USC-SIPI REPORT # 408

Charles Ragin's Fuzzy Set Qualitative Comparative Analysis (fsQCA) Applied to Linguistic Summarization

by

Jerry M. Mendel and Mohammad Korjani

December 2010

Signal and Image Processing Institute
UNIVERSITY OF SOUTHERN CALIFORNIA
Viterbi School of Engineering
Department of Electrical Engineering
3740 McClintock Avenue, Suite 400
Los Angeles, CA 90089-2564 U.S.A.

ABSTRACT

Fuzzy Set Qualitative Comparative Analysis (fsQCA) is a methodology for obtaining linguistic summarizations from data that are associated with cases. It was developed by the eminent sociologist Prof. Charles C. Ragin, but has, as of this date, not been applied by engineers or computer scientists. Unlike more quantitative methods that are based on correlation, fsQCA seeks to establish logical connections between combinations of causal conditions (*conjunctural causation*) and an outcome, the result being rules that summarize (describe) the sufficiency between subsets of all of the possible combinations of the causal conditions (or their complements) and the outcome. The rules are connected by the word OR to the output. Each rule is a possible path from the causal conditions to the outcome and represents *equifinal causation*.

This report, for the first time, explains fsQCA in a very quantitative way, something that is needed if engineers and computer scientists are to use fsQCA.

There can be multiple results from fsQCA, i.e. collections of combinations of causal conditions each of which can be interpreted as a linguistic summary, ranging from the most "complex" summary to "intermediate" summaries to the most "parsimonious" summary. The method that is used to obtain the intermediate linguistic summaries is called *counterfactual analysis*; it is very important to fsQCA and is also described in this report.

This report also provides examples that illustrate every step of fsQCA, guidelines for the number of causal conditions that can be used as a function of the number of cases that are available, comparisons of fsQCA with two existing approaches to linguistic summarization that also use fuzzy sets, and descriptions of a method for obtaining the membership functions that are needed in order to implement fsQCA.

Table of Contents

List of Figures

List of Tables

I. INTRODUCTION

A *linguistic summarization* is a sentence or a group of sentences that describes a pattern in a database. There are different kinds of linguistic summarizations, ranging from a library of pre-chosen sentences (summarizers), from which the most representative one (or group) is chosen and is then declared to be *the* linguistic summarization, to a collection of if-then rules, some or all of which are chosen to be *the* linguistic summarization. Each of the different kinds of linguistic summarizations has its useful place; however, in this report we are interested only in linguistic summarizations that are in the form of *if-then rules*.

According to Kacprzyk and Zadrożny (2010) linguistic data (base) summaries using type-1 fuzzy sets were introduced by Yager [(1982), (1991), (1995), (1996)], advanced by Kacprzyk and Yager (2001), Kacprzyk, et al. (2000) and Zadrożny and Kacprzyk (1995), implemented in Kacprzyk and Zadrożny [(2000a-2000c), (2001a-2001f), (2002), (2003)], and extended to type-2 fuzzy sets by Niewiadomski (2008a,b). Linguistic summarizations of time series that use type-1 fuzzy sets has been studied by Kacprzyk and Wilbik, e.g. [2010 (see, also, 13 other references by these authors, including Zadrożny, that are in this article]. Because all of these summarizations are for a library of pre-defined summarizers, and are not in the form of if-then rules, they are not elaborated upon in this report; however, detailed comparisons of three different summarization methods are given in Section VI.

Linguistic summarization using if-then rules and type-1 fuzzy sets had its origins in Zadeh's classical 1973 paper. Although these if-then rules are the foundation for the developments of many kinds of quantitative rule-based systems, such as fuzzy logic control, rule-based classification, etc., until recently very few people, other than perhaps Zadeh [e.g., (1996), (1999), (2001), (2008)], thought of them any longer as linguistic summarizations. This is because it is the mathematical implementations of the set of rules that has become important in such applications, rather than the rules themselves. In essence, the rules have become the means to the end, where the end is a mathematical formula that produces a numerical output. Only since Zadeh's pioneering works on computing with words has there been a return to the understanding that a collection of if-then rules, by themselves, is indeed a linguistic summarization.

Wang and Mendel (1991) developed the first method to extract if-then rules from time-series data (the WM method). Many improvements to the WM method have been published since the original method. All of these works use the if-then rules as a *predictive model*, which according to Hand, et al. (2001) "*… has the specific objective of allowing us to predict the value of some target characteristic of an object on the basis of observed values of other characteristics of the object.*"

Recently, Wu and Mendel [(2010), (2011)] developed a different way to extract rules from data in which the rules use interval type-2 fuzzy sets to model the words in both their antecedents and consequents. Their rules construct a *descriptive model*, which according to Hand, et al. (2001) "*... presents, in convenient form, the main features of the data. It is essentially a summary of the data, permitting us to study the most important aspects of the data without them being obscured by the sheer size of the data set*[1].

The linguistic summarization method that is described in this report also leads to a descriptive

[1] The linguistic summarizations mentioned above, due to Yager, Kacprzyk, Zadrożny, Niewiadomski and Wilbik also fall into the class of descriptive models.

model, and is called *Qualitative Comparative Analysis* (QCA). It is not a method originated by the authors of this report, but is a method discovered by the first author that has been used mainly in the fields of social and political sciences and that does not seem to have been used (prior to our works) in engineering or computer science. Consequently, this report should be viewed as a *conduit* of fsQCA from the less mathematically oriented social and political sciences literatures into the more mathematically oriented engineering and computer science literatures. This report also includes new results.

According to Ragin (2008, p. 183): "The goal of QCA is to derive a logically simplified statement describing the different combinations of conditions linked to an outcome." Each combination of conditions and same outcome is sometimes referred to as a *type* or a *typological configuration* [Fiss (2010)]. According to Rihoux and Ragin (2009, p. 33 and p. 66):

> Crisp set Qualitative Comparative Analysis (csQCA) was the first QCA technique, developed in the late 1980s, by Professor Charles Ragin[2] and programmer Kriss Drass. Ragin's research in the field of historical sociology led him to search for tools for the treatment of complex sets of binary data that did not exist in the mainstream statistics literature[3]. He adapted for his own research, with the help of Drass, Boolean algorithms that had been developed in the 1950s by electrical engineers to simplify switching circuits, most notably Quine[4] (1952) and McCluskey (1966). In these so-called minimization algorithms, he had found an instrument for identifying patterns of[5] multiple-conjunctural causation and a tool to "simplify complex data structures in a logical and holistic manner [Ragin (1987), p. viii]. … csQCA is based on Boolean algebra, which uses only binary data (0 or 1), and is based on a few simple logical operations[6] [union, intersection and complement]. … [In csQCA,] it is important to follow a sequence of steps, from the construction of a binary data table to the final 'minimal formulas.' … Two key challenges in this sequence, before running the minimization procedure, are: (1) implementing a useful and meaningful dichotomization of each variable, and (2) obtaining a 'truth table' (table of configuration) that is free of 'contradictory configurations.' … The key csQCA procedure is 'Boolean minimization.'

csQCA was extended by Ragin to fuzzy sets, because he realized that categorizing social science causes and effects as black or white was not realistic. Fuzzy sets let him get around this. According to [Rihoux and Ragin (2009), p. 120]:

> fsQCA retains key aspects of the general QCA approach, while allowing the analysis of phenomena that vary by level or degree. … The fsQCA procedure … provides a bridge between fuzzy sets and conventional truth table analysis by constructing a Boolean truth table summarizing the results of multiple fuzzy-set analyses. … Fuzzy membership scores (i.e., the varying degree to which cases belong to sets) combine qualitative and quantitative assessments. … The key set theoretic relation in the study of causal complexity is the *subset relation*; cases can be precisely assessed in terms of their degree of consistency [subsethood] with the subset relation, usually with the goal of establishing that a combination of conditions is sufficient for a given outcome.

Both csQCA and fsQCA are set-theoretic methods. They differ from conventional quantitative variable-based methods (e.g., correlation and regression) in that they [Fiss (2010)] "… do not disaggregate cases into independent, analytically separate aspects but instead treat configurations as different types of cases." Additionally, [Fiss (2010)] "The basic intuition underlying QCA[7] is that cases are best understood as configurations of attributes resembling overall types and that a

---

[2] He is now a professor of sociology and political science at the University of Arizona. In the 1980's he was a professor of sociology and political science at Northwestern University.

[3] See, also, Appendix B.

[4] Quine was not an electrical engineer; he was a famous American philosopher and logician.

[5] The Glossary, in Appendix A of this report, explains many terms that either may be new to the reader or are used in a context that may be different from the ones they are used to.

[6] Bracketed phrases, inserted by the present authors, are meant to clarify quoted materials.

[7] It is quite common to refer to both csQCA and fsQCA as "QCA" letting the context determine which QCA it is. More recently, the phrase *Configurational Comparative Methods* is used to cover all QCA methods, e.g., Rihoux and Ragin (2009).

comparison across cases can allow the researcher to strip away attributes that are unrelated to the outcome in question."

According to Ragin (2008, p. 183), "… QCA summarizes the truth table in a logically shorthand manner." This is *linguistic summarization*.

Kacprzyk and Zadrożny (2010, Table I) classify linguistic summaries into five forms, type 5 being the most general form, about which they state:

> Type 5 summaries represent the most general form … fuzzy rules describing the dependencies between specific values of particular attributes. … Two approaches to Type 5 summaries have been proposed. First, a subset of such summaries may be obtained by analogy with association rules concept and employing their efficient algorithms. Second, genetic algorithms may be used to search the summaries' space.

fsQCA provides a type 5 summary and is a new approach for engineers and computer scientists to obtain such a summary.

Although there are two kinds of fsQCA, one for establishing sufficient conditions and one for establishing necessary conditions, our emphasis in this report is on the former[8], because it is only the sufficient conditions for a specific outcome that are in the form of if-then rules; hence, our use of the term "fsQCA" implies the phrase "fsQCA for sufficient conditions."

One may ask: "Why is this report needed, since Ragin et al. have already published so much about fsQCA?" Our answer to this rhetorical question is: *This report, for the first time, explains fsQCA in a very quantitative way, something that is not found in the existing literature, and something that is needed if engineers and computer scientists are to use fsQCA*.

The rest of this report is organized as follows: Section II describes the steps of fsQCA and puts them on an analytical footing. Section III provides numerical examples that illustrate the steps of fsQCA and also show how the results from those steps can be collected together in summary tables. Section IV is about Counterfactual Analysis (CA), which is a way to overcome the limitation of a lack of empirical instances, i.e. the problem of *limited diversity*. CA leads to so-called *intermediate solutions* (summarizations). Section V provides new theoretical results that either help to improve the computations fsQCA or help to explain the results from fsQCA. Section VI provides very detailed comparisons between fsQCA and two other methods for linguistic summarization that also use fuzzy sets. This is needed so that it is absolutely clear how fsQCA differs from those methods. Section VII is about the connections between the number of cases and number of causal conditions that can be used in fsQCA. Section VIII describes a method for establishing the membership functions that are needed in order to perform fsQCA. Section IX provides our conclusions and also includes a list of limitations of fsQCA that should be viewed in a positive way as a list of *research opportunities*.

There are six appendixes to this report. Appendix A is a glossary that explains many terms used in this report that either may be new to the reader or are used in a context that may be different from the ones they are used to. Appendix B explains why Ragin became disillusioned with conventional quantitative analysis and developed fsQCA. Appendix C contains an algorithm for determining intermediate solutions and was provided to the senior author by Ragin. Appendix D contains examples for five causal conditions (Section III contains similar examples for three causal conditions). Appendix E explains how fsQCA can also be used to determine if a causal condition or its complement is a necessary condition. Because necessary conditions are very rare in engineering and computer science applications, they are not emphasized in this report. Appendix F provides a detailed example of how minimal prime implicants can be

[8] How to determine necessary conditions is explained in Appendix E.

computed by hand. It is meant for explanatory purposes only, because the QM algorithm is used in practice to make such calculations.

II. FSQCA

A. Introduction

fsQCA seeks to establish logical connections between combinations of causal conditions and a desired outcome, the result being rules (typological configurations) that summarize[9] (describe) the sufficiency between subsets of all of the possible combinations of the causal conditions (or their complements) and the outcome. It is not a methodology that is derived through mathematics, e.g. as the solution to an optimization problem (nor are the linguistic summarization methods of Yager, Kacprzyk, Zadrożny, Niewiadomski and Wilbik that are mentioned in Section I), although, as will be seen below, it uses mathematics. Our mathematical description of fsQCA does not appear in the existing literatures about fsQCA. It is needed, though, if engineers and computer scientists are to use fsQCA.

To begin, one must choose a desired outcome, e.g. *high* (*low*) *180-day cumulative oil production rate*. One then postulates a subset of *k* possible causes for the desired outcome, using the substantive knowledge of the researcher or domain experts. These causes or their complements become the antecedents in the fsQCA causal combinations, and, in each causal combination, they are always connected by the logical AND operation that is modeled using the minimum operation.

Let C_j denote the j^{th} causal condition ($j = 1,...,k'$) and O denote the desired outcome. For the k' possible causal conditions, it is assumed that each has n_C terms[10] describing it (e.g., *low permeability*, *moderate permeability* and *high permeability*, in which case $n_C = 3$) and these terms are denoted C_j^v ($v = 1,...,n_C$), respectively. In Ragin's works [e.g., Rihoux and Ragin (2009), Chapter 5] it is quite common for each causal condition to be described just by one term, e.g., in a study of the breakdown of democratic systems for 18 countries in Europe between World Wars 1 and 2, he chose the following five (k') candidate causal conditions: [country is] *developed*, *urban*, *literate*, *industrial* and *unstable*. Observe that none of these words has adjectives appended to it, whereas in most engineering applications for fsQCA it is very common to have adjectives appended to a causal condition, e.g., *low permeability* and *high permeability*. If a causal condition has more than one term describing it (e.g., *low permeability* and *high permeability*), then *each term is treated as an independent causal condition*[11]; hence, $C_j^v \equiv C_j$ ($j = 1,2,...,n_C k'$). These causes or their complements become the antecedents in the fsQCA causal combinations, and, in each causal combination, they are always connected by the

---

[9] Ragin does not think of fsQCA as linguistic summarization; he thinks of it as describing what's going on between a collection of causal conditions and an outcome. It is only in Rihoux and Ragin [2009, page 15, Box 1.4] that "summarizing data" is acknowledged as one of the five types of uses of QCA techniques. Consequently, it now seems legitimate to use fsQCA for linguistic summarization. The other four uses for QCA are: check coherence of data, check hypotheses of existing theories, quickly test conjectures, and develop new theoretical arguments.

[10] It may be that n_C is different for each C_j in which case n_C could be changed to n_{C_j}. In this report it is assumed that n_C is the same for each C_j.

[11] Initially, it was thought that a separate fsQCA would have to be performed for each term of a causal condition. Doing this would not only have led to an explosion in the number of fsQCAs, but it would also have led to possible conflicting summaries (e.g., one for *low permeability* and another for *high permeability*) that would somehow have to be resolved. Prof. Peer Fiss (of the Marshall School of Business, at the University of Southern California) told us that when he used a causal condition that had two terms associated with it (e.g., *low income* and *high income*) he treated each term as a separate causal condition. We have adopted his approach.

logical AND operation that is modeled using the minimum operation. In the rest of this report, we assume that $n_C k' \equiv k$, and talk about *k* causal conditions.

It is also assumed that the desired outcome has n_O terms describing it (e.g., *low 180-day cumulative oil production rate* and *high 180-day cumulative oil production rate*, in which case $n_O = 2$) and these terms are denoted O^w ($w = 1,...,n_O$).

fsQCA is usually performed separately for each of the n_0 outcome terms (as well as for their complements). For Ragin's study of democratic systems there are two fsQCAs (typological configurations), one for survival of democratic systems and one for non-survival—breakdown—of democratic systems. Generally there is *causal asymmetry* between the two fsQCAs, meaning that it is generally not the complements of the causal combinations associated with survival of democratic systems that are associated with breakdown of democratic systems. How many outcome terms fsQCA is performed for is up to the researcher. So, without loss of generality, it is assumed that $n_0 = 1$; hence, $O^w \equiv O$.

B. Membership Functions and Derived Membership Functions

fsQCA needs membership functions (MFs) for the *k* possible causal conditions and the desired outcome. Let $\xi_j \in \Xi_j$ denote the domain for the *j*^th^ causal condition, and $\omega \in \Omega$ denote the domain for the desired outcome. The MFs for C_j and *O* are respectively denoted, as:

$$\mu_{C_j}(\xi_j),\ \xi_j \in \Xi_j,\ j = 1,2,...,k$$

and

$$\mu_O(\omega),\ \omega \in \Omega$$

In this section, it is assumed that these membership functions, which are continuous functions of independent variables, ξ_j or ω, are known. How to obtain them is non-trivial, but those details (described in Section VIII) are not needed now in order to understand the major computations of fsQCA, which are the focus of this section.

Before proceeding to a description of the fsQCA computations, it is important to understand that they can only be performed for the available *N* cases (data) (e.g., the 18 European countries in Ragin's study of the survival or breakdown of democratic systems, or the 60 oil wells in a fracture optimization data set, etc.). These cases have no natural ordering (e.g., 18 countries, 60 oil wells), but instead each case is identified by an integer, so that by knowing the integer one also knows the case. The integers *x* = 1, 2, …, *N* are used to represent the *N* cases, and in this way the cases are ordered[12].

Next, one computes *derived membership functions* for the *k* possible causal conditions and the desired outcome as functions of the ordered cases[13]. Let $\xi_j(x)$ denote the *j*^th^ causal variable for causal condition C_j evaluated for case-*x*, and $\omega(x)$ denote the desired outcome variable for outcome O evaluated for case-*x*, where $x = 1,2,...,N$. The derived membership functions are

---

[12] For a person to repeat someone else's fsQCAs, and compare their intricate details with someone else's intricate details, they need to know the ordering of the *N*; hence, it is assumed that this information is provided to them.

[13] Ragin does not use the phrase "derived membership functions" nor does he interpret such a calculation as a MF of another fuzzy set. The latter is important for the subsethood calculation that is performed below. Instead, he provides tables [Ragin (2008), Rihoux and Ragin (2009)] that list the membership in each causal condition and in the causal combinations for each of the *N* cases. We provide such tables in Section III.

$\mu_{C_j}(x)$ and $\mu_O(x)$, where

$$\mu_{C_j}(x) = \mu_{C_j}(\xi_j(x)) \quad x = 1,2,...,N \tag{1}$$

$$\mu_O(x) = \mu_O(\omega(x)) \quad x = 1,2,...,N \tag{2}$$

Generally, $\mu_{C_j}(x)$ and $\mu_O(x)$ are neither normal nor unimodal functions[14] of x.

C. Methodology of fsQCA

C.1 Candidate rules (causal combinations): fsQCA begins by establishing a set of 2^k candidate rules, one rule for each possible causal combination of the k causal conditions or its complement. Each such causal combination is an ordered combination of the k causal conditions, and is denoted herein as F_i (which we shall call a *firing-level fuzzy set*[15]) where

$$F_i = A_1^i \wedge A_2^i \wedge ... \wedge A_k^i, \tag{3}$$

in which $\wedge$ is the logical AND operation that is implemented using minimum;

$$A_j^i = C_j \text{ or } c_j \quad j = 1,...,k \text{ and } i = 1,...,2^k \tag{4}$$

in which c_j denotes the complement of C_i, so that

$$\mu_{A_j^i}(x) = \mu_{C_j}(x) \text{ or } \mu_{c_j}(x); \tag{5}$$

and the derived MF for c_j, $\mu_{c_j}(x)$, is computed, as:

$$\mu_{c_j}(x) = 1 - \mu_{C_j}(x) \tag{6}$$

It is useful to think about fsQCA as establishing one rule (typology) for the desired outcome O that has the form:

$$\text{IF } F_1 \text{ or } F_2 \text{ or ... or } F_{2^k}, \text{ THEN } O \tag{7}$$

where the logical OR operation (a disjunction) is implemented using the maximum. This rule can also be expressed as a collection of 2^k rules (one for each causal combination) each having the same consequent, i.e.:

---

14 This is okay, because in a traditional type-1 fuzzy logic system (T1 FLS), when fired-rule output fuzzy sets are combined by the union operation, the resulting fuzzy set is also non-normal and is also frequently non-unimodal; so, such fuzzy sets are already in wide use.

15 Ragin does not use the term "firing-level fuzzy set"; instead, he uses the phrase "causal combination." Here terms and phrases are used that are already well established in the T1 FLS literature, however, the terms "firing-level fuzzy set" and "causal combination" are used by us interchangeably.

$$\begin{cases} \text{IF } F_1, \text{THEN } O; \text{or} \\ \text{IF } F_2, \text{THEN } O; \text{or} \\ \quad \ldots \\ \text{IF } F_{2^k}, \text{THEN } O \end{cases} \tag{8}$$

We shall refer to the 2^k rules in (8) as *candidate rules*. In the rest of fsQCA these rules are either deleted or simplified.

C.2 From candidate rules to subset of firing-level surviving rules: The first major computation is the *firing level*[16] for each case [e.g., Mendel (2001)], i.e. the membership value of the i^{th} causal combination in its k causal conditions, $\mu_{F_i}(x)$, but for each of the N cases, i.e.:

$$\mu_{F_i}(x) = \min\left\{\mu_{A_1^i}(x), \mu_{A_2^i}(x), \ldots, \mu_{A_k^i}(x)\right\} \quad x = 1,2,\ldots,N \text{ and } i = 1,2,\ldots,2^k \tag{9}$$

Clearly, each firing level can be a number whose value may range from 0 to 1, and is associated with F_i in (3). The universe of discourse X for F_i contains exactly N elements, one for each of the available data cases.

Each of the 2^k causal combinations is thought of by Ragin as one vertex in a 2^k- dimensional hypercube [Kosko (1992)]. In crisp set QCA, a candidate rule is either fully supported (i.e., it's firing level MF value equals 1) or is not supported at all (i.e., it's firing level MF value equals 0), and only the fully supported candidate rules survive. fsQCA backs off from the stringent requirement of crisp set QCA by replacing the vertex membership value of "1" with a vertex membership value of > 0.5, meaning that if the firing level is greater than 0.5 then the causal combination is closer to its vertex than it is away from its vertex. Only those cases whose firing levels are greater than 0.5 are said to support the existence of a candidate rule.

Let N_{F_i} denote the number of cases for which $\mu_{F_i}(x) > 0.5$. A candidate rule is kept only if its firing level is greater than 0.5 for "enough" cases, i.e. if $N_{F_i} \geq f$. Quantifying "enough" is subjective and is done by choosing a number (f) that is called the "frequency threshold," whose value depends on how many cases are available. f is an fsQCA tuning parameter and must be chosen by the user. Some guidance on how to choose f is provided by Ragin, i.e. for a small number of cases f is set at 1 or 2 [Ragin (2007)], whereas for a large number of cases f is at least[17] 10 [Ragin (2008), p. 197]. Unfortunately the words "small" and "large" are fuzzy, so the

[16] Ragin does not use this terminology; he calls this the "fuzzy set membership of cases in the causal conditions." In order to be consistent with terms that are already used in the FS literature, the term "firing level" is used herein.

[17] Ragin and Fiss (Fiss is the co-author of Chapter 11 of [Ragin (2008, p. 197)]) state (for large N): "The fuzzy-set analysis that follows uses a frequency threshold of at least ten strong instances. This value was selected because it captures more than 80 percent of the [more than N = 700] cases assigned to [causal] combinations [in their works]." Setting the threshold at 10 and then deleting all causal combinations where there are fewer than 10 cases left them with 80% of the cases. In other words, there is a tradeoff between requiring more cases per causal combination and covering the whole population. In a later work [Fiss (2010)] f was chosen to be 3 when N = 205.

user must vary f until acceptable results are obtained. At this time, there does not seem to be a way to quantify what "acceptable results" means.

One (new) idea for choosing f is based on counting the number of cases for which the MF of the desired outcome is > 0.5. This is very different than thinking about choosing f in terms of all of the cases. For example, the number of oil wells for which 180-day cumulative oil production rate is high is quite small, whereas the number of oil wells for which 180-day cumulative oil production rate is low is large; hence, f can be different for these different desired outcomes.

The outcome of this first major computation of fsQCA is a small subset of R_S surviving rules of the original 2^k candidate rules that could be called the "subset of firing-level surviving rules"[18]. The firing levels for these R_S surviving rules are denoted F_l^S with associated re-numbered membership functions $\mu_{F_l^S}(x)$, $l = 1,2,...,R_S$.

C.3 From subset of firing-level surviving rules to subset of actual rules: So far the calculations of fsQCA have focused exclusively on the antecedents of a candidate rule (which means that they do not have to be repeated for different desired outcomes). The next major calculation involves both the antecedents and the consequent of a firing-level surviving rule.

A traditional fuzzy logic (FL) rule assumes that its antecedents are sufficient for its consequent, by virtue of the a priori construct of that if-then rule, e.g., "if D, then E" means "D implies E," i.e. "D is sufficient for E." One does not usually question the existence of the stated FL rule; however, at this point in fsQCA it is not known if the antecedents of a firing-level surviving rule are indeed sufficient for the consequent, i.e. *one questions the existence of the rule*![19]

If the antecedents are sufficient for the consequent, then the rule actually exists; however, if they are not then the rule does not exist. So, the next calculation of fsQCA establishes whether or not a rule exists. This calculation is a quantification of the fact that a causal combination is *sufficient* for an outcome if the outcome always occurs when the causal combination is present (however, the outcome may also occur when a different causal combination is present), i.e. the causal combination (the antecedents) is a *subset* of the outcome. Ragin uses Kosko's [1992] subsethood formula in order to compute the subsethood—*consistency*[20]—of the antecedents in the outcome for each of the R_S firing-strength surviving rules.

In fuzzy set theory, subsethood is a computation involving two fuzzy sets, e.g., if one wants to compute the subsethood of type-1 fuzzy set D in type-1 fuzzy set E, then, using Kosko's subsethood formula, one computes[21]

$$ss_K(D,E) = \frac{\sum_{i=1}^{N} \min(\mu_D(x_i), \mu_E(x_i))}{\sum_{i=1}^{N} \mu_D(x_i)} \qquad (10)$$

[18] Ragin does not use this terminology.

[19] Such questioning seems to be related to *causality*, i.e. which causal combinations are the causes of the desired outcome? Hence, fsQCA can also be viewed as a methodology for establishing causality, not of a single causal condition but of combinations of such causal conditions, i.e. *conjuntural causation*.

[20] "Consistency" is the term favored by Ragin [Ragin (2008) and Rihoux and Ragin (2009)], although he also uses the term "subsethood." Here we use the two terms interchangeably.

[21] There are other formulas for subsethood and to-date no research has been conducted to study if one formula is more preferable for fsQCA than another.

Subsethood of the antecedents in the outcome of each firing-level surviving rule requires computing

$$ss_K(F_l^S, O) = \frac{\sum_{x=1}^{N} \min(\mu_{F_l^S}(x), \mu_O(x))}{\sum_{x=1}^{N} \mu_{F_l^S}(x)} \quad l = 1,2,...,R_S \tag{11}$$

All of the MFs needed to compute (11) are available at this point; hence, (11) can be computed. Some important geometric interpretations of subsethood (consistency) are given in Section E.

For F_l^S to be a subset of O, Ragin (2007) requires that[22,23]

$$ss_K(F_l^S, O) \geq 0.80 \tag{12}$$

The result is an even smaller subset of rules, namely[24] R_A *actual rules*. The firing levels for the actual rules are denoted F_n^A with associated (further re-ordered) membership functions $\mu_{F_n^A}(x)$, $n = 1,2,...,R_A$. Each of the actual rules is associated with a sub-set of the original *N* data cases, so that by knowing the specific actual rule one also knows the best instances for that rule. Connecting the best instances to each of the actual rules is very important in fsQCA because it allows the end-user to better understand what is happening in those specific cases.

C.4. From subset of actual rules to subsets of complex and parsimonious rules: It is quite possible that there are still too many rules, but now for a totally different reason than before. When the R_A actual rules are combined using the logical OR (disjunction) operation, then, because all of these rules share the same consequent, they can be re-expressed, as:

$$\text{IF } (F_1^A \vee F_2^A \vee ... \vee F_{R_A}^A), \text{ THEN } O \tag{13}$$

Each F_n^A in (13) is the conjunction of one of the 2^k combinations of the *k* original causal conditions, e.g., $F_2^A \equiv C_1 \wedge c_2 \wedge c_3 \wedge ... \wedge C_{k-1} \wedge c_k$, $F_5^A \equiv c_1 \wedge C_2 \wedge c_3 \wedge ... \wedge c_{k-1} \wedge C_k$, etc. It should be obvious to anyone who is familiar with set theory that there can be a lot of redundancy in $F_1^A \vee F_2^A \vee ... \vee F_{R_A}^A$, e.g., if $k = 3$, $R_A = 2$, $F_1^A \equiv c_1 \wedge c_2 \wedge c_3$ and $F_2^A \equiv c_1 \wedge c_2 \wedge C_3$, then by using simple set theory calculations, $F_1^A \vee F_2^A$ can be simplified, i.e.:

$$F_1^A \vee F_2^A = (c_1 \wedge c_2 \wedge c_3) \vee (c_1 \wedge c_2 \wedge C_3) = (c_1 \wedge c_2) \wedge (c_3 \vee C_3) = c_1 \wedge c_2 \tag{14}$$

because $c_3 \vee C_3 = \Xi_3$ and $\mu_{\Xi_3}(\xi_3) = 1$.

---

[22] Ragin (2007) also advocates "… looking for gaps in the upper range of consistency [subsethood] that might be useful for establishing a threshold, keeping in mind that it is always possible to examine several different thresholds and assessing the consequences of lowering or raising the consistency [subsethood] cut-off."

[23] Fiss (2010) uses a threshold of 0.80, which is the threshold we use.

[24] Ragin calls them "*primitive Boolean expressions*" or "*truth table rows*." Sometimes we also use the former.

While it was easy to perform the set-theoretic computations in (14) by hand, it is difficult to do this for larger values of k and R_A. Instead, Ragin uses the Quine-McCluskey (QM) algorithm[25] to do this automatically. In order to use the QM algorithm it is important to step back from all of the details just presented, that made use of fuzzy sets, and realize that the R_A actual rules can be re-interpreted as knowing exactly which of the R_A vertices of the 2^k possible vertices contributed antecedents to the rules. There no longer is anything fuzzy about this, because either a vertex has or has not contributed antecedents. Consequently, each of the remaining R_A vertices is obtained by converting their fuzzy membership values into crisp—binary—values, by computing ($n = 1,2,...,R_A$)

$$B_i^n = \begin{cases} 1 & \text{if } C_i^n \text{ is the component of the causal combination} \\ 0 & \text{if } c_i^n \text{ is the component of the causal combination} \end{cases} \quad i = 1,...,k \tag{15}$$

the results of which are collected into R_A $k \times 1$ binary vectors $\mathbf{b}_n$, where

$$\mathbf{b}_n = col(B_1^n, B_2^n, ..., B_k^n) \quad n = 1,2,...,R_A \tag{16}$$

Each $\mathbf{b}_n$ may be said to be a binary version of F_n^A.

The QM algorithm computes the *prime implicants*[26] as well as the *minimal prime implicants*[27] of $\{\mathbf{b}_n\}_{n=1}^{R_A}$, both of which are used by fsQCA (see Section D below). Ragin equates the prime implicants with a *complex solution* (linguistic summarization, containing R_C terms), the minimal prime implicants with a *parsimonious solution* (linguistic summarization, containing R_P terms), and interprets these two solutions as the end-points of a countable continuum of solutions, where the *intermediate solutions* (linguistic summarizations containing R_I terms) have to be established using a methodology called *counterfactual analysis* (CA) [Ragin (2008, Chapters 8 and 9)]. He believes that the most useful linguistic summarization is an intermediate summarization.

Each of the R_C terms (still referred to as a *causal combination*) in the complex solution usually is comprised of fewer than the original k causal conditions. Each of the R_P terms in the parsimonious solution is almost always comprised of much fewer than the original k causal conditions. Each of the R_I terms in an intermediate solution is almost always comprised of fewer causal conditions than are in the complex solutions, but more than the causal conditions that are in the parsimonious solutions.

