International Journal of Social Research Methodology
Vol. 12, No. 1, February 2009, 33–50

When dichotomisation becomes a problem for the analysis of middle-sized datasets

Andrea Monika Herrmann and Lasse Cronqvist

(*Received 15 September 2005; Accepted 15 May 2007*)

This article aims at illustrating the circumstances in which Qualitative Comparative Analysis (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: the problem of contradicting observations. In QCA analysis, contradictions arise from the sheer number of cases and the problem of dichotomisation. In order to handle contradictions, the method for analysing middle-sized-N situations should therefore be chosen according to two parameters: the size of a dataset, and the need to preserve raw-data information. 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.

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* Munck, 2004; Muno, 2003). 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 e.g. King, Keohane, & Verba, 1994). From our 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 she/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 around.

Several attempts have been made to bridge the methodological divide between qualitative and quantitative analyses (see e.g. Campbell, 1975; Eckstein, 1975;

Corresponding author. Email: a.herrmann@geo.uu.nl

ISSN 1364-5579 print/ISSN 1464-5300 online

DOI: 10.1080/13645570701708543
http://www.informaworld.com

Lijphart, 1971, 1975; Smelser, 1976). Probably the most renowned proposal has been advanced by King et al. (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 Collier, Seawright, & Munck, 2004; McKeown, 1999; Munck, 1998; Ragin, 2000, p. 14). 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 e.g. Katz & Beck, 2004). But only a few tools have been developed for the analysis of 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, 2005a). It was, however, not later than 1990 that QCA started to encounter harsh, and often unfounded, criticism (see e.g. Markoff, 1990, p. 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 & Rihoux, 2002, pp. 119–144).

This article 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. 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 and the necessity to preserve the richness of raw-data information. 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. Instead, we aim to present QCA, fs/QCA and MVQCA as genuine alternatives to the more traditional qualitative (small-*N*) and quantitative (large-*N*) methods.

To illustrate our argument, the article is organised as follows: The second section briefly introduces the logic of QCA and the problem of contradicting observations which can notably limit the explanatory power of QCA. The third section illustrates how fs/QCA addresses the problem of contradictions and points to the limits of this method. The fourth section shows in which circumstances MVQCA succeeds in striking a balance between QCA and fs/QCA. Finally, the last section concludes by summarising our argument.

QCA – A powerful tool for analysing middle-sized datasets

Various scholars (see Ragin, 2000, p. 25; see also Bollen, Entwisle, & Alderson, 1993; Ragin, 1989; Sigelman & Gadbois, 1983) have illustrated that research in the social sciences is dominated by the analysis of either small-*N* or large-*N* situations, whereas very little research is carried out on the basis of a middle-sized number of cases. Until 1987, when Charles Ragin invented QCA (Ragin, 1987), this bias towards the use of qualitative and 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.

Importantly, though, we will show that QCA can be a particularly fruitful method when two criteria are fulfilled. First, the middle-sized dataset to be analysed is comparatively small.[4] Second, raw data can be recoded into dichotomous variables without a loss of important cluster information. The reason is that the risk of contradictions is minimised in these situations. A contradictory observation is made whenever the combination of causal combinations leads to different outcomes (see Ragin, 1987, pp. 113–118). Since cases involved in contradictions are often excluded from the analysis, a high proportion of such instances entails a situation in which a parsimonious solution only covers a small number of studied cases. 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 & De Meur, 1994).

Analysing the causes of breakdown of democratic regimes in the interwar period, Vanhanen constructed three socio-economic indices which he identified as pillars of democratisation. The first index (*Index of Occupational Diversification – IOD*) reports the arithmetic mean of urban population[5] and non-agricultural population[6] in a country. The second measure (*Index of Knowledge Distribution – IKD*) combines a population's literacy rate[7] and university education[8] accordingly. The third measure (*Family Farm – FF*) indicates the percentage of arable land which is owned and cultivated by families, thereby offering employment for not more than four people – including family members (see Vanhanen, 1984, pp. 33–37). In line with Berg-Schlosser and De Meur (1994), we focus our analysis on 16 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 & De Meur, 1994, p. 254).[9] 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.

It is not the aim of this article to review how QCA, fs/QCA and MVQCA analyses are carried out in detail. Instead, we limit our illustrations to the four analytical steps which are central to the understanding of our argument and illustrate them by using the Vanhanen dataset as an example.

In a nutshell, QCA consists in applying the logic of Mill's method of difference so as to reduce causal complexity (see Mill, 1872, pp. 451–452). Once a researcher has determined which cases she/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

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 darms) [%]	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, pp. 142–149).
Note: IOD = (% of population living in cities or towns + % of active population working outside agricultural sector)/2; IKD = (% of literate population + (no. of students per 100,000 inhabitants/5,000 students))/2; FF = % of arable land owned and cultivated by families.

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 statistically most representative. Table 2 reports the results obtained from recoding Vanhanen's raw dataset into dichotomous scores.

The *second step* consists in converting the obtained dataset into a so-called 'truth table' which lists all logically possible combinations of causes. Hence, a QCA truth table contains 2^k rows of possible causal combinations, whereby k stands for the number of causal conditions (Ragin, 1987, pp. 87–89). Converting the dichotomous Vanhanen dataset into a truth table leads to the results presented in Table 3. The latter reports whether and, if so, which outcomes have been observed for each causal combination. To give a few examples, row 3 recapitulates that democracy breakdown in Romania and Poland resulted from low occupational diversification, low knowledge distribution and a high percentage of arable land owned and cultivated by families. Row 9 reports the contradictory observation that a high share of occupational diversification, knowledge distribution and family-owned land can lead to democracy breakdown (like in Austria and Germany), or to democracy continuity (like in 'the Netherlands' case). Row 4 tells us that no case with the causal combination of low occupational diversification, high knowledge distribution and little family-owned land could be observed. Hence, it is unclear whether, or not, this causal combination entails democracy breakdown. It

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 darms)	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
Meaning	0 = causal condition absent (raw-data value ≤ 45%) 1 = causal condition present (raw-data value > 45%)		0 = causal condition absent (raw-data value ≤ 38%) 1 = causal condition present (raw-data value > 38%)	0 = observed outcome negative 1 = observed outcome positive (as reported by Berg-Schlosser and & De Meur, (1994))

Source: Vanhanen (1984, pp. 142–149), recoded as described in the text.

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)
POR, GRE, SPA, ITA	0	0	0	1
ROM, POL	0	0	1	1
	0	1	0	?
FIN, CZE, SWE, HUN	0	1	1	1 : 0 (Contradiction)
	1	0	0	?
	1	0	1	?
FRA, BEL, UK	1	1	0	0
NET, AUS, GER	1	1	1	1 : 0 (Contradiction)

Source: Vanhanen (1984, pp. 142–149), recoded and summarised as described in the text.
Note: 0 = causal condition absent, or respectively: observed outcome negative; 1 = causal condition present, or respectively: observed outcome positive; ? = unobservable outcome; : = as well as.

should be noted that the predominant concern of QCA is akin to qualitative methods in that they only consider *whether* a causal combination has been observed and *which result* the latter has produced. However, no analytical attention is paid to the *number of times* a certain combination has occurred.

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, p. 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, p. 93). If we apply this rule to the Vanhanen dataset – thereby including all logical remainders[10] into, and excluding all contradictions from the minimisation procedure of our analysis – we obtain the following Boolean equation:

$$1\,(\text{Breakdown of Democracy}) = ikd$$

In (other) words, breakdown of democratic regimes results from an unequal knowledge distribution. The fact that *low* knowledge distribution entailed democracy *breakdown* in the interwar period might not be surprising to the extent that Table 2 has already suggested that *high* knowledge diversification is consistently linked to democracy *survival*. However, only the systematic comparison of all – observed and unobserved – causal combinations confirms that democracy breakdown is caused by low knowledge distribution *alone*.

The *last step* of a QCA analysis consists in interpreting the obtained result with regard to its sufficiency and/or necessity (see Ragin, 1987, pp. 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 minimisation procedure. It is therefore worthwhile to study the sources of contradictions more closely.

The problem of contradicting observations in QCA

While QCA has encountered much unfounded criticism (see De Meur & Rihoux, 2002, pp. 123–141), the difficulty of dealing with contradicting cases often keeps researchers from using this method on a more than experimental level. Studying the sources of contradictions, we find that QCA is a particularly useful method only for comparatively small middle-sized datasets[11] which can be transformed into dichotomous scores without a loss of important (cluster-) information.

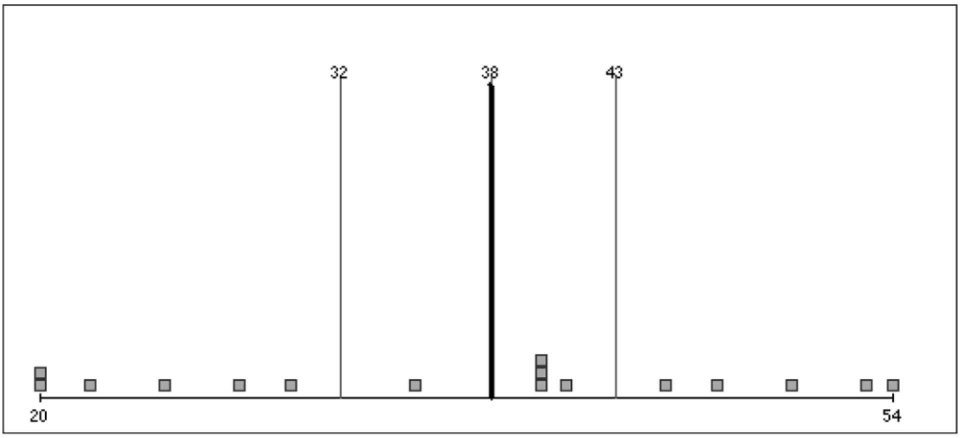
To understand this argument, it is first important to note that the occurrence of contradictions does not constitute a problem *per se*. On the contrary, the fact that a researcher needs to take a decision about how to deal with contradicting observations constitutes a particular strength of QCA (see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

Yet, one theoretical and one practical difficulty are related to this procedure. 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 increases accordingly. Thus, 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 faces both financial and time constraints. At some point, a researcher simply must end the empirical phase and make sense of the data she/he has gathered thus far. 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, she/he often deals with contradicting cases by excluding them (Ragin, 1987, p. 116).[12] We chose this solution when analysing the Vanhanen dataset (see above). However, applying such a 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 the analysis, we found that democracy breakdown results from an unequal distribution of knowledge ('ikd'). While this solution is very parsimonious, 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 particularly revealing in situations where the probability of contradictions is minimal. So, when is this the case? Importantly, contradictions in QCA analysis arise from two sources: first, from the sheer number of observed cases. Since QCA does not consider *how often* but only *that* a causal combination occurs, deviant cases are not identified as such and excluded from the analyses. 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. Yet, 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. In sum, the higher the number of cases, the higher the probability that deviants are included which entail contradictions. Hence, whenever a middle-sized dataset contains a comparatively large number of cases, the risk of contradictory observations is high. Accordingly, a QCA analysis is better avoided.

The second source of contradictions 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 e.g. Bollen et al., 1993; Goldthorpe, 1997). Since QCA can only operate on the basis of dichotomous scores, it obliges a researcher to choose one threshold according to which she/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.



38 = Most representative threshold for dividing cases on FF variable in two clusters.
32 & 43 = Most representative thresholds for dividing cases on FF variable into three clusters.

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

The ‘FF’ variable of the Vanhanen dataset exemplifies this argument. Figure 1 shows that the 16 cases form roughly three clusters on this variable. Even though we performed a cluster analysis so as to choose the single 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 row of Table 3.

In sum, 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. Similarly, contradicting observations are to be avoided by applying QCA only to small middle-sized datasets.

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.

Fs/QCA – Preserving rich raw-data information has its price

Like QCA, fs/QCA has been designed as a tool for analysing middle-sized-*N* situations (see Ragin, 2000). But contrary to QCA, fs/QCA is a particularly insightful method for analysing middle-sized datasets whose dichotomisation entails a loss of important (cluster-) information. And, even more importantly, fs/QCA should only be applied to comparatively large middle-sized datasets.[13] Otherwise, this method bears the risk of not revealing all causal conditions which provoke the studied outcome. We will illustrate this argument by briefly reviewing the most important steps of an fs/QCA analysis.

Like QCA, fs/QCA proceeds in four steps. However, the analytical procedure is different apart from the *first step* which also consists in transforming a raw dataset, this time, 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 raw data into dichotomous scores. It allows one to retain the data's richness due to the use of decimal membership scores (Ragin, 2000, pp. 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). A researcher's decision about *how* to transform raw data into membership scores can be based on various grounds: theoretical considerations, empirical insights, or a mathematical procedure. Depending on the chosen approach, the researcher will set, for each variable, a threshold for 'zero' membership (0.00) on the one hand, and for full membership (1.00) on the other. Furthermore, the chosen approach will also tell her/him whether to assign in-between membership scores in regular steps.

We decided to transform the Vanhanen raw dataset (see Table 1) into membership scores on the basis of a simple average linkage method. By calculating the distance between arithmetic means of various case groups, this method reveals the most pronounced case clusters and, hence, statistically meaningful thresholds for each variable in a sample. Based on this analysis, we transformed the Vanhanen raw data into five-step fuzzy membership scores, using the threshold values reported in Table 4.

Table 5 reports the fuzzy membership scores which we obtained by transforming the raw dataset, as reported in Table 1, according to the conversion measures summarised in Table 4.[14]

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 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.

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).

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, pp. 142–149), recoded into fuzzy membership scores as described in the text.

To reduce causal complexity, fs/QCA and QCA basically proceed in opposite directions. We have seen above that QCA *first* uses the minimisation rule for reducing 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, pp. 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*. Importantly, Ragin resorts to probabilistic criteria in order to identify all necessary and sufficient conditions (Ragin, 2000, pp. 107–115). More precisely, Ragin suggests 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, pp. 111–112). The so-obtained results are interpreted in the *fourth* and *final step* (see Ragin, 2000, pp. 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, as suggested by Ragin, namely a 0.05 significance level and a benchmark proportion of 0.65. Furthermore, we decided to apply an adjustment factor of 0.3 because the latter roughly represents the size of one step in our five-step 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:

$$1 \text{ (Breakdown of Democracy)} = \sim IKD$$

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 on 'QCA'). This suggests that the same fs/QCA solution also covers only a limited number of cases. In this regard, it is crucial to note 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, pp. 113–115, in particular Table 4.9). For example, if a researcher uses a 0.10 significance level and a benchmark proportion of 0.50, a case set must contain at least four cases with the same causal condition to make the latter qualify as a necessary/sufficient predictor of the outcome (see Ragin, 2000, p. 114, Table 4.9).[15]

Therefore, 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 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 shown 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, she/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.

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). MVQCA strikes a balance between QCA and fs/QCA because a researcher can preserve as much raw-data cluster information as necessary for the avoidance of contradictions. On the other hand, she/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.[16] 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, this time, '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. A researcher can divide raw data of each variable into as many value-groups as necessary for preserving all essential cluster information. At the same time, MVQCA requires the retention of as few clusters as possible in order to facilitate the reduction of causal complexity. In sum, 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, *FF*: here, the 16 cases form roughly three clusters (see Figure 1). 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% (akin to section on 'QCA'). For variable *FF*, by contrast, we use *two* thresholds, which we place at a cut-off value of 32%, and of 43%.[17] Table 6 reports the outcome of such conversion. In doing so, it differs from the above QCA summary table (see Table 2) only in its use of multi-value scores for variable *FF*.

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, p. 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.

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.

Contradicting observations can be prevented by increasing the number of thresholds on one (or more) variable(s). By depicting more case clusters a researcher can eliminate all contradictions. However, two good reasons exist why a researcher might accept (a few) contradictory observations – usually with the result that the obtained solution does not consider all observed cases. First, an increasing number of multi-value scores entails an exponential increase in the number of causal combinations. This makes it often more difficult to obtain a parsimonious solution. Hence, a researcher may prefer a more simple solution which does not consider all observed cases to a very complex solution which considers all cases.

Second, if additional thresholds are introduced with the aim of preventing contradictions, this can entail a distorted representation of case clusters contained in the raw dataset. Consider our Vanhanen example: the two cases which are still involved in a

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
Meaning	0 = causal condition absent (raw-data value ≤ 45%) 1 = causal condition present (raw-data value > 45%)		0 = condition absent (≤ 32%) 1 = condition partly present (32% < raw value ≤ 43%) 2 = condition present (> 43%)	0 = observed outcome negative 1 = observed outcome positive (as reported by Berg-Schlosser & De Meur (1994))

Source: Vanhanen (1984, pp. 142–149), recoded as described in the text.

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 an analytical manipulation in that such a conversion does not reflect the case clusters of the raw dataset. Abstaining from this manipulation, we preferred to accept one contradiction and to exclude Hungary and Czechoslovakia from the minimisation procedure.

The difficulty of setting most representative thresholds also shows that MVQCA should only be used for genuinely middle-sized case sets. The reason is that the larger a case set, the higher the possibility that contradicting cases (such as Hungary and Czechoslovakia) are included. To prevent contradictions, more thresholds need to be introduced, which 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. 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, pp. 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:

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)
POR, GRE, SPA, ITA	0	0	0	1
ROM	0	0	1	1
POL	0	0	2	1
	0	1	0	?
HUN, CZE	0	1	1	1:0 (Contradiction)
FIN, SWE	0	1	2	0
	1	0	0	?
	1	0	1	?
	1	0	2	?
BEL, UK	1	1	0	0
FRA, NET	1	1	1	0
AUS, GER	1	1	2	1

Source: Vanhanen (1984, pp. 142–149), recoded and summarised as described in the text.
Note: 0 = causal condition absent, or respectively: observed outcome negative; 1 = causal condition present, or respectively: observed outcome positive; ? = unobservable outcome; : = as well as.

$$1\ (\text{Breakdown of Democracy}) = IKD_0 + IOD_1 \times FF_2$$

Akin to QCA, the *last analytical step* consists in interpreting the obtained findings with regard to their necessity and sufficiency. Interestingly, the MVQCA result agrees with the above QCA and fs/QCA outcome in that *ikd* (unequal knowledge distribution) emerges as a sufficient condition for democracy breakdown in the interwar period. But contrary to the previous QCA and fs/QCA result, a further term is retained in the solution. The combination of high occupational diversification and a high share of family-owned farms qualifies as a second sufficient condition for democracy breakdown. 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 pp. 129–136). Furthermore, this solution is revealing as it shows that Austria and Germany follow a different democracy breakdown pattern. While democratic regimes in Greece, Italy, Portugal, Spain, Poland and Romania came to an end due to low knowledge distribution, Austria and Germany experienced democratic failure following high occupational diversification combined with a high share of family-owned land. In other words, rather backward economies with little educated citizens, as well as advanced countries with an economically active population both in cities and on the country-side, were equally susceptible to democracy breakdown.

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. This 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. Importantly, though, an MVQCA analysis

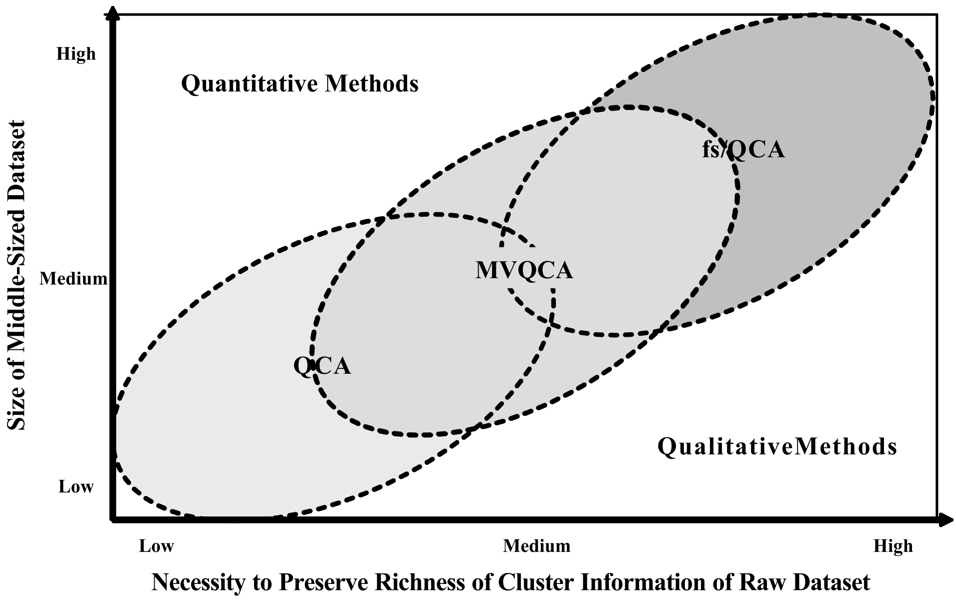


Figure 2. When (Not) to prefer a QCA, fs/QCA and MVQCA analysis.

should not be carried out on the basis of a raw dataset which contains and requires the preservation of ample cluster information as a researcher would find it difficult to obtain a parsimonious outcome.

Conclusion

This article has shown that QCA and its ramifications, fs/QCA and MVQCA, are very useful methods for analysing middle-sized datasets. In so doing, we have demonstrated that the problems related to handling contradictory observations guide a researcher in her/his 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 overall size of her middle-sized case set and the need to preserve cluster information contained in the raw dataset.

Following this logic, we have shown that QCA should only be used for analysing small middle-sized-*N* situations with a reduced necessity to preserve rich raw-data information. 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 require 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 illustrated 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 summarises our argument.

To conclude, 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 research puzzle. We therefore hope that our discussion helps to understand under which conditions QCA, fs/QCA or MVQCA become particularly helpful tools of analysis.

Acknowledgements

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 article.

Notes

1. Following the suggestion of Charles Ragin (see Ragin, 2003, p. 13), we use the notion of 'small-*N*' for samples which include one to four cases. Examples of small-*N* methods are hermeneutics, in-depth interviews and long-term observations (process-tracing).
2. Following the suggestion of Charles Ragin (see Ragin, 2003, p. 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.
3. Consequent to the remarks of Notes 1 and 2, the notion of 'middle-sized-*N*' refers to samples which include 5 to 50 cases (see Ragin, 2003, p. 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.
5. The *urban population* indicator refers to that percentage of a population which lives in cities or towns.
6. The measure of *non-agricultural population* reports the percentage of an economically active population which works in sectors outside agriculture.
7. The *literacy rate* refers to that share of population above 15 years which is able to read and write.
8. The variable on *university education* describes the number of students per 100,000 inhabitants who are enrolled in institutes of higher education. Setting the level for 100% university education at 5,000 students per 100,000 inhabitants, the indicator is calculated as follows:

$$\frac{\text{Country's number of students per 100,000 inhabitants}}{\text{5,000 students}}$$

9. 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).
10. Logical remainders constitute all those combinations of causal conditions which are not, or cannot be, observed so that their outcome is unknown to the researcher (see Ragin, 1987, pp. 104–113).
11. See Note 4.
12. 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, pp. 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, p. 118).
13. See Note 4.
14. We wish to emphasise 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:

$$\textit{Membership}\text{ score } = \frac{\text{Individual value raw data} - \text{Minimum value raw data}}{\text{Maximum value raw data} - \text{Minimum value raw data}}$$

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.

15. It should be noted that such probabilistic criteria are fairly lax. Usually, a researcher would choose more conventional criteria, such as a 0.05 significance level and a benchmark proportion of 0.65. In this situation, a case set needs to contain at least seven consistent cases to make a cause qualify as a necessary/sufficient condition.
16. See Note 4.
17. 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.

Notes on contributor

Dr. Andrea Monika Herrmann is Assistant Professor at the Innovation Studies Group of Utrecht University (Faculty of Geosciences; Utrecht University; P.O. Box 80115; 3508 TC Utrecht; e-mail: a.herrmann@geo.uu.nl; webpage: www.geo.uu.nl/staff/herrmann). She was a post doctoral research fellow at the Max-Planck-Institut für Gesellschaftsforschung (Cologne) from 2006 to 2008. She holds a PhD from the European University Institute (Florence) and an MSc from the London School of Economics (London). Her work includes publications on industrial relations and methodology, and a book (2008) entitled: *One Political Economy, One Competitive Strategy?* Oxford; Oxford University Press.

Dr. Lasse Cronqvist was awarded the degree of a Dr. Phil. from the University of Marburg, Germany in July 2007. He is currently a senior lecturer at the University of Trier (FB III – Political Science), 54286 Trier, Germany. Email: cronqvist@uni-trier.de. He is the originator of Multi-Value Qualitative Comparative Analysis (MVQCA), a methodological tool designed for the analysis of middle-size datasets, and the TOSMANA software, one of the main tools for QCA-style research (http://www.tosmana.org). His most recent publications include *Democratization and Political Culture in Comparative Perspective* (edited with Norbert Kersting) (Wiesbaden: VS Verlag, 2005) and ‘Multi-Value QCA (mvQCA)’ (with Dirk Berg Schlosser, 2008) in B. Rihoux and Ragin, C. (Eds.) *Configurational Comparative Methods* (pp. 87–122). Thousands Oaks: Sage.

References

Berg-Schlosser, D., & De Meur, G. (1994). Conditions of democracy in interwar Europe: A Boolean test of major hypotheses. *Comparative Politics, 26,* 253–279.

Bollen, K., Entwisle, B., & Alderson, A. (1993). Macrocomparative research methods. *Annual Review of Sociology, 19,* 321–351.

Campbell, D.T. (1975). ‘Degrees of freedom’ and the case study. *Comparative Political Studies, 8,* 178–193.

Collier, D., Seawright, J., & Munck, G. L. (2004). The quest for standards: King, Keohane, and Verba's designing social inquiry. In H.E. Brady & D. Collier (Eds.), *Rethinking social inquiry: Diverse tools, shared standards* (pp. 21–50). Lanham, MD: Rowman & Littlefield.

Cronqvist, L. (2003, September 18–21). *Using multi-value logic synthesis in social science.* Paper presented at the ECPR Conference in Marburg (unpublished document).

Cronqvist, L. (2004). *Presentation of TOSMANA: Adding multi-value variables and visual aids to QCA* (Working Paper 2004 (20) of the Compasss Research Group). Retrieved August, 2005 from http://www.compasss.org/Cronqvist2004.PDF

Cronqvist, L. (2005a). *Introduction to multi-value qualitative comparative analysis (MVQCA).* (Didactical Paper 2005 (4) of the Compasss Research Group). Retrieved August, 2005 from http://www.compasss.org/Cronqvist%20introduction_to_mvqca% 202005.pdf

Cronqvist, L. (2005b). *Tosmana – tool for small-N analysis.* Marburg: University of Marburg. Retrieved from http://www.tosmana.net

De Meur, G., & Rihoux, B. (2002). *L'Analyse Quali-Quantitative Comparée: AQQC – QCA.* Louvain-La-Neuve: Bruylant-Academia.

Eckstein, H. (1975). Case study and theory in political science. In F.I. Greenstein & N.W. Polsby (Eds.), *Handbook of political science* (pp. 79–137). Reading, MA: Addison-Wesley.

Goldthorpe, J.H. (1997). Current issues in comparative macrosociology: A debate on methodological issues. *Comparative Social Research, 16,* 1–26.

Katz, J., & Beck, N. (2004). *Random coefficient models for time-series-cross-section data* (EconPapers, Working Paper 1205). Retrieved August, 2005 from http://www.hss.caltech.edu/SSPapers/wp1205.pdf

King, G., Keohane, R.O., & Verba, S. (1994). *Designing social inquiry – scientific inference in qualitative research.* Princeton, NJ: Princeton University Press.

Lijphart, A. (1971). Comparative politics and comparative method. *American Political Science Review, 65,* 682–693.

Lijphart, A. (1975). The comparable-cases strategy in comparative research. *Comparative Political Studies, 8,* 158–177.

Markoff, J. (1990). A comparative method: Reflections on Charles Ragin's innovations in comparative analysis. *Historical Methods, 23,* 177–181.

McKeown, T. J. (1999). Case studies and the statistical worldview. *International Organization, 53,* 161–190.

Mill, J. S. (1872). *A system of logic: Ratiocinative and inductive* (8th ed.). London: Longmans, Green Reader and Dyer.

Munck, G.L. (1998). Canons of research design in qualitative research. *Studies in Comparative International Development, 33,* 18–45.

Munck, G.L. (2004). Tools for qualitative research. In H.E. Brady & D. Collier (Eds.), *Rethinking social inquiry: Diverse tools, shared standards* (pp. 105–121). Oxford: Rowman & Littlefield.

Muno, W. (2003). Fallstudien und die vergleichende Methode [case studies and the comparative method]. In S. Pickel, G. Pickel, H.-J. Lauth, & D. Jahn (Eds.), *Vergleichende Politikwissenschaftliche Methoden: Neue Entwicklungen und Diskussionen* [comparative methods in political sciences: new developments and discussions] (pp. 19–36). Wiesbaden: Westdeutscher Verlag.

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

Ragin, C.C. (1989). New directions in comparative research. In M.L. Kohn (Ed.), *Cross-national research in sociology* (pp. 57–76). Newbury Park, CA: Sage.

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

Ragin, C. C. (2003). *Making comparative analysis count* (Working Paper 2003 (10) of the Compasss Research Group). Retrieved August, 2005 from http://www.compasss.org/RaginDayOne.PDF

Sigelman, L., & Gadbois, G.H. (1983). Contemporary comparative politics: An inventory and assessment. *Comparative Political Studies, 16,* 275–305.

Smelser, N.J. (1976). *Comparative methods in the social sciences.* Englewood Cliffs, NJ: Prentice Hall.

Vanhanen, T. (1984). *The emergence of democracy: A comparative study of 119 states, 1850–1979.* Helsinki: Societas Scientiarum Fennica.