the extent that independent variables, and/or the outcome variable require further elaboration and redefinition. Accordingly, one way in which a researcher can solve contradictions consists in going back to her case set so as to complete the analysis by increasing the number of independent variables (Ragin 1987: 113-116).

While we do not wish to question the exploratory potential of QCA, it should be noted that one theoretical and one practical difficulty persist relating to the aforementioned procedure of addressing contradictions. On a theoretical level, each added variable leads to an exponential increase in possible causal combinations. Hence, for a constant number of cases, the number of unobservable causal combinations also increases accordingly. This constitutes a problem to the extent that it often becomes difficult to obtain parsimonious causal explanations. Taking this argument to its extreme, a researcher can end up with an individual explanation for each considered case. On a practical level, any research project is faced with financial and time constraints. At some point, a researcher simply must end the empirical phase and make sense of the hitherto gathered data. Similarly, a researcher may wish to analyse an already existing dataset for which no further data can be gathered. It is, precisely, in these situations that our argument becomes relevant.

Once a researcher has to come to terms with the existing data, s/he often deals with contradicting cases by excluding them (Ragin 1987: 116)[6]. We chose this solution when analysing the Vanhanen dataset (see above). However, applying such procedure usually means that only a limited number of cases can be explained. Taking the Vanhanen dataset as an example, we see that 7 out of 16 cases were involved in contradicting observations. Excluding these cases from our QCA analysis, we found that democracy breakdown results from an unequal distribution of knowledge ('ikd'). Admittedly, this solution is very parsimonious. Importantly, though, it has been obtained from considering only 9 out of 16 countries, or 56% out of all observed cases.

Given the difficulties related to handling contradictory observations, it becomes clear that QCA is a particularly useful tool of analysis in situations where the probability of contradictions is minimal. But, which situations are relevant in this respect? To answer this question, it is important to note that in QCA analysis contradictions arise from two sources. Firstly, contradictions result from the sheer number of considered cases. Since QCA does not consider *how often* but only *that* a causal combination occurs, deviant cases are not identified as such. Taking the Vanhanen dataset as an example, we see that low occupational diversification, high knowledge distribution, and a high share of family-sized landholdings led to the persistence of democracy in Finland, Czechoslovakia and Sweden, whilst the same causal combination entailed the breakdown of the democratic regime in Hungary. This suggests that the Hungarian case deviates from the norm and, hence, requires a special explanation. Thus, the higher the number of considered cases, the higher the probability that deviant cases are included which, in turn, lead to contradictions. Hence, whenever a middle-sized case set contains a comparatively large number of cases, the risk of contradictory observations is rather high.

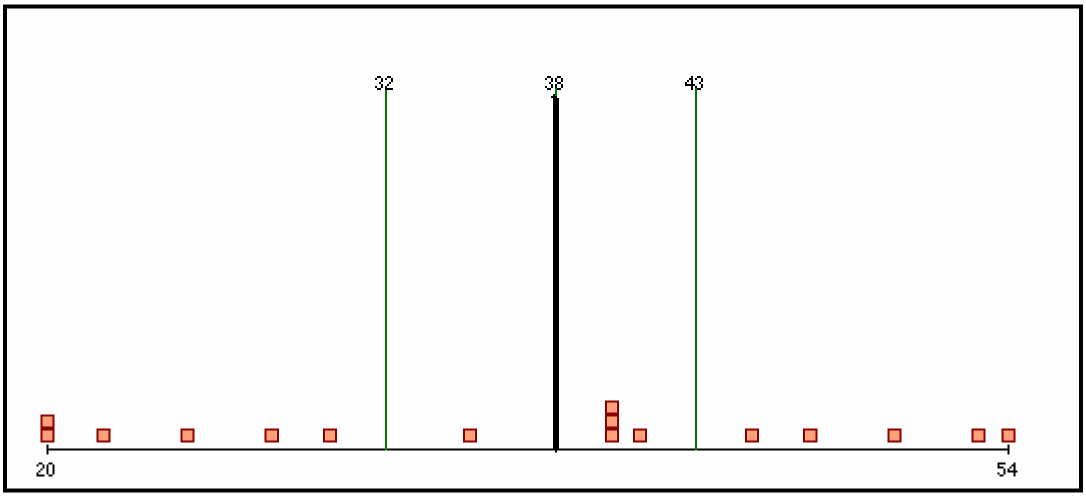
The second source of contradictions in QCA analysis consists in the loss of information whenever rich raw data is transformed into dichotomous scores. This problem, which is acute for ordinal and scale variables, has been criticised in the literature as the 'problem of dichotomisation' (see for example Bollen et al. 1993;

[6] Ragin points out that further possibilities exist to deal with contradictions. That is, a researcher can also decide to assign an outcome score of '0', or respectively '1' to all contradicting cases (Ragin 1987: 116-117). These procedures are, however, problematic in that they 'violate the spirit of case-oriented qualitative research. [Accordingly they] should be used only when it is impossible to return to the original cases and construct a better truth table.' (Ragin 1987: 118).

Goldthorpe 1997). Since QCA can only operate on the basis of dichotomous scores, it obliges a researcher to choose one threshold according to which s/he assigns a score of '0', or respectively '1'to the various cases. Yet, along a variable's scale, cases often cluster together in several groups. In these situations, the introduction of merely one threshold can lead to the loss of important cluster information because a suboptimal division between cases has to be made.

The 'family-farm' variable of the Vanhanen dataset exemplifies this argument. Figure 1 shows that the sixteen cases form roughly three clusters on this variable. Even though we performed a cluster analysis so as to choose the most representative threshold (namely 38), the division of cases into two groups cannot represent the richness of information contained in the raw dataset. Therefore, different scores are attributed to countries with close scale values on the FF-index. Consider for example France with a value of 35 in the raw dataset which was transformed into a dichotomous score of 0, and the Netherlands with a raw value of 40 which was transformed into a score of 1. On the other hand, cases such as Germany and the Netherlands are assigned the same dichotomous score (namely 1), even though the original values (namely 54 and 40) are rather distant. This suboptimal dichotomisation of raw data can be held responsible for the contradictions reported in the last line of table 3.

Figure 1: Distribution of raw data on the family farm index (FF)



Thus, to avoid contradictions in QCA analysis, the latter should only be used if the dichotomisation of raw data allows the preservation of cluster information contained in the original dataset. In a similar vein, contradicting observations are to be avoided by analysing only small middle-sized datasets. We therefore argue that QCA is a well-suited method for analysing rather small middle-sized datasets for which the dichotomisation of raw data comes without a loss of essential information.

Both fs/QCA and MVQCA have been developed as a response to the problem of contradictions. While both methods allow the minimisation, or even the elimination, of the risk of contradicting observations, we argue that they should not be used for the analysis of *any* middle-sized dataset. Akin to their forerunner, the explanatory power of an fs/QCA and MVQCA analysis depends on a dataset's size on the one hand, and on the necessity to conserve the richness of raw-data information on the other. By outlining the most important steps of an fs/QCA and an MVQCA analysis, the next sections seek to illustrate our argument.

3. Fs/QCA – Preserving rich raw-data information has its price [7]

Like QCA, fs/QCA has been designed as a tool for analysing middle-sized-N situations (see Ragin 2000). However, in line with our previous argument, we hold that fs/QCA is a more limited analytical tool than Ragin's textbook might suggest. More precisely, we agree that fs/QCA is a particularly useful method for analysing middle-sized datasets whose dichotomisation entails a loss of important (cluster-) information. But even more importantly, we argue that fs/QCA should only be used for analysing comparatively large middle-sized datasets. Otherwise, this method bears the risk of not revealing all causal conditions which provoke the studied outcome.

To illustrate this point, we will briefly review those steps of an fs/QCA analysis which are important for the understanding of our argument. Like QCA, fs/QCA proceeds in four steps. However, the analytical procedure is different apart from the first step. Akin to QCA, the **first step** of an fs/QCA analysis consists in transforming a raw dataset into so-called 'membership scores' in order to draw up a summary table. In contrast to QCA, fs/QCA does not require the transformation of the raw dataset into dichotomous scores. It allows the retention of the richness of data due to the use of decimal membership scores (Ragin 2000: 153-171). This, in turn, makes fs/QCA a particularly useful method for analysing middle-sized datasets which contain one (or more) ordinal and/or scale variable(s). It is important to note that membership scores are qualitative measures as they result from a researcher's deliberate decision about *how* to transform raw data into membership scores: Once a researcher has collected all necessary empirical evidence, s/he has to decide for each variable at which level to set the threshold for 'zero' membership (0.00) on the one hand, and for full membership (1.00) on the other. Furthermore, s/he also must decide whether to assign in-between membership scores in regular steps.

In order to transform the Vanhanen raw dataset into meaningful membership scores, we carried out a cluster analysis on the basis of a simple average linkage method. By calculating the distance between arithmetic means of various case groups, this method allows the determination of the most pronounced case clusters in a sample. Based on this analysis, we decided to transform the Vanhanen raw dataset into five-stepped fuzzy membership scores as reported in table 4.

Table 4: Conversion table

Raw-Data Values on…					
- IOD	<26	26 - 43	43 - 57	57 - 71	> 71
- IKD	<30	30 - 35	35 - 43	43 - 47,5	> 47,5
- FF	<23	23 - 33	33 - 43	43 - 48	> 48
…converted into the following…					
…fuzzy Membership Score	0	0.25	0.5	0.75	1

Source: Own Calculations based on Cluster Analysis (Simple Average Linkage Method)

[7] It should be noted that our argument refers to the characteristics of fs/QCA as presented in Ragin's textbook 'Fuzzy-Set Social Science' (Ragin 2000). To avoid unnecessary confusion, we deliberately decided not to consider most recent amendments (see Ragin 2004; Ragin 2006).

Table 5 reports the fuzzy membership scores which we obtained by transforming the Vanhanen raw dataset according to the conversion measures summarised in table 4.[8]

Table 5: Fs/QCA summary table on causes of democracy breakdown in the interwar period

Case	**IOD** (Index of Occupational Diversification)	**IKD** (Index of Knowledge Distribution)	**FF** (Family Farms)	**Outcome** (Breakdown of Democracy)
AUS	0.5	1	0.75	1
BEL	0.75	1	0.25	0
CZE	0.25	1	0.5	0
FIN	0	0.75	0.75	0
FRA	0.5	1	0.5	0
GER	0.5	1	1	1
GRE	0.25	0	0.25	1
HUN	0.25	0.75	0.5	1
ITA	0.25	0.5	0	1
NET	0.75	1	0.5	0
POL	0	0.5	1	1
POR	0.25	0	0	1
ROM	0	0	0.5	1
SPA	0.25	0.25	0	1
SWE	0.25	1	1	0
UK	1	1	0.25	0

Source: Vanhanen (1984), recoded into fuzzy membership scores as described in the text

Contrary to QCA, the second step of an fs/QCA analysis does not consist in drawing up a truth table. As fs/QCA allows for the preservation of the richness of raw-data information, this would be a futile enterprise. The use of decimal membership scores makes it very unlikely that two cases show exactly the same causal combination. Accordingly, it is not only inherently difficult to draw up a truth

[8] We wish to emphasize that the results obtained from data transformation as described in table 4 are stable. In this respect, it is important to note that different ways exist in which the original (Vanhanen) dataset can be transformed into membership scores. Another, statistically neutral way consists in assigning zero membership (0.00) to the lowest observed value, while full membership (1.00) is assigned to the highest value of each variable. All intermediary values are then converted proportionately: The lowest case value is deduced from each individual case-value; the so obtained figure is then divided by the difference between the highest and the lowest score. In other words, the following equation is applied to each variable:

$$\text{Membership Score} = \frac{\text{(Individual Value Raw Data – Minimum Value Raw Data)}}{\text{(Maximum Value Raw Data – Minimum Value Raw Data)}}$$

This way of determining membership scores is however susceptible to outliers, because the obtained membership scores will depict a distorted image if the case sample includes outliers which provide extreme maximum or minimum values. For this reason, we preferred determining membership scores on the basis of a cluster analysis using the simple average linkage method. Yet, we cross checked our results. In so doing, we found that the results reported in the remainder of this section are stable in that they do not change if membership scores are determined according to the aforementioned standardisation formula.

table; the probability that two cases are involved in a contradictory observation is also close to zero. Therefore, fs/QCA is not affected by the problem of contradicting cases.

In order to reduce causal complexity, fs/QCA and QCA basically proceed in opposite directions. We have seen above that QCA *first* uses the minimisation rule to reduce causal complexity, and *then* interprets the findings in light of their necessity and/or sufficiency. Fs/QCA, by contrast, *first* identifies all necessary and/or sufficient causal conditions, and *then* eliminates more complex expressions, covered by less complex expressions, with the help of the so-called containment rule (see Ragin 2000: 238-242). Accordingly, the **second step** of an fs/QCA consists in identifying all *necessary causes*, while the **third** step is concerned with isolating all *sufficient conditions*. It is important to note that, in order to identify all necessary and sufficient conditions, Ragin resorts to the use of probabilistic criteria (Ragin 2000: 107-115). More precisely, Ragin suggest to apply a binominal probability test for case sets of less than 30 cases, and a simple z-test for case sets of more than 30 cases (Ragin 2000: 111-112). The so obtained results are interpreted in the **fourth and final step** (see Ragin 2000: 238-246).

Let us apply these analytical steps to the Vanhanen dataset (as reported in table 5). Since this dataset contains less than 30 cases, we use a binominal test to identify first all necessary and then all sufficient conditions for democracy breakdown. In so doing, we use conventional probabilistic criteria, namely a .05 significance level and a benchmark proportion of .65. Furthermore, we decided to apply an adjustment factor of 0.3 which roughly represents the size of one step in our five-stepped membership scale. Interestingly, the result obtained from this fs/QCA analysis shows that democracy breakdown in the interwar period results from an uneven distribution of knowledge. Expressed in a Boolean equation, we find that:

1 (Breakdown of Democracy) = ~IKD

At first sight, this outcome seems reassuring as it is identical to the result obtained from the above QCA analysis. In other words, both a QCA and an fs/QCA analysis show that an uneven distribution of knowledge is both a necessary and sufficient condition for the breakdown of a democratic regime in the interwar period. But, let us pause for a moment to contemplate the reliability of this result.

Let us remember that the solution 'ikd', obtained from our QCA analysis, merely considered 9 out of 16 cases (see section 2). This suggests that the same fs/QCA solution also covers only a limited number of cases. In this regard, it is important to notice that the use of probabilistic criteria entails that certain causal combinations only qualify as necessary and/or sufficient conditions if a minimum number of consistent cases exist which pass the respective probabilistic test (see Ragin 2000: 113-115, in particular table 4.9.). For example, if a researcher uses a .10 significance level and a benchmark proportion of .50, a case set must contain at least 4 cases with a certain causal condition to make the latter qualify as a necessary/sufficient predictor of the outcome (see Ragin 2000: 114, table 4.9.).[9]

As a result, an fs/QCA analysis is unlikely to reveal all causes leading to the observed outcome if it is carried out on the basis of a small number of cases. This is, precisely, the reason for which '~IKD' qualifies as the only solution of our fs/QCA

[9] It should be noted that such probabilistic criteria are fairly lax. Usually, a researcher would choose more conventional criteria, such as a .05 significance level and a benchmark proportion of .65. In this situation, a case set needs to contain at least seven consistent cases to make a cause qualify as a necessary/sufficient condition.

analysis. We will demonstrate below that, apart from '~IKD', another causal combination explains democracy breakdown in the interwar period. However, this solution merely applies to a rather limited number of (deviant) cases. Since the Vanhanen dataset does not contain enough instances of this solution, the latter does not qualify as a predictor of the outcome in fs/QCA. Furthermore, it should be noted that we would not have obtained any solution from an fs/QCA analysis if we had used stricter probabilistic criteria, or no adjustment factor. This is, exactly, the reason why we argue that an fs/QCA analysis should only be carried out on the basis of a comparatively large middle-sized case set.

In sum, our exemplary illustrations show that an fs/QCA analysis is a particularly adequate method whenever the necessity to preserve richness of raw-data information is high. More importantly, we have argued that fs/QCA should only be used for analysing comparatively large middle-sized datasets. Otherwise, a researcher runs the risk of not revealing all causes which lead to the observed outcome. If a researcher wants to analyse a comparatively small middle-sized dataset for which the necessity to preserve raw-data richness is pronounced, s/he is better advised to resort to other methods. While traditional qualitative methods can constitute a fruitful tool in these situations, another methodological option is provided by MVQCA, the second ramification of QCA. In line with the present and the previous section, the following part illustrates the opportunities and constraints related to an MVQCA analysis.

4. MVQCA – The challenge of preserving rich raw-data information without preventing the reduction of causal complexity

Like fs/QCA, MVQCA has been designed as a response to the problem of contradicting observations in general, and the problem of dichotomisation in particular (see Cronqvist 2005a). Accordingly, QCA and its ramifications differ most notably in the extent to which they allow for the preservation of cluster information contained in a raw dataset. In that, MVQCA strikes a balance between QCA and fs/QCA because a researcher can preserve as much information as necessary for the avoidance of contradictions. On the other hand, however, s/he must take care to preserve as little information as possible in order to obtain parsimonious causal explanations. Therefore, we argue that MVQCA is a particularly adequate method for analysing genuinely middle-sized case sets which require the retention of some raw-data richness. In line with our above illustrations, we will outline those steps of an MVQCA analysis which are important for the understanding of this argument.

Overall, MVQCA is very similar to QCA as it is carried out in the same four steps. In line with QCA and fs/QCA, the **first step** of an MVQCA analysis consists in converting the collected raw data into more handy, multi-value scores so as to draw up a summary table. In contrast to QCA, raw data does not necessarily need to be converted into dichotomous values. Furthermore, the richness of raw data is not preserved entirely by resorting to fuzzy membership scores. Instead, MVQCA allows a researcher to transform each causal expression into multi-value scores, i.e. into as many value-groups as necessary for preserving all essential cluster information of the raw dataset. At the same time, MVQCA requires the retention of as few clusters as possible so as to facilitate the reduction of causal complexity. Accordingly, a researcher must pay attention to select thresholds in such a way that a raw dataset is converted into as many value-groups as necessary, and as few groups as possible (Cronqvist 2005b).

Studying the cluster distribution in the Vanhanen case set with the aid of a simple average linkage method, we find that cases are distributed fairly evenly on the scale of the *IOD* and *IKD* variable. This is, however, different for the third causal variable, family farm ('FF'), on which the sixteen cases form roughly three clusters (see figure 1 above). This suggests that the raw-data scores of this variable should be converted in such a way that this cluster information is preserved. Accordingly, we decided to transform variables *IOD* and *IKD* into dichotomous scores by placing just *one* threshold at a cut-off value of 45. For variable *FF*, by contrast, we preserve the cluster information by using *two* thresholds, which we place at a cut-off value of 32, and of 43.[10] Table 6 reports the outcome of such conversion. In that, it differs from the above QCA summary table (see table 2) only in its use of multi-value scores for variable *FF*.

Table 6: MVQCA summary table on causes of democracy breakdown in the interwar period

Case	**IOD** (Index of Occupational Diversification)	**IKD** (Index of Knowledge Distribution)	**FF** (Family Farms)	**Outcome** (Breakdown of Democracy)
AUS	**1**	**1**	**2**	**1**
BEL	**1**	**1**	**0**	**0**
CZE	**0**	**1**	**1**	**0**
FIN	**0**	**1**	**2**	**0**
FRA	**1**	**1**	**1**	**0**
GER	**1**	**1**	**2**	**1**
GRE	**0**	**0**	**0**	**1**
HUN	**0**	**1**	**1**	**1**
ITA	**0**	**0**	**0**	**1**
NET	**1**	**1**	**1**	**0**
POL	**0**	**0**	**2**	**1**
POR	**0**	**0**	**0**	**1**
ROM	**0**	**0**	**1**	**1**
SPA	**0**	**0**	**0**	**1**
SWE	**0**	**1**	**2**	**0**
UK	**1**	**1**	**0**	**0**

Source: Vanhanen (1984), recoded as described in the text

Akin to QCA, the **second step** of an MVQCA analysis consists in converting the summary table into a truth table. The latter contains as many rows as there are logically possible combinations of causes which, in turn, depend on the number of values assigned to each variable (Cronqvist 2003: 7). Accordingly, the truth table obtained from summarising the above Vanhanen dataset contains $2 \times 2 \times 3 = 12$ rows. Table 7 presents the outcome obtained from converting the Vanhanen summary table into a truth table.

[10] Hence, we assign a score of '0' to those cases with a raw-data value between 0 and 32, and a score of '1' to cases with a value from 32.1 to 43. Finally, we assign a score of '2' to all cases with an original value above 43.

Table 7: MVQCA truth table on causes of democracy breakdown in the interwar period

Case	IOD (Index of Occupational Diversification)	IKD (Index of Knowledge Distribution)	FF (Family Farms)	Outcome (Breakdown of Democracy)	Number of observed cases
Por, Gre, Spa, Ita	0	0	0	1	4
Rom	0	0	1	1	1
Pol	0	0	2	1	1
	0	1	0	?	0
Hun CZE	0	1	1	(1:0) C	2
Fin, Swe	0	1	2	0	2
	1	0	0	?	0
	1	0	1	?	0
	1	0	2	?	0
Bel, Uk	1	1	0	0	2
Fra, Net	1	1	1	0	2
Aus, Ger	1	1	2	1	2

Source: Vanhanen (1984), recoded and summarised as described in the text

The attentive reader will have noticed that table 7 still contains two contradicting cases. How can this be reconciled with our previous statement that MVQCA has been designed to remedy the problem of contradictions? Importantly, MVQCA is similar to QCA in that it only recapitulates *whether or not* a causal combination is observed and which *result* the latter produces. However, no attention is paid to the *number of times* a combination occurs. Therefore, MVQCA is susceptible to the emergence of contradicting observations. Or better, contradicting observations can be prevented in MVQCA by increasing the number of thresholds on one (or more) variable(s), thereby depicting more case clusters. In so doing, a researcher can eliminate all contradictions.

That said, two good reasons exist why a researcher might accept (a few) contradictory observations - usually with the result that the obtained MVQCA solution does not consider all analysed cases. Firstly, an increasing number of multi-value scores entails an exponential increase in the number of causal combinations. This, in turn, makes it often more difficult to obtain a parsimonious solution. Accordingly, a researcher may prefer a more simple solution, which does not consider all cases, to a very complex solution, which considers all cases of the dataset. Secondly, if additional thresholds are introduced with the aim of preventing contradictions, this can lead to a distorted representation of case clusters contained in the raw dataset. Consider our Vanhanen example: The two cases which are still involved in a contradiction, Hungary and Czechoslovakia, score very similarly on all three variables (see table 1). In order to eliminate this contradiction, we would have to place one threshold in such a way that it separates the two cases explicitly, either on the IOD or the IKD variable. This, however, would mean a data manipulation to the extent that such conversion does not reflect the case clusters of the raw dataset. Abstaining from this manipulation, we preferred to accept one contraction, and to exclude Hungary and Czechoslovakia from the minimisation procedure of our MVQCA analysis.

The difficulty of setting most representative thresholds also shows that MVQCA should only be used for genuinely middle-sized case sets. The reason, simply, is that the larger a case set, the higher the possibility that contradicting cases (such as Hungary and Czechoslovakia) are included. In order to prevent contradictions, more thresholds need to be introduced which, in turn, makes it increasingly difficult to obtain parsimonious solutions.

In line with QCA, the **third step** of an MVQCA analysis resorts to the minimisation rule in order to reduce causal complexity. Thus, whenever two or more causal combinations differ in only one condition, the latter can be excluded as a causally relevant factor if *all* possible values of this condition are covered by the expression (Cronqvist 2005a: 5-7). If we apply this logic to the multi-value Vanhanen dataset - including all logical remainders into, and excluding all contradictions from the minimisation procedure - we obtain the following summary equation:

1 (Democracy Breakdown) $=$ $IKD_0 + IOD_1 * FF_2$

Akin to QCA, the **last analytical step** consists in interpreting the obtained findings with regard to their necessity and sufficiency. Interestingly, the results obtained from our MVQCA analysis, agree with the above QCA and fs/QCA outcome to the extent that *ikd* (unequal knowledge distribution) emerges as a sufficient condition for democracy breakdown. But contrary to the previous QCA and fs/QCA result, a further condition is retained in the solution: The combination of a high level of occupational diversification and a high share of family-owned farms qualifies as a second sufficient condition for breakdown of democratic regimes in the interwar period. This result is particularly interesting as it disagrees with the analyses of Vanhanen who finds that a high level of occupational diversification *supports* democracy (see Vanhanen 1984: in particular 129-136).

This more complete solution illustrates that MVQCA is the most appropriate method for analysing a genuinely middle-sized dataset which requires the retention of some raw-data information. On the one hand, and in contrast to fs/QCA, MVQCA succeeds in revealing all causal conditions which lead to the observed outcome – in our case democracy breakdown. Only Hungary and Czechoslovakia are still involved in one contradictory observation which, in turn, indicates that an in-depth analysis of these two countries is unavoidable. On the other hand, and contrary to QCA, the MVQCA solution considers a high number of observed cases, namely 14 out of 16 countries which equals 88% of all observations.

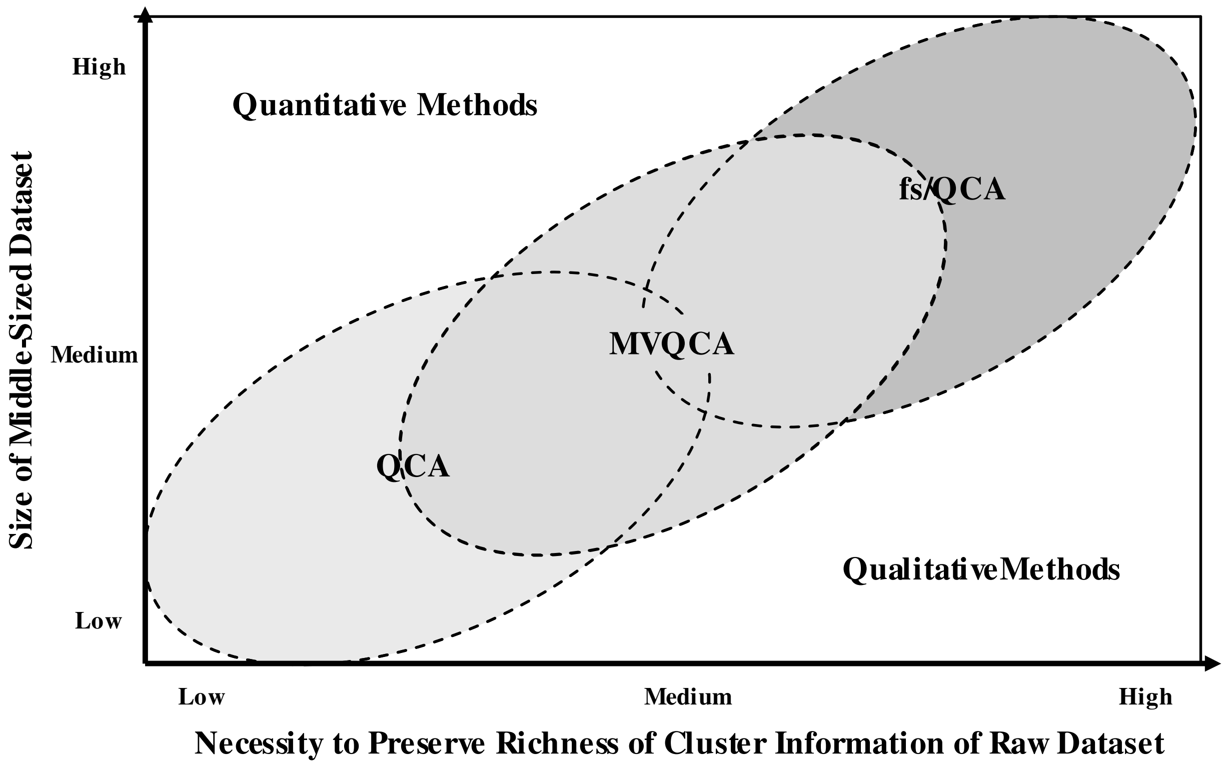
That said, we want to stress that an MVQCA analysis should not be carried out on the basis of a raw dataset which contains, and requires the preservation of ample cluster information. For our example, the introduction of *only one* additional threshold has not only led to a more complete, but also to a less simple solution. This shows that a researcher, who performs an MVQCA analysis on the basis of a very rich dataset, will find it difficult to obtain a concise outcome. Therefore, we argue that MVQCA should only be used for the analysis of genuinely middle-sized datasets where the retention of few raw-data information is required.

5. Conclusion

This paper has shown that QCA and its ramifications, fs/QCA and MVQCA, are very useful methods for analysing middle-sized datasets. That said, we have demonstrated that the problems related to handling contradictory observations guide a researcher in her choice of method. More precisely, we have argued that a researcher who wants to avoid causal explanations which *cover only a limited number of cases*, which are *not complete*, or *not parsimonious* should choose the method according to two parameters: The first is the overall size of her middle-sized case set, while the second is the need to preserve cluster information contained in the raw dataset needs to be preserved.

Following this logic, we have shown that a QCA analysis should only be carried out for small middle-sized-N situations for which (the necessity to preserve) rich raw-data information is reduced. Otherwise, the solutions obtained from a QCA analysis risk to cover only a limited number of cases. The opposite holds true for an fs/QCA analysis, as this method is most opportune for analysing comparatively large middle-sized datasets which necessitate the retention of rich raw-data cluster information. Most importantly, we have shown that fs/QCA bears the risk of not detecting all causal explanations if it is carried out on the basis of a comparatively small middle-sized dataset. Finally, we have argued that MVQCA strikes a balance between QCA and fs/QCA in that it is most adequate for the analysis of genuinely middle-sized datasets which necessitate the preservation of some cluster information. Otherwise, the risk is high that the solution obtained from an MVQCA analysis is not parsimonious. Figure 2 provides an overview over these arguments.

Figure 2: When (not) to prefer a QCA, fs/QCA and MVQCA analysis



To conclude our discussion, we want to stress that we perceive neither QCA and its ramifications, nor any other method, as superior *per se*. Instead, we believe that the superior explanatory power of any method varies from one research scenario to another as it depends on the research question to be studied. Accordingly, the use of a certain method should not be perceived as an aim in itself, but rather as a tool that helps to shed light on the puzzle in question. We therefore hope that our discussion helps to understand under which conditions QCA, fs/QCA or MVQCA become particularly helpful tools of analysis.

References

1. Berg-Schlosser, Dirk and Gisèle De Meur. 1994. "Conditions of Democracy in Interwar Europe: A Boolean Test of Major Hypotheses." *Comparative Politics* 26(3): 253-79.

2. Bollen, Kenneth, Barbara Entwisle, and Artur Alderson. 1993. "Macrocomparative Research Methods." *Annual Review of Sociology* 19: 321-51.

3. Campbell, Donald T. 1975. ""Degrees of Freedom" and the Case Study." *Comparative Political Studies* 8(2): 178-93.

4. Collier, David, Jason Seawright, and Gerardo L. Munck. 2004. "The Quest for Standards: King, Keohane, and Verba's Designing Social Inquiry." Pp. PAGES in *Rethinking Social Inquiry: Diverse Tools, Shared Standards*, Eds. Henry E. Brady and David Collier. Lanham: Rowman&Littlefield.

5. Cronqvist, Lasse. 2003. *Using Multi-Value Logic Synthesis in Social Science. Paper Presented at the ECPR-Conference in Marburg, 18th to 21st September 2003*. Marburg: unpublished document.

6. ———. 2004. "Presentation of TOSMANA. Adding Multi-Value Variables and Visual Aids to QCA." *Working Paper 2004 (20) of the Compasss Research Group, available at: http://www.compasss.org/Cronqvist2004.PDF*.

7. ———. 2005a. "Introduction to Multi-Value Qualitative Comparative Analysis (MVQCA)." *Didactical Paper 2005 (4) of the Compasss Research Group, available at: http://www.compasss.org/Cronqvist%20introduction_to_mvqca%202005.pdf*.

8. ———. 2005b. *Tosmana - Tool for Small-N Analysis*. Marburg: University of Marburg, available at: http://www.tosmana.net.

9. De Meur, Gisèle and Benoît Rihoux . 2002. *L'Analyse Quali-Quantitative Comparée: AQQC - QCA*. Louvain-La-Neuve: Bruylant-Academia.

10. Eckstein, Harry. 1975. "Case Study and Theory in Political Science." Pp. 79-137 in *Handbook of Political Science*, vol. 7, Eds. Fred I. Greenstein and Nelson W. Polsby. Massachusetts: Addison-Wesley.

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

12. Katz, Jonathan and Nathaniel Beck. 2004. "Random Coefficient Models for Time-Series-Cross-Section Data." *EconPapers; Working Paper 1205, available at: http://www.hss.caltech.edu/SSPapers/wp1205.pdf*.

13. King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social*

Inquiry - Scientific Inference in Qualitative Research. Princeton: Princeton University Press.

14. Lijphart, Arend. 1971. "Comparative Politics and Comparative Method." *American Political Science Review* 65(3): 682-93.

15. ———. 1975. "The Comparable-Cases Strategy in Comparative Research." *Comparative Political Studies* 8(2): 158-77.

16. Markoff, John. 1990. "A Comparative Method: Reflections on Charles Ragin's Innovations in Comparative Analysis." *Historical Methods* 23(4): 177-81.

17. McKeown, Timothy J. 1999. "Case Studies and the Statistical Worldview." *International Organization* 53(1): 161-90.

18. Mill, John S. 1872. *A System of Logic: Ratiocinative and Inductive*. 8th ed. London: Longmans, Green Reader, and Dyer.

19. Munck, Gerardo L. 1998. "Canons of Research Design in Qualitative Research ." *Studies in Comparative International Development* 33(3): 18-45.

20. ———. 2004. "Tools for Qualitative Research." Pp. 105-21 in *Rethinking Social Inquiry: Diverse Tools, Shared Standards*, Eds. Henry E. Brady and David Collier. Oxford: Rowman & Littlefield.

21. Muno, Wolfgang. 2003. "Fallstudien und die vergleichende Methode." Pp. 19-36 in *Vergleichende Politikwissenschaftliche Methoden: Neue Entwicklungen und Diskussionen*, Eds. Susanne Pickel, Gert Pickel, Hans-Joachim Lauth, and Detlef Jahn. Wiesbaden: Westdeutscher Verlag.

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

23. Ragin, Charles C. 1989. "New Directions in Comparative Research." Pp. 57-76 in *Cross-National Research in Sociology*, Ed. Melvin L. Kohn. Newbury Park: Sage Publications.

24. ———. 2000. *Fuzzy-Set Social Science*. Chicago: University of Chicago Press.

25. ———. 2003. "Making Comparative Analysis Count." *Working Paper 2003 (10) of the Compasss Research Group, available at: http://www.compasss.org/RaginDayOne.PDF*.

26. ———. 2004. "From Fuzzy Sets to Crisp Truth Tables." *Working Paper 2004 (28) of the Compasss Research Group, available at: http://www.compasss.org/Raginfztt_April05.pdf.*

27. ———. 2006. "Set Relations in Social Research: Evaluating Their Consistency and Change." *Political Analysis* 14 (3): 291-310.

28. Sigelman, Lee and George H. Gadbois. 1983. "Contemporary Comparative Politics: An Inventory and Assessment." *Comparative Political Studies* 16(3): 275-305.

29. Smelser, Neil J. 1976. *Comparative Methods in the Social Sciences*. Englewood Cliffs, N.J.: Prentice Hall.

30. Vanhanen, Tatu. 1984. *The Emergence of Democracy: A Comparative Study of 119 States, 1850-1979*. Helsinki: Societas Scientiarum Fennica.