---

[25] The Quine-McCluskey algorithm is used to minimize Boolean functions; see, e.g., sub-section D, http://en.wikipedia.org/wiki/Quine-McCluskey_algorithm, or Mendelson (1970).

[26] A *prime implicant* is a combination of primitive Boolean expressions that differ on only one cause and have the same output.

[27] *Minimal prime implicants* (also called *essential prime implicants* by Ragin (1987)), cover as many of the primitive Boolean expressions as possible with a logically minimal number of prime implicants. For an example of how minimal prime implicants are determined from prime implicants, see Ragin (1987, pp. 95-98). Other examples can be found in Mendelson (1970).

C.5 From subsets of complex and parsimonious rules to subset of intermediate rules: Counterfactual analysis (CA) involves performing thought experiments in which the substantive knowledge of a domain expert is used. Recall that $2^k - R_S$ causal combinations were eliminated early on in fsQCA because either there were too few cases to support them or their firing levels were below the threshold of 0.5. CA begins with both the complex and parsimonious solutions and modifies the complex solutions subject to the constraint that the parsimonious solution must always be present (in some form) in the final intermediate solutions. The modifications use causal combinations for which there either were no cases or not enough cases, and require that the user bring a lot of substantive knowledge about the cases into the modification process. Each modified complex solution is called a *counterfactual*, and each counterfactual is usually less complex in its structure than is the complex solution, unless the complex term does not change as a result of CA, in which case it becomes the counterfactual. Once all of the counterfactuals have been obtained for all of the complex terms, they are combined using the set theory operation union. This result is called the (set of) *intermediate solutions*, and it contains R_I terms. Because CA is so important to fsQCA, more details about it are given in Section IV.

C.6 From subset of intermediate rules to subset of simplified intermediate rules: Because CA leads to a new set of solutions, it is possible that their union can be simplified. This is accomplished by subjecting the intermediate solutions to the QM algorithm in which the remainders are set to false. The result of doing this are a set of R_{SI} *simplified intermediate summarizations* (solutions)[28].

C.7 From subset of simplified intermediate rules to subset of believable simplified intermediate rules: It is important to re-compute the consistencies of the R_{SI} simplified intermediate summarizations, because CA also makes no use of the fuzzy nature of the causal conditions and outcome, and so this connection back to fuzziness has to be re-established. It can happen, for example, that one or more of the simplified intermediate solutions have a consistency that is well below 0.80, in which case they are not to be believed. Further discussions about this are given in Section E. The end of all of this is a collection of R_{BSI} *believable simplified intermediate summarizations* (solutions) for a desired outcome.

C.8 Summary: Because of the many steps in fsQCA in which the original 2^k causal combinations have been reduced to R_{BSI} believable simplified intermediate solutions, we summarize fsQCA in a new mnemonic way in Fig. 1. The emphasis in this diagram is on the reduction of causal combinations from 2^k to R_{BSI}. Also shown are the "fuzzy" and "crisp" computations. In addition, we provide a flowchart for fsQCA in Figs. 2 and 3.

It is also quite common to compute the *best instances* for the R_{BSI} believable simplified intermediate solutions[29] as well as the *coverage* of the cases by them. Coverage is discussed in Section F.

It is also frequently very useful to also summarize the calculations for each fsQCA in a collection of tables, as is illustrated by the examples in Section III.

---

[28] Ragin does not use different naems for the intermediate and simplified intermediate solutions, and it was only in e-mail to the senior author that he mentioned his using QM to obtain the simplified intermediate solutions.

[29] This can also be done for the complex and parsimonious solutions.

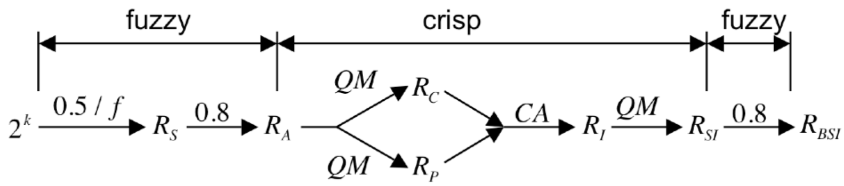

Figure 1. Mnemonic summary of fsQCA.

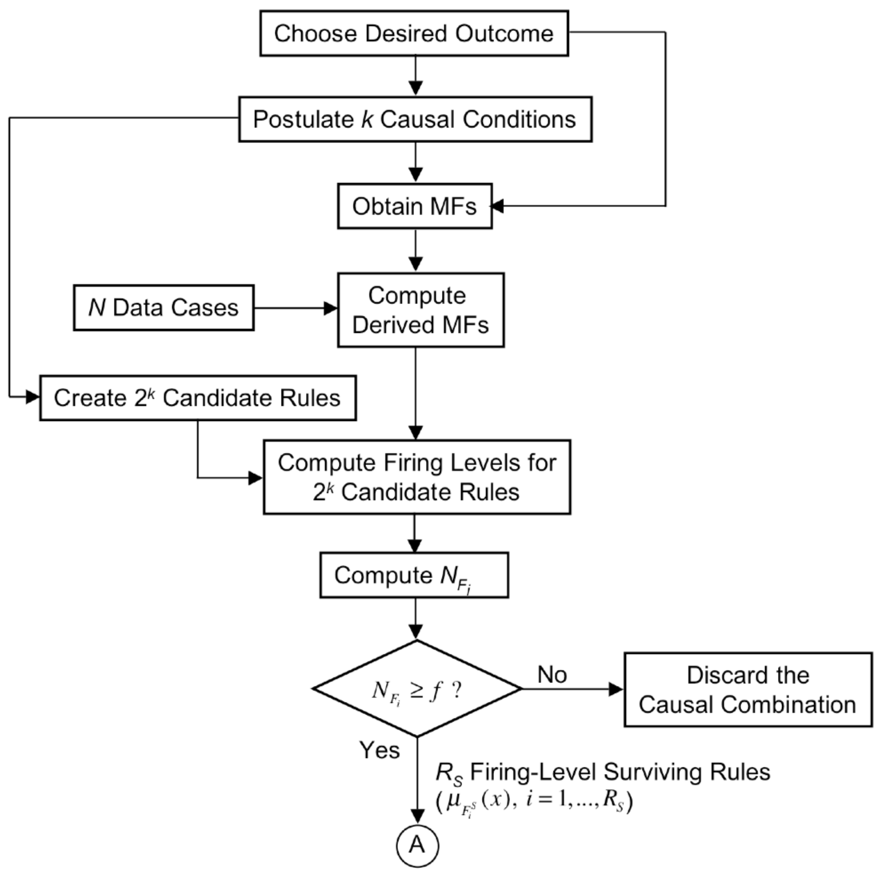

Figure 2. Flowchart for fsQCA; it is continued in Fig. 3.

D. Prime Implicants and Minimal Prime Implicants

Fig. 4 shows how the original 2^k causal combinations are partitioned into three non-overlapping subsets by fsQCA.

- Subset X_1 contains the causal combinations whose firing levels are < 0.5. This group of causal combinations never made it to the fsQCA sufficiency test, and can be thought of as the causal combinations that do not have cases, and are the ones associated with *limited diversity*. Limited diversity refers to limited knowledge about a causal combination. The removed causal combination is called a "remainder," which means either its presence or absence can cause the desired outcome to happen, but we just don't know.

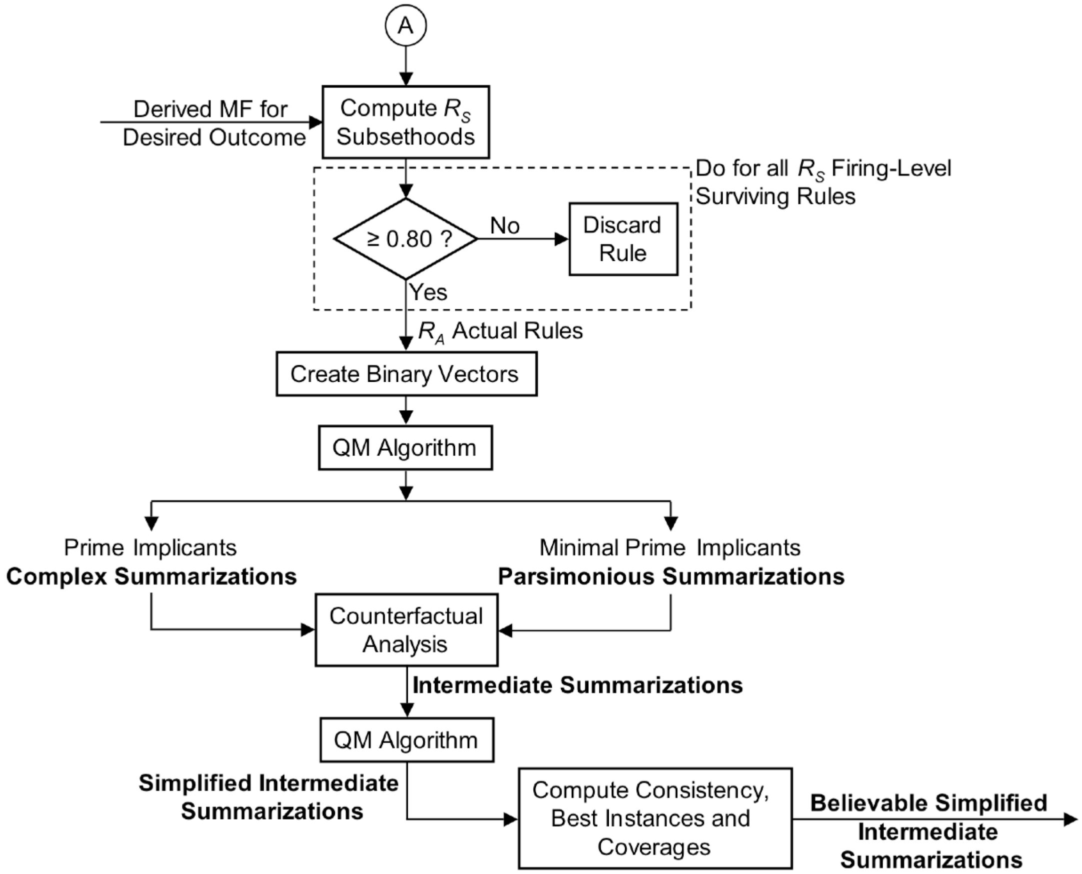

Figure 3. Flowchart for fsQCA, continued.

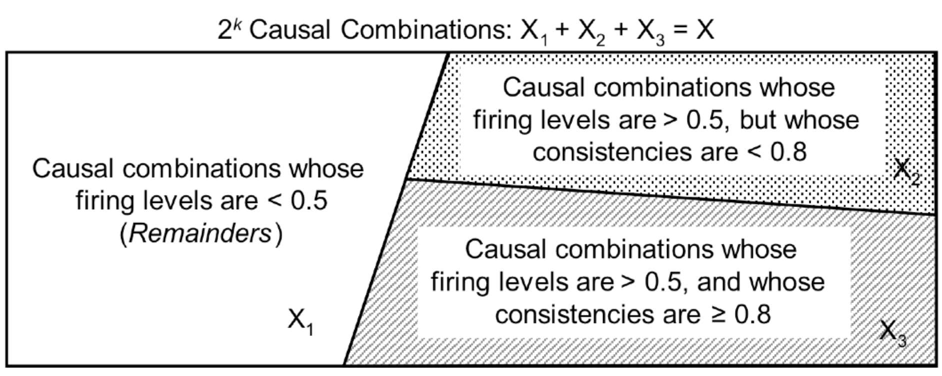

Figure 4. fsQCA partitions the original 2^k causal combinations into three subsets.

- Subset X_2 contains the causal combinations whose firing levels are > 0.5 but that failed the sufficiency test. Because these causal combinations made it all the way through fsQCA but failed the sufficiency test, they are forever after discarded.

- Subset X_3 contains the causal combinations whose firing levels are > 0.5 and passed the sufficiency test. These causal combinations are the short-term winners of fsQCA, the so-called *actual rules* or *primitive Boolean expressions* [Ragin (1987)].

The primitive Boolean expressions can be simplified in two different ways, leading to two sets of sufficient conditions—the R_C prime implicants (referred to by Ragin as the *complex sufficient conditions*) and the R_P minimal prime implicants (referred to by Ragin as the *parsimonious sufficient conditions*).

Prime implicants are obtained from the primitive Boolean expressions by using Boolean algebra reduction techniques (that are equivalent to set theoretic reduction techniques) that simplify (reduce) those expressions until no further simplifications are possible. The following are examples of reduction techniques that are used frequently: $ABC + ABC = ABC$, $A + a = 1$ and $ABC + AB = AB$. The latter, known as the *absorption rule*, is true because:

$$ABC + AB = ABC + AB(C + c) = ABC + ABC + ABc = ABC + ABc = AB(C + c) = AB \quad (17)$$

Sometimes it is possible to perform these reductions by hand; however, when there are many causal conditions and combinations it is very tedious (and next to impossible) to do this by hand. The Quine-McCluskey (QM) minimization method can be used to obtain the prime implicants automatically. This requires setting the causal conditions in X_3 as *present* and the causal conditions in both X_1 and X_2 as *absent*.

We used free software called "Logic Friday" that is available at: http://sontrak.com/.

Many times there are too many prime implicants, i.e., they are not all needed in order to cover the primitive Boolean expressions. A second running of the QM algorithm, in which subsets X_1 and X_3 are combined by the union operation and are then simplified (reduced), produces the *minimal prime implicants*. This requires setting the causal conditions in X_3 as *present*, the causal conditions in X_2 as *absent*, and the causal conditions in X_1 as *don't care*. In other words, remainders are set to be present for the desired outcome if and only if they result in simplifications of the primitive Boolean expressions; otherwise, they are treated as absent. An example that illustrates the calculation of minimal prime implicants is given in Appendix F.

When X_2 is vacuous (i.e., it contains no elements), a situation that we have encountered but have not seen reported on in Ragin's works, then there will be no minimal prime implicants, and consequently no parsimonious solutions. This is because of the following:

Theorem 1. If $X_2 \neq \varnothing$ (see Fig. 4), then minimal prime implicants exist. If, on the other hand, $X_2 = \varnothing$, then minimal prime implicants do not exist.

Proof: Consider *k* causal conditions $\{C_1, C_2, ..., C_k\}$. In order to find minimal prime implicants the causal combinations in X_3 are set to *present*, causal combinations in X_2 are set to *absent*, and causal combinations in X_1 are set to *don't care*.

If $X_2 = \varnothing$ then $X_1 + X_2 + X_3 = X_1 + X_3 = X$ so that $X_1 + X_3$ always equals *X*. One can therefore set all don't care combinations to *present*. Because no causal combinations have been removed (i.e., $X_1 + X_3 = X$), all possible combinations of the causal conditions can be generated by expanding the following product: $(C_1 + c_1)(C_2 + c_2)(C_3 + c_3)...(C_k + c_k)$; but, this equals 1. Consequently, it is obvious that by using Boolean algebra reduction techniques, all combinations of causal conditions must combine. This is the simplest result for $X_1 + X_3$ and it does not contain any combination; hence, minimal prime implicants do not exist if $X_2 = \varnothing$.

If, on the other hand, $X_2 \neq \varnothing$ then $X_1 + X_3 \neq X$, so $X_1 + X_3$ does not contain all combinations and simplification in $X_1 + X_3$ results in the minimal prime implicants. ■

E. Geometry of Consistency and Best Instances

Prior to QM and CA it is easy to connect each case to a causal combination because each surviving causal combination that has a MF value greater than 0.5 can be directly connected to a case, and each actual causal combination whose consistency is greater than 0.8 can also be directly connected to a case. This will be made very clear in the examples of Section III.

Unfortunately, the causal combinations used in QM are not revealed to the end-user because they are internal to the QM processing. Because CA begins with the results from QM, CA also does not provide a direct connection to the causal combinations. Consequently, after QM and CA it is no longer possible to directly connect cases to each of the R_{BSI} believable simplified intermediate solutions (see Fig. 1). In this section we explain how Ragin establishes the best instances for the believable simplified intermediate solutions.

In order to understand how Ragin does this, it is important to first discuss the *geometry of consistency*, something that is emphasized in Ragin's 2000 and 2008 books, but is not mentioned in Rihoux and Ragin (2009). The geometry of consistency lets us reconnect the R_{BSI} believable simplified intermediate solutions to the fuzzy natures of the desired outcome and the causal conditions that appear in each of these solutions.

Fig. 5 is modeled after Fig. 3.1 in Ragin (2008). The 45-degree line in Fig. 5 is very important and derives from the consistency formula (11), because maximum consistency is 1, and this can only occur when $\min(\mu_{F_l^S}(x), \mu_O(x)) = \mu_{F_l^S}(x)$, for $\forall x = 1,2,...,N$, which will be true if $\mu_O(x) \geq \mu_{F_l^S}(x)$ for $\forall x = 1,2,...,N$. So, if $\mu_O(x)$ lies above the 45-degree line $\mu_O(x) = \mu_{F_l^S}(x)$, for $\forall x = 1,2,...,N$ (anywhere in the region $A \cup B$), then $ss_K(F_l^S, O) = 1$. This leads to:

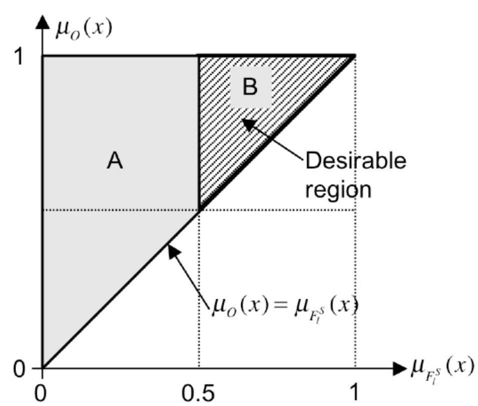

Figure 5. Consistency regions. Regions A and B are where maximum consistency can occur.

Geometrical Fact #1: For *maximum consistency* the pairs $(\mu_O(x), \mu_{F_l^S}(x))$ ($x = 1,...,N$) must be above the 45-degree line, i.e., they must be in the upper triangle $A \cup B$ on the plot of $\mu_O(x)$ versus $\mu_{F_l^S}(x)$. ■

Next we explain why, if all cases hypothetically were to lie in $A \cup B$, it is impossible for all of them to lie in Region *A*, but that it is possible for *all* cases to lie in Region *B*.

1. *The MF of each of the* R_A *actual causal combinations is > 0.5 for at least one case.* This is because only those R_S causal combinations whose firing levels are > 0.5 for at least one case make it to the consistency test, and that test reduces the number of causal combinations from R_S to R_A (Fig. 1). Each of the R_A causal combinations still has a firing level that is > 0.5 for at least one case.
2. *During QM and CA causal conditions are only* removed *from a causal combination; hence, the number of causal conditions in a causal combination can never increase as a result of QM and CA.* For QM, this is obvious from the fact that, for both the primary and minimal primary implicants, set theory (e.g., absorption) is used to combine terms, meaning that the number of causal conditions in a term, after QM, is never larger than the number of causal conditions in a term before QM. For CA, it should also be clear, from the fact that an intermediate solution has a number of causal conditions between the number in the complex and parsimonious solutions, that statement 2 is also true.
3. In Section III, we will prove that *if* $\mu_{F_{i*}}(x \mid C_1, C_2, ..., C_{k_1})$ *has been computed for* k_1 *causal conditions* $C_1, C_2, ..., C_{k_1}$ *and one now considers* k_2 *causal conditions* $C_1, C_2, ..., C_{k_2}$*, where* $k_2 > k_1$*, then (for all x*)

$$\mu_{F_{i*}}(x \mid C_1, C_2, ..., C_{k_2}) \le \mu_{F_{i*}}(x \mid C_1, C_2, ..., C_{k_1}) \tag{18}$$

4. *Viewing (18) from right-to-left observe that when causal conditions are* removed *from an existing causal combination, the firing level for the resulting causal combination can never be smaller than the prior firing level, i.e. firing levels tend to become strengthened when fewer causal conditions are included in a causal combination.*
5. **Geometrical Fact #2:** *There must be at least one case for which* $\mu_{F_l^{BSI}}(x) > 0.5$*, which means that it is not possible for all cases to be in Region A in Fig. 5.* Prior to QM and CA we began with R_A causal combinations, which, from Item 1 had MF values greater than 0.50 for at least one case. The R_A causal combinations are reduced to R_{SI} causal combinations by the QM-CA-QM calculations (Fig. 1). From Items 2 and 4, it must therefore be true that $\mu_{F_l^{BSI}}(x) > 0.5$ for at least one case.
6. **Geometrical Fact #3:** *A case can only be in Region B if* $\mu_{F_l^{BSI}}(x) > 0.5$*; hence, if all cases lie in* $A \cup B$ *this is only possible if all cases lie in Region B.*

Item 6 does not mean that all cases will lie in Region *B*. It means that Region *B* is the most desirable region for a case to lie in to obtain the largest possible value for consistency, which is why it is labeled "Desirable region" in Fig. 5. Ragin (in e-mail to the first author, on October 26, 2010) suggests that Region *B* should extend a bit below ($-\delta$) the 45-degree line to account for the subjectivity of MF values.

We are now able to state *Ragin's procedure* (provided to the first author in e-mail, dated October 19, 2010) *for choosing the best instances for each of the R_{BSI} believable simplified intermediate solutions*[30]. Let F_q^{BSI} denote one of the R_{BSI} terms in the believable simplified intermediate solutions ($q = 1,...,R_{BSI}$):

1. For each case and each F_q^{BSI} compute $\mu_{F_q^{BSI}}(x)$ ($x = 1,...,N$ and $q = 1,...,R_{BSI}$). The result of doing this is an $N \times R_{BSI}$ matrix of $\mu_{F_q^{BSI}}(x)$ numbers in which each row is for a case and each column is for one of the believable simplified intermediate solutions.
2. For each row of the $N \times R_{BSI}$ matrix, find the subset of the believable simplified intermediate solutions that have the largest $\mu_{F_q^{BSI}}(x) \geq 0.50$, i.e., find $q'(x)$ [and, subsequently $F_{q'(x)}^{BSI}$] such that

$$q'(x) = arg\left\{ \max_{q=1,...,R_{BSI}} \mu_{F_q^{BSI}}(x) \geq 0.50 \right\} \quad x = 1,...,N \tag{19}$$

   Often there is only one solution of (19); however, it is possible that there could be more than one solution of (19) because there can be ties, i.e.

$$q'(x) = \{q_1'(x), q_2'(x),...,q_Q'(x)\} \tag{20}$$

   where

$$q_j'(x) \in [1,...,R_{BSI}],\ j = 1,...,R_Q \tag{21}$$

   Consequently, more than one believable simplified intermediate solution may be retained for each case.
3. Focus on each $F_{q'(x)}^{BSI}$, and examine the ordered pair $(\mu_{F_{q'(x)}^{BSI}}(x), \mu_O(x))$. If $\mu_O(x) \geq \mu_{F_{q'(x)}^{BSI}}(x) - \delta$ (by Step 2, $\mu_{F_{q'(x)}^{BSI}}(x) \geq 0.50$) then Case *x* is declared a *best instance* for the believable simplified intermediate solution $F_{q'(x)}^{BSI}$.

At the end of this procedure, each of the believable simplified intermediate solutions will have best instances attached to it.

F. Coverage

Coverage is an assessment of the way respective terms in the believable simplified intermediate solution[31] "cover" observed cases [Rihoux and Ragin (2009)]. Ragin [2008, Ch. 3] mentions three kinds of coverage and Rihoux and Ragin [2009, p. 64] define them as: (1) *solution coverage*, C_s, which is the proportion of cases that are covered by *all* of the terms; (2) *raw coverage*, C_r, which is the proportion of cases that are covered by *each* term one at a time; and, (3) *unique coverage*, C_u, which is the proportion of cases that are uniquely covered by a

[30] This same procedure can be used for the complex, parsimonious and intermediate solutions.
[31] Coverage can also be computed for the complex and parsimonious solutions.

specific term (no other terms cover those cases). Each measure of coverage provides a different insight into the believable simplified intermediate solutions. The formulas for these three coverages are obtained by extending their comparable formulas from crisp to fuzzy sets, something that is commonly done by the fuzzy set community [e.g., see Mendel (1995) or Mendel (2001, Ch. 1)].

There are several different definitions of coverage for crisp sets (e.g., Rihoux and Ragin [2009], Kacprzyk and Zadrożny [2005], Niewiadomski [2008] and Wu and Mendel [2011]). For an if-then rule, these definitions are:

$$C_1 = \frac{\text{Number of cases covered by both antecedents \& consequents}}{\text{Number of cases covered by antecedents}} \tag{22}$$

$$C_2 = \frac{\text{Number of cases covered by both antecedents \& consequents}}{\text{Number of cases covered by consequent}} \tag{23}$$

$$C_3 = \frac{\text{Number of cases covered by both antecedents \& consequents}}{\text{Total number of cases}} \tag{24}$$

Observe that the numerators of (22)-(23) are the same, and that they count only the number of cases for which both the antecedent and consequent occur simultaneously, i.e. for which their MFs equal 1. The denominator of C_1 counts only the number of cases covered by the antecedents; the denominator of C_2 counts only the number of cases covered by the consequent; but, the denominator of C_3 counts the total number of cases, which causes C_3 to be much smaller than C_1 or C_2 and has led to the use of an amplification function [Wu and Mendel (2011)] to rescale C_3 to a meaningful degree of sufficient coverage.

No theory exists for determining which coverage definition is the best one to use. Presently, this can only be determined by means of experiments. Observe, however, that C_1 in (22) is very similar to the definition of consistency, and has a high correlation with it, so it is not very useful, and, C_3 in (24) needs a scaling function that may differ for different experiments. Consequently, C_2 in (23) seems the most reasonable coverage to use, which agrees with how Ragin computes coverage.

For fuzzy sets, (22)-(24) cannot be used as is because for such sets a term is associated with a case only to the degree of its MF for each case (for crisp sets, that degree is always either 1 or 0). From fuzzy set theory [Klir and Yuan (1995)], the number of cases covered by a single fuzzy set is a simple summation of membership scores in that fuzzy set; and, the number of cases simultaneously covered by two fuzzy sets is the size of the overlap of the two fuzzy sets. Consequently, if one wants to compute the coverage of T1 FS D in T1 FS E, $C(D,E)$, then, using the extension of (23) to T1 FSs, one computes

$$C(D,E) = \frac{\sum_{i=1}^{N} \min(\mu_D(x_i), \mu_E(x_i))}{\sum_{i=1}^{N} \mu_E(x_i)} \tag{25}$$

Although the numerator of this coverage is identical to the numerator in the formula for consistency in (11) the denominator of (25) is different. (25) is the coverage formula that is used by Ragin [2008].

The believable simplified intermediate solutions often contain several terms, F_l^{BSI} ($l = 1,2,...,R_{BSI}$), connected by the logical OR (modeled by the maximum). We shall refer to the union of the terms of the believable simplified intermediate solutions as the *composite solution*, and denote it as F^{BSI}. The firing level of F^{BSI} is the maximum of the firing levels of each of its terms. Consequently, *solution coverage*, C_s, which is the proportion of cases that are covered by *all* of the terms in F^{BSI}, is obtained from (25), as:

$$C_s(F^{BSI},O) = \frac{\sum_{x=1}^{N} \min\left(\max_i\left(\mu_{F_i^{BSI}}(x)\right),\mu_O(x)\right)}{\sum_{x=1}^{N}\mu_O(x)} \quad (26)$$

Raw coverage, C_r, is the proportion of cases that are covered by *each* term F_l^{BSI} separately, and is computed directly from (25), as:

$$C_r(F_l^{BSI},O) = \frac{\sum_{x=1}^{N} \min(\mu_{F_l^{BSI}}(x),\mu_O(x))}{\sum_{x=1}^{N}\mu_O(x)}, \quad l = 1,2,...,R_{BSI} \quad (27)$$

Unique coverage, C_u, is the proportion of cases that are uniquely covered by a specific term, $F_{l_i}^{BSI}$ ($l_i = 1,2,...,R_{BSI}$), and is calculated by subtracting the solution coverage of $F^{BSI} \wedge \neg F_{l_i}^{BSI}$ from $C_s(F^{BSI},O)$, i.e.:

$$C_u(F_{l_i}^{BSI},O) = C_s(F^{BSI},O) - \frac{\sum_{x=1}^{N} \min\left(\max_{\substack{j \\ j\neq i}}\left(\mu_{F_{l_j}^{BSI}}(x)\right),\mu_O(x)\right)}{\sum_{x=1}^{N}\mu_O(x)}, \quad l_i = 1,2,...,R_{BSI} \quad (28)$$

All of these coverages provide different insights into the believable simplified intermediate solutions. Presently, there is no threshold for coverage, as there is on consistency, because, in general, coverage is only used descriptively, although sometimes Ragin uses it to exclude a solution. Consequently, there are no guidelines given regarding what is "good coverage" because coverage depends on the nature of the evidence (Ragin, private e-mail to first author, Sep. 23, 2010). Of the three kinds of coverage unique coverage is almost always very low in solutions with several terms (private e-mail from Ragin to the first author sent on Sept. 23, 2010).

III. EXAMPLES

In order to illustrate fsQCA, we present two very simple examples. For desired outcome O there are three candidate causal conditions, $C_1 = A$, $C_2 = B$ and $C_3 = C$, each of which has only one fuzzy set associated with it, e.g. when it is possible to model a causal condition using an s-shaped or reverse s-shaped MF (see Section VIII). In Example 1, the desired outcome is *Breakdown of Democracy*, whereas in Example 2, the desired outcome is *Likely*[32] *Breakdown of Democracy*.

A. Example 1. Three Causal Conditions and O = Breakdown of Democracy

The data in Table I are taken from Table 5.2 in Rihoux and Ragin (2009), for which the desired outcome is *O = Breakdown of Democracy* (of European countries between World Wars 1 and 2) and the three causal conditions are *A = developed* (country), *B = urban* (country) and *C = literate* (country).

TABLE I
DATA- AND FUZZY-MEMBERSHIP-MATRIX (SHOWING ORIGINAL VARIABLES AND THEIR FUZZY-SET MEMBERSHIP FUNCTION SCORES)[a]

Case	*Outcome*		*Condition and MF scores*					
	o	*MF(O)*	*A*	*MF(A)*	*B*	*MF(B)*	*C*	*MF(C)*
1	-9	0.95	720	0.81	33.4	0.12	98	0.99
2	10	0.05	1098	0.99	60.5	0.89	94.4	0.98
3	7	0.11	586	0.58	69	0.98	95.9	0.98
4	-6	0.88	468	0.16	28.5	0.07	95	0.98
5	4	0.23	590	0.58	22	0.03	99.1	0.99
6	10	0.05	983	0.98	21.2	0.03	96.2	0.99
7	-9	0.95	795	0.89	56.5	0.79	98	0.99
8	-8	0.94	390	0.04	31.1	0.09	59.2	0.13
9	-1	0.58	424	0.07	36.3	0.16	85	0.88
10	8	0.08	662	0.72	25	0.05	95	0.98
11	-9	0.95	517	0.34	31.4	0.10	72.1	0.41
12	10	0.05	1008	0.98	78.8	1	99.9	0.99
13	-6	0.88	350	0.02	37	0.17	76.9	0.59
14	-9	0.95	320	0.01	15.3	0.02	38	0.01
15	-4	0.79	331	0.01	21.9	0.03	61.8	0.17
16	-8	0.94	367	0.03	43	0.30	55.6	0.09
17	10	0.05	897	0.95	34	0.13	99.9	0.99
18	10	0.05	1038	0.98	74	0.99	99.9	0.99

[a] This table is modeled after Table 5.2 in Rihoux and Ragin (2009), and the numbers in it are the same as the ones in that table.

Using knowledge and techniques from social science, numerical values were obtained for A, B and C for 18 European countries that in Table I are called[33] "Cases 1–18." Numerical values were initially obtained by Ragin for *o = Survival of Democracy*, which was assumed to be the complement[34] of *Breakdown of Democracy*; hence, $MF(O)$ was computed from $MF(o)$ as

[32] A synonym for *Likely* is *Promising*. It does not have the same probabilistic connotation that *Likely* has, so one could use *Promising* instead of *Likely*.

[33] The numbered cases correspond to the following countries: 1-Austria, 2-Belgium, 3-Czechoslovakia, 4-Estonia, 5-Finland, 6-France, 7-Germany, 8-Greece, 9-Hungary, 10-Ireland, 11-Italy, 12-Netherlands, 13-Poland, 14-Portugal, 15-Romania, 16-Spain, 17-Sweden, and 18-United Kingdom.

[34] Breakdown of Democracy actually is an *antonym* of Survival of Democracy, and the MF of an antonym is very different from the MF of the complement. Let B = Breakdown of Democracy and S = Survival of Democracy; then,

$1 - MF(o)$. S-shaped MFs were obtained for *Survival of Democracy*, developed (country), urban (country), and literate (country) using a method that is described in Ragin (2008) and in Section VIII, the details of which are not important for this example. Using these MFs, Ragin obtained the MF scores that are also given in Table I. These MFs implement (1) and (2).

From this point on, *A*, *B* and *C* are viewed as generic causal conditions for a generic outcome *O*, because there are more actual causal conditions that are associated with *Breakdown of Democracy* than are shown in Table I, and because tables for three causal conditions are easy to display. A more comprehensive version of this example, with five causal conditions, is presented in Appendix D.

For three causal conditions there are eight causal combinations, all of which are given in Table II, along with their memberships. These memberships are the firing levels in (9), e.g.

$$\mu_{F_1}(x) = \min\left\{\mu_a(x), \mu_b(x), \mu_c(x)\right\} = \min\left\{1-\mu_A(x), 1-\mu_B(x), 1-\mu_C(x)\right\} \tag{29}$$

TABLE II

FUZZY SET MEMBERSHIP OF CASES IN CAUSAL COMBINATIONS[a]

Case	*Membership in causal conditions*			*Membership in corners of vector space formed by causal conditions: Firing Levels*							
	MF(A)	*MF(B)*	*MF(C)*	F_1 *abc*	F_2 *abC*	F_3 *aBc*	F_4 *aBC*	F_5 *Abc*	F_6 *AbC*	F_7 *ABc*	F_8 *ABC*
1	0.81	0.12	0.99	0.01	0.19	0.01	0.12	0.12	**0.81**	0.01	0.12
2	0.99	0.89	0.98	0.01	0.01	0.01	0.01	0.01	0.11	0.02	**0.89**
3	0.58	0.98	0.98	0.02	0.02	0.02	0.42	0.02	0.02	0.02	**0.58**
4	0.16	0.07	0.98	0.02	**0.84**	0.02	0.07	0.07	0.16	0.02	0.07
5	0.58	0.03	0.99	0.01	0.42	0.01	0.03	0.03	**0.58**	0.01	0.03
6	0.98	0.03	0.99	0.01	0.02	0.01	0.02	0.02	**0.97**	0.01	0.03
7	0.89	0.79	0.99	0.01	0.11	0.01	0.11	0.11	0.21	0.01	**0.79**
8	0.04	0.09	0.13	**0.87**	0.13	0.09	0.09	0.09	0.04	0.04	0.04
9	0.07	0.16	0.88	0.12	**0.84**	0.12	0.16	0.16	0.07	0.07	0.07
10	0.72	0.05	0.98	0.02	0.28	0.02	0.05	0.05	**0.72**	0.02	0.05
11	0.34	0.1	0.41	**0.59**	0.41	0.10	0.10	0.10	0.34	0.10	0.10
12	0.98	1	0.99	0	0	0.01	0.02	0	0	0.01	**0.98**
13	0.02	0.17	0.59	0.41	**0.59**	0.17	0.17	0.17	0.02	0.02	0.02
14	0.01	0.02	0.01	**0.98**	0.01	0.02	0.01	0.02	0.01	0.01	0.01
15	0.01	0.03	0.17	**0.83**	0.17	0.03	0.03	0.03	0.01	0.01	0.01
16	0.03	0.30	0.09	**0.70**	0.09	0.30	0.09	0.30	0.03	0.03	0.03
17	0.95	0.13	0.99	0.01	0.05	0.01	0.05	0.05	**0.87**	0.01	0.13
18	0.98	0.99	0.99	0.01	0.01	0.01	0.02	0.01	0.01	0.01	**0.98**
Number of memberships > 0.5 (N_{F_i})				5	3	0	0	0	5	0	5

[a] This table is modeled after Table 5.6 in Rihoux and Ragin (2009), and the numbers in it are the same as the ones in that table.

The bold-faced numbers in Table II indicate memberships that are greater than 0.5. The numbers of such memberships (N_{F_i}) are listed in the last row of the table for each of the eight causal combinations. Using a frequency threshold of three (i.e., $f = 3$, which is the smallest non-zero value that N_{F_i} has in this example), only four of the eight causal combinations survive, i. e. $R = 4$. Those firing-level surviving rules are summarized in Table III; they constitute the

---

a MF for an antonym is $\mu_B(x) = \mu_S(10 - x)$ [Kim et al. (2000) and Zadeh (2005)]. This $\mu_B(x)$ is very different from using the complement of *S*; however, for the purposes of this example, we use the complement because it is widely used by Ragin.

elements in $X_2 + X_3$ that are depicted in Fig. 4. The four causal combinations that did not pass the frequency threshold test are in X_1.

The first column of Table III is called "Best Instances." It lists the cases that are associated with each surviving causal combination. This is a very important column because it directly connects the fsQCA back to the original cases. The next three columns of this table are for the three causal conditions and their entries are listed as 0 or 1, where a 0 occurs if the complement of the causal condition appears in the causal combination, and a 1 appears if the causal condition appears in the causal combination [e.g., $abC \to (0,0,1)$]. The next column in this table states the causal combination (the corresponding vector space corner) using set notation (e.g., abC). The last column in this table gives the count (from Table II) of the number of MF entries that are > 0.5.

TABLE III

DISTRIBUTION OF CASES ACROSS CAUSAL CONDITIONS[a]

Best Instances	*Causal Conditions*			*Corresponding Vector Space Corner*	*Number of cases with > 0.5 membership*
	A	*B*	*C*		
8, 11, 14, 15, 16	0	0	0	*abc*	5
4, 9, 13	0	0	1	*abC*	3
1, 5, 6, 10, 17	1	0	1	*AbC*	5
2, 3, 7, 12, 18	1	1	1	*ABC*	5

[a] This table is modeled after a combination of Tables 5.7 and 5.8 in Rihoux and Ragin (2009) and the numbers in it are the same as the ones in Table 5.7.

Next, the consistencies (subsethoods) are computed using (11). Note that these calculations use the MFs for all 18 cases. Results are summarized in Table IV, which looks like Table III, except that it has one more column called "Set theoretic Consistency." The rows of Table III are re-ordered so that the first row of Table IV has the largest value for Consistency and the last row has the smallest value for Consistency; however, for this example, no reordering of the rows was necessary in going from Table III to Table IV.

TABLE IV

DISTRIBUTION OF CASES ACROSS CAUSAL CONDITIONS AND SET-THEORETIC CONSISTENCY OF CAUSAL COMBINATIONS[a,]

Best Instances	*Causal Conditions*			*Corresponding Vector Space Corner*	*Number of cases with > 0.5 membership*	*Set-theoretic Consistency*
	A	*B*	*C*			
8, 11, 14, 15, 16	0	0	0	*abc*	5	0.98
4, 9, 13	0	0	1	*abC*	3	0.84
1, 5, 6, 10, 17	1	0	1	*AbC*	5	0.44
2, 3, 7, 12, 18	1	1	1	*ABC*	5	0.34

[a] This table is modeled after a combination of Tables 5.7 and 5.8 in Rihoux and Ragin (2009) and the numbers in it are the same as the ones in Table 5.7.

Using a consistency threshold of 0.80 only two of the four rules become actual rules, i.e. $R_A = 2$. These are the first two rules in Table V, *abc* and *abC*, and they constitute the elements of X_3 in Fig. 4. Observe that *abC*, which has fewer cases with > 0.5 membership than do *AbC* or *ABC*, survives, whereas *AbC* and *ABC* do not.

The prime implicant for $abc+abC$ is easy to obtain, because $abc+abC=ab(c+C)=ab$. The minimal prime implicant, found from the QM algorithm, is *a*. These solutions can be expressed linguistically, as:

$$\begin{cases} \text{Complex solution} & \text{IF } C_1 = a \text{ and } C_2 = b, \text{THEN } O \\ \text{Parsimonious solution} & \text{IF } C_1 = a, \text{THEN } O \end{cases} \tag{30a}$$

In words, these solutions are:

$$\begin{cases} \text{Complex solution} & \begin{array}{l} \textit{Not developed} \text{ and } \textit{not urban} \text{ (rural) is a sufficient causal combination} \\ \text{for } \textit{Breakdown of Democracy} \end{array} \\ \text{Parsimonious solution} & \begin{array}{l} \textit{Not developed} \text{ is a sufficient condition} \\ \text{for } \textit{Breakdown of Democracy} \end{array} \end{cases} \tag{30b}$$

To complete fsQCA, counterfactual analysis has to be performed. This is done in Section IV.

B. Example 2. Three Causal Conditions and O = Likely Breakdown of Democracy

Examining $MF(O)$ in Table I, observe that there are eight cases for which $MF(O)<0.5$; so, it seems plausible that these cases do not contribute much useful knowledge about *Breakdown of Democracy*. In order to examine this conjecture, we now focus on the modified desired outcome of *Likely Breakdown of Democracy*. Our interpretation of "*Likely Breakdown of Democracy*" is that only those cases for which $MF(O)>0.5$ should be kept for its fsQCA. It is very important to understand that we are not equating *Likely Breakdown of Democracy* and *Breakdown of Democracy*; instead, we are treating each as possible desired outcomes.

Tables V–VIII are analogous to Tables I–IV. The numbering of the cases in the former tables corresponds to the numbering of the cases in the latter tables. Although the same four rules survive in Table VI (compare with Table II), and constitute the elements in X_2+X_3, the number of cases supporting *AbC* and *ABC* has decreased from five to one, because four cases that previously supported these rules have been omitted from this example since $MF(O)<0.5$ for them. Whereas the frequency threshold for Example 1 was 3, for this example it is 1, a number that is supported by having only 10 cases. The four causal combinations that did not pass the frequency threshold test are in X_1. The consistencies in Table IV were computed for all 18 cases, whereas the consistencies in Table VIII are computed only for the 10 cases for which $MF(O)>0.5$.

Comparing Tables IV and VIII, observe that for *Likely Breakdown of Democracy*: (1) The consistencies for all four causal combinations have increased, (2) all four of the causal combinations pass the 0.8 consistency test, which means that $X_2=\varnothing$, and (3) causal conditions *AbC* and *ABC*, which were much less than 0.80 when all 18 cases were used, and were therefore discarded, have now achieved the maximum possible consistency values of 1 when only the 10 cases were used. It seems that the cases that were eliminated, for which their $MF(O)<0.5$, dragged down the consistencies of *AbC* and *ABC* by very considerable amounts. A theoretical explanation of this is given in Example 8 in Section V.

TABLE V
DATA- AND FUZZY-MEMBERSHIP-MATRIX (SHOWING ORIGINAL VARIABLES AND THEIR FUZZY-SET MEMBERSHIP FUNCTION SCORES) ONLY FOR THE CASES FOR WHICH $MF(O) > 0.5$

Case	*Outcome*		*Condition and MF scores*					
	o	*MF(O)*	*A*	*MF(A)*	*B*	*MF(B)*	*C*	*MF(C)*
1	-9	0.95	720	0.81	33.4	0.12	98	0.99
4	-6	0.88	468	0.16	28.5	0.07	95	0.98
7	-9	0.95	795	0.89	56.5	0.79	98	0.99
8	-8	0.94	390	0.04	31.1	0.09	59.2	0.13
9	-1	0.58	424	0.07	36.3	0.16	85	0.88
11	-9	0.95	517	0.34	31.4	0.10	72.1	0.41
13	-6	0.88	350	0.02	37	0.17	76.9	0.59
14	-9	0.95	320	0.01	15.3	0.02	38	0.01
15	-4	0.79	331	0.01	21.9	0.03	61.8	0.17
16	-8	0.94	367	0.03	43	0.30	55.6	0.09

TABLE VI
FUZZY SET MEMBERSHIP OF CASES IN CAUSAL COMBINATIONS ONLY FOR THE CASES FOR WHICH $MF(O) > 0.5$

Case	*Membership in causal conditions*			*Membership in corners of vector space formed by causal conditions: Firing Levels*							
	MF(A)	*MF(B)*	*MF(C)*	F_1 *abc*	F_2 *abC*	F_3 *aBc*	F_4 *aBC*	F_5 *Abc*	F_6 *AbC*	F_7 *ABc*	F_8 *ABC*
1	0.81	0.12	0.99	0.01	0.19	0.01	0.12	0.12	**0.81**	0.01	0.12
4	0.16	0.07	0.98	0.02	**0.84**	0.02	0.07	0.07	0.16	0.02	0.07
7	0.89	0.79	0.99	0.01	0.11	0.01	0.11	0.11	0.21	0.01	**0.79**
8	0.04	0.09	0.13	**0.87**	0.13	0.09	0.09	0.09	0.04	0.04	0.04
9	0.07	0.16	0.88	0.12	**0.84**	0.12	0.16	0.16	0.07	0.07	0.07
11	0.34	0.1	0.41	**0.59**	0.41	0.10	0.10	0.10	0.34	0.10	0.10
13	0.02	0.17	0.59	0.41	**0.59**	0.17	0.17	0.17	0.02	0.02	0.02
14	0.01	0.02	0.01	**0.98**	0.01	0.02	0.01	0.02	0.01	0.01	0.01
15	0.01	0.03	0.17	**0.83**	0.17	0.03	0.03	0.03	0.01	0.01	0.01
16	0.03	0.30	0.09	**0.70**	0.09	0.30	0.09	0.30	0.03	0.03	0.03
Number of memberships > 0.5 (N_{F_i})				5	3	0	0	0	1	0	1
Sums[a]				4.54	3.38	0.87	0.95	1.17	1.70	0.32	1.26

[a] These are used in Table XIII.

TABLE VII
DISTRIBUTION OF CASES ACROSS CAUSAL CONDITIONS FOR THE CASES FOR WHICH $MF(O) > 0.5$

Best Instances	*Causal Conditions*			*Corresponding Vector Space Corner*	*Number of cases with > 0.5 membership*
	A	*B*	*C*		
8, 11, 14, 15, 16	0	0	0	*abc*	5
4, 9, 13	0	0	1	*abC*	3
1	1	0	1	*AbC*	1
7	1	1	1	*ABC*	1

The primitive Boolean expressions (from Table VII) are: $abc + abC + AbC + ABC$. The prime implicants are also easy to obtain, as follows:

$$abc + abC + AbC + ABC = ab\,(c+C) + AC(b+B) = ab + AC \tag{31}$$

There are no minimum prime implicants, because $X_2 = \varnothing$ (see Theorem 1).

The complex solution can be expressed linguistically, as:

$$\text{IF } (C_1 = a \text{ and } C_2 = b) \text{ or } (C_1 = A \text{ and } C_2 = C), \text{THEN } O \tag{32}$$

In words, these solutions are:

$$\begin{array}{l}(\textit{Not developed} \text{ and } \textit{not urban} \text{ (rural)) OR } (\textit{Developed} \text{ and } \textit{Literate}) \\ \text{are sufficient causal combinations for } \textit{Likely Beakdown of Democracy}\end{array} \tag{33}$$

TABLE VIII
DISTRIBUTION OF CASES ACROSS CAUSAL CONDITIONS AND SET-THEORETIC CONSISTENCY OF CAUSAL COMBINATIONS FOR THE CASES FOR WHICH $MF(O) > 0.5$

Best Instances	*Causal Conditions*			*Corresponding Vector Space Corner*	*Number of cases with > 0.5 membership*	*Set-theoretic Consistency*
	A	*B*	*C*			
8, 11, 14, 15, 16	0	0	0	*abc*	5	0.985
4, 9, 13	0	0	1	*abC*	3	0.923
1	1	0	1	*AbC*	1	1
7	1	1	1	*ABC*	1	1

C. Comments

(1) Example 1 demonstrates that, by including the eight countries for which $MF(O) < 0.5$, the rule AC in (32) can be made to vanish, and Example 2 demonstrates that the results from fsQCA for *Likely Breakdown of Democracy* are different than for *Breakdown of Democracy*.

(2) See Appendix D for two examples that parallel Examples 1 and 2, but for five causal conditions. Regarding the *Breakdown of Democracy* problem, the examples in Appendix D are much more realistic than the ones in this section because they include two other causal conditions (*industrial country* and *stable country*) that are closely aligned with the desired outcome. In order to display results for five causal conditions (for which there would be $2^5 = 32$ causal combinations), a totally different way of computing the surviving causal combinations is presented in Section V, one that does not necessitate computing the firing levels for all 32 causal combinations but requires doing this only for the subset of causal combinations in (Fig. 4) $X_2 + X_3$.

IV. Counterfactual Analysis (CA)

A. Introduction

As we have seen, by using the QM algorithm one obtains two solutions (linguistic summarizations), the prime implicants—the complex (least parsimonious) solution—and the minimal prime implicants—the parsimonious solution. Ragin (2008) stresses that these solutions should be viewed as the two end-points of a continuum[35] of solutions that range from the most complicated to the least complicated. The methodology for providing "intermediate solutions" is what he calls "counterfactual analysis (CA)," and such solutions are considered to be the most useful ones by him.

CA [Fiss (2010)] offers a way to overcome the limitations of a lack of empirical instances, i.e. the problem of *limited diversity* [Ragin (2000, pp. 81 ff.)], and involves thought experiments. Recall that diversity refers to whether or not a case actually exists for a particular combination of causal conditions. In most applications it is very common for no cases to exist for many combinations of causal conditions, and this is referred to as "limited diversity."

CA begins with both the complex and parsimonious solutions and modifies the complex solutions subject to the constraint that the parsimonious solution must always be present (in some form) in the final intermediate solutions. The modifications use causal combinations for which there either were no cases or not enough cases, and require that the user bring a lot of substantive knowledge about the cases into the modification process. Each modified complex solution is called a *counterfactual*, and each counterfactual is usually less complex in its structure than is the complex solution, unless the complex term does not change as a result of CA, in which case it becomes the counterfactual. Once all of the counterfactuals have been obtained for all of the complex terms, they are combined using the set theory operation union. This result is called the (set of) *intermediate solutions*, and it contains R_I terms (Fig. 1).

Recall that, after CA, the intermediate solutions are simplified using set theory simplifications (or one last QM to obtain only the primary implicants of the union of counterfactuals) in order to remove redundancies. This result is called (by us) the (set of) *simplified intermediate solutions*, and it contains R_{SI} terms (Fig. 1). Finally, the consistencies of the R_{SI} simplified intermediate solutions are computed and only those solutions whose consistencies that are close to or above 0.80 are retained. The result is the collection of R_{BSI} *believable simplified intermediate solutions* for a desired outcome (Fig. 1).

According to Ragin (Appendix C):

> The **procedure** [for CA] evaluates each term in the parsimonious solution against each term in the complex solution. The number of terms in a solution is the number of combinations of causal conditions joined by "+". For example, if the parsimonious solution is $AB + CD$, then there are two parsimonious terms. If there are R_P parsimonious terms and R_C complex terms, the [CA] procedure cycles $R_C \times R_P$ times, once for each possible pairing.

Furthermore, according to Ragin (private e-mail to the first author on September 15, 2010):

> In essence, you can examine each element in a complex solution recipe and ask: Can I drop this element? If doing so would contradict the parsimonious solution, the answer is no. For the remaining elements, the answer is yes if this elimination is consistent with substantive knowledge.

Although it is possible to explain and illustrate CA using generic symbols for causal conditions, as is done next in Example 3, substantive knowledge about adding or removing a

[35] The word "continuum" is used in the following sense (Merriam-Webster's Online Dictionary): a coherent whole characterized as a collection, sequence, or progression of values or elements varying by minute degrees.

causal condition—a thought experiment—is only possible (meaningful) for non-generic causal conditions for which the symbols have actual meanings.

B. Example 3. Illustration of CA

Ragin (Appendix C, when letters are used to replace its linguistic causal conditions) begins with the following complex and parsimonious solutions for a situation in which there are five causal conditions:

Complex solution ($R_C = 5$): *ABCd* + *ABCE* + *ABCF* + *AbdEf* + *BCdEf*
Parsimonious solution ($R_P = 2$): *B* + *E*

In addition, he assumes the following:
Substantive knowledge: The desired outcome could have occurred if *C*, *D*, *E*, or *F* occurred.

Because the substantive knowledge is silent about the first two causal conditions (*A* and *B*), meaning that this knowledge is neutral with respect to those causal conditions, one or both must be included in an intermediate solution if they already appear in a complex solution.

Based on the just-quoted procedure, CA for this example requires $2 \times 5 = 10$ cycles of computation. Note that the order in which the parsimonious and complex terms are examined is unimportant, because the same results are obtained regardless of that ordering.

A complete CA is summarized in Table IX, in which first the parsimonious term *B* is used and then parsimonious term *E* is used. The notation, e.g. $\cdot|_{CA(D)}$ means replace *d* in the complex solution term by *D*. After a counterfactual has been produced, it is unioned with its associated complex solution; this is shown in the second line of the two lines that are in the braces.

In some situations (Cycles 2, 3 and 7), the complex term becomes its own counterfactual because the substantive knowledge is already contained in that term. In other situations (Cycles 4, 6 and 8), a complex term is excluded from CA because it does not contain the parsimonious term.

Seven of the 10 possible cycles produce an intermediate solution, and two pairs of counterfactuals are the same (Cycles 2 and 7 and Cycles 5 and 10), so that each of their unions is that term. The final result is that there are five counterfactuals in the union of the counterfactuals—the intermediate solutions—*ABC* + *ABCE* + *ABCF* +*BCE* + *AbE*. The footnote to Table IX demonstrates that this union reduces to three terms, *ABC* + *BCE* + *AbE*, and it is these three terms that constitute the simplified intermediate solutions.

Observe that each of these simplified intermediate solutions is more complex than the parsimonious solutions and is less complex than the complex solutions. ■

It should be clear from this example that there are some rules that can be extracted from the CA process that can greatly simplify that process.

C. CA Rules

Five CA rules are:

CA Rule 1: *If a complex solution term does not contain the parsimonious solution term, then there is no counterfactual for that combination of parsimonious and complex solution terms* (e.g., see Cycle 4 in Table IX, for which the parsimonious solution term is *B*, and the complex solution term is *AbdEf*).

TABLE IX

THE PROCESS OF COUNTERFACTUAL ANALYSIS (CA); PARSIMONIOUS SOLUTIONS ARE SHOWN UNDERLINED

Cycle	*Parsimonious Solution*	*Complex Solution*	*Results: Counterfactuals*
1	B	$A\underline{B}Cd$	$\left\{\begin{array}{c} A\underline{B}Cd\big\vert_{CA(D)} \rightarrow A\underline{B}CD \\ A\underline{B}Cd + A\underline{B}CD = A\underline{B}C \end{array}\right\} \rightarrow A\underline{B}C$
			It is unnecessary to include E and F in $A\underline{B}C$ because both will be absorbed into $A\underline{B}C$ (see first part of footnote a).
2		$A\underline{B}CE$	Because C and E are already in $A\underline{B}CE$, the counterfactual is $A\underline{B}CE$
3		$A\underline{B}CF$	Because C and F are already in $A\underline{B}CF$, the counterfactual is $A\underline{B}CF$
4		$AbdEf$	Because B is not in $AbdEf$, counterfactual analysis is not done for this term.
5		$\underline{B}CdEf$	$\left\{\begin{array}{c} \underline{B}CdEf\big\vert_{CA(D)} \rightarrow \underline{B}CDEf \\ \underline{B}CdEf + \underline{B}CDEf = \underline{B}CEf \end{array}\right\} \xrightarrow{\underline{B}CEf} \left\{\begin{array}{c} \underline{B}CEf\big\vert_{CA(F)} \rightarrow \underline{B}CEF \\ \underline{B}CEf + \underline{B}CEF = \underline{B}CE \end{array}\right\} \rightarrow \underline{B}CE$
6	E	$ABCd$	Because E is not in $ABCd$, counterfactual analysis is not done for this term.
7		$ABC\underline{E}$	Because C and E are already in $A\underline{B}CE$, the counterfactual is $A\underline{B}CE$
8		$ABCF$	Because E is not in $ABCF$ counterfactual analysis is not done for this term.
9		$Abd\underline{E}f$	$\left\{\begin{array}{c} Abd\underline{E}f\big\vert_{CA(D)} \rightarrow AbD\underline{E}f \\ Abd\underline{E}f + AbD\underline{E}f = Ab\underline{E}f \end{array}\right\} \xrightarrow{Ab\underline{E}f} \left\{\begin{array}{c} Ab\underline{E}f\big\vert_{CA(F)} \rightarrow Ab\underline{E}F \\ Ab\underline{E}f + Ab\underline{E}F = Ab\underline{E} \end{array}\right\} \rightarrow Ab\underline{E}$
			It is unnecessary to include C in $Ab\underline{E}$ because it will be absorbed into it.
10		$BCd\underline{E}f$	$\left\{\begin{array}{c} BCd\underline{E}f\big\vert_{CA(D)} \rightarrow BCD\underline{E}f \\ BCd\underline{E}f + BCD\underline{E}f = BC\underline{E}f \end{array}\right\} \xrightarrow{BC\underline{E}f} \left\{\begin{array}{c} BC\underline{E}f\big\vert_{CA(F)} \rightarrow BC\underline{E}F \\ BC\underline{E}f + BC\underline{E}F = BC\underline{E} \end{array}\right\} \rightarrow BC\underline{E}$
Union of Counterfactuals– Intermediate Solutions			$ABC + ABCE + ABCF + BCE + AbE$
Simplified Intermediate Solutions[a]			$ABC + BCE + AbE$

[a] $ABC + ABCE = ABC(E+e) + ABCE = ABCE + ABCe = ABC(E+e) = ABC$ (absorption); similarly, $ABC + ABCF = ABC$; hence, $ABC + ABCE + ABCF + BCE + AbE = ABC + BCE + AbE$.

CA Rule 2: *If the substantive knowledge is silent about a causal condition (or its complement) that already appears in a complex solution term, then no change is made to that causal condition in a counterfactual term* (e.g., see Cycle 1 in Table IX, for which the complex solution term is *ABCd* and the substantive knowledge is silent about *A* and *B*).

CA Rule 3: *If the substantive knowledge contains a causal condition (or its complement) that already appears in a complex solution term, then the counterfactual for that complex solution term is the same as that complex solution term* (e.g., see Cycle 2 in Table IX, for which the parsimonious solution term is *B*, the complex solution term is *ABCE*, and the substantive knowledge includes *C* and *E*).

CA Rule 4: *If the substantive knowledge contains a causal condition (or its complement) that does not appear in a complex solution term, then that causal condition (or its complement) does not contribute anything to that complex solution term, and such substantive knowledge can be bypassed during the CA for that complex solution term, i.e. such a causal condition (or its complement) is absorbed into the complex solution term* (e.g., see Cycle 1 in Table IX, for which the parsimonious solution term is *B*, the complex solution term is *ABCd*, and the substantive knowledge includes *E* and *F*; see, also, footnote a to that table).

CA Rule 5: *If the substantive knowledge contains the complement of a causal condition (or its complement) that appears in a complex solution term, then the counterfactual for that complex solution term no longer contains that causal condition (or its complement), i.e. it is absorbed* *into the remaining causal combination of that complex solution term* (e.g., see Cycle 1 in Table IX, for which the parsimonious solution term is *B*, the complex solution term is *ABCd*, and the substantive knowledge includes *D*).

Each of these rules is quite useful. They let us simplify a complex solution term like (see Cycle 9 in Table IX) *AbdEf*, when the substantive knowledge is *CDEF*, in *one shot* to *AbE*.

D. Comments

1. Ragin distinguishes between *easy* and *difficult* counterfactuals. According to Fiss (2010):

   > "Easy" counterfactuals refer to situations in which a redundant causal condition is *added* [by means of disjunction] to a causal combination that by itself already led to the outcome in question. As an example, assume we have evidence that the combination of conditions *ABc* leads to the presence of the outcome. We do not have evidence as to whether the combination *ABC* would also lead to the outcome, but theoretical or substantive knowledge links the presence (not the absence [the complement]) of *C* to the outcome. In such a situation, an "easy" CA indicates that both *ABc* and *ABC* will lead to the outcome, and the expression can be reduced to *AB* because whether *C* is absent or present has no effect on the outcome. In "easy" CA, the researcher thus asks: would *adding* another causal combination make a difference? If the answer is "no," we can proceed with the simplified expression.
   >
   > In contrast, "difficult" counterfactuals refer to situations in which a causal condition is *removed* from a set of causal conditions leading to the outcome on the assumption that this causal condition is redundant. For instance, we might have evidence that the causal combination *ABC* leads to the outcome, but we do not have evidence as to whether the causal combination *ABc* would also lead to the outcome in question. This case is of course the inverse of the situation above. In a "difficult" CA, the researcher asks: would removing a causal condition make a difference? This question is more difficult to answer. Theoretical or substantive knowledge links the *presence*, not the absence [the complement], of *C* to the outcome, and since we have no empirical instance of *ABc*, it is much harder to determine whether *C* is in fact a redundant causal condition that may be dropped, thus simplifying the solution to merely *AB*. [If evidence is available, then it is okay to remove a causal condition.]

2. The distinction between the two kinds of counterfactuals is not needed in order to perform CA. Ragin also states that difficult counterfactuals are to be avoided at all costs. In a private e-mail to the senior author (September 15, 2010) he states:

   > The idea of a difficult counterfactual is simply that you shouldn't remove an element from a complex solution if that element makes sense! The parsimonious solution sometimes does exactly that, and the purpose of CA is to put it back in! The parsimonious solution doesn't care which counterfactuals are easy and which are difficult (from the perspective of existing knowledge).

3. In a private (2009) e-mail, the first author asked Ragin:

   > The extreme situation for counterfactuals would be if we began with no cases. Then the final results from QCA would be based entirely on one's substantive knowledge. I doubt anyone would take those results seriously. So, in your experience what is the "breakpoint," if you will, between this extreme situation and situations where counterfactuals are used, i.e. what percentage of the total number of causal combinations can be treated as counterfactuals with people still believing the results?

   He replied:

   > This is a great question. You are way ahead of my colleagues in social science. I like to think of it in terms of the width of the interval between the parsimonious and the complex solution. If the parsimonious solution is *A* and the complex solution is *AbCDe*, then the width is great, perhaps too great. It also depends on the number of initial conditions in the truth table, of course. This can be formalized and I think I've done it in my notes somewhere. Still, the short answer to your question is that *I don't have a breakpoint*. My guess also is that it depends in part on the nature of the research question

4. Surprisingly, there is nothing written about CA in Rihoux and Ragin (2009), which suggests to this author that CA is problematic for the social scientists who wrote the chapters of that textbook. In a private (2009) e-mail to the senior author, Ragin states:

   > In general, social scientists have been very slow to catch on to the idea of CA and to the fact that some are easy and some are difficult. I published the ideas about five years ago because (almost) everyone was reporting parsimonious solutions and never checking to see which remainders had been incorporated and then assessing them. What I sketched then was how to do CA by hand, with a little algebra, comparing the parsimonious and complex solutions. Still, no one really did it, because it seemed all too technical. So I had it programmed into fsQCA. Finally, social scientists are starting to try it. Still, there are very few who (previously) checked all the remainders (one at a time), even though this was my original recommendation. It's just too demanding for most social scientists. They are accustomed to privileging parsimony, which I think is a mistake.

   Some references for CA are Ragin (2008, Chapters 8 and 9) and Ragin and Sonnett (2004). Also, Appendix C provides a "New Intermediate Solution Algorithm," sent to the senior author by Ragin in December of 2009.

5. In August, 2010 the first author asked Prof. Peer Fiss why CA, that seems to fill in missing cases, is not as objectionable as what is done in conventional methods (e.g., regression), in which interpolation or extrapolation can be used after-the-fact to obtain results for cases that were not originally available. His answer to this question is [expressed in our own words]: The causal combinations that are included as part of CA are based on and supported by the substantive knowledge of an expert, and these causal combinations are used to lead to intermediate solutions, whereas the interpolated or extrapolated conclusions drawn from, e.g. a regression occur after the regression is completed, and, therefore are not supported by such knowledge, and are meant to be a replacement for such knowledge, but this can not be verified because such knowledge does not exist. This answer is best embodied by Item 3 in Appendix B, and is so important that it is repeated here:

   > ***Conventional methods have a de facto dependence on simplifying assumptions about kinds of cases not found in the data set.*** Limited diversity is the rule, not the exception, in the study of naturally occurring social phenomena. Once researchers identify relevant causal variables, they typically find that many regions of the vector space defined by those variables[36] are void, or virtually void, of cases. When conventional researchers estimate statistical models using data that are limited in their diversity, the additive-linear techniques they typically apply to their data assume, in essence, that if there were cases in the vacant regions in the vector space, they would conform to the patterns exhibited by cases in the regions that are well populated with cases. Thus, these models incorporate de facto assumptions about kinds of cases that are absent or virtually absent from the researcher's data set. Unfortunately, these assumptions are invisible, not only to most researchers, but also to the audience for social research. In the fuzzy set approach, by contrast, the consideration of limited diversity is an explicit part of the analysis. Not only do researchers identify simplifying assumptions, it is possible, as well, to evaluate the plausibility of these assumptions and then selectively incorporate those that seem plausible. The process of incorporating simplifying assumptions is explicit and visible. Furthermore, the audiences for social research are free to challenge such assumptions and to construct alternate representations of the same evidence.

E. CA Applied to Examples 1 and 2

In the rest of this section CA is applied to the results for Examples 1 and 2 (Section III).

E.1 Example 1–Continued: Three causal conditions and O = Breakdown of Democracy: Recall, from Section III.A, that:

[36] Recall, that if there are k causal variables, then there are 2^k corners in the Boolean vector space.

- *Complex solution obtained by hand* [see first line of (30a)] ($R_C = 1$): *ab*
- *Parsimonious solution obtained from QM* [see second line of (30a)] ($R_P = 1$): *a*
- *Substantive knowledge* (these were made up by us, but seemed reasonable): The desired outcome could have occurred if *a* (not developed), *b* (not urban (rural)) or *c* (not literate) occurred.
- *Counterfactual Analysis*: Using CA Rules 2 and 3, one finds that the intermediate solution is *ab*, or, in words:

Not developed and *not urban* (rural) is a sufficient causal combination for *Breakdown of Democracy* (34)

According to Wagemann and Schneider (2007), it is a good practice to summarize the intermediate solutions in a table such as the one in Table X.

TABLE X
SUMMARY FOR THE BELIEVABLE SIMPLIFIED INTERMEDIATE SOLUTION (BSIS) OF EXAMPLE 1

Case	*MFs for Outcome and Causal Conditions*			*BSIS MF*	*In the Desirable Region (Fig. 5)— Best Instance?*
	MF(O)	*MF(A)*	*M(B)*	*MF(ab)*	
1	0.95	0.81	0.12	0.19	No
2	0.05	0.99	0.89	0.01	No
3	0.11	0.58	0.98	0.02	No
4	0.88	0.16	0.07	0.84	**Yes**
5	0.23	0.58	0.03	0.42	No
6	0.05	0.98	0.03	0.02	No
7	0.95	0.89	0.79	0.11	No
8	0.94	0.04	0.09	0.91	**Yes**
9	0.58	0.07	0.16	0.84	No
10	0.08	0.72	0.05	0.28	No
11	0.95	0.34	0.10	0.66	**Yes**
12	0.05	0.98	1.00	0	No
13	0.88	0.02	0.17	0.83	**Yes**
14	0.95	0.01	0.02	0.98	**Yes**
15	0.79	0.01	0.03	0.97	No
16	0.94	0.03	0.30	0.70	**Yes**
17	0.05	0.95	0.13	0.05	No
18	0.05	0.98	0.99	0.01	No

- *Consistency*: The set theoretic consistency of *ab* was computed to be [using (11)] 0.837, which is greater than 0.80, so this solution is retained, and therefore *ab* is the believable simplified intermediate solution (and $R_{BSI} = 1$).
- *Best Instances*: Referring to the three-step procedure that is given in Section II. E: Step 1 leads to the numbers that are in the column of Table X called "BSIS MF;" Step 2 does not have to be performed because $R_{BSI} = 1$; and, Step 3 leads to the best instances that are labeled "**Yes**" in the last column of Table X. We do not show a plot of $\mu_O(x)$ versus $\mu_{ab}(x)$ for $x = 1,...,18$ because it is very messy since many of the cases lie almost on top of one another. Observe that the Best Instances for *ab* are Cases 4, 8, 11, 13, 14, and 16.

- *Coverage*: Using $MF(O)$ and $MF(ab)$ that are given in Table X it is straightforward to compute the raw coverage[37] in (27), as $C_r(ab, O) = 0.736$.

E.2 Example 2–Continued: Three causal conditions and O = Likely Breakdown of Democracy: Recall, from Section III.B, that:

- *Complex solution obtained by hand* [see (32)] ($R_C = 2$): *ab* + *AC*
- *Parsimonious solution*: As mentioned in Example 2 (just below (31)), there are no minimum prime implicants, because $X_2 = \varnothing$ (see Theorem 1); hence, there is no parsimonious solution.
- *Counterfactual Analysis*: No CA is possible because there is no parsimonious solution; hence, the final solution is the same as the complex solution, and is *ab* + *AC*, or, in words:

(*Not developed* and *not urban (rural)*) OR (*Developed* and *Literate*)
are sufficient causal combinations for *Likely Breakdown of Democracy* (35)

Results for all cases are summarized in Table XI.

TABLE XI
SUMMARY FOR THE BELIEVABLE SIMPLIFIED INTERMEDIATE SOLUTION (BSIS) OF EXAMPLE 2

Case	*MFs for Outcome and Causal Conditions*				*BSIS #1 MF*	*BSIS #2 MF*	*Maximum MF of BSISs*		*In the Desirable Region (Fig. 5)—Best Instance?*
	MF(O)	*MF(A)*	*MF(B)*	*MF(C)*	*MF(ab)*	*MF(AC)*	*MF*	*BSIS*	
1	0.95	0.81	0.12	0.99	0.09	0.81	0.81	*AC*	**Yes**
4	0.88	0.16	0.07	0.98	0.84	0.16	0.84	*ab*	**Yes**
7	0.95	0.89	0.79	0.99	0.11	0.89	0.89	*AC*	**Yes**
8	0.94	0.04	0.09	0.13	0.91	0.13	0.91	*ab*	**Yes**
9	0.58	0.07	0.16	0.88	0.84	0.07	0.84	*ab*	No
11	0.95	0.34	0.10	0.41	0.66	0.34	0.66	*ab*	**Yes**
13	0.88	0.02	0.17	0.59	0.83	0.02	0.83	*ab*	**Yes**
14	0.95	0.01	0.02	0.01	0.98	0.01	0.98	*ab*	**Yes**
15	0.79	0.01	0.03	0.17	0.97	0.01	0.97	*ab*	No
16	0.94	0.03	0.30	0.09	0.70	0.03	0.70	*ab*	**Yes**

- *Consistency*: The consistency for *ab* is 0.93 and for *AC* is 1 (both were computed using (11)), and since they are both greater than 0.80 they are both retained, and both are believable simplified intermediate solutions ($R_{BSI} = 2$).
- *Best Instances*: Referring to the three-step procedure that is given in Section II. E: Step 1 leads to the numbers that are in the columns of Table XI called "BSIS#1MF" and "BSIS#2MF;" Step 2 has to be performed because $R_{BSI} = 2$, and the results of doing this are in the column of Table XI called "Maximum MF of BSISs" [in this example only one believable simplified intermediate solution is retained for each case, i.e. $Q = 1$ for all of the cases (see (20)]; and, Step 3 leads to the best instances that are labeled "**Yes**" in the last column of Table XI. Observe that the Best Instances for *ab* are Cases 4, 8, 11, 13, 14, and 16, and the Best Instances for *AC* are Cases 1 and 7.

[37] Because there is only one intermediate term, raw coverage, unique coverage and solution coverage are all the same [Rihoux and Ragin (2009, p. 64)].

- *Coverage*: Using $MF(O)$ and $MF(ab)$ that are given in Table XI, it is straightforward to compute the raw coverage in (27), as $C_r(ab,O) = 0.736$ and $C_r(AC,O) = 0.270$.

V. New Theoretical Results

A. Introduction

In this section, new theoretical results are presented that not only expedite the fsQCA calculations in two different ways, one for the computation of firing levels of the 2^k causal combinations, and the other for computation of consistency, but also establish when cases for which $MF(O) < 0.5$ can cause rules to be eliminated. None of these results could have been obtained without the mathematical formalization of fsQCA that has been presented in Section II.

B. A Simplified Way to Determine the Winning Causal Combination for Each Case

Examination of Tables I and V reveals that each of their rows contains only one element (in boldface) whose membership is greater than 0.50. This was already observed in Ragin (2008, p. 131) who states: "This table demonstrates an important property of intersections of fuzzy sets, namely, that each case can have (at most) only a single membership score greater than 0.5 in the logical possible combinations from a given set of causal conditions." Because no mathematical proof of this fact is given by Ragin, this section begins with such a proof. The proof includes a direct way for computing the one causal combination (out of 2^k possible causal combinations) for each case whose membership score is greater than 0.5. Theorem 2 leads to a major simplification of the way in which computations are performed in fsQCA, something that is discussed later in this section.

Theorem 2. Given k causal conditions, $C_1, C_2, ..., C_k$ and their respective complements, $c_1, c_2, ..., c_k$. Form the 2^k causal combinations $F_i = A_1^i \wedge A_2^i \wedge ... \wedge A_k^i$ ($i = 1, ..., 2^k$) where $A_j^i = C_j$ or c_j and $j = 1, ..., k$. Let $\mu_{F_i}(x) = \min\left\{\mu_{A_1^i}(x), \mu_{A_2^i}(x), ..., \mu_{A_k^i}(x)\right\}$, $x = 1, 2, ..., N$. Then for each x there is only one i, i^*, for which $\mu_{F_{i^*}}(x) > 0.5$ and $\mu_{F_{i^*}}(x)$ can be computed as:

$$\mu_{F_{i^*}}(x) = \min\left\{\max\left(\mu_{C_1}(x), \mu_{c_1}(x)\right), \max\left(\mu_{C_2}(x), \mu_{c_2}(x)\right), ..., \max\left(\mu_{C_k}(x), \mu_{c_k}(x)\right)\right\} \tag{36}$$

$F_{i^*}(x)$ is determined from the right-hand side of (36), as:

$$F_{i^*}(x) = arg\max\left(\mu_{C_1}(x), \mu_{c_1}(x)\right) arg\max\left(\mu_{C_2}(x), \mu_{c_2}(x)\right) \; ... \; arg\max\left(\mu_{C_k}(x), \mu_{c_k}(x)\right) \tag{37}$$

where $arg\max\left(\mu_{C_j}(x), \mu_{c_j}(x)\right)$ denotes the winner of $\max\left(\mu_{C_j}(x), \mu_{c_j}(x)\right)$, namely C_j or c_j.

Proof[38]: If $A_j^i = C_j$ and $\mu_{C_j}(x) = \min\left\{\mu_{C_j}(x), \mu_{c_j}(x)\right\}$, then it is true that $\mu_{c_j}(x) = \max\left\{\mu_{C_j}(x), \mu_{c_j}(x)\right\}$. If, instead, $\mu_{C_j}(x) = \max\left\{\mu_{C_j}(x), \mu_{c_j}(x)\right\}$ then it is true that $\mu_{c_j}(x) = \min\left\{\mu_{C_j}(x), \mu_{c_j}(x)\right\}$. Consequently, choosing $A_j^i = C_j$ or c_j is equivalent to choosing

[38] This proof was provided by Ms. Jhiin Joo.

$\mu_{A_j^i}(x)$ as either $\min\left\{\mu_{C_j}(x), \mu_{c_j}(x)\right\}$ or $\max\left\{\mu_{C_j}(x), \mu_{c_j}(x)\right\}$, so that rather than thinking about the 2^k causal combinations for each case in terms of C_j and c_j, with their associated MFs $\mu_{C_j}(x)$ and $\mu_{c_j}(x)$, one can think about the 2^k causal combinations for each case in terms of the following new ordering of the F_i and their MFs, $\mu_{F_i}(x)$:

$$\begin{cases} \mu_{F_1}(x) = \min\left\{\min\left(\mu_{C_1}(x), \mu_{c_1}(x)\right), \min\left(\mu_{C_2}(x), \mu_{c_2}(x)\right), ..., \min\left(\mu_{C_k}(x), \mu_{c_k}(x)\right)\right\} \\ \mu_{F_2}(x) = \min\left\{\min\left(\mu_{C_1}(x), \mu_{c_1}(x)\right), \min\left(\mu_{C_2}(x), \mu_{c_2}(x)\right), ..., \max\left(\mu_{C_k}(x), \mu_{c_k}(x)\right)\right\} \\ \qquad \qquad \qquad \qquad \ldots \\ \mu_{F_{2^k}}(x) = \min\left\{\max\left(\mu_{C_1}(x), \mu_{c_1}(x)\right), \max\left(\mu_{C_2}(x), \mu_{c_2}(x)\right), ..., \max\left(\mu_{C_k}(x), \mu_{c_k}(x)\right)\right\} \end{cases} \tag{38}$$

Because it is always true that

$$\min\left\{\mu_{C_j}(x), \mu_{c_j}(x)\right\} \leq 0.5\,, \tag{39}$$

when such terms are evaluated in (38) the first $2^k - 1$ terms will always have a MF value that is also ≤ 0.5. It is only the last term in (38) that can have a MF value that is > 0.5, and that term is the one in (36).

Observe that (37) is an immediate consequence of the last line of (38) [which is the same as (36)], and that $i^* = 2^k$. ■

Example 4. In order to illustrate (36) and (37), we shall focus on Case 1 in Table II, for which $F_{i^*}(1) = AbC$; hence,

$$\begin{aligned} \mu_{F_{i^*}}(1) &= \min\left\{\max\left(MF(A), MF(a)\right), \max\left(MF(B), MF(b)\right), \max\left(MF(C), MF(c)\right)\right\} \\ &= \min\left\{\max\left(0.81, 0.19\right), \max\left(0.12, 0.88\right), \max\left(0.99, 0.01\right)\right\} \end{aligned} \tag{40}$$

$$\mu_{F_{i^*}}(1) = \min\left\{0.81, 0.88, 0.99\right\} = 0.81 \tag{41}$$

Observe, also, that:

$$F_{i^*}(1) = arg\max\left(0.81, 0.19\right) arg\max\left(0.12, 0.88\right) arg\max\left(0.99, 0.01\right) = AbC \tag{42}$$

This agrees with the boldface item in the Case-1 row of Table II. ■

The procedure that has just been used in this example, which can be referred to as the *min-max procedure* (for simplicity), is very different from the one that was used to create Table II. It only computes one number for each case instead of 2^k numbers, which is a huge savings in computation. Of course, one can no longer fill in all of the numbers in Table II, but most of those numbers are irrelevant to succeeding steps of fsQCA. In addition, this new procedure directly provides the subset of firing-level surviving rules and their associated cases. Because only this subset of rules is used in remaining fsQCA computations, it is much easier to display the results for just the surviving rules. A table such as Table II is no longer needed to do this, as is demonstrated in our next example.

This is also important because of another reason. While it is easy to display Table II for a small number of causal conditions and cases, it is difficult-to-impossible to do this for even five causal conditions, for which there are $2^5 = 32$ causal combinations. A table with 32 causal condition columns would have to be printed in such small type that it would not be readable.

Example 5. Applying Theorem 1 to all 18 cases in Table I, one obtains (calculations are left to the reader): Case $1 \rightarrow AbC$, Case $2 \rightarrow ABC$, Case $3 \rightarrow ABC$, Case $4 \rightarrow abC$, Case $5 \rightarrow AbC$, Case $6 \rightarrow AbC$, Case $7 \rightarrow ABC$, Case $8 \rightarrow abc$, Case $9 \rightarrow abC$, Case $10 \rightarrow AbC$, Case $11 \rightarrow abc$, Case $12 \rightarrow ABC$, Case $13 \rightarrow abC$, Case $14 \rightarrow abc$, Case $15 \rightarrow abc$, Case $16 \rightarrow abc$, Case $17 \rightarrow AbC$ and Case $18 \rightarrow ABC$. Examining these 18 results, one observes that only four causal combinations have survived, namely: *AbC* (Cases 1, 5, 6, 10, and 17), *ABC* (Cases 2, 3, 7, 12, and 18), *abC* (Cases 4, 9, and 13) and *abc* (Cases 8, 11, 14, 15 and 16), which agree with our previous results in Example 1. Table III collects these results; hence, Table II is no longer needed, and one can go directly from Table I to Table III. ■

Based on Theorem 1, one is able to modify the Figs. 2 and 3 fsQCA flowcharts to the ones that are depicted in Figs. 6 and 7, respectively. Comparing Figs. 2 and 6, observe that no firing levels are now computed, whereas previously 2^k firing intervals were computed. Comparing Figs. 3 and 7, observe that firing intervals are computed but only after the R_S surviving rules have been established, and then only R_S (and not 2^k) firing intervals are computed for all of the cases.

C. A Simplified Way to Determine the Winning Causal Combination for Each Case When the Number of Causal Conditions Changes

Sometimes one wants to perform fsQCA for different combinations of causal conditions, by either including more causal conditions into the mix of the original *k* causal conditions, or by removing some of the original *k* causal conditions. Presently, doing any of these things requires treating each modified set of causal conditions as a totally new fsQCA. The results in this section show that there are much easier ways to perform fsQCA once it has already been performed for *k* causal conditions.

Observe in (37) that, e.g., $arg\max\left(\mu_{C_1}(x), \mu_{c_1}(x)\right)$ is unchanged whether there are one, two, three, etc. causal conditions. This means that, for each case, the winning causal combination for *k* causal conditions includes the winning causal combination for k' causal conditions, when $k' < k$, and is contained in the winning causal combination for k'' causal conditions, when $k'' > k$. It also means that if one knows the winning causal combination for k'' causal conditions, where $k'' > k$, and one wants to know the winning causal combination for *k* causal conditions, one simply deletes the undesired $k'' - k$ causal conditions from the winning causal combination of k'' causal conditions.

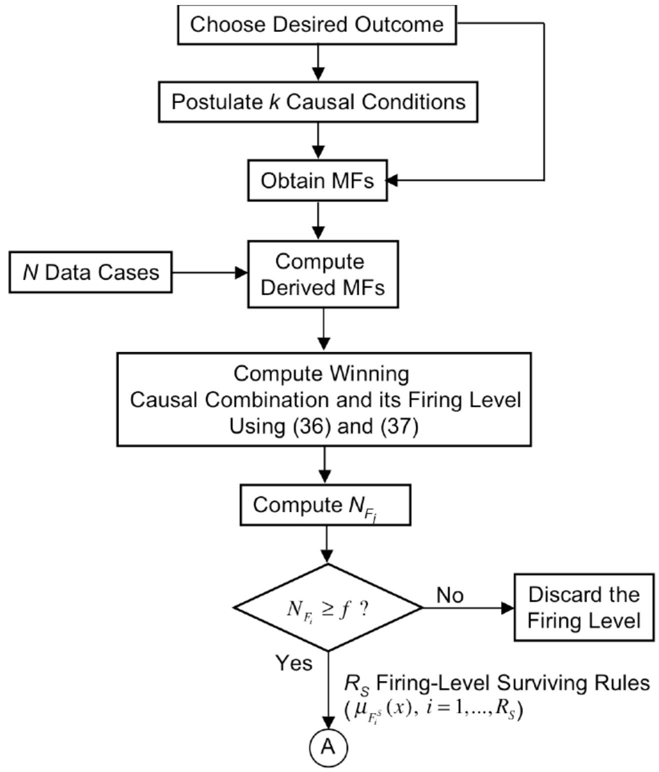

Figure 6. Modified Fig. 2 flowchart for fsQCA; it is continued in Fig. 7.

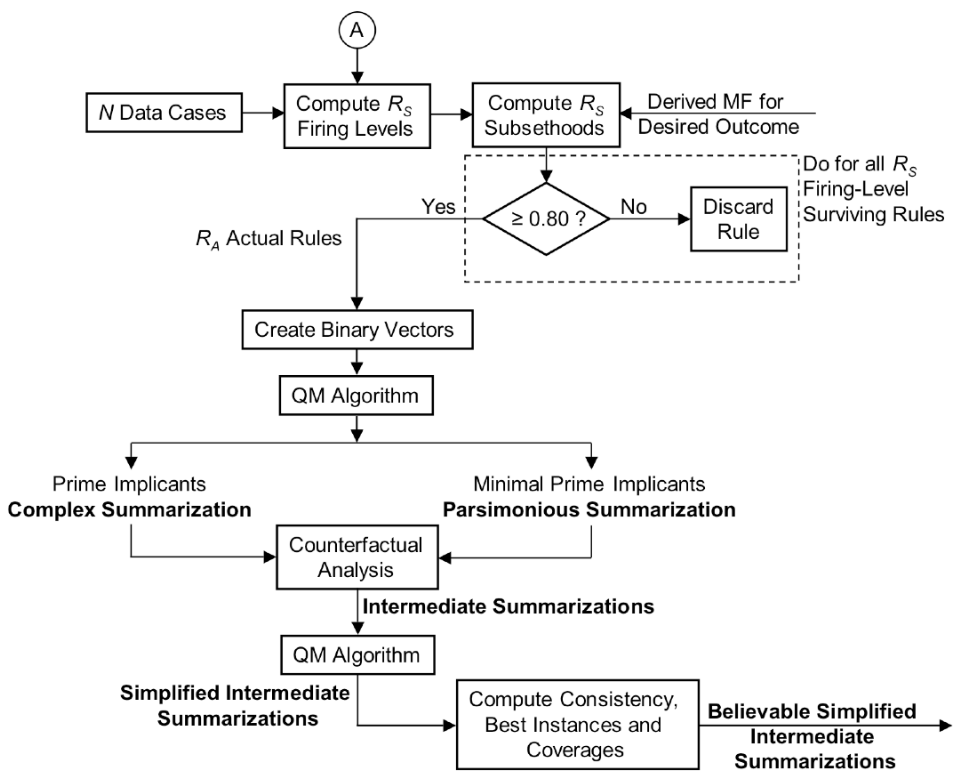

Figure 7. Modified Fig. 3 flowchart for fsQCA.

For example, if $AbcdE$ is a winning causal combination for case 1 when five causal conditions are used, and if one wants to eliminate causal conditions B and D, then AcE is the winning causal combination for case 1 when the three causal conditions A, C and E are used. No new computations have to be performed to obtain this result, because the winning causal combination for case 1 when the three causal conditions A, C and E are used is *contained* in the winning causal combination for case 1 when five causal conditions are used.

These observations suggest that there are both a *forward recursion* and a *backward recursion* for (36) and (37).

In what follows, it is assumed that the smallest number of causal conditions for which an fsQCA is performed is two.

Corollary 2-1 (Forward Recursion). For each case, it is true that ($k = 3,4,\ldots$):

$$F_{i*}(x \mid C_1, C_2, \ldots, C_k) = F_{i*}(x \mid C_1, C_2, \ldots, C_{k-1}) arg \max\left(\mu_{C_k}(x), \mu_{c_k}(x)\right) \tag{44}$$

$$\mu_{F_{i*}}(x \mid C_1, C_2, \ldots, C_k) = \min\left\{\mu_{F_{i*}}(x \mid C_1, C_2, \ldots, C_{k-1}), \max\left(\mu_{C_k}(x), \mu_{c_k}(x)\right)\right\} \tag{45}$$

These results, arguably for the first time, connect fsQCA firing-level calculations for k and k - 1 causal conditions. They *provide an entirely new way to perform fsQCA computations* when one wishes to study different combinations of causal conditions on a desired outcome, and should lead to a vast reduction in computation time for such a study.

Proof: It is easy to prove both (44) and (45) by using mathematical induction, and this is left to the reader.

Example 6: Here the proofs of (44) and (45) are illustrated for $k = 3$.

(a) *Proof of (44)*: From (37), it follows that:

$$F_{i*}(x \mid C_1, C_2) = arg \max\left(\mu_{C_1}(x), \mu_{c_1}(x)\right) arg \max\left(\mu_{C_2}(x), \mu_{c_2}(x)\right) \tag{46}$$

$$F_{i*}(x \mid C_1, C_2, C_3) = arg \max\left(\mu_{C_1}(x), \mu_{c_1}(x)\right) arg \max\left(\mu_{C_2}(x), \mu_{c_2}(x)\right) arg \max\left(\mu_{C_3}(x), \mu_{c_3}(x)\right) \tag{47}$$

Comparing (46) and (47), it is easy to see that:

$$F_{i*}(x \mid C_1, C_2, C_3) = F_{i*}(x \mid C_1, C_2) arg \max\left(\mu_{C_3}(x), \mu_{c_3}(x)\right) \tag{48}$$

which is (44).

(b) *Proof of (45)*: From (36), it follows that:

$$\mu_{F_{i*}}(x \mid C_1, C_2) = \min\left\{\max\left(\mu_{C_1}(x), \mu_{c_1}(x)\right), \max\left(\mu_{C_2}(x), \mu_{c_2}(x)\right)\right\} \tag{49}$$

$$\mu_{F_{i*}}(x \mid C_1, C_2, C_3) = \min\left\{\max\left(\mu_{C_1}(x), \mu_{c_1}(x)\right), \max\left(\mu_{C_2}(x), \mu_{c_2}(x)\right), \max\left(\mu_{C_3}(x), \mu_{c_3}(x)\right)\right\} \quad (50)$$

(50) can also be expressed as:

$$\mu_{F_{i*}}(x \mid C_1, C_2, C_3) = \min\left\{\min\left\{\max\left(\mu_{C_1}(x), \mu_{c_1}(x)\right), \max\left(\mu_{C_2}(x), \mu_{c_2}(x)\right)\right\}, \max\left(\mu_{C_3}(x), \mu_{c_3}(x)\right)\right\} \quad (51)$$

which can then be expressed as:

$$\mu_{F_{i*}}(x \mid C_1, C_2, C_3) = \min\left\{\mu_{F_{i*}}(x \mid C_1, C_2), \max\left(\mu_{C_3}(x), \mu_{c_3}(x)\right)\right\} \quad (52)$$

which is (45). ■

Corollary 2-2 (Backward Recursion). Let $\not{C}_j$ denote the suppression of causal condition C_j. Then it is true that:

$$F_{i*}(x \mid C_1, C_2, ..., \not{C}_j, ..., C_k) = F_{i*}(x \mid C_1, C_2, ..., C_{j-1}, C_{j+1}, ..., C_k) \quad (53)$$

Proof: Obvious from (37).

This backward recursion can also lead to a vast reduction in computation time. For example, if the winning causal combination $F_{i*}(C_1, C_2, ..., C_k)$ has been determined for six causal conditions ($k = 6$), then it can be used to establish the winning causal combination for any combination of five, four, three, or two of the causal conditions, *by inspection*!

No way has yet been determined for computing $\mu_{F_{i*}}(x \mid C_1, C_2, ..., \not{C}_j, ..., C_k)$ from $\mu_{F_{i*}}(x \mid C_1, C_2, ..., C_j, ..., C_k)$. It seems that once $F_{i*}(C_1, C_2, ..., \not{C}_j, ..., C_k)$ has been determined from (53), $\mu_{F_{i*}}(x \mid C_1, C_2, ..., \not{C}_j, ..., C_k)$ must be computed directly from (36), as:

$$\begin{aligned}\mu_{F_{i*}}(x \mid C_1, C_2, ..., \not{C}_j, ..., C_k) = \min\Big[\Big\{&\max\left(\mu_{C_1}(x), \mu_{c_1}(x)\right), ..., \max\left(\mu_{C_{j-1}}(x), \mu_{c_{j-1}}(x)\right), \\ &\max\left(\mu_{C_{j+1}}(x), \mu_{c_{j+1}}(x)\right), ..., \max\left(\mu_{C_k}(x), \mu_{c_k}(x)\right)\Big\}\Big]\end{aligned} \quad (54)$$

Corollary 2-3 (Firing Levels are Bounded). If $\mu_{F_{i*}}(x \mid C_1, C_2, ..., C_{k_1})$ has been computed for k_1 causal conditions, and one now considers k_2 causal conditions, where $k_2 > k_1$, then

$$\mu_{F_{i*}}(x \mid C_1, C_2, ..., C_{k_2}) \le \mu_{F_{i*}}(x \mid C_1, C_2, ..., C_{k_1}) \quad (55)$$

This means that when new causal conditions are added to an existing set of causal conditions, the firing level for the new winning causal combination (which, by Corollary 2-1, contains the prior winning causal combination) can never be larger than the prior firing level, i.e. firing levels tend to become weakened when more causal conditions are included.

Proof: (55) is obvious from (45). ■

D. Recursive Computation of Consistency

The formula for subsethood of the antecedents (causal combinations) in the outcome of each firing-level surviving rule is given in (11), which is repeated here for the convenience of the readers:

$$ss_K(F_l^S, O) = \frac{\sum_{x=1}^{N} \min(\mu_{F_l^S}(x), \mu_O(x))}{\sum_{x=1}^{N} \mu_{F_l^S}(x)} \quad l = 1,2,...,R_S \tag{56}$$

Consider two populations, one of size N_1 and the other of size N_2. In order to show the dependency of $ss_K(F_l^S, O)$ on the population size, we use a conditioning notation, i.e.:

$$ss_K(F_l^S, O \mid N_1) = \frac{\sum_{x=1}^{N_1} \min(\mu_{F_l^S}(x), \mu_O(x))}{\sum_{x=1}^{N_1} \mu_{F_l^S}(x)} \quad l = 1,2,...,R_S \tag{57}$$

$$ss_K(F_l^S, O \mid N_2) = \frac{\sum_{x=1}^{N_2} \min(\mu_{F_l^S}(x), \mu_O(x))}{\sum_{x=1}^{N_2} \mu_{F_l^S}(x)} \quad l = 1,2,...,R_S \tag{58}$$

Theorem 3. Suppose $N_2 > N_1$. $ss_K(F_l^S, O \mid N_2)$ can be computed recursively from $ss_K(F_l^S, O \mid N_1)$, as:

$$ss_K(F_l^S, O \mid N_2) = \frac{1}{1 + \dfrac{\sum_{x=N_1+1}^{N_2} \mu_{F_l^S}(x)}{\sum_{x=1}^{N_1} \mu_{F_l^S}(x)}} \times ss_K(F_l^S, O \mid N_1) + \frac{\sum_{x=N_1+1}^{N_2} \min(\mu_{F_l^S}(x), \mu_O(x))}{\sum_{x=1}^{N_2} \mu_{F_l^S}(x)} \tag{59}$$

After proving this result, we shall show how it can be used to provide an understanding of when or if a rule that has survived fsQCA based on N_1 cases can be obliterated (or not) by an additional set of $N_2 - N_1$ cases.

Proof: Beginning with (58) and $N_2 > N_1$, it follows that:

$$ss_K(F_l^S, O \mid N_2) = \frac{\sum_{x=1}^{N_1} \min(\mu_{F_l^S}(x), \mu_O(x)) + \sum_{x=N_1+1}^{N_2} \min(\mu_{F_l^S}(x), \mu_O(x))}{\sum_{x=1}^{N_2} \mu_{F_l^S}(x)}$$

$$ss_K(F_l^S,O\,|\,N_2)=\frac{\sum_{x=1}^{N_1}\min(\mu_{F_l^S}(x),\mu_O(x))}{\sum_{x=1}^{N_2}\mu_{F_l^S}(x)}+\frac{\sum_{x=N_1+1}^{N_2}\min(\mu_{F_l^S}(x),\mu_O(x))}{\sum_{x=1}^{N_2}\mu_{F_l^S}(x)}$$

$$ss_K(F_l^S,O\,|\,N_2)=\frac{\sum_{x=1}^{N_1}\min(\mu_{F_l^S}(x),\mu_O(x))}{\sum_{x=1}^{N_1}\mu_{F_l^S}(x)+\sum_{x=N_1+1}^{N_2}\mu_{F_l^S}(x)}+\frac{\sum_{x=N_1+1}^{N_2}\min(\mu_{F_l^S}(x),\mu_O(x))}{\sum_{x=1}^{N_2}\mu_{F_l^S}(x)}$$

$$\begin{aligned}ss_K(F_l^S,O\,|\,N_2)=&\frac{1}{1+\dfrac{\sum_{x=N_1+1}^{N_2}\mu_{F_l^S}(x)}{\sum_{x=1}^{N_1}\mu_{F_l^S}(x)}}\times\frac{\sum_{x=1}^{N_1}\min(\mu_{F_l^S}(x),\mu_O(x))}{\sum_{x=1}^{N_1}\mu_{F_l^S}(x)}\\&+\frac{\sum_{x=N_1+1}^{N_2}\min(\mu_{F_l^S}(x),\mu_O(x))}{\sum_{x=1}^{N_2}\mu_{F_l^S}(x)}\end{aligned}\qquad(60)$$

Using the formula for $ss_K(F_l^S,O\,|\,N_1)$ that is given in (57) it is easy to see that (60) can be expressed as in (59). ■

Note that $\sum_{x=1}^{N_2}\mu_{F_l^S}(x)$ can also be computed recursively, but we do not pursue that here.

Equation (59) allows us to examine the relative sizes of $ss_K(F_l^S,O\,|\,N_2)$ and $ss_K(F_l^S,O\,|\,N_1)$. Let

$$ss_K(F_l^S,O\,|\,N_2-N_1)\equiv\frac{\sum_{x=N_1+1}^{N_2}\min(\mu_{F_l^S}(x),\mu_O(x))}{\sum_{x=N_1+1}^{N_2}\mu_{F_l^S}(x)}\qquad(61)$$

$ss_K(F_l^S,O\,|\,N_2-N_1)$ is the subsethood of F_l^S in *O*, but only for the new N_2-N_1 cases.

Corollary 3-1. If

$$ss_K(F_l^S,O\,|\,N_2-N_1)\substack{\geq\\<}ss_K(F_l^S,O\,|\,N_1)\qquad(62)$$

then

$$ss_K(F_l^S,O\,|\,N_2)\substack{\geq\\<}ss_K(F_l^S,O\,|\,N_1)\qquad(63)$$

Proof: When $ss_K(F_l^S,O\,|\,N_2)\geq ss_K(F_l^S,O\,|\,N_1)$, then, using (59), it follows that:

$$\frac{1}{1+\dfrac{\sum_{x=N_1+1}^{N_2}\mu_{F_l^S}(x)}{\sum_{x=1}^{N_1}\mu_{F_l^S}(x)}}\times ss_K(F_l^S,O\,|\,N_1)+\frac{\sum_{x=N_1+1}^{N_2}\min(\mu_{F_l^S}(x),\mu_O(x))}{\sum_{x=1}^{N_2}\mu_{F_l^S}(x)}\geq ss_K(F_l^S,O\,|\,N_1) \quad (64)$$

This inequality can be reorganized as follows:

$$\begin{aligned}
\frac{\sum_{x=N_1+1}^{N_2}\min(\mu_{F_l^S}(x),\mu_O(x))}{\sum_{x=1}^{N_2}\mu_{F_l^S}(x)} &\geq ss_K(F_l^S,O\,|\,N_1)\left[1-\frac{1}{1+\dfrac{\sum_{x=N_1+1}^{N_2}\mu_{F_l^S}(x)}{\sum_{x=1}^{N_1}\mu_{F_l^S}(x)}}\right]\\
&\geq ss_K(F_l^S,O\,|\,N_1)\left[1-\frac{\sum_{x=1}^{N_1}\mu_{F_l^S}(x)}{\sum_{x=1}^{N_1}\mu_{F_l^S}(x)+\sum_{x=N_1+1}^{N_2}\mu_{F_l^S}(x)}\right]\\
&\geq ss_K(F_l^S,O\,|\,N_1)\left[\frac{\sum_{x=N_1+1}^{N_2}\mu_{F_l^S}(x)}{\sum_{x=1}^{N_1}\mu_{F_l^S}(x)+\sum_{x=N_1+1}^{N_2}\mu_{F_l^S}(x)}\right]
\end{aligned}$$

$$\frac{\sum_{x=N_1+1}^{N_2}\min(\mu_{F_l^S}(x),\mu_O(x))}{\sum_{x=1}^{N_2}\mu_{F_l^S}(x)}\geq ss_K(F_l^S,O\,|\,N_1)\left[\frac{\sum_{x=N_1+1}^{N_2}\mu_{F_l^S}(x)}{\sum_{x=1}^{N_2}\mu_{F_l^S}(x)}\right] \quad (65)$$

A further simplification of (65) leads to:

$$\sum_{x=N_1+1}^{N_2}\min(\mu_{F_l^S}(x),\mu_O(x))\geq ss_K(F_l^S,O\,|\,N_1)\times\sum_{x=N_1+1}^{N_2}\mu_{F_l^S}(x) \quad (66)$$

This can be expressed as:

$$\frac{\sum_{x=N_1+1}^{N_2}\min(\mu_{F_l^S}(x),\mu_O(x))}{\sum_{x=N_1+1}^{N_2}\mu_{F_l^S}(x)}\geq ss_K(F_l^S,O\,|\,N_1) \quad (67)$$

Because the left-hand side of this last equation is $ss_K(F_l^S,O\,|\,N_2-N_1)$ [see (61)], it follows that (67) can be expressed as in the top part of (63).

The proof that $ss_K(F_l^S,O\,|\,N_2)<ss_K(F_l^S,O\,|\,N_1)$ is so similar to the one just given for $ss_K(F_l^S,O\,|\,N_2)\geq ss_K(F_l^S,O\,|\,N_1)$, that we leave it to the reader. ■

The next corollary to Theorem 3 introduces the consistency threshold into the analysis, and is more useful than Corollary 3-1 because it can be used to establish when a causal combination can be obliterated by a new set of $N_2 - N_1$ cases.

Corollary 3-2. Suppose $ss_K(F_l^S, O \mid N_1) = 0.8 + \Delta(N_1)$, where[39] $0 \le \Delta(N_1) \le 0.2$. Let

$$\frac{\sum_{x=1}^{N_1} \mu_{F_l^S}(x)}{\sum_{x=N_1+1}^{N_2} \mu_{F_l^S}(x)} \equiv \rho \tag{68}$$

Then

$$ss_K(F_l^S, O \mid N_2) < 0.8, \tag{69}$$

if and only if

$$ss_K(F_l^S, O \mid N_2 - N_1) < 0.8 - \Delta(N_1)\rho \tag{70}$$

which means that F_l^S *is obliterated*. Another way to express (70) is:

$$\rho < \frac{0.8 - ss_K(F_l^S, O \mid N_2 - N_1)}{ss_K(F_l^S, O \mid N_1) - 0.8} \tag{71}$$

Observe that (70) or (71) provide constructive tests to establish if a rule that has survived the consistency threshold based on N_1 cases will be obliterated by the additional $N_2 - N_1$ cases. This requires computing both sides of (70) [or (71)] to see if it is (or is not) satisfied. If it is satisfied, then (69) will be true and the l^{th} rule will be obliterated. If it is not satisfied, then the l^{th} rule will not be obliterated.

ρ is a very interesting parameter; it is the ratio of the sums of firing levels, one for the original N_1 cases and the other for the additional $N_2 - N_1$ cases. Example 7 below examines ρ in order to learn something about it for rules that are obliterated.

Proof: *(a) Sufficiency of (70)*: To begin, we express $ss_K(F_l^S, O \mid N_2)$ in (59) in terms of ρ, i.e.:

$$\begin{aligned} ss_K(F_l^S, O \mid N_2) &= \frac{\rho}{\rho+1} \times ss_K(F_l^S, O \mid N_1) + \frac{\sum_{x=N_1+1}^{N_2} \min(\mu_{F_l^S}(x), \mu_O(x))}{\sum_{x=N_1+1}^{N_2} \mu_{F_l^S}(x)} \times \frac{\sum_{x=N_1+1}^{N_2} \mu_{F_l^S}(x)}{\sum_{x=1}^{N_2} \mu_{F_l^S}(x)} \\ &= \frac{\rho}{\rho+1} \times ss_K(F_l^S, O \mid N_1) + ss_K(F_l^S, O \mid N_2 - N_1) \times \frac{\sum_{x=N_1+1}^{N_2} \mu_{F_l^S}(x)}{\sum_{x=1}^{N_2} \mu_{F_l^S}(x)} \end{aligned} \tag{72}$$

[39] The maximum value of consistency is 1.

Beginning with the identity

$$\sum_{x=1}^{N_2} \mu_{F_l^S}(x) = \sum_{x=1}^{N_1} \mu_{F_l^S}(x) + \sum_{x=N_1+1}^{N_2} \mu_{F_l^S}(x) \tag{73}$$

it is straightforward to show that

$$\frac{\sum_{x=1}^{N_2} \mu_{F_l^S}(x)}{\sum_{x=N_1+1}^{N_2} \mu_{F_l^S}(x)} = \frac{\sum_{x=1}^{N_1} \mu_{F_l^S}(x)}{\sum_{x=N_1+1}^{N_2} \mu_{F_l^S}(x)} + 1 = \rho + 1 \tag{74}$$

so that

$$\frac{\sum_{x=N_1+1}^{N_2} \mu_{F_l^S}(x)}{\sum_{x=1}^{N_2} \mu_{F_l^S}(x)} = \frac{1}{\rho+1} \tag{75}$$

Substituting (75) into (72), it follows that:

$$ss_K(F_l^S, O \mid N_2) = \frac{\rho}{\rho+1} \times ss_K(F_l^S, O \mid N_1) + \frac{1}{\rho+1} \times ss_K(F_l^S, O \mid N_2 - N_1) \tag{76}$$

If (70) is true, and making use of the fact that $ss_K(F_l^S, O \mid N_1) = 0.8 + \Delta(N_1)$, then (76) becomes:

$$ss_K(F_l^S, O \mid N_2) < \frac{1}{\rho+1}\left[\rho \times (0.8 + \Delta(N_1)) + (0.8 - \Delta(N_1)\rho)\right] = \frac{1}{\rho+1}\left[0.8(\rho+1)\right] = 0.8 \tag{77}$$

which completes the proof of the sufficiency of (70).

(b) Necessity of (70): Substituting $ss_K(F_l^S, O \mid N_1) = 0.8 + \Delta(N_1)$ into the right-hand side of (59), and assuming that $ss_K(F_l^S, O \mid N_2) < 0.8$, it follows that:

$$\frac{1}{1 + \frac{\sum_{x=N_1+1}^{N_2} \mu_{F_l^S}(x)}{\sum_{x=1}^{N_1} \mu_{F_l^S}(x)}} \times (0.8 + \Delta(N_1)) + \frac{\sum_{x=N_1+1}^{N_2} \min(\mu_{F_l^S}(x), \mu_O(x))}{\sum_{x=1}^{N_2} \mu_{F_l^S}(x)} < 0.8 \tag{78}$$

Performing straightforward algebra on (78), it follows that:

$$\frac{\sum_{x=1}^{N_1}\mu_{F_l^S}(x)}{\sum_{x=1}^{N_2}\mu_{F_l^S}(x)}\times(0.8+\Delta(N_1))+\frac{\sum_{x=N_1+1}^{N_2}\min(\mu_{F_l^S}(x),\mu_O(x))}{\sum_{x=1}^{N_2}\mu_{F_l^S}(x)}<0.8$$

$$(0.8+\Delta(N_1))\sum_{x=1}^{N_1}\mu_{F_l^S}(x)+\sum_{x=N_1+1}^{N_2}\min(\mu_{F_l^S}(x),\mu_O(x))<0.8\sum_{x=1}^{N_2}\mu_{F_l^S}(x)$$

$$\Delta(N_1)\sum_{x=1}^{N_1}\mu_{F_l^S}(x)+\sum_{x=N_1+1}^{N_2}\min(\mu_{F_l^S}(x),\mu_O(x))<0.8\sum_{x=N_1+1}^{N_2}\mu_{F_l^S}(x)$$

$$\Delta(N_1)\frac{\sum_{x=1}^{N_1}\mu_{F_l^S}(x)}{\sum_{x=N_1+1}^{N_2}\mu_{F_l^S}(x)}+\frac{\sum_{x=N_1+1}^{N_2}\min(\mu_{F_l^S}(x),\mu_O(x))}{\sum_{x=N_1+1}^{N_2}\mu_{F_l^S}(x)}<0.8$$

$$ss_K(F_l^S,O\,|\,N_2-N_1)<0.8-\Delta(N_1)\frac{\sum_{x=1}^{N_1}\mu_{F_l^S}(x)}{\sum_{x=N_1+1}^{N_2}\mu_{F_l^S}(x)}=0.8-\Delta(N_1)\rho \tag{79}$$

which is (70). This completes the proof of the necessity of (70).

It is straightforward to obtain (71) from (70), using the additional fact that

$$\Delta(N_1)=ss_K(F_l^S,O\,|\,N_1)-0.8\geq 0 \;\blacksquare \tag{80}$$

Example 7. In this example, we return to Examples 1 and 2 (Section III) in order to explain (predict) why the four rules in Example 2, for *Likely Breakdown of Democracy*, were reduced to two rules in Example 1 for *Breakdown of Democracy*. The importance of this example is it alerts us to choose the desired outcome in fsQCA very carefully.

We begin with Example 2 for *Likely Breakdown of Democracy*, for which $N_1=10$, and treat the eight cases that were eliminated from Example 1 [when the requirement that $MF(O)>0.5$ was imposed in Example 2] as our N_2-N_1 additional cases; hence, $N_2=18$. Results are summarized in Table XII.

The eight cases in this table are the ones that were eliminated from Example 1 and do not appear in Example 2. The memberships in the output of *Breakdown of Democracy* (O) for these eight cases were taken from Table I. The memberships of the firing levels for causal combinations *AbC, ABC, abc* and *abC* were taken from Table II. The top portion of Table XII provides the minima that are needed to compute the consistency $ss_K(F_l^S,O\,|\,N_2-N_1)$ [the left-hand side of (70)] for the eight cases, using (61). $ss_K(F_l^S,O\,|\,N_1)$ was taken from the last column of Table VIII. $\Delta(N_1)$ was computed from $\Delta(N_1)=ss_K(F_l^S,O\,|\,N_1)-0.8$. $\sum_{x=1}^{N_1}\mu_{F_l^S}(x)$, in the numerator fraction on the right-hand side of ρ, was found from Table VI in its last row for *AbC, ABC, abc* and *abC*; and, $\sum_{x=N_1+1}^{N_2}\mu_{F_l^S}(x)$, in the denominator of ρ, is found in the "Sums" row in the top portion of the present table. The right-hand side of (70) could then be computed.

Inequality (70) could then be tested, because both of its sides have been computed. Observe from its row that it is satisfied for *AbC* and *ABC*, but is not satisfied for *abc* and *abC*, which means that *AbC* and *ABC* are obliterated because of these eight cases, whereas *abc* and *abC* are not.

Examining ρ, observe from its row in Table XII that when it is smaller than 1, rules get obliterated, whereas when it is larger than 1 they do not. We do not claim that this is a general rule. It is true for this example, and is worthy of further study. ■

TABLE XII
COMPUTATIONS ASSOCIATED WITH INEQUALITY (70) FOR EXAMPLES 1 AND 2. THE NUMBERS IN THE TOP PORTION OF THE TABLE ARE FIRING LEVELS, AND FOR NOTATIONAL SIMPLICITY, E. G., *A* IS SHORT FOR *MF*(*A*).

Case	*O*	*AbC*	*Min (O, AbC)*	*ABC*	*Min (O, ABC)*	*abc*	*Min (O, abc)*	*abC*	*Min (O, abC)*
2	0.05	0.11	0.05	0.89	0.05	0.01	0.01	0.01	0.01
3	0.11	0.02	0.02	0.58	0.11	0.02	0.02	0.02	0.02
5	0.23	0.58	0.23	0.03	0.03	0.01	0.01	0.42	0.23
6	0.05	0.97	0.05	0.03	0.03	0.01	0.01	0.02	0.02
10	0.08	0.72	0.08	0.05	0.05	0.02	0.02	0.28	0.08
12	0.05	0	0	0.98	0.05	0	0	0	0
17	0.05	0.87	0.05	0.13	0.05	0.01	0.01	0.05	0.05
18	0.05	0.01	0.01	0.98	0.05	0.01	0.01	0.01	0.01
Sums		3.28	0.49	3.67	0.42	0.09	0.09	0.81	0.42
Consistency for 8 cases [lhs of (70)] $ss_K(F_l^S,O\mid N_2-N_1)$		0.49/3.28 = 0.149		0.42/3.67 = 0.114		0.09/0.09 = 1		0.42/0.81 = 0.519	
$ss_K(F_l^S,O\mid N_1)$		1		1		0.985		0.923	
$\Delta(N_1)$		0.20		0.20		0.185		0.123	
$\sum_{x=1}^{N_1}\mu_{F_l^S}(x)$		1.70		1.26		4.54		3.38	
$\sum_{x=N_1+1}^{N_2}\mu_{F_l^S}(x)$		3.28		3.67		0.09		0.81	
ρ		1.70/3.28 = 0.518		1.26/3.67 = 0.343		4.54/0.09 = 50.444		3.38/0.81 = 4.173	
Rhs of (70)		0.8 -0.2(0.518) = 0.696		0.8 -0.2(0.343) = 0.731		0.8 - 0.2(50.444) = -8.532		0.8 -0.2(4.173) = 0.287	
Is Inequality (70) satisfied?		0.149 < 0.696 YES		0.114 < 0.731 YES		1 > -8.532 NO		0.519 > 0.287 NO	
Conclusion		*AbC is obliterated*		*ABC is obliterated*		*abc is retained*		*abC is retained*	

E. Consistency and it Dependence on the Number of Causal Conditions

In Section C we showed that when the number of causal conditions changes (increases or decreases) then there is a simplified way to determine the winning causal combination for the changed number of causal conditions. Unfortunately, as we show next, similar results are not obtained for consistency calculations.

Starting with (55), which we repeat here for the convenience of the reader,

$$\mu_{F_{i*}}(x\mid C_1,C_2,...,C_{k_2})\le\mu_{F_{i*}}(x\mid C_1,C_2,...,C_{k_1}) \quad (81)$$

it follows that:

$$\min\left(\mu_{F_{i*}}(x \mid C_1, C_2, ..., C_{k_2}), \mu_O(x)\right) \le \min\left(\mu_{F_{i*}}(x \mid C_1, C_2, ..., C_{k_1}), \mu_O(x)\right) \tag{82}$$

Consequently,

$$\frac{\sum_{x=1}^{N} \min\left(\mu_{F_{i*}}(x \mid C_1, C_2, ..., C_{k_2}), \mu_O(x)\right)}{\sum_{x=1}^{N} \mu_{F_{i*}}(x \mid C_1, C_2, ..., C_{k_2})} \le \frac{\sum_{x=1}^{N} \min\left(\mu_{F_{i*}}(x \mid C_1, C_2, ..., C_{k_1}), \mu_O(x)\right)}{\sum_{x=1}^{N} \mu_{F_{i*}}(x \mid C_1, C_2, ..., C_{k_2})} \tag{83}$$

which can also be expressed, as:

$$\begin{aligned} \frac{\sum_{x=1}^{N} \min\left(\mu_{F_{i*}}(x \mid C_1, C_2, ..., C_{k_2}), \mu_O(x)\right)}{\sum_{x=1}^{N} \mu_{F_{i*}}(x \mid C_1, C_2, ..., C_{k_2})} \le & \frac{\sum_{x=1}^{N} \min\left(\mu_{F_{i*}}(x \mid C_1, C_2, ..., C_{k_1}), \mu_O(x)\right)}{\sum_{x=1}^{N} \mu_{F_{i*}}(x \mid C_1, C_2, ..., C_{k_1})} \\ & \times \frac{\sum_{x=1}^{N} \mu_{F_{i*}}(x \mid C_1, C_2, ..., C_{k_1})}{\sum_{x=1}^{N} \mu_{F_{i*}}(x \mid C_1, C_2, ..., C_{k_2})} \end{aligned} \tag{84}$$

or,

$$ss_K(F_{i*}, O \mid k_1) \le ss_K(F_{i*}, O \mid k_2) \times \frac{\sum_{x=1}^{N} \mu_{F_{i*}}(x \mid C_1, C_2, ..., C_{k_1})}{\sum_{x=1}^{N} \mu_{F_{i*}}(x \mid C_1, C_2, ..., C_{k_2})} \tag{85}$$

Starting with (81) it is also true that

$$\sum_{x=1}^{N} \mu_{F_{i*}}(x \mid C_1, C_2, ..., C_{k_2}) \le \sum_{x=1}^{N} \mu_{F_{i*}}(x \mid C_1, C_2, ..., C_{k_1}) \tag{86}$$

which can be expressed as:

$$\frac{\sum_{x=1}^{N} \mu_{F_{i*}}(x \mid C_1, C_2, ..., C_{k_1})}{\sum_{x=1}^{N} \mu_{F_{i*}}(x \mid C_1, C_2, ..., C_{k_2})} \ge 1 \tag{87}$$

When (87) is applied to (85) we reach the disappointing conclusion that it is possible for $ss_K(F_{i*}, O \mid k_1)$ to be larger or smaller than $ss_K(F_{i*}, O \mid k_2)$, depending upon how much larger than 1 the ratio $\sum_{x=1}^{N} \mu_{F_{i*}}(x \mid C_1, C_2, ..., C_{k_1}) / \sum_{x=1}^{N} \mu_{F_{i*}}(x \mid C_1, C_2, ..., C_{k_2})$ is.

Perhaps another inequality exists between $ss_K(F_{i*}, O \mid k_1)$ and $ss_K(F_{i*}, O \mid k_2)$ that will resolve this issue, but to-date we have not found it.

VI. Comparisons of Linguistic Summarization Methods That Use Fuzzy Sets

In this section comparisons are provided between fsQCA and two other methods that use fuzzy sets for linguistic summarization, namely Yager, Kacprzyk, Zadrożny, et al.'s summarizers (see Paragraph 2 in Section I for the references) and Wu and Mendel's (2010, 2011) if-then rules. This is needed so that it is clear how fsQCA differs from those methods. Our comparisons are given in Table XIII, which has been organized according to the chronology in which we have presented fsQCA. The comparisons should be self-explanatory, and so no additional comments about them are provided.

Note that fsQCA is not meant to be a replacement for these existing linguistic summarization methods; it is meant to provide a different kind of if-then linguistic summarization, one that establishes the rules from data that are available about cases.

One may also wonder about a comparison between fsQCA and the if-then rules that are obtained from the Wang-Mendel (WM) method [Wang and Mendel (1992)]. Because the WM-rules are *predictive* whereas the fsQCA rules are *descriptive* (see Section I for explanations of what this means), the two are non-competitive and are therefore not compared in this report.

TABLE XIII
COMPARISONS OF THREE LINGUISTIC SUMMARIZATION (LS) METHODS THAT USE FUZZY SETS

Does the LS Method Include	*Linguistic Summarization Method*		
	fsQCA	*Yager, Kacprcyk, Zadrozny, et al.'s Summarizers*	*Wu & Mendel's If-then rules*
Focusing on a specific desired outcome?	Yes (fsQCA must be repeated for each such outcome and usually its complement, one at a time)	Yes and No [one or more desired outcomes (summarizers) are chosen, but each can have more than one linguistic term associated with it; summaries are pre-established for all of these linguistic terms]	Yes and No (a desired outcome is chosen, and is the consequent of the if-then rules; it has m_O linguistic terms associated with it, and rules are pre-established for all of these linguistic terms)
Pre-establishing a library (codebook) of summaries?	No	Yes	Yes
A specific structure for the summaries?	Yes (summaries are multiple-antecedent if-then rules, all for the *same* consequent term; rules connected by OR)	Yes (there are two canonical forms for the summaries[a]; multiple summary connections are not specified explicitly[b])	Yes (summaries are multiple antecedent if-then rules, for *all* of the consequent terms; multiple summary connections are not specified explicitly[b])
Pre-chosen summarizers?	Yes (called *causal conditions*)	Yes (called *summarizers*)	Yes (called *antecedents*)
A summarizer that is described by more than one linguistic term?	Yes (each linguistic term is considered to be a separate causal condition in the same causal combination of the causal conditions)	Yes (usually each of the k summarizers is described by m_i terms; but, each of these terms is not thought of as a separate summarizer in the same summary, i.e. there is only one linguistic term per summarizer for each summary)	Yes (usually each of the k antecedents is described by m_i linguistic terms; but, each of these linguistic terms is not thought of as a separate antecedent in the same rule, i.e. there is only one linguistic term per antecedent for each rule)
The complement of a summarizer?	Always (both the causal condition and its complement are used)	Usually not	Never
The concept of combinations of summarizers?	Yes (each is called a *causal combination* in which the causal conditions are connected by AND; also, called *conjunctural causation*)	Yes (*multiple summarizers* are connected by AND, but they do not constitute a causal combination in the sense of fsQCA)	Yes (*multiple antecedents* connected by AND constitute a causal combination in the sense of fsQCA)
A collection of 2^k summaries that are constructed from k summarizers or their complements?	Yes (to begin, there are 2^k candidate causal combinations—rules)	No (if there are k summarizers and each is described by the same number of m linguistic terms, then there can be m^k summarizations in the library of pre-established summaries)	No (if there are k antecedents and each is described by the same number of m linguistic terms, then there can be m^k rules in the library of pre-established summary rules)

[a] The two forms are: (1) Q (a linguistic quantifier) objects from a given database are/have summarizers 1 through N at truth level T (e.g., Many automobile models have heavy weight and low MPG [T = 0.60]); and, (2) Q (a linguistic quantifier) objects from a given database with a pre-specified linguistically-qualified summarizer are/have additional summarizers 1 through N-1 at truth level T (e.g., Many automobile models with heavy weights have low MPG [T = 0.58]).

[b] "Multiple summary connections are not specified explicitly" means that when more than one summarizer (or if-then rule) is used, it is not specified if the summarizers (or if-then rule) are connected by the words OR, AND, or ELSE.

TABLE XIII (CONTINUED)

Does the LS Method Include	*Linguistic Summarization Method*		
	fsQCA	*Yager, Kacprcyk, Zadrozny, et al.'s Summarizers*	*Wu & Mendel's If-then rules*
Interpreting a candidate summarization as a corner in a vector space?	Yes (with 2^k dimensions)	No	No
Removing a subset of the candidate summarizations based on computing firing levels and a frequency threshold test?	Yes	No	No
Computing subsethood?	Yes (called *consistency*)	Yes (called *truth level;* the formulas for truth level and consistency are the same)	Yes (called *truth level;* the formulas for truth level and consistency are the same)
Discarding additional candidate summarizations based on a subsethood threshold?	Yes (usually the threshold is ≥ 0.80)	No (in addition to truth level, other summarization measures are computed, with the objective usually being to choose a best summarization)	No (in addition to truth level, other summarization measures are computed, with the objective usually being to choose a best summarization)
Further processing?	Yes (*QM algorithm* used to compute prime and minimal prime implicants, after which *Counterfactual Analysis* is performed)	No	No
Accounting for *limited diversity*?	Yes (done during *Counterfactual Analysis*, by means of thought experiments and using the substantive knowledge of an expert)	No	No
Multiple summaries for the same outcome?	Yes (this occurs automatically, and is called *equifinal causation—equifinality*)	Maybe (it depends on the user, and occurs only if the user decides he/she wants more than one summary; the user must choose how many summaries and the structure of the summaries)	Maybe (it depends on the user, and occurs only if the user decides he/she wants more than one summary; the user must choose how many summaries and the structure of the summaries)
A direct connection to *best instances*?	Yes	No	No
Collections of summaries?	Yes (ranging from *most complex*, to *intermediate* to *parsimonious*; intermediate summaries are considered to be the most useful ones)	No (no reason why they could not be obtained, but this would require creating a library of summarizations with numbers of summarizers ranging from 1 to *k*)	No (no reason why they could not be obtained, but this would require creating a library of rules with numbers of antecedents ranging from 1 to *k*)

VII. On the Number of Causal Conditions

In his early works on QCA, Ragin limited QCA to "small *N*" because in many social science studies having access to only a small number of cases is the norm; however, in later works he relaxed this to "moderate *N*," and eventually, as the general applicability of QCA became more evident, he removed all limitations on *N*.

Of course, there must be some connection between the number of cases that are available and the number of causal conditions that can be used in QCA. By analogy to a variable-oriented approach (e.g., a regression model, in which one cannot assume too many independent variables when only a limited amount of data are present), in the QCA approach there must be some connection between the number of causal conditions (i.e., variables) one can safely use relative to the number of available cases; but, what is the connection?

Marx (2005) provides an answer to this question (however, as pointed out below, this answer is not necessarily applicable to fsQCA). His work was motivated by the attack on QCA by Lieberson (2004) who hypothesized: "QCA is unable to distinguish real from random data and generated 'valid' models and explanations based on random data." If this were true, results obtained from QCA would be meaningless.

Marx (2005) "… addresses this issue through a methodological experiment [that] uses randomly created data matrices to show that QCA [csQCA] can make a distinction between real and random data." His focus is on *the number of contradictions that should occur when using QCA on random data*. Contradictions [Marx (2005, p. 4)] "… occur in QCA when an identical configuration of independent variables accounts for both the presence and absence of an outcome. In QCA-terms, a contradiction occurs when, [e.g.] $AbC \Rightarrow D$ and $AbC \Rightarrow d$."

Marx postulates [assumes] (2005, p. 4, and p. 5) " … a QCA analysis on random data should result in many contradictions. … If many contradictions occur, QCA is not able to produce a valid model on random data." He further notes (2005, p. 4): " … the question of under what conditions it is safe to make this assumption has never been addressed." His paper addresses it.

In short, what Marx's postulate means is that when random data matrices are used one should expect to be inundated with many contradictions. His results show that this postulate depends in a *very quantitative way* on the relationship of the number of variables to the number of cases. More specifically, he shows [Marx (2005, p.21)]: " … a QCA-application is restricted by the proportion of variables on cases and by an upper limit of variables which can be used in an analysis [this limit occurs when contradictions do not occur]. If both restrictions are not taken into account QCA cannot make a distinction between random and real data." For 10 cases [Marx (2005, Table 5)], the number of variables [causal conditions] has to be limited to 4; for 15 cases, the number of variables has to be limited to 5; for 25 cases, the number of variables has to be limited to 6; for 30 cases, the number of variables has to be limited to 7; and, for 45 cases[40], the number of variables has to be limited to 8." He also states [Marx (2005, p. 18)]: " … for 50 cases the upper-limit of variables is 8."

Marx summarizes these quantitative results by stating [Marx (2005, p. 2)]: " … a QCA-analysis should be performed with care. … if the research-design takes the limiting conditions into consideration a QCA-analysis can produce valid models which contribute to model/theory development."

[40] He limited his study to 50 cases, because in social science [Marx (2005, p. 5)] "knowing more than 50 cases more or less in depth becomes difficult."

In a private correspondence with Ragin, by the first author, about this (June 6, 2010), he states: “The problem of ‘too few’ cases (which really means cases that are limited in their diversity) is partially ameliorated via ‘easy’ counterfactuals [see Section IV], something that Marx has not addressed. … A typical fsQCA analysis has 4-8 causal conditions, almost regardless of the number of cases. I think the real reason for this has more to do with our ability to decipher with complexity than anything else. …”

VIII. Membership Functions

A. Introduction

Ragin (2000, pp. 165-171) has interesting general discussions about measuring MFs; however, it is in Ragin (2008, pp. 85-97) where one finds details about two methods for "calibrating fuzzy sets." In the *direct method* "… the researcher specifies the values of an interval scale that correspond to the three qualitative breakpoints that structure a fuzzy set: full membership, full non-membership, and the crossover point. These three qualitative breakpoints are then used to transform the original interval scale to fuzzy membership scores." In the *indirect method* the researcher has to provide a "qualitative assessment of the degree to which cases with given scores on an interval scale are members of the target set. The researcher assigns each case into one of six categories and then uses a simple estimation technique to rescale the original measure so that it conforms to these qualitative assessments."

Both of Ragin's methods are limited to fuzzy sets that can be described using either a left-shoulder or a right shoulder MF. While it may be true that many causal conditions and desired outcomes in social science and political science can be described by such fuzzy sets, it is unfortunately not true that such sets are the only ones that are used in engineering and computer science, because in those fields causal conditions and desired outcomes frequently include adjectives, such as *low*, *small*, *moderate amount*, *some*, *high*, *large*, etc. A term that uses the adjective *low* or *high* may indeed be modeled using a shoulder MF; however, terms that use the adjectives *moderate* or *some* are always modeled using a fuzzy set whose MF is an interior (non-shoulder) MF.

Because some terms used in engineering and computer science can be modeled using shoulder MFs, the rest of this section provides details, but only for Ragin's direct method, since it is useful when quantitative data are available for a target set[41], which is the situation that we always have.

B. Details of Ragin's Direct Method

To begin, data must be collected from an expert about three qualitative breakpoints for a particular target set by answering the following questions[42,43]:

Based on your knowledge of the range for a specific causal condition or desired outcome:

(Q1) What is the threshold for full non-membership in this set?

(Q2) What is the crossover point in this set (the value of the interval-scale variable where there is maximum ambiguity as to whether a case is more in or more out of the target set)?

(Q3) What is the threshold for full membership in this set?

For example, for the set of developed countries based on average per capita national income (for which it is known that in US$, average per capita national income ranges from $110 for Burundi to $40,110 for Switzerland [Ragin (2008, Table 5.2)]), Ragin chooses $2,500, $5,000, and $20,000 and as the answers to these three questions.

---

[41] A *target set* is one of Ragin's terms used for either a causal condition or the desired outcome.

[42] Of course, one could ask more than three questions, but to do so would most likely be even more confounding to a subject than asking just these three questions.

[43] These questions are not stated in the order in which they would actually be asked; that order is (3), (1) and (2). Our ordering, which is in an increasing order of the primary variable, is more useful for the rest of the discussions in this section.

As another example, Fiss (2010) has the following method for establishing the three answers for the target set *high-performing firms*, that are based on knowing the return on assets (ROA) for 205 high-technology manufacturing firms located in the United Kingdom: (Q1) ROA ≤7.8%, (about the 50th percentile), (Q3) ROA ≥ 16.3% (the 75th percentile or higher), and (Q2) ROA = 12% (approximately the average of 7.8% and 16.3%). When the target set was *very-high-performing firms*, he used: (Q1) ROA ≤7.8% (about the 50th percentile), (Q3) ROA ≥ 25% ("arguably very-high performance in the eyes of most analysts") and (Q2) ROA =16.3% ("the 75th percentile in full membership of *high-performing firms*).

Let $\mu_A(\xi)$ denote the MF for fuzzy set *A*. The answers to the above three questions provide ξ_1, ξ_2 and ξ_3, respectively for which $\mu_A(\xi_1)=0$, $\mu_A(\xi_2)=0.5$ and $\mu_A(\xi_3)=1$. One approach (arguably the easiest approach) to obtain $\mu_A(\xi)$ from this data, is to use these three MF values and the breakpoints to approximate $\mu_A(\xi)$ by a piecewise linear function with one straight line segment going from $(\xi_1,0)$ to $(\xi_2,0.5)$ and another straight line segment going from $(\xi_2,0.5)$ to $(\xi_3,1)$. $\mu_A(\xi)$ is 0 for all $\xi \le \xi_1$ and $\mu_A(\xi)$ is 1 for all $\xi \ge \xi_3$.

Another approach to approximating $\mu_A(\xi)$, the one that is advocated by Ragin (2008, pp. 87-94), is to smooth out the MF by using a *log-odds transformation*. To be specific, let $y_A(\xi) \equiv \text{odds of membership}$, where[44]

$$y_A(\xi) = \frac{\mu_A(\xi)}{1-\mu_A(\xi)} \tag{88}$$

and $z_A(\xi) \equiv \log \text{ odds of membership}$, where

$$z_A(\xi) = \ln y_A(\xi) \tag{89}$$

Then, from (89) and (88), it follows that:

$$\mu_A(\xi) = \frac{\exp\left(z_A(\xi)\right)}{1+\exp\left(z_A(\xi)\right)} \tag{90}$$

The usefulness of the log odds is that it is completely symmetric about $z_A(\xi)=0$ and it is automatically bounded between 0 and 1. The drawback to the log odds is that it is incapable of producing values of $\mu_A(\xi)$ that are exactly equal to either 1 or 0, due to the logarithmic transformation.

Ragin defines the *threshold of full membership* as $z_A(\xi)=3$, for which $\mu_A(\xi) \simeq 0.95$, and the *threshold of full non-membership* as $z_A(\xi)=-3$, for which $\mu_A(\xi) \simeq 0.05$. A plot of $z_A(\xi)$ versus $\xi - \xi_2$ is depicted in Fig 8. The straight lines are approximations and connect $z_A(\xi)=-3$ to $z_A(\xi)=0$, and $z_A(\xi)=0$ to $z_A(\xi)=3$. Note that $z_A(\xi)=0$ occurs when $y_A(\xi)=1$ which occurs when $\mu_A(\xi)=0.5$.

[44] (88) is the ratio of the membership of being in *A* over the membership of not being in *A*.

Once ξ_1, ξ_2 and ξ_3 are known (by answering the above three questions), then it is easy to determine (approximate) $z_A(\xi)$ for all values of ξ above or below ξ_2. Observe, from Fig. 8, that the slope of $z_A(\xi)$ for all ξ above ξ_2 (to the right of the origin) is $3/(\xi_3-\xi_2)$, whereas the slope of $z_A(\xi)$ for all ξ below ξ_2 (to the left of the origin) is $3/(\xi_2-\xi_1)$. Consequently, for all $\xi' \ni \xi'-\xi_2>0$,

$$z_A(\xi)=\frac{3}{\xi_3-\xi_2}\times(\xi'-\xi_2), \tag{91}$$

whereas for all $\xi'' \ni \xi''-\xi_2<0$,

$$z_A(\xi)=\frac{3}{\xi_2-\xi_1}\times(\xi''-\xi_2). \tag{92}$$

Once $z_A(\xi)$ has been computed by either of these formulas, then $\mu_A(\xi)$ is solved for from (90).

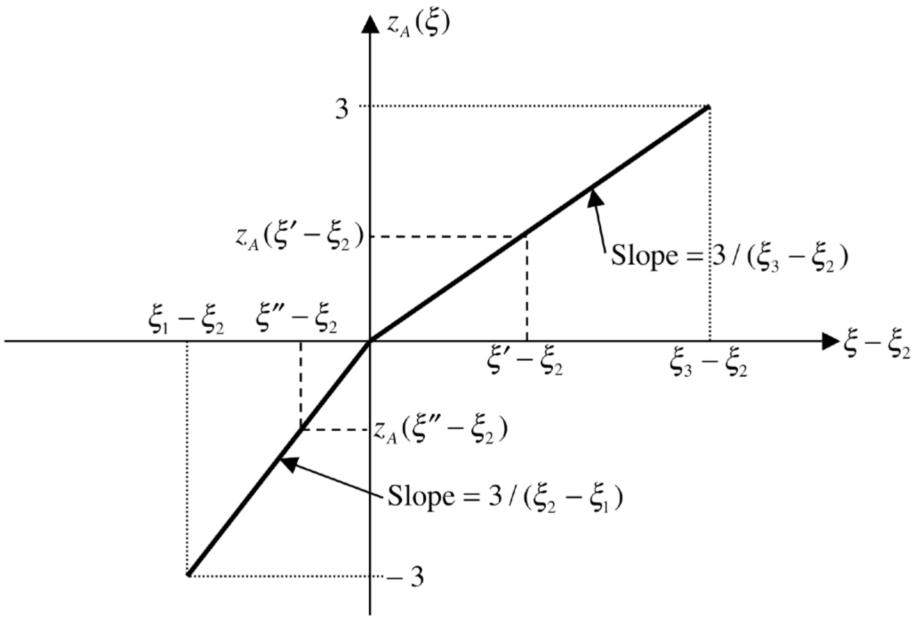

Figure 8. Log-odds, $z_A(\xi)$, shown as a function of $\xi-\xi_2$ rather than as a function of ξ.

It should be obvious that the Achilles' heel in Ragin's direct method is providing crisp answers to the above three questions. An expert is more likely to feel comfortable with providing intervals of numbers as the answers rather than a single number. And, different experts will most likely provide different intervals, reflecting their personal substantive knowledge about the causal condition or desired outcome. This suggests that type-1 fuzzy sets may not be a good model for fsQCA and that interval type-2 fuzzy sets, that can make use of the uncertainties in the answers to the three questions, may be a better model. We plan to study the application of IT2 FSs to fsQCA and will summarize those results in another report.

IX. Conclusions

It is quite common these days for people who work in the general field of computational intelligence (CI), which includes fuzzy sets as one of its major pillars (the others being neural networks and evolutionary computing), to inquire how a CI technique can be used to solve problems in interdisciplinary or non-traditional (i.e., non-engineering or non-computer-science) fields. The expectations are there will be a flow from CI into these fields. Rarely, does the flow occur in the other direction. Charles Ragin's fsQCA is one of those remarkable exceptions and represents a flow from social science and political science into CI.

As is stated in the Introduction, *this report has for the first time explained fsQCA in a very quantitative way, something that is not found in the existing literature, and something that is needed if engineers and computer scientists are to use fsQCA*. This report has also provided new theoretical results for fsQCA, results that could only have been obtained after a formal quantification of fsQCA had occurred.

Stepping back from the details, fsQCA for sufficient conditions—each of which is a linguistic summarization—involves the following steps: (1) Choose a desired outcome; (2) Choose k causal conditions (if a condition is described by more than one term, treat each term as an independent causal condition[45]); (3) Treat the desired outcome and causal conditions as fuzzy sets, and determine MFs for all of them; (4) Evaluate these MFs for all available cases, the results being *derived MFs*; (5) Create 2^k candidate causal combinations (rules) and view each as a possible corner in a 2^k-dimensional vector space; (6) Compute the MF of each of the 2^k candidate causal combinations in all of the available cases, and keep only the ones—the R_S surviving causal combinations (firing-level surviving rules)—whose MF values are > 0.5, i.e., keep the causal combinations that are closer to corners and not the ones that are farther away from corners; (7) Compute the consistencies (subsethoods) of these R_S surviving causal combinations, and keep only those R_A causal combinations—the actual causal conditions (actual rules)—whose consistencies are[46] > 0.80; (8) Binarize the R_A actual causal conditions and use the QM algorithm to obtain prime implicants (the complex solutions) and minimal prime implicants (the parsimonious solutions); (9) Use substantive knowledge from an expert to perform Counterfactual Analysis on the complex solutions, constrained by the parsimonious solutions, to obtain the intermediate solutions; (10) Perform QM on the intermediate solutions to obtain the simplified intermediate solutions; (11) Retain only those simplified intermediate solutions whose consistencies are approximately greater than 0.80, the believable simplified intermediate solutions; (12) Connect each solution (complex, intermediate and parsimonious) with its best instances; and, (13) Compute the coverage of each solution. It is a good practice to perform fsQCA not only for the desired outcome but also for the complement of that outcome, because of the asymmetrical nature of this logic-based approach to linguistic summarization.

fsQCA for necessary conditions can also be performed as explained in Appendix E, but because such conditions are very rare in engineering and computer science applications, they are not emphasized in this report.

---

[45] Although we have not illustrated this in the present report, it has been successfully used by us in applications of fsQCA. These will be reported on in future publications.

[46] Sometimes people use a lower threshold, e.g., 0.75, instead of 0.80.

fsQCA is very different from other linguistic summarization methods, and is compared with two existing LS methods in Section VI that also use fuzzy sets. As mentioned in that section, fsQCA is not meant as a replacement for either of these methods. It is, instead, meant as a different way to obtain a linguistic summarization.

We believe that fsQCA has the potential to be widely applicable in engineering and computer science. In another publication we shall report on its use for an oilfield application.

We conclude this report with a list of some *limitations of fsQCA* that should be viewed in a positive way as a list of *research opportunities*, many of which are already under study.

- The results from fsQCA depend on the frequency cutoff and consistency threshold. If intervals of numbers were used for both of these instead of crisp numbers, then more robust results could be obtained.
- Counterfactual Analysis (CA), while very important, is complicated to understand, and consequently is often ignored. CA is totally dependent on substantive knowledge that is known to an expert. Sometimes expert knowledge is not available and must be extracted from case-data. How to do this and use it in CA is very important. The robustness of fsQCA results to such knowledge also needs to be studied.
- It is very common in engineering and computer science applications of fuzzy logic to use more than two terms for a causal condition (e.g., *low permeability*, *moderate permeability* and *high permeability*). The more terms that are used, the larger k becomes, and subsequently the larger the number of causal combinations, 2^k, becomes, which worsens the problem of limited diversity. How to use fsQCA (or if it can be used) for very large 2^k needs to be studied.
- It would seem that there should be some connection between the numbers of causal conditions one can safely use in QCA relative to the number of available cases. Although some work has been done to study this for crisp set QCA [Marx (2005)] no work has yet been done to study this for fsQCA. Such a study must, of course, also account for the handling of limited diversity by counterfactual analysis.
- Related to using more than two terms for a causal condition, is the problem of how to obtain the MFs for them. MFs depend on the size of the vocabulary used to describe a causal condition. For example, the MF of *low permeability* will be quite different if the vocabulary for permeability includes two terms (*low permeability* and *high permeability*) or three (or more) terms (e.g., *low permeability*, *moderate permeability* and *high permeability*). The impact of this on fsQCA needs to be studied.
- Also related to using more than two terms for a causal condition, is the problem of simplifying the final results from fsQCA, because each causal combination (with up to k terms in it) will be long and therefore so complicated that it will be very difficult to understand. As an example of what we mean by simplifying the final results from fsQCA, suppose the MFs for *low permeability* and *high permeability* have been established for a vocabulary of permeability that includes three terms (*low permeability*, *moderate permeability* and *high permeability*), but only these two terms are used in fsQCA. Then, a summary that contains, e.g. *not low* permeability AND *not high* permeability can be simplified to *moderate* permeability. This kind of simplification can be done by using similarity notions from fuzzy set theory, and needs to be studied, to obtain a library of simplifications.

- Coverage, as defined by Ragin, needs further study because exactly what to do with it is vague.
- There are other linguistic summarization measures that have been defined in the linguistic summarization literature to assist in deciding which summary to use. Perhaps some of them could also be used in fsQCA.
- Validation of fsQCA needs to be examined because engineers and computer scientists are so used to doing this for, e.g. classification problems. Validation experiments will provide a level of confidence to engineers and computer scientists so that they will try fsQCA on other problems. We will report on some validation experiments in a future publication.
- Because *words mean different things to different people* [e.g., Mendel and Wu (2010)] the fuzzy sets that are used in fsQCA should be interval type-2 fuzzy sets rather than type-1 fuzzy sets. This means that fsQCA needs to be re-examined for IT2 FSs.

Finally, the reader needs to be aware that Ragin presented his first ideas about crisp QCA in his 1987 book, but after he realized that things are not black and white in social science applications he extended (abandoned) crisp QCA to fsQCA that has been presented by him in two versions. The first version is in his 2000 book and the second version is in his 2008 book. The major difference between the two versions is in the way he computes consistency. In his 2008 book (p. 48) he refers to the 2000 approach as "simplistic." It does not use the subsethood formula, but instead uses a crisp counting technique. We mention all of this because if a reader only reads the 2000 book he/she will be implementing an out-of-date version of fsQCA.

Acknowledgements

The first author is very pleased to acknowledge the generous help provided to him by Prof. Charles Ragin in explaining the many ramifications of fsQCA. This was done through an extensive e-mail dialog with Prof. Ragin, beginning on October 8, 2009, and continuing even to this date. Both authors would like to acknowledge help provided to them by Prof. Peer Fiss, who met with them a number of times to discuss some ramifications of fsQCA.

APPENDIX A. GLOSSARY

Because many terms in this report may either be new to the reader or are used in a context that may be different from the ones they are used to, they are explained in this Glossary, which is modeled very closely after the Glossary in Rihoux and Ragin (2009). A non-boldfaced italicized item in the description of each term means that such an item also appears in the Glossary. The terms in this Glossary are arranged in alphabetical order.

Case

A ***case*** is an object for which data (measurements) are available, i.e. there are empirical instances (e.g., oil well, ground vehicle, kind of wine, patient, country, village, library, etc.). The cases dealt with are (or should be) well known rather than anonymous.

Causal Combination

It is a combination of the causal conditions or their complements (a *configuration*). If there are k causal conditions, then one begins fsQCA with 2^k candidate causal combinations.

Causal Complexity

The more causal *conditions* are to be considered, the more combinations of them can occur, and therefore the more ***causally complex*** is the situation.

Coherence of Data

Data are ***coherent*** [Rihoux and Ragin (2009), p. 15] when there are no *contradictory configurations*, i.e. there are no *cases* that are identical with respect to causal *conditions*, but different in *outcome*. Checking the coherence of data means detecting contradictory configurations.

Condition

A ***condition*** (also referred to as a <u>condition variable</u> or a <u>causal condition</u>) [Rihoux and Ragin (2009), p. 182] is an explanatory variable that may affect the outcome (it is not an 'independent variable' in the statistical sense).

Configuration

A ***configuration*** [Rihoux and Ragin (2009), p. 182] is a combination of *conditions* relevant to a given *outcome*; it is analogous to a multiple-antecedent rule, and may correspond to one, more than one, or no empirical *case*(s).

Conjunctural Causation

Conjunctural causation refers to causation being due not necessarily to a single causal condition but instead to a group of causal conditions that are combined using conjunction (AND).

Consistency

Consistency [Rihoux and Ragin (2009), p. 182] is the degree to which empirical evidence supports the claim that a set theoretic relation exists. A *subset relation* may signal a *necessary* or a *sufficient condition*, depending on which is the subset, the cause (*sufficiency*) or the *outcome* (*necessity*).

Contradictory Configuration

A ***contradictory configuration*** [Rihoux and Ragin (2009), p. 182] is a *configuration* whose *outcome* value is 1 for some *cases* and 0 for other cases; it, therefore, covers a set of empirical cases, which, although they share the same set of *condition* values (e.g., antecedents), display different outcome values (e.g., consequent).

Counterfactual Analysis

Counterfactual analysis involves evaluating the outcome that a *counterfactual case* would exhibit if, in fact, it existed. It is like a thought experiment. In counterfactual analysis *remainders* are treated as *don't care* combinations; it results in new (often) simpler causal combinations called counterfactuals.

Counterfactual Case (see also, **Logical Remainder, Difficult Counterfactual Case**, and **Easy Counterfactual Case**)

A ***counterfactual case*** is a substantively relevant combination of causal *conditions* that does not exist empirically.

Coverage

Coverage is an assessment of the way respective terms of the minimal formulas (see *prime implicants*) "cover" observed cases. There can be three kinds of coverage: (1) raw coverage which is the proportion of outcome cases that are covered by a given term; (2) unique coverage which is the proportion of outcome cases that are uniquely covered by a given term (no other terms cover those cases); and, (3) solution coverage, which is the proportion of cases that are covered by all the terms.

Difficult Counterfactual Case

Difficult counterfactuals attempt to remove a contributing causal *condition* from a *configuration* displaying the *outcome*.

Diversity

Diversity refers to whether or not a *case* actually exists for a particular combination of causal *conditions*. In social science applications it is very common for no cases to exist for many combinations of causal conditions, and this is referred to as "limited diversity."

Easy Counterfactual Case

Easy counterfactuals assume that adding a redundant causal *condition* to a *configuration* known to produce the *outcome* (e.g., condition *D* to *ABC*, so that the result is *ABCD*) would still produce the outcome.

Equifinality

Equifinality refers to different causal combinations leading to the same outcome.

Holistic perspective

The ***holistic perspective*** [Rihoux and Ragin (2009), p. 6] means that each individual *case* is considered as a complex combination of properties, a specific 'whole' that should not be lost or obscured in the course of the analysis.

Limited Diversity (*see* Diversity)

Logical Remainder
A ***logical remainder*** (also called a counterfactual or a non-observed case) [Rihoux and Ragin (2009), p. 182] is a *configuration* (combination of conditions) that lacks empirical instances.

Multiple-Conjunctural Causation
The phrase ***multiple-conjunctural causation*** [Rihoux and Ragin (2009), p. 8] means that different causal 'paths'—each path being relevant, in a distinct way—may lead to the same *outcome*. The term 'multiple' refers to the number of paths, while the term 'conjunctural' conveys the notion that each path consists of a combination of *conditions*.

Necessary Condition
A *condition* is ***necessary*** for an *outcome* [Rihoux and Ragin (2009), p. 183] if it is always present when the outcome occurs. In short, the outcome is a *subset* of the cause (the same cause may also affect other outcomes).

Net Effects
According to Ragin (2008, pp. 112-114): To estimate the ***net effect*** of a given independent variable, the researcher offsets the impact of competing causal conditions by subtracting from the estimate of the effect of each variable any explained variation in the dependent variable it shares with other causal variables. This is the core meaning of ***net effects***—the calculation of the non-overlapping contribution of each independent variable to explained variation in the dependent variable. Degree of overlap is a direct function of correlation. Generally, the greater the correlation of an independent variable is with its competitors the less is its net effect.

Outcome
An ***outcome*** (also referred to as an outcome variable) [Rihoux and Ragin (2009), p. 183] is the variable to be explained by the *conditions*; usually the outcome is the main focus of a study, and it is analogous to the consequent in a rule.

Prime Implicants
Prime implicants [Rihoux and Ragin (2009), p. 183] are reduced expressions derived in the course of Boolean minimization. Each prime implicant is usually a set of *conditions* joined by the Boolean "AND" operator. A subset of the derived prime implicants constitutes a minimal formula, the endpoint of a Boolean minimization. Each prime implicant in a minimal formula covers a collection of *configurations* from the *truth table* for a given *outcome*.

Qualitative Comparative Analysis
Qualitative Comparative Analysis (QCA) [Rihoux and Ragin (2009), pp. xix and xx] is an umbrella term that captures the three main types (Boolean, multi-value and fuzzy set) as a group. It is now common to refer to the original Boolean version of QCA as csQCA, where "cs" denotes "crisp set," the version that allows multiple-category conditions to be used as mvQCA, where "mv" denotes "multi-value," and to the fuzzy set version of QCA as fsQCA, where "fs" denotes "fuzzy set.

Remainder (*see* **Logical Remainder**)

Subset Relation

With crisp sets, a ***subset relation*** [Rihoux and Ragin (2009), p. 184] exists between two sets whenever all the members of one set are contained within the other set; with fuzzy sets a subset relation exists between two sets whenever membership scores in one set are consistently less than or equal to membership scores in the other set.

Sufficient Condition

A *condition* (or combinations of conditions) is ***sufficient*** for an *outcome* [Rihoux and Ragin (2009), p. 184] if the outcome always occurs when the condition is present (however, the outcome can occur for other reasons as well). In short the cause is a *subset* of the outcome.

Truth Table

A ***truth table*** [Rihoux and Ragin (2009), p. 184] is a synthetic display of all *configurations* (combinations of *conditions*) joined by the Boolean "AND" operator (set intersection).

APPENDIX B. ON WHY RAGIN DEVELOPED QCA

In the 1980's, Ragin became disillusioned with "conventional quantitative methods" (also called "conventional quantitative analysis" or "variable-oriented methods"), i.e. mathematical modeling methods that use correlation, such as regression and multivariate statistical analyses. This led him to invent QCA. Since then, he has expressed some very strong negative opinions about "conventional quantitative analysis" in three books [Ragin (2000), Ragin (2008) and Rihoux and Ragin (2009)]. Although he has done this in order to justify QCA, his negative opinions about "conventional quantitative analysis" are very illuminating. Although we don't necessarily agree with all of his opinions, we feel that he has given them a great deal of thought, and are worth thinking about by all researchers who rely on mathematical models of data.

There are lots of duplications in his statements about conventional quantitative analysis that appear in the three books. The most concise summary of his views, one that contrasts the fuzzy-set approach (which he advocates) and the conventional quantitative analysis approach (which he is very critical of), are found in Ragin (2000, pp. 311-315) and are given next verbatim.

> The many contrasts between the fuzzy-set approach and conventional quantitative methods highlight the assumptions and judgments that are embedded in these methods—at least as they are conventionally used. I summarize these embedded features here.
>
> 1. ***The dependence of conventional methods on fixed preferably "given" populations.*** Before researchers using conventional methods can compute a single statistic, much less a correlation between two variables, they must demarcate and fix boundaries of their sample or population. Once established, such boundaries are rarely questioned or revised. Instead, these boundaries reinforce the embedded assumption of case homogeneity and thus pose a barrier to the recognition of heterogeneity and diversity. The fuzzy set approach problematizes population boundaries and permits great heterogeneity, as manifested, for example, in causal complexity. This approach also allows population boundaries to be fuzzy rather than crisp. That is, cases can vary in their degree of membership in the set of cases relevant to a research question, and this varying degree of domain membership can be made an explicit part of the analysis.
> 2. ***The dependence of conventional methods on an "accounting" approach to difference***. From the viewpoint of conventional methods, cases are similar if they have many aspects that are the same and different if they have many aspects that are different. In additive-linear models[47], similarities and differences are tallied, and cases with mostly similar scores on the independent variables receive mostly the same predicted values on the dependent variable. The embedded assumptions that permit this approach to cases are the following: (1) that all cases included in an analysis are members of the same population (i.e., case homogeneity); (2) that aspects of cases are non-interactive—that is, a case's score on one aspect usually does not modify the meaning or relevance or causal impact of its score on other aspects (i.e., additivity); and, (3) that a case can offset its low score on some variables with a high score on others (i.e., compensation). From a fuzzy-set perspective, however, cases are not always as similar or comparable as they may seem. Sometimes a single difference between two cases can provide a basis for establishing a difference in kind—a qualitative distinction. The fuzzy-set approach implements its concern for potential differences in kind through its configurational approach to cases. Cases are evaluated in terms of their degree of membership in specific property-space locations, which in turn are conceived as ideal-type membership profiles. Two cases are considered similar only if they both have strong membership in the same ideal type property-space location.
> 3. ***The de facto dependence of conventional methods on simplifying assumptions about kinds of cases not found in the data set.*** Limited diversity is the rule, not the exception, in the study of naturally occurring social phenomena. Once researchers identify relevant causal variables, they typically find that many regions of the vector space defined by those variables[48] are void, or virtually void, of cases. When conventional researchers estimate statistical models using data that are limited in their diversity, the

[47] A conventional regression model, without interaction terms, is an example of an *additive-linear model*.
[48] If there are k causal variables (conditions), then there are 2^k corners in the Boolean vector space.

additive-linear techniques they typically apply to their data assume, in essence, that if there were cases in the vacant regions in the vector space, they would conform to the patterns exhibited by cases in the regions that are well populated with cases. Thus, these models incorporate de facto assumptions about kinds of cases that are absent or virtually absent from the researcher's data set. Unfortunately, these assumptions are invisible, not only to most researchers, but also to the audience for social research. In the fuzzy set approach, by contrast, the consideration of limited diversity is an explicit part of the analysis. Not only do researchers identify simplifying assumptions, it is possible, as well, to evaluate the plausibility of these assumptions and then selectively incorporate those that seem plausible. The process of incorporating simplifying assumptions is explicit and visible. Furthermore, the audiences for social research are free to challenge such assumptions and to construct alternate representations of the same evidence.

4. ***The dependence of conventional methods on a correlational understanding of causation, an approach that is insensitive to necessity and sufficiency.*** Bivariate correlation, in one way or another, constitutes the backbone of most conventional techniques of quantitative analysis. Even a complex multi-variable model involving estimates of the net effects of many causal variables on an outcome variable is based, in the end, on a table of bivariate correlations. While very powerful as an analytical tool, the correlational approach to causation is incapable of addressing set-theoretic relationships. Is the outcome a subset of one or more causes (necessity)? Is a cause or causal combination a subset of the outcome (sufficiency)? While it is possible to discern set-theoretic relationships in cross-tabulations of categorical data, few statistical techniques focus on such relationships. The correlational approach to causation equates all prediction errors, whether they constitute violations of necessity or sufficiency, and counts all cases in the ambiguous null-location, where neither cause nor the outcome is present, as correct predictions. More telling is the simple fact that the study of set-theoretic relationships is completely outside the scope of techniques that use interval and ratio scale variables. With these measures, the usual focus is on covariance; set-theoretic relationships, the core of the fuzzy-set approach, are inaccessible. In the correlational view, causes do not delimit possibilities (necessity), nor do they combine in different ways to generate outcomes (sufficiency). Rather, each cause increases or decreases the level or probability of an outcome, net of the effects of other relevant causes.
5. ***The dependence of conventional methods on additive, linear models and their consequent inability to unravel causal complexity.*** If no single cause is either necessary or sufficient for an outcome (i.e., maximum causal complexity), then there is little point in examining bivariate relationships or assessing the 'net effects' of 'independent' variables. To assess causal complexity with conventional multivariate techniques, researchers must use interaction models. Generally, such models work well when: (1) the number of cases is very large, (2) the cases are fully diverse (i.e., limited diversity is not present), and (3) the number of independent variables is relatively small. When these conditions are not met, interaction terms tend to be highly collinear, and many different interaction models fit a given data set equally well. In short, when faced with causal complexity, conventional techniques are also confounded. For this reason, researchers are always warned to start with models that assumes perfect additivity (i.e., extreme causal simplicity) and then to add interaction terms (two-way terms first, then three-way terms if several two-way terms pass, and so on) only if the inclusion of interaction terms is strongly supported by both theory and evidence. In general, this is very good advice for users of conventional methods. However, causation is not always as simple as it seems, and it is likely that assuming simplicity merely relegates causal complexity to the error vector of additive models.
6. ***The reliance of conventional methods on open-ended "variables" containing unspecified and often unknown quantities of irrelevant variation***. The 'variable' is central to social research. The ideal variable is usually portrayed as an interval- or ratio-scale measure, and researchers are usually advised to maximize the variation that each variable exhibits. When variables display a wide range of values (and thus more variation), it is possible to derive better estimates of their effects in multivariate analyses. Generally, the advice to maximize variation is sound. Imagine trying to assess the impact on [political] party identification using incomes values that range from $20,000 to $25,000. While it is generally true that more variation is better then less, it is *not* true that all variation is meaningful. Thus, the assumption that 'variation must be maximized' should not be accepted uncritically. The meaningfulness of variation depends on the research question being asked and the concepts that are relevant to that question. Consider, for example, the researcher interested in studying the impact of having 'too little' income on food consumption. Social scientists know how to conduct a survey asking respondents about their income and how to find out who has more and who has less. That is, they are very good at making relative assessments. But, what's *a lot* of income? How much is *too little*? How many people have *enough*? How many have *far too little*? If the

research question concerns the impact of having too little income, then most of the variation in the upper range of income is simply irrelevant. To answer the questions that address specific ranges of values (e. g., 'too little' or 'enough'), it is necessary to go outside of the data set and to find out about people—how they live and how they spend their money. This knowledge can be used to develop qualitative anchors that distinguish relevant from irrelevant variation, a central concern of the fuzzy-set approach.

7. ***The dependence of conventional methods on mechanistically derived anchors to structure the interpretation of scores.*** Most conventional methods of data analysis focus on co-variation and correlation—an assessment of the degree to which the values of variables go together. If high values on one variable tend to go with high values of another, and low values with low values, then there is a positive correlation. The definition of 'high' versus low,' however, is determined by the means and standard deviation of the relevant-variables, which, in turn, are derived mechanistically from the sample in question. For example, a researcher might compute the mean level of income and its standard deviation for a sample of individuals and use these values to define 'low' and 'high income levels. Note, however, that the resulting evaluation of income is strictly in terms of *relative* levels within that sample. Note also that these statistics equate very different kinds of income gaps. For example, one single standard deviation unit might separate the well off from the rich on the one hand, and destitute from the working poor, on the other. Thus this 'ready-made' measure is only a measure of *relative* income; it is dependent on mechanistically derived anchors, and it treats all variation the same, no matter where in the range of values it occurs. With fuzzy sets, by contrast, 'high' and 'low' are defined in terms of strong and weak membership in sets, using criteria specified by the researcher. These criteria, in turn, are based on the researcher's knowledge. They are not sample specific, nor are they mechanistically derived.

Although Ragin is very critical of conventional quantitative methods, this does not mean he feels they should be replaced by QCA; instead, he believes there is a time and a place to use each methodology, e.g., Rihoux and Ragin (2009, p. 171) state:

> … the intension of QCA techniques is certainly not to supplant regression and related analyses, especially since the underlying logic and goals of the respective methods display stark differences. … one of the key differences is that regression-based methods focus primarily on the problem of estimating the net, independent effect of each variable included in an analysis of the outcome. By contrast, it would be a serious mistake to apply QCA techniques to this task, as the latter focuses on combinations of conditions. From the perspective of QCA, the idea of isolating the net, independent effect of each condition variable makes no sense. Fundamentally, QCA techniques attempt to explain specific outcomes in particular cases (hopefully also producing 'modest' generalizations); statistical analysis, by contrast, tries to generalize about averages across all cases in a population without attention to any specific case. … Probably a useful way to combine QCA techniques and other formal (typically statistical) techniques is to consider them sequentially.

Finally, Ragin (2008, pp. 176-177) states:

> While conventional quantitative methods are clearly rigorous, they are organized around a specific kind of rigor. That is, they have their own rigor and their own discipline, not a *universal* rigor. … They are typically centered on the task of estimating the 'net effects' of 'independent' variables on outcomes. … *net effects thinking* … this feature of conventional methods limits their usefulness. … the argument presented here is not that conventional analysis techniques are flawed—in fact, they are powerful and rigorous. Rather, the argument is that they are not well suited for analyzing causal complexity. Indeed, the assessment of net effects requires that the researcher assume that causation is uncomplicated.

Appendix C. New Intermediate Solution Algorithm

Prepared by Charles Ragin

(Sent to Jerry M. Mendel on Dec, 13, 2009)

The basic sketch of the new intermediate solution algorithm follows.

The procedure evaluates each term in the parsimonious solution against each term in the complex solution. The number of terms in a solution is the number of combinations of causal conditions joined by "+". For example, if the parsimonious solution is *AB* + *CD*, then there are two parsimonious terms. If there are *k* [our R_P] parsimonious terms and *m* [our R_C] complex terms, the procedure cycles $k \times m$ times, once for each possible pairing.

1. Check to see if the complex term is a subset of the parsimonious term. All the elements of the parsimonious term must appear in the complex term. For example, if the parsimonious term is *AB* and the complex term is *ABcD*, then the subset relation is satisfied. However, if the parsimonious term is *AB* and the complex term is *aBcD*, then the subset relation is *not* satisfied. If the subset relation is not satisfied, then the procedure stops for this parsimonious/complex pairing and proceeds to the next pairing.

Complex terms:

```
1 GLOBAL*ECOVALPOL*CULTZONE*pwelreg+
2 GLOBAL*ECOVALPOL*CULTZONE*DEMLONG+
3 GLOBAL*ECOVALPOL*CULTZONE*HISTDEV+
4 GLOBAL*ecovalpol*pwelreg*DEMLONG*histdev +
5 ECOVALPOL*CULTZONE*pwelreg*DEMLONG*histdev
```

Parsimonious:

```
6 ECOVALPOL+
7 DEMLONG
```

Are the elements in #6 a subset of the elements in #1? Yes, proceed because #6 is a superset of #1.

. . .

Are the elements in #6 a subset of the elements in #4? No, stop.

. . .

2. Check to see if the user-inputted theoretical term [the user's substantive knowledge] is *consistent* with the parsimonious term. If it is not, then the element(s) in the theoretical term that differ from the parsimonious term are "trumped" by the parsimonious term.

The theoretical term, as inputted by the user [for this example], is:

```
CULTZONE (present)
PWELREG (present)
DEMLONG (present)
HISTDEV (present)

(The other two may be present or absent.)
```

In this example, the parsimonious term (#6) does not contradict with the theoretical term. However, we still need to complexify the theoretical term, yielding, for the #6#1 combination:

```
8 ECOVALPOL*CULTZONE*PWELREG*DEMLONG*HISTDEV
(ECOVALPOL added to theoretical case)
```

3. Compare the complex term (#1) and the theoretical term (using the theoretical term that has been altered in step 2). Use the theoretical term to eliminate any extraneous elements from the complex term (pwelreg/PWELREG):

```
1 GLOBAL*ECOVALPOL*CULTZONE*pwelreg+
8 ECOVALPOL*CULTZONE*PWELREG*DEMLONG*HISTDEV
```

Which yields [see our CA rules]:

```
9 GLOBAL*ECOVALPOL*CULTZONE
```

4. Save this intermediate term, and then proceed to the next pairing. Save all intermediate terms that are valid.

5. List all valid intermediate terms in the truth table spreadsheet with outcome set to 1. Only these terms should appear in the truth table.

For example, GLOBAL*ECOVALPOL*CULTZONE would appear as

```
GLOBAL   ECOVALPOL   CULTZONE   PWELREG   DEMLONG   HISTDEV
1        1           1          -         -         -
```

(the dash signals “don’t care”/“already eliminated”)

6. Simplify this truth table using Quine, setting all remainders to false. The result of this analysis is the full intermediate solution.

APPENDIX D. BREAKDOWN OF DEMOCRACY EXAMPLES FOR FIVE CAUSAL CONDITIONS

In this appendix, Examples 1 and 2 are redone by increasing their three causal conditions to five: A = *developed* (country), B = *urban* (country), C = *literate* (country), D = *industrial* (country) and E =*stable* (country).

D.1 Example D-1: Five Causal Conditions and O = Breakdown of Democracy

In this example, the desired outcome is O = *Breakdown of Democracy* (of European countries between World Wars 1 and 2). The data in Table D.I are taken from Table 5.2 in Rihoux and Ragin (2009).

Using knowledge and techniques from social science, numerical values were obtained for A, B, C, D, and e for 18 European countries that in Table D.I are called[49] "Cases 1–18." Numerical values were initially obtained by Ragin for o = *Survival of Democracy*, which was assumed to be the complement of *Breakdown of Democracy*; hence, $MF(O)$ was computed from $MF(o)$ as $1-MF(o)$ (see footnote 34). Similarly, numerical values were obtained by him for e = *Unstable Country*, which was assumed to be the complement of *Stable Country*; hence, $MF(E)$ was computed from $MF(e)$ as $1-MF(e)$. S-shaped MFs were obtained for o, A, B, C, D and e using a method that is described in Ragin (2008) and in Section VIII.B, the details of which are not important for this example. Using these MFs, Ragin obtained the MF scores that are also given in Table D.I. These MFs implement Eqs. (1) and (2).

TABLE D. I
DATA- AND FUZZY-MEMBERSHIP-MATRIX (SHOWING ORIGINAL VARIABLES AND THEIR FUZZY-SET MEMBERSHIP FUNCTION SCORES)[a]

Case	*Outcome*		*Condition and MF scores*									
	o	*MF(O)*	*A*	*MF(A)*	*B*	*MF(B)*	*C*	*MF(C)*	*D*	*MF(D)*	*e*	*MF(E)*
1	-9	0.95	720	0.81	33.4	0.12	98	0.99	33.4	0.73	10	0.43
2	10	0.05	1098	0.99	60.5	0.89	94.4	0.98	48.9	1	4	0.98
3	7	0.11	586	0.58	69	0.98	95.9	0.98	37.4	0.90	6	0.91
4	-6	0.88	468	0.16	28.5	0.07	95	0.98	14	0.01	6	0.91
5	4	0.23	590	0.58	22	0.03	99.1	0.99	22	0.08	9	0.58
6	10	0.05	983	0.98	21.2	0.03	96.2	0.99	34.8	0.81	5	0.95
7	-9	0.95	795	0.89	56.5	0.79	98	0.99	40.4	0.96	11	0.31
8	-8	0.94	390	0.04	31.1	0.09	59.2	0.13	28.1	0.36	10	0.43
9	-1	0.58	424	0.07	36.3	0.16	85	0.88	21.6	0.07	13	0.13
10	8	0.08	662	0.72	25	0.05	95	0.98	14.5	0.01	5	0.95
11	-9	0.95	517	0.34	31.4	0.10	72.1	0.41	29.6	0.47	9	0.58
12	10	0.05	1008	0.98	78.8	1	99.9	0.99	39.3	0.94	2	0.99
13	-6	0.88	350	0.02	37	0.17	76.9	0.59	11.2	0	21	0
14	-9	0.95	320	0.01	15.3	0.02	38	0.01	23.1	0.11	19	0.01
15	-4	0.79	331	0.01	21.9	0.03	61.8	0.17	12.2	0	7	0.84
16	-8	0.94	367	0.03	43	0.30	55.6	0.09	25.5	0.21	12	0.20
17	10	0.05	897	0.95	34	0.13	99.9	0.99	32.3	0.67	6	0.91
18	10	0.05	1038	0.98	74	0.99	99.9	0.99	49.9	1	4	0.98

[a] This table is modeled after Table 5.2 in Rihoux and Ragin (2009), and the numbers in it are the same as the ones in that table.

Because with five causal conditions there are $2^5 = 32$ causal combinations, it is not possible to display a table like the one in Table II—it would be too long. Instead, the new min-max formulas

[49] The numbered cases correspond to the following countries: 1-Austria, 2-Belgium, 3-Czechoslovakia, 4-Estonia, 5-Finland, 6-France, 7-Germany, 8-Greece, 9-Hungary, 10-Ireland, 11-Italy, 12-Netherlands, 13-Poland, 14-Portugal, 15-Romania, 16-Spain, 17-Sweden, and 18-United Kingdom.

in (36) and (37) were used to establish the winning causal combination for each case. The following Table D.II (which is a new table) summarizes the min-max computations and is for all 18 cases. It uses MF values for the five causal conditions taken from Table D.I.

TABLE D.II
MIN-MAX CALCULATIONS AND ASSOCIATED CAUSAL COMBINATIONS

Case	*Maximum (MF, complement of MF)/Winner (W)*					*Minimum calculation [Using (36)]*	*Causal combination [Using (37)]*
	Max(A,a)/ W	*Max(B,b)/ W*	*Max(C,c)/ W*	*Max(D,d)/ W*	*Max(E,e)/ W*		
1	0.81/A	0.88/b	0.99/C	0.73/D	0.57/e	0.57	*AbCDe*
2	0.99/A	0.89/B	0.98/C	1/D	0.98/E	0.89	*ABCDE*
3	0.58/A	0.98/B	0.98/C	0.90/D	0.91/E	0.58	*ABCDE*
4	0.84/a	0.93/b	0.98/C	0.99/d	0.91/E	0.84	*abCdE*
5	0.58/A	0.97/b	0.99/C	0.92/d	0.58/E	0.58	*AbCdE*
6	0.98/A	0.97/b	0.99/C	0.81/D	0.95/E	0.81	*AbCDE*
7	0.89/A	0.79/B	0.99/C	0.96/D	0.69/e	0.69	*ABCDe*
8	0.96/a	0.91/b	0.87/c	0.64/d	0.57/e	0.57	*abcde*
9	0.93/a	0.84/b	0.88/C	0.93/d	0.87/e	0.84	*abCde*
10	0.72/A	0.95/b	0.98/C	0.99/d	0.95/E	0.72	*AbCdE*
11	0.66/a	0.90/b	0.59/c	0.53/d	0.58/E	0.53	*abcdE*
12	0.98/A	1/B	0.99/C	0.94/D	0.99/E	0.94	*ABCDE*
13	0.98/a	0.83/b	0.59/C	1/d	1/e	0.59	*abCde*
14	0.99/a	0.98/b	0.99/c	0.80/d	0.90/e	0.89	*abcde*
15	0.99/a	0.97/b	0.83/c	1/d	0.84/E	0.83	*abcdE*
16	0.97/a	0.70/b	0.91/c	0.70/d	0.80/e	0.70	*abcde*
17	0.95/A	0.87/b	0.99/C	0.67/D	0.91/E	0.67	*AbCDE*
18	0.98/A	0.99/B	0.99/C	1/D	0.98/E	0.98	*ABCDE*

The firing-level surviving rules (obtained from the last column in Table D.II) are summarized in Table D.III when all 18 cases are used. Observe that only nine out of the 32 possible causal combinations have survived. Their *Best Instances* were obtained from the last and first columns of Table D.II.

TABLE D.III
DISTRIBUTION OF CASES ACROSS CAUSAL CONDITIONS

Best Instances	*Causal Conditions*					*Corresponding Vector Space Corner*	*Number of cases with > 0.5 membership*
	A	*B*	*C*	*D*	*E*		
1	1	0	1	1	0	*AbCDe*	1
2, 3, 12, 18	1	1	1	1	1	*ABCDE*	4
4	0	0	1	0	1	*abCdE*	1
5, 10	1	0	1	0	1	*AbCdE*	2
6,17	1	0	1	1	1	*AbCDE*	2
7	1	1	1	1	0	*ABCDe*	1
8, 14, 16	0	0	0	0	0	*abcde*	3
9, 13	0	0	1	0	0	*abCde*	2
11, 15	0	0	0	0	1	*abcdE*	2

In this example, we used a frequency threshold of 1; hence, all nine of the surviving causal combinations in Table D.III had subsethoods (consistencies) computed using (11). Note that these calculations used the MFs for all 18 cases. In order to compute these subsethoods, one must first compute the *firing levels* for all 18 cases, but only for the nine surviving causal combinations (rather than for all 32 causal combinations). These firing levels are summarized in Table D.IV. MFs for the five causal conditions are in Table D.I. Results for the subsethoods are summarized in Table D.V, which looks like Table D.III, except that it has one more column called "Set theoretic Consistency."

TABLE D.IV
FIRING LEVELS FOR NINE SURVIVING CAUSAL COMBINATIONS AND 18 CASES. MFS FOR THE FIVE CAUSAL CONDITIONS ARE IN TABLE D.I.

Case	*Memberships of Surviving Causal Combinations (minimum of five causal-condition-MFs): Firing Levels*								
	AbCDe	*ABCDE*	*abCdE*	*AbCdE*	*AbCDE*	*ABCDe*	*abcde*	*abCde*	*abcdE*
1	0.57	0.12	0.19	0.27	0.43	0.12	0.01	0.19	0.01
2	0.02	0.89	0	0	0.11	0.02	0	0	0
3	0.02	0.58	0.02	0.02	0.02	0.09	0.02	0.02	0.02
4	0.01	0.01	0.84	0.16	0.01	0.01	0.02	0.09	0.02
5	0.08	0.03	0.42	0.58	0.08	0.03	0.01	0.42	0.01
6	0.05	0.03	0.02	0.19	0.81	0.03	0.01	0.02	0.01
7	0.21	0.31	0.04	0.04	0.21	0.69	0.01	0.04	0.01
8	0.04	0.04	0.13	0.04	0.04	0.04	0.57	0.13	0.43
9	0.07	0.07	0.13	0.07	0.07	0.07	0.12	0.84	0.12
10	0.01	0.01	0.28	0.72	0.01	0.01	0.02	0.05	0.02
11	0.34	0.10	0.41	0.34	0.34	0.10	0.42	0.41	0.53
12	0.00	0.94	0	0	0	0.01	0	0	0
13	0.00	0	0	0	0	0	0.41	0.59	0
14	0.01	0.01	0.01	0.01	0.01	0.01	0.89	0.01	0.01
15	0.00	0	0.17	0.01	0	0	0.16	0.16	0.83
16	0.03	0.03	0.09	0.03	0.03	0.03	0.70	0.09	0.20
17	0.09	0.13	0.05	0.33	0.67	0.09	0.01	0.05	0.01
18	0.01	0.98	0	0	0.01	0.02	0	0	0
Sum	1.56	4.28	2.80	2.81	2.85	1.36	3.38	3.09	2.23

TABLE D.V
DISTRIBUTION OF CASES ACROSS CAUSAL CONDITIONS AND SET-THEORETIC CONSISTENCY OF CAUSAL COMBINATIONS[a]

Best Instances	*Causal Conditions*					*Corresponding Vector Space Corner*	*Number of cases with > 0.5 membership*	*Set-theoretic Consistency*
	A	*B*	*C*	*D*	*E*			
1	1	0	1	1	0	***AbCDe***	1	**0.974**
2, 3, 12, 18	1	1	1	1	1	*ABCDE*	4	0.250
4	0	0	1	0	1	***abCdE***	1	**0.861**
5, 10	1	0	1	0	1	*AbCdE*	2	0.498
6,17	1	0	1	1	1	*AbCDE*	2	0.495
7	1	1	1	1	0	***ABCDe***	1	**0.971**
8, 14, 16	0	0	0	0	0	***abcde***	3	**1**
9, 13	0	0	1	0	0	***abCde***	2	**0.855**
11, 15	0	0	0	0	1	***abcdE***	2	**0.982**

[a] Bold-faced entries are for consistencies > 0.8.

Examining Table D.V, observe that of the nine causal combinations only six have consistency values ≥ 0.8 (the ones that are in bold face). It follows, therefore, that the $R_A = 6$ actual rules are:

$$AbCDe + abCdE + ABCDe + abcde + abCde + abcdE \rightarrow O \quad \text{(D-1)}$$

It is very interesting to observe that causal condition *ABCDE* that had the most cases supporting it—four— has now vanished from the analysis. Referring to Table D.I, observe that $MF(O)$ for Cases 2, 3, 12 and 18, which are the best instances of *ABCDE*, are 0.05, 0.11, 0.05

and 0.05, respectively, which suggests that the low MF of these cases in *O* may have led to the demise of *ABCDE*.

The prime implicants for (D-1) are easy to obtain by first recognizing that:

$$\begin{aligned} abCdE + abCde &= abCd(e+E) = abCd \\ abcde + abcdE &= abcd(e+E) = abcd \end{aligned} \tag{D-2}$$

The two surviving terms in (D-2) can be combined, i.e.:

$$abCd + abcd = abd(c+C) = abd \tag{D-3}$$

Substituting (D-2) and (D-3) into (D-1), it follows that:

$$\begin{aligned} &AbCDe + abCdE + ABCDe + abcde \\ &\quad + abCde + abcdE \\ &\qquad = (abCdE + abCde) + (abcde + abcdE) \\ &\qquad\quad + AbCDe + ABCDe \\ &\qquad = abd + AbCDe + ABCDe \\ &\qquad = abd + ACDe(b+B) \\ &\qquad = abd + ACDe \end{aligned} \tag{D-4}$$

This is exactly the same solution that Ragin obtained, that is given on the bottom of page 115 in Rihoux and Ragin (2009).

The minimal prime implicants, found from the QM algorithm are $a + e$. These parsimonious solutions also agree with one ones that are given in Rihoux and Ragin (2009, p. 117).

The complex and parsimonious solutions can be expressed linguistically, as:

$$\begin{cases} \text{Complex solution} & \begin{array}{l} \text{IF } (C_1 = a \text{ and } C_2 = b \text{ and } C_3 = d) \\ \text{OR } (C_1 = A \text{ and } C_2 = C \text{ and } C_3 = D \text{ and } C_4 = e), \\ \text{THEN } O \end{array} \\ \text{Parsimonious solution} & \qquad \text{IF } C_1 = a \text{ OR } C_2 = e, \text{THEN } O \end{cases} \tag{D-5}$$

In words, these solutions are:

$$\begin{cases} \text{Complex solution} & \begin{array}{l} (\textit{Not developed} \text{ and } \textit{not urban} \text{ (rural) and } \textit{not industrial}) \\ \text{OR } (\textit{Developed} \text{ and } \textit{literate} \text{ and } \textit{industrial} \text{ and } \textit{not stable}), \\ \text{are sufficient causal combinations for } \textit{Breakdown of Democracy} \end{array} \\ \text{Parsimonious solution} & \begin{array}{c} \textit{Not developed} \text{ OR } \textit{Not stable} \\ \text{are sufficient conditions for } \textit{Breakdown of Democracy} \end{array} \end{cases} \tag{D-6}$$

Next this example is continued by proceeding to its counterfactual analysis. To that end, the rest of this example is structured like the continuation of Example 1 that is given in Section IV.E.

- *Complex solution obtained by hand* [see first line of (D-5)] ($R_C = 2$): *abd+ACDe*
- *Parsimonious solution obtained from QM* [see second line of (D-5)] ($R_P = 2$): *a+e*
- *Substantive knowledge* (these were made up by us, but seemed reasonable): The desired outcome could have occurred if *a* (not developed), *b* (not urban (rural)), *c* (not literate), or *d* (not industrial).
- *Counterfactual Analysis*: There are four cycles in this CA. The details are worked out in Table D.VI using our new CA rules. The simplified intermediate solutions are *abd* + *e*. These solutions agree with one ones that are given in Rihoux and Ragin (2009, p. 117).

TABLE D.VI
THE PROCESS OF COUNTERFACTUAL ANALYSIS (CA) FOR EXAMPLE D.1

Cycle	*Parsimonious Solution*	*Complex Solution*	*Results (parsimonious solution is underlined): Counterfactuals*
1	*a*	$\underline{a}bd$	$\underline{a}bd$ (by CA Rules #2 and 3)
2		*ACDe*	No counterfactual (by CA Rule #1)
3	*e*	*abd*	No counterfactual (by CA Rule #1)
4		$ACD\underline{e}$	$\underline{e}$ (by CA Rules # 3and 4)
Union of Counterfactuals– Intermediate Solutions			*abd* + *e*

Translating these *simplified intermediate solutions* into words, we have the following:

(*Not developed* and *not urban (rural)* and *not industrial*) OR (*unstable*) are sufficient causal combinations for *Breakdown of Democracy*. (D-7)

Results for all cases are summarized in Table D.VII.

- *Consistency*: The consistency of *abd* was computed to be [using (11)] 0.886, and of *e* is 0.902, and, both are greater than 0.80, so both solutions are retained and both are believable simplified intermediate solutions ($R_{BSI} = 2$).
- *Best Instances*: Referring to the three-step procedure that is given in Section II. E: Step 1 leads to the numbers that are in the columns of Table D.VII called “BSIS#1MF” and “BSIS#2MF;” Step 2 has to be performed because $R_{BSI} = 2$, and the results of doing this are in the column of Table D.VII called “Maximum MF of BSISs” (in this example only one believable simplified intermediate solution is retained for each case, except for Case 5, for which there is a tie); and, Step 3 leads to the best instances that are labeled “**Yes**” in the last column of Table D.VII. Observe that the Best Instances for *abd* are Cases 4, 8, and 11, and the Best Instances for *e* are Cases 1, 7, 14 and 16.
- *Coverage*: Using $MF(O)$, $MF(abd)$ and $MF(e)$ that are given in Table D.VII, it is straightforward to compute the raw coverage in (27), as $C_r(abd,O) = 0.678$ and $C_r(e,O) = 0.657$.

Observations: (1) When *D* (or *d*) and *E* (or *e*) are deleted from the causal combinations that survive the frequency threshold, as given in Table D.III [*AbCDe*, *ABCDE*, *abCdE*, *AbCdE*, *AbCDE*, *ABCDe*, *abcde*, *abCde* and *abcdE*], one obtains *AbC*, ABC, *abC* and *abc* (some of these are repeated more than one time) which are the same as the causal combinations that survived the frequency threshold, as given in Table III for three causal conditions. This illustrates Corollary 2-2, that the winning causal combinations for a smaller number of causal conditions are contained within the winning causal combinations for a larger number of causal conditions.

(2) By comparing the firing levels in Table II with the ones in Table D.IV for the causal combinations in Table II that are contained within the causal combinations of Table D.IV (e.g., compare *abc* with *abcde* and *abcdE*), numerical confirmation of Corollary 2-3 is obtained, i.e. it is confirmed that firing levels tend to become weakened when more causal conditions are included.

TABLE D.VII

SUMMARY FOR THE BELIEVABLE SIMPLIFIED INTERMEDIATE SOLUTIONS (BSIS) OF EXAMPLE D.1

Case	*MFs for Outcome and Causal Conditions*					*BSIS #1 MF*	*BSIS #2 MF*	*Maximum MF of BSISs*		*Best Instance?*[a]
	MF(O)	*MF(A)*	*MF(B)*	*MF(D)*	*MF(E)*	*MF(abd)*	*MF(e)*	*MF*	*BSIS*	
1	0.95	0.81	0.12	0.73	0.43	0.09	0.57	0.57	*e*	**Yes**
2	0.05	0.99	0.89	1	0.98	0	0.02	0.02	*e*	No
3	0.11	0.58	0.98	0.90	0.91	0.02	0.09	0.09	*e*	No
4	0.88	0.16	0.07	0.01	0.91	0.84	0.09	0.84	*abd*	**Yes**
5	0.23	0.58	0.03	0.08	0.58	0.42	0.42	0.42	*abd/e*	No
6	0.05	0.98	0.03	0.81	0.95	0.02	0.05	0.05	*e*	No
7	0.95	0.89	0.79	0.96	0.31	0.04	0.69	0.69	*e*	**Yes**
8	0.94	0.04	0.09	0.36	0.43	0.64	0.57	0.64	*abd*	**Yes**
9	0.58	0.07	0.16	0.07	0.13	0.84	0.87	0.87	*e*	No
10	0.08	0.72	0.05	0.01	0.95	0.28	0.05	0.28	*abd*	No
11	0.95	0.34	0.10	0.47	0.58	0.53	0.42	0.53	*abd*	**Yes**
12	0.05	0.98	1	0.94	0.99	0	0.01	0.01	*e*	No
13	0.88	0.02	0.17	0	0	0.83	1	1	*e*	No
14	0.95	0.01	0.02	0.11	0.01	0.89	0.99	0.99	*e*	**Yes**
15	0.79	0.01	0.03	0	0.84	0.97	0.16	0.97	*abd*	No
16	0.94	0.03	0.30	0.21	0.20	0.70	0.80	0.80	*e*	**Yes**
17	0.05	0.95	0.13	0.67	0.91	0.05	0.09	0.09	*e*	No
18	0.05	0.98	0.99	1	0.98	0	0.02	0.02	*e*	No

[a] (*MF*(*IS*), *MF*(*O*)) has to be in the Desirable Region that is depicted in Fig. 5.

D.2 Example D-2: Five Causal Conditions and O = Likely Breakdown of Democracy

Examining $MF(O)$ in Table D. VII (as we did in Example 2 for Table I), observe that there are eight cases for which $MF(O) < 0.5$; so, it seems plausible that these cases do not contribute much useful knowledge about *Breakdown of Democracy*. In order to examine this conjecture, we now focus, as we did in Example 2, on the modified desired outcome of *Likely Breakdown of Democracy*. Our interpretation of *Likely Breakdown of Democracy* is that only those cases for which $MF(O) > 0.5$ should be kept for its fsQCA; hence, Table D.VIII is obtained from Table D.I by retaining only those cases for which $MF(O) > 0.5$. Again, it is very important to understand that we are not equating *Likely Breakdown of Democracy* and *Breakdown of Democracy*; instead, we are treating each as possible desired outcomes.

As in Example 2, which was for three causal conditions, only 10 cases survive, because this test depends only on the MF of the desired outcome.

TABLE D.VIII
DATA- AND FUZZY-MEMBERSHIP-MATRIX (SHOWING ORIGINAL VARIABLES AND THEIR FUZZY-SET MEMBERSHIP FUNCTION SCORES) ONLY FOR THE CASES FOR WHICH $MF(O) > 0.5$

Case	Outcome		Condition and MF scores									
	o	*MF(O)*	*A*	*MF(A)*	*B*	*MF(B)*	*C*	*MF(C)*	*D*	*MF(D)*	*e*	*MF(E)*
1	-9	0.95	720	0.81	33.4	0.12	98	0.99	33.4	0.73	10	0.43
4	-6	0.88	468	0.16	28.5	0.07	95	0.98	14	0.01	6	0.91
7	-9	0.95	795	0.89	56.5	0.79	98	0.99	40.4	0.96	11	0.31
8	-8	0.94	390	0.04	31.1	0.09	59.2	0.13	28.1	0.36	10	0.43
9	-1	0.58	424	0.07	36.3	0.16	85	0.88	21.6	0.07	13	0.13
11	-9	0.95	517	0.34	31.4	0.10	72.1	0.41	29.6	0.47	9	0.58
13	-6	0.88	350	0.02	37	0.17	76.9	0.59	11.2	0	21	0
14	-9	0.95	320	0.01	15.3	0.02	38	0.01	23.1	0.11	19	0.01
15	-4	0.79	331	0.01	21.9	0.03	61.8	0.17	12.2	0	7	0.84
16	-8	0.94	367	0.03	43	0.30	55.6	0.09	25.5	0.21	12	0.2

Table D.IX is obtained from Table D.II by retaining only those cases (listed in the first column of Table D.VIII) for which $MF(O) > 0.5$. The firing-level surviving rules are summarized in Table D.X when only the 10 cases are used. Now only six of the 32 possible causal combinations have survived. Surprisingly, *ABCDE*, which had the most number of cases (4) supporting it when all 18 cases were included (Table D.III), is no longer a surviving causal combination.

TABLE D.IX
MIN-MAX CALCULATIONS AND ASSOCIATED CAUSAL COMBINATIONS ONLY FOR THE CASES FOR WHICH $MF(O) > 0.5$

Case	Maximum (MF, complement of MF) /Winner (W)					Minimum calculation (Using (36))	Causal combination (Using(37))
	Max(A,a)/ W	*Max(B,b)/ W*	*Max(C,c)/ W*	*Max(D,d)/ W*	*Max(E,e)/ W*		
1	0.81/A	0.88/b	0.99/C	0.73/D	0.57/e	0.57	*AbCDe*
4	0.84/a	0.93/b	0.98/C	0.99/d	0.91/E	0.84	*abCdE*
7	0.89/A	0.79/B	0.99/C	0.96/D	0.69/e	0.69	*ABCDe*
8	0.96/a	0.91/b	0.87/c	0.64/d	0.57/e	0.57	*abcde*
9	0.93/a	0.84/b	0.88/C	0.93/d	0.87/e	0.84	*abCde*
11	0.66/a	0.9/b	0.59/c	0.53/d	0.58/E	0.53	*abcdE*
13	0.98/a	0.83/b	0.59/C	1/d	1/e	0.59	*abCde*
14	0.99/a	0.98/b	0.99/c	0.80/d	0.9/e	0.89	*abcde*
15	0.99/a	0.97/b	0.83/c	1/d	0.84/E	0.83	*abcdE*
16	0.97/a	0.70/b	0.91/c	0.70/d	0.80/e	0.7	*abcde*

TABLE D.X
DISTRIBUTION OF CASES ACROSS CAUSAL CONDITIONS ONLY FOR THE CASES FOR WHICH $MF(O) > 0.5$

Best Instances	Causal Conditions					Corresponding Vector Space Corner	Number of cases with > 0.5 membership
	A	*B*	*C*	*D*	*E*		
1	1	0	1	1	0	*AbCDe*	1
4	0	0	1	0	1	*abCdE*	1
7	1	1	1	1	0	*ABCDe*	1
8, 14, 16	0	0	0	0	0	*abcde*	3
9, 13	0	0	1	0	0	*abCde*	2
11, 15	0	0	0	0	1	*abcdE*	2

Subsethoods (consistencies) were computed for the six surviving causal combinations. These calculations used the MFs for only the 10 cases for which $MF(O) > 0.5$. In order to compute these subsethoods, one must first compute the firing levels for all 10 cases but just for the six surviving causal combinations. These firing levels are summarized in Table D.XI, which was obtained directly from Table D.IV by removing three columns (*ABCDE*, *AbCdE*, and *AbCDE*) and eight rows (cases 2, 3, 5, 6, 10, 12, 17, and 18). Results for the subsethoods are summarized in Table D.XII, which looks like Table D.X except that it has one more column called "Set theoretic Consistency."

TABLE D.XI
FIRING LEVELS FOR SIX SURVIVING CAUSAL COMBINATIONS AND 10 CASES FOR WHICH $MF(O) > 0.5$. MFS FOR THE FIVE CAUSAL CONDITIONS ARE IN TABLE D.VIII.

Cases	*Membership of Surviving Causal Combination (minimum of five causal condition MFs): Firing Levels*					
	AbCDe	*abCdE*	*ABCDe*	*abcde*	*abCde*	*abcdE*
1	0.57	0.19	0.12	0.01	0.19	0.01
4	0.01	0.84	0.01	0.02	0.09	0.02
7	0.21	0.04	0.69	0.01	0.04	0.01
8	0.04	0.13	0.04	0.57	0.13	0.43
9	0.07	0.13	0.07	0.12	0.84	0.12
11	0.34	0.41	0.10	0.42	0.41	0.53
13	0	0	0	0.41	0.59	0
14	0.01	0.01	0.01	0.89	0.01	0.01
15	0	0.17	0	0.16	0.16	0.83
16	0.03	0.09	0.03	0.70	0.09	0.20
Sum	1.28	2.10	1.07	3.31	2.55	2.16

TABLE D.XII
DISTRIBUTION OF CASES ACROSS CAUSAL CONDITIONS AND SET-THEORETIC CONSISTENCY OF CAUSAL COMBINATIONS[a]

Best Instances	*Causal Conditions*					*Corresponding Vector Space Corner*	*Number of cases with > 0.5 membership*	*Set-theoretic Consistency*
	A	*B*	*C*	*D*	*E*			
1	1	0	1	1	0	***AbCDe***	1	**1**
4	0	0	1	0	1	***abCdE***	1	**1**
7	1	1	1	1	0	***ABCDe***	1	**1**
8, 14, 16	0	0	0	0	0	***abcde***	3	**1**
9, 13	0	0	1	0	0	***abCde***	2	**0.898**
11, 15	0	0	0	0	1	***abcdE***	2	**0.981**

[a] Bold-faced entries are for consistencies > 0.8.

All six causal combinations have survived the consistency test, i.e., $R_A = 6$. After all of this work, *we have obtained exactly the same results that were obtained by using all 18 cases in Example D.1*, namely:

$$\begin{aligned} &AbCDe + abCdE + ABCDe + abcde + abCde + abcdE \\ &= abd + ACDe \rightarrow O \end{aligned} \tag{D-8}$$

Next, we continue this example by proceeding to counterfactual analysis. To that end, the rest of this example is structured like the continuation of Example 2 that is given in Section IV.E.

- *Complex solution obtained by hand* [see second line of (D-8)] ($R_C = 2$): *abd+ACDe*

- *Parsimonious solution obtained from* QM: There are no prime implicants because $X_2 = \varnothing$ (see Theorem 1); hence, there is no parsimonious solution.
- *Counterfactual Analysis*: No CA is possible because there is no parsimonious solution; hence, the final solution is the same as the complex solution *abd+ACDe*; or, in words:

(*Not developed* and *not urban (rural)* and *not industrial*) OR (*Developed* and *Literate* and*Industrial* and *Not stable*) are sufficient conditions for *Breakdown of Democracy*. (D-9)

Results for all cases are summarized in Table D.XIII.

- *Consistency*: The consistency of *abd* was computed to be [using (11)] 0.932 and of *ACDe* is 1, and, both are greater than 0.80, so both solutions are retained and both are believable simplified intermediate solutions ($R_{BSI} = 2$).
- *Best Instances*: Referring to the three-step procedure that is given in Section II. E: Step 1 leads to the numbers that are in the columns of Table D.XIII called "BSIS#1MF" and "BSIS#2MF;" Step 2 has to be performed because $R_{BSI} = 2$, and the results of doing this are in the column of Table D.XIII called "Maximum MF of BSISs" (in this example only one believable simplified intermediate solution is retained for each case); and, Step 3 leads to the best instances that are labeled "**Yes**" in the last column of Table D.XIII. Observe that the Best Instances for *abd* are Cases 4, 8, 11, 13, 14 and 16, and the Best Instances for *ACDe* are Cases 1 and 7.
- *Coverage*: Using $MF(O)$, $MF(abd)$ and $MF(ACDe)$ that are given in Table D.XIII, it is straightforward to compute the raw coverasge in (27), as The raw coverage is $C_r(abd,O) = 0.684$ and $C_r(ACDe,O) = 0.20$.

TABLE D. XIII
SUMMARY FOR BELIEVABLE SIMPLIFIED INTERMEDIATE SOLUTIONS (BSIS) OF EXAMPLE D.2

Case	*MFs for Outcome and Causal Conditions*						*BSIS#1 MF*	*BSIS#2 MF*	*Maximum MF of BSISs*		*Best Instance?*[a]
	MF(O)	*MF(A)*	*M(B)*	*MF(C)*	*MF(D)*	*MF(E)*	*MF(abd)*	*MF(ACDe)*	*MF*	*BSIS*	
1	0.95	0.81	0.12	0.99	0.73	0.43	0.19	0.57	0.57	*ACDe*	**Yes**
4	0.88	0.16	0.07	0.98	0.01	0.91	0.84	0.01	0.84	*abd*	**Yes**
7	0.95	0.89	0.79	0.99	0.96	0.31	0.04	0.69	0.69	*ACDe*	**Yes**
8	0.94	0.04	0.09	0.13	0.36	0.43	0.64	0.04	0.64	*abd*	**Yes**
9	0.58	0.07	0.16	0.88	0.07	0.13	0.84	0.07	0.84	*abd*	No
11	0.95	0.34	0.10	0.41	0.47	0.58	0.53	0.34	0.53	*abd*	**Yes**
13	0.88	0.02	0.17	0.59	0	0	0.83	0	0.83	*abd*	**Yes**
14	0.95	0.01	0.02	0.01	0.11	0.01	0.89	0.01	0.89	*abd*	**Yes**
15	0.79	0.01	0.03	0.17	0	0.84	0.97	0	0.97	*abd*	No
16	0.94	0.03	0.30	0.09	0.21	0.20	0.70	0.03	0.70	*abd*	**Yes**

[a] (*MF*(*IS*), *MF*(*O*)) has to be in the Desirable Region that is depicted in Fig. 5.

Comparing the consistency numbers for the six causal combinations in Table D.XII with the respective ones in Table D.V, observe that when the number of cases was reduced from 18 in Example D.1 to 10 in Example D.2, this strengthened the consistencies for all six causal combinations. Observe, also, that no new causal combinations have occurred for the present situation of five causal conditions, when we focused on the desired outcome of *Likely Breakdown of Democracy* instead of *Breakdown of Democracy*, whereas one new causal combination appeared in Example 2, for *Likely Breakdown of Democracy*, when there were only three causal conditions. More comparisons are needed to better understand this. It is quite possible that having too many weak cases (the other eight) can indeed cause a sufficient condition for the strong 10 cases to disappear, as happened in Example 1.

In order to try and better understand this, we created Table D.XIV that is comparable to Table XII (for Example 7). For the present case of five causal conditions, the test for the obliteration of a rule is failed by all six of the surviving causal combinations; hence, they are all retained.

TABLE D. XIV
COMPUTATIONS ASSOCIATED WITH INEQUALITY (70) FOR EXAMPLES D.1 AND D.2. THE NUMBERS IN THE TOP PORTION OF THE TABLE ARE FIRING LEVELS, AND FOR NOTATIONAL SIMPLICITY, E. G. A IS SHORT FOR $MF(A)$.

Case	*O*	*AbCDe*	*Min (O, AbCDe)*	*abCdE*	*Min (O, abCdE)*	*ABCDe*	*Min (O, ABCDe)*	*abcde*	*Min (O, abcde)*	*abCde*	*Min (O, abCde)*	*abcdE*	*Min (O, abcdE)*
2	0.05	0.02	0.02	0	0	0.05	0.05	0	0	0	0	0	0
3	0.11	0.02	0.02	0.02	0.02	0.09	0.09	0.02	0.02	0.02	0.02	0.02	0.02
5	0.23	0.08	0.08	0.42	0.23	0.03	0.03	0.01	0.01	0.42	0.23	0.01	0.01
6	0.05	0.05	0.05	0.02	0.02	0.03	0.03	0.01	0.01	0.02	0.02	0.01	0.01
10	0.08	0.01	0.01	0.28	0.08	0.01	001	0.02	0.02	0.05	0.05	0.02	0.02
12	0.05	0	0	0	0	0.01	0.01	0	0	0	0	0	0
17	0.05	0.09	0.05	0.05	0.05	0.09	0.05	0.01	0.01	0.05	0.05	0.01	0.01
18	0.05	0.01	0.01	0	0	0.02	0.02	0	0	0	0	0	0
Sums		0.28	0.24	0.79	0.40	0.33	0.29	0.07	0.07	0.56	0.37	0.07	0.07
Consistency for 8 cases [lhs of (70)] $ss_K(F_l^S,O \mid N_2-N_1)$		0.24/0.28 = 0.857		0.40/0.79 = 0.506		0.29/0.33 = 0.879		0.07/0.07 = 1		0.37/0.56 = 0.661		0.07/0.07 = 1	
$ss_K(F_l^S,O \mid N_1)$		1		1		1		1		0.898		0.981	
$\Delta(N_1)$		0.20		0.20		0.20		0.20		0.098		0.181	
$\sum_{x=1}^{N_1}\mu_{F_l^S}(x)$		1.28		2.01		1.07		3.31		2.55		2.16	
$\sum_{x=N_1+1}^{N_2}\mu_{F_l^S}(x)$		0.28		0.79		0.33		0.07		0.56		0.07	
ρ		4.571		2.544		3.242		47.286		4.554		30.857	
Rhs of (70)		0.8 - 0.2(4.571) = - 0.114		0.8 – 0.2(2.544) = 0.291		0.8 – 0.2(3.242) = 0.152		0.8 – 0.2(47.286) = - 8.657		0.8 – 0.098(4.554) = 0.354		0.8 – 0.181(30.857) = - 4.785	
Is Inequality (70) satisfied?		0.857 < -0.114? NO		0.506 < 0.291? NO		0.879 <0.152? NO		1 < - 8.657? NO		0.661 < 0.354? NO		1 < - 4.785? NO	
Conclusion		*AbCDe is retained*		*abCdE is retained*		*ABCDe is retained*		*abcde is retained*		*abCde is retained*		*abcdE is retained*	

The eight cases in this table are the ones that were eliminated from Example D.1 and do not appear in Example D.2. The memberships in the output of *Breakdown of Democracy* (O) for these eight cases were taken from Table D.I. The memberships of the firing levels for causal combinations *AbCDe, abCdE, ABCDe*, *abcde*, *abCde* and *abcdE* were taken from Table D.IV. The top portion of Table D.XIV provides the minima that are needed to compute the consistency $ss_K(F_l^S,O \mid N_2-N_1)$ [the left-hand side of (70)] for the eight cases, using (61). $ss_K(F_l^S,O \mid N_1)$ was taken from the last column of Table D.XII. $\Delta(N_1)$ was computed from $\Delta(N_1)=ss_K(F_l^S,O \mid N_1)-0.8$. $\sum_{x=1}^{N_1}\mu_{F_l^S}(x)$, in the numerator fraction on the right-hand side of ρ, was found from Table D.XI in its last row, for *AbCDe, abCdE, ABCDe*, *abcde*, *abCde* and *abcdE*; and, $\sum_{x=N_1+1}^{N_2}\mu_{F_l^S}(x)$, in the denominator of ρ, is found in the "Sums" row in the top portion of the present table. The right-hand side of (70) could then be computed. Inequality (70) could then be tested, because both of its sides have been computed. Observe from its row that it is satisfied for all six of the causal combinations, which means that none of them are obliterated.

Examining Table XII, observe that the two causal combinations that are obliterated have values of ρ that are less than one. The two surviving causal combinations in Table XII, as well as all six of the surviving casual combinations in Table D.XIV have values of ρ that are greater than one. These observations suggest that perhaps this could be a simple test for survival of a causal combination, something that needs further study.

For crisp sets, in order to determine if a causal condition is *necessary*, one focuses on the desired outcome and establishes all of the cases that are associated with it. If a causal condition appears for *all of those cases*, then it is a necessary condition.

When querying the necessity of the k causal conditions Ragin (2000, p. 211) states:

> Jointly necessary conditions can be identified one at a time. … If a single condition is necessary, it will be present in all instances of the outcome; if several conditions are necessary, then each condition will be present in all instances of the outcome. After identifying all necessary conditions one at a time, the researcher can view them as jointly necessary conditions and evaluate their plausibility as a combination or conjuncture.

Finally, Ragin (2000, p. 295) states that both C_j and c_j should be tested for necessity.

The details for establishing if C_j or c_j (or neither of them) is a necessary condition are much simpler than the details for establishing the sufficient conditions. This is due in part to the fact that if C_j is a necessary condition then c_j cannot be, and vice-versa.

To begin, one computes $\mu_{C_j}(x)$ and $\mu_{c_j}(x)$ for $x = 1,...,N$ and $j = 1,...,k$. Then, one counts the number of cases for which $\mu_{C_j}(x) \geq 0.5$ and $\mu_{c_j}(x) \geq 0.5$, calling these counts N_{C_j} and N_{c_j}, respectively. Although Ragin (2008) suggests selecting a frequency threshold, f_n, against which N_{C_j} and N_{c_j} are compared in order to determine whether or not C_j or c_j should be treated as *candidate necessary conditions*, no guidelines are given in the literature for how to choose f_n; hence, the following ad hoc procedure was developed (which, in a private e-mail to the first author, has met with Ragin's approval):

$$\begin{cases} \text{IF } N_{C_j} > N_{c_j}\text{, THEN choose } C_j \text{ as a candidate necessary condition} \\ \text{IF } N_{c_j} < N_{C_j}\text{, THEN choose } c_j \text{ as a candidate necessary condition} \\ \text{IF } N_{C_j} = N_{c_j}\text{, choose neither } C_j \text{ or } c_j \text{ as a candidate necessary condition} \end{cases} \quad j = 1,...,k \qquad \text{(E-1)}$$

The result is a set of at most k candidate necessary conditions, C_j^c, where $C_j^c = C_j$ or $C_j^c = c_j$, depending on (E-1). Associated with C_j^c is its derived MF $\mu_{C_j^c}(x)$, where

$$\mu_{C_j^c}(x) = \mu_{C_j^c}(\xi_j(x)) \quad x = 1,2,...,N \qquad \text{(E-2)}$$

The next major calculation involves both the k candidate necessary conditions and the desired outcome. Because a condition is *necessary* for an outcome [Rihoux and Ragin (2009), p. 183] if it is always present when the outcome occurs, necessity requires viewing the outcome as a *subset* of the cause (the same cause may also lead to other outcomes). Consequently, the next calculation is subsethood, but it is a different subsethood than the one in (11), i.e. for necessity, one computes:

$$ss_K(O, C_j^c) = \frac{\sum_{x=1}^{N} \min[\mu_O(x), \mu_{C_j^c}(x)]}{\sum_{x=1}^{N} \mu_O(x)} \quad j = 1,2,...,k \tag{E-3}$$

For O to be considered a subset of C_j^c, Ragin requires that

$$ss_K(O, C_j^c) \geq 0.90 \tag{E-4}$$

The result is a subset of R_N "necessary" causal conditions or their complements, where R_N is frequently zero or is an integer much smaller than *k*. It is also useful to summarize the above calculations in a table, as will be illustrated in Example E-1 below.

Observe that no Quine-McCluskey algorithm is needed for necessity, because each causal condition or its complement stands alone.

Ragin [Rihoux and Ragin (2009)] advocates examining each of the just obtained R_N necessary causal conditions or their complements and accepting (or not accepting) each only if it is not "trivial." An example of a trivial causal condition for an oil-field data set is "The data comes from Section 32 of the oil field." It is trivial because all of the data comes from this section.

Comments: (1) Rules in fuzzy systems are formulated as, e.g. IF $x_1 = F_1$ and ... and $x_p = F_p$, THEN *y* = *G*. Such rules focus only on *sufficiency*. They are "activated" by measured values of x_1, …, x_p using Zadeh's sup-min composition, or by similarity computations between each fuzzified x_i and its corresponding antecedent FS F_i. For fuzzy sytems we usually do not focus on *necessity*. If we did, then such a focus would lead to a rule stated as, e.g., *y* = *G* ONLY IF $x_1 = F_1$. Such a rule requires knowledge of the consequent prior to knowledge about the antecedent, which in engineering terms is associated with an "inverse problem." It is not clear how such a rule would be activated, which may be why they are avoided in a FLS.

(2) This does not mean that necessary conditions should be avoided for linguistic summarization. Perhaps, though, a necessary condition should form a linguistic summarization that is a pre-cursor to the linguistic summarization that is derived from fsQCA for the sufficient conditions, i.e. the necessary conditions should be stated prior to the statement of the sufficient conditions.

(3) Necessary conditions are more rare than are sufficient conditions; however, Ragin [(2000), (2007), (2008)] and Rihoux and Ragin (2009) advise checking for necessary conditions *before* conducting fsQCA for sufficiency. Wagemann and Schneider (2007) feel that a stronger emphasis should be given to necessary conditions.

(4) What does one do with necessary conditions in fsQCA? Rihoux and Ragin (2009, Box 5.5, p. 118) provide the following "good practice" about necessary conditions: "If you explicitly hypothesize necessary conditions, test them before conducting fsQCA for sufficiency; set a high consistency threshold for necessary conditions and eliminate any condition that is found to be necessary from fsQCA for sufficiency (i.e., address such conditions separately, as necessary conditions)." All of the non-trivial necessary conditions are used *as is* in each of the causal combinations that are tested for sufficiency. Suppose that k' of the *k* causal conditions have been found to be necessary. This leaves $k - k'$ causal conditions that have to be tested for sufficiency.

Each causal combination still has k components, but k' of them are fixed; hence, this reduces the number of possible causal combinations from 2^k to $2^{k-k'}$.

(5) In a private e-mail, Ragin states that his position about necessary conditions has changed over the years. One approach is the one that is explained in Item (4). Another approach is to ignore necessity. In a private e-mail to the first author, Ragin states:

> What you have described is in fact what I routinely do. Ignore necessity. … That being said (ignore necessity), there is still a lot of interest in necessary conditions in the social sciences. Is 'prior state breakdown' (e.g., fiscal crises) a necessary condition for 'social revolution'? If it is, then the outcome will be a subset of the cause when we look across historical cases. It takes the form: can you identify a case of social revolution not preceded by state breakdown? … Also, sometimes a student will be baffled because the truth table [fsQCA sufficiency] analysis results are weak. Often, if we go back and look at the data in terms of necessary conditions, we find a lot to talk about. Necessary conditions, however, are almost always one at a time. If you start compounding them via the min, you almost always degrade your consistency score [for the sufficiency analysis].

Example E-1. Breakdown of Democracy for Three or Five Causal Conditions

Table E.I contains all of the computations associated with examining which (if any) of the five causal conditions (or their complements) in Ragin's Breakdown of Democracy example (our Examples 1 and D-1) are necessary conditions.

TABLE E.I
CALCULATIONS FOR NECESSARY CONDITIONS

Case	*MF Scores*										
	MF(O)	*MF(A)*	*MF(a)*	*MF(B)*	*MF(b)*	*MF(C)*	*MF(c)*	*MF(D)*	*MF(d)*	*MF(E)*	*MF(e)*
1	0.95	**0.81**	0.19	0.12	**0.88**	**0.99**	0.01	**0.73**	0.27	0.43	**0.57**
2	0.05	**0.99**	0.01	**0.89**	0.11	**0.98**	0.02	**1**	0	**0.98**	0.02
3	0.11	**0.58**	0.42	**0.98**	0.02	**0.98**	0.02	**0.90**	0.10	**0.91**	0.09
4	0.88	0.16	**0.84**	0.07	**0.93**	**0.98**	0.02	0.01	**0.99**	**0.91**	0.09
5	0.23	**0.58**	0.42	0.03	**0.97**	**0.99**	0.01	0.08	**0.92**	**0.58**	0.42
6	0.05	**0.98**	0.02	0.03	**0.97**	**0.99**	0.01	**0.81**	0.19	**0.95**	0.05
7	0.95	**0.89**	0.11	**0.79**	0.21	**0.99**	0.01	**0.96**	0.04	0.31	**0.69**
8	0.94	0.04	**0.96**	0.09	**0.91**	0.13	**0.87**	0.36	**0.64**	0.43	**0.57**
9	0.58	0.07	**0.93**	0.16	**0.84**	**0.88**	0.12	0.07	**0.93**	0.13	**0.87**
10	0.08	**0.72**	0.28	0.05	**0.95**	**0.98**	0.02	0.01	**0.99**	**0.95**	0.05
11	0.95	0.34	**0.66**	0.10	**0.90**	0.41	**0.59**	0.47	**0.53**	**0.58**	0.42
12	0.05	**0.98**	0.02	**1**	0	**0.99**	0.01	**0.94**	0.06	**0.99**	0.01
13	0.88	0.02	**0.98**	0.17	**0.83**	**0.59**	0.41	0	**1**	0	**1**
14	0.95	0.01	**0.99**	0.02	**0.98**	0.01	**0.99**	0.11	**0.89**	0.01	**0.99**
15	0.79	0.01	**0.99**	0.03	**0.97**	0.17	**0.83**	0	**1**	**0.84**	0.16
16	0.94	0.03	**0.97**	0.30	**0.70**	0.09	**0.91**	0.21	**0.79**	0.20	**0.80**
17	0.05	**0.95**	0.05	0.13	**0.87**	**0.99**	0.01	**0.67**	0.33	**0.91**	0.09
18	0.05	**0.98**	0.02	**0.99**	0.01	**0.99**	0.01	**1**	0	**0.98**	0.02
Counts	NA	10	8	5	13	13	5	8	10	11	7
Candidate Necessary Condition	NA	*A*		*b*		*C*		*d*		*E*	
Subsethood	NA	0.32		0.86		0.57		0.72		0.47	
Necessary Condition	NA	No		No (Close)		No		No		No	

The MF values for the desired outcome and the five causal conditions were taken from Table D.I. The MF values for the complements of each of the causal conditions were computed by using (6). The numbers that are in boldface are greater than 0.5, and let us compute the counts N_{C_j} and N_{c_j}, both of which are given in the row called *Counts*. By using

(E-1), one obtains the *Candidate Necessary Conditions*, namely *A, b, C, d* and *E*. Subsethood $ss_K(O, C_j^c)$ was then computed, with the results given in the row called *Subsethood*. Finally, the test in (E-4) was carried out for each of the five subsethoods, and a decision was made as to whether or not *A, b, C, d* or *E* are necessary conditions for *Breakdown of Democracy*. Observe form the last row of the table that none of these conditions passes the test in (E-4); hence, there are no necessary conditions for *Breakdown of Democracy*.

Observe, also that the subsethood for *b* (*not urban—rural*), which is 0.86, is rather close to the threshold of 0.90. One might, therefore, be tempted to say that *b* is a necessary condition for *Breakdown of Democracy*. It turns out, though, that a country that experienced a *Breakdown of Democracy* was Germany and it was not rural. By adhering to the very high threshold of 0.90, one would not have made the mistake of saying that *b* is a necessary condition for *Breakdown of Democracy*, and also supports having such a high threshold in (E-4).

Final Comment: It is our experience that necessary conditions are very rare. As evidenced above, there is no consensus on what to do with necessary conditions, even if they exist. Hence, we advocate the position similar to the one in Comments (2) and (5) above, i.e. a necessary condition should form a linguistic summarization that is a pre-cursor to the linguistic summarization that is derived from fsQCA for the sufficient conditions (meaning that the necessary conditions should be stated prior to the statement of the sufficient conditions); however, they should be ignored during the computation of sufficient conditions. This is why necessary conditions have not been emphasized in this report.

APPENDIX F. EXAMPLE OF COMPUTING MINIMAL PRIME IMPLICANTS

From Table III of Example 1 (Section III.A), observe that two causal combinations pass both the consistency and frequency thresholds, and must therefore be *present* in our result, namely *abC* and *abc*; and, two causal combinations only pass the frequency threshold and must therefore be *absent* in our result, namely *AbC* and *ABC*. The remaining four causal combinations, *aBc, aBC, Abc,* and *ABc*, which are called *remainders*, must be *absent* to produce the prime implicants and must be treated as *don't care* to produce minimal prime implicants. Using the notation in Fig. 4, this means: $X_1 = \{aBc, aBC, Abc, ABc\}$, $X_2 = \{AbC, ABC\}$ and $X_3 = \{abC, abc\}$. Just above (30a), we showed that the *prime implicant* is $abC + abc = ab(C+c) = ab$. Consequently *complex solution* is equal to *ab*.

To find the *minimal prime implicants*, remainder combinations must be taken into account. Remainders are set to be present if and only if they result in the simplest Boolean expressions. To find out which remainders do this, one can combine all possible combinations of causal conditions and see which of those produce the simplest expression, i.e. one can combine all possible causal combinations in X_1 with the causal combinations in X_3 and use Boolean algebra reduction techniques to get the minimal prime implicants. Details for doing this are given in Table F.I. Observe, from this table, that the *simplest Boolean expression* is *a*; hence, the minimum prime implicant, which is the parsimonious solution, is *a*. This result was obtained (shown in bold face in Table F.I) by setting *aBc* and *aBC* in X_1 to be present and remaining two don't care combinations, *Abc* and *ABc*, to be absent.

Note that, using Boolean reduction techniques one may combine different combinations of the causal conditions in X_1 and X_3, which may lead to different Boolean expressions. However, the simplest Boolean expression does not change.

TABLE F.I
PROCEDURE FOR CALCULATING MINIMAL PRIME IMPLICANTS

Number of causal combinations selected from X_1	$X_1 + X_3$
1	$(abc + abC) + aBc = ab + aBc$
	$(abc + abC) + aBC = ab + aBC$
	$(abc + abC) + Abc = ab + Abc$
	$(abc + abC) + ABc = ab + ABc$
2	$\boldsymbol{(abc + abC) + aBc + aBC = ab + aB = a}$
	$(abc + abC) + aBc + Abc = ab + aBc + Abc$
	$(abc + abC) + aBc + ABc = ab + Bc$
	$(abc + abC) + aBC + Abc = ab + aBC + Abc$
	$(abc + abC) + Abc + ABc = ab + Ac$
	$(abc + abC) + aBC + ABc = ab + aBC + ABc$
3	$(abc + abC) + aBc + aBC + Abc = ab + aB + Abc = a + Abc$
	$(abc + abC) + aBc + Abc + ABc = ab + aBc + Ac$
	$(abc + abC) + aBC + Abc + ABc = ab + aBC + Ac$
	$(abc + abC) + aBc + aBC + ABc = ab + aB + ABc = a + ABc$
4	$(abc + abC) + aBc + aBC + Abc + ABc = ab + aB + Ac = a + Ac$

REFERENCES

P. C. Fiss, "Building better causal theories: a fuzzy set approach to typologies in organization research," *Academy of Management J.*, in press, 2010.

D. J. Hand, H. Mannila, and P. Smyth, *Principles of Data Mining*, Boston, MA: MIT Press, 2001.

J. Kacprzyk and A. Wilbik, "A comprehensive comparison of time series described by linguistic summaries and its application to the comparison performance of a mutual fund and its benchmark," in *Proc. FUZZ-IEEE'10*, Barcelona, Spain, pp. 25-32, 2010.

J. Kacprzyk and R. R. Yager, "Linguistic summaries of data using fuzzy logic," *Int'l. J. of General Systems*, vol. 30, pp. 133–154, 2001.

J. Kacprzyk and S. Zadrożny, "On combining intelligent querying and data mining using fuzzy logic concepts," in G. Bordogna and G. Pasi (Eds.): *Recent Research Issues on the Management of Fuzziness in Databases*, Physica-Verlag, Heidelberg, pp. 67-81, 2000a.

J. Kacprzyk and S. Zadrożny, "Data mining via fuzzy querying over the Internet," in Pons O., Vila M.A. and J. Kacprzyk, (eds.): *Knowledge Management in Fuzzy Databases*, Physica-Verlag, Heidelberg and New York, pp. 211-233, 2000b.

J. Kacprzyk and S. Zadrożny, "On a fuzzy querying and data mining interface," *Kybernetika*, vol. 36, pp. 657-670, 2000c.

J. Kacprzyk and S. Zadrożny, "CWW in intelligent database querying: standalone and Internet-based applications," *Information Sciences*, vol. 34, pp. 71-109, 2001a.

J. Kacprzyk and S. Zadrożny, "Using fuzzy querying over the Internet to browse through information resources," in B. Reusch and K.-H. Temme (Eds.): *Computational Intelligence in Theory and Practice*, Physica-Verlag, Heidelberg and New York, pp. 235-262, 2001b.

J. Kacprzyk and S. Zadrożny, "On linguistic approaches in flexible querying and mining of association rules," in H.L. Larsen, J. Kacprzyk, S. Zadrożny, T. Andreasen and H. Christiansen (Eds.) *Flexible Query Answering Systems: Recent Advances*, Springer-Verlag, Heidelberg and New York, pp. 475- 484, 2001c.

J. Kacprzyk and S. Zadrożny, "Data mining via linguistic summaries of databases: an interactive approach," in L. Ding (Ed.): *A New Paradigm of Knowledge Engineering by Soft Computing*, World Scientific, Singapore, pp. 325-345, 2001d.

J. Kacprzyk and S. Zadrożny, "Fuzzy linguistic summaries via association rules," in A. Kandel, M. Last and H. Bunke (Eds.): *Data Mining and Computational Intelligence*, Physica-Verlag (Springer-Verlag), Heidelberg and New York, pp. 115-139, 2001e.

J. Kacprzyk and S. Zadrożny, "Protoforms of linguistic data summaries: towards more general natural-language-based data mining tools," in A. Abraham, J. Ruiz del Solar, M. Koeppen (Eds.): *Soft Computing Systems*, IOS Press, Amsterdam, pp. 417-425, 2002.

J. Kacprzyk and S. Zadrożny, "Linguistic summarization of data sets using association rules," in *Proc. FUZZ-IEEE'03 (St. Louis, USA)*, pp. 702-707, 2003.

J. Kacprzyk and S. Zadrożny, "Linguistic database summaries and their protoforms: towards natural language based knowledge discovery tools," *Information Sciences*, vol. 173, pp. 281-304, 2005.

J. Kacprzyk and S. Zadrożny, "Computing with words is an implementable paradigm: fuzzy queries, linguistic data summaries and natural language generation," *IEEE Trans. on Fuzzy Systems*, vol. 18, pp.461-472, June 2010.

J. Kacprzyk J., R. R. Yager, and S. Zadrożny, "A fuzzy logic based approach to linguistic summaries of databases," *Int'l. J. of Applied Mathematics and Computer Science*, vol. 10, pp. 813-834, 2000.

C. S. Kim, D. S. Kim and J. S. Park, "A new fuzzy resolution principle based on the antonym," *Fuzzy Sets and Systems*, vol. 113, pp. 299-307, 2000.

B. Kosko, *Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Modeling Machine Intelligence*, Prentice-Hall, Englewood Cliffs, NJ, 1992.

S. Lieberson, "Comments on the use and utility of QCA," in *Qualitative Methods. Newsletter of the American Political Science Association Organized Section on Qualitative Methods*, vol. 2, no. 2, pp. 13-14, 2004.

A. Marx, "Towards more robust model specification in QCA: Results from a methodological experiment," Compass—Working Paper (footnote 1 in this paper indicates that an earlier version of this paper was presented at the 100th ASA meeting in 2005).

E. J. McCluskey, *Introduction to the Theory of Switching Circuits*, McGraw-Hill, New York, 1966.

J. M. Mendel, "Fuzzy logic systems for engineering: a tutorial," *IEEE Proc.*, vol. 83, pp. 345-377, March 1995.

J. M. Mendel, *Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions*, Prentice-Hall, Upper-Saddle River, NJ, 2001.

J. M. Mendel and D. Wu, *Perceptual Computing: Aiding People in Making Subjective Judgments*, John Wiley and IEEE Press, Hoboken, NJ, 2010.

E. Mendelson, *Schaum's Outline of Theory and Problems on Boolean Algebra and Switching Circuits*, McGraw-Hill, New York, 1970.

A. Niewiadomski, “A type-2 fuzzy approach to linguistic summarization of data,” *IEEE Trans. on Fuzzy Systems*, vol. 16, no. 1, pp. 198- 212, 2008a.

A. Niewiadomski, *Method for the Linguistic Summarization of Data: Application of Fuzzy Sets and Their Extensions*, Warszawa, Wydawnicza EXIT, 2008b.

W. V. Quine, “The problem of simplifying truth functions,” *American Mathematical Monthly*, vol. 59, pp. 521-531, 1952.

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

C. C. Ragin, *Fuzzy-set Social Science*, Univ. of Chicago Press, Chicago, IL, 2000.

C. C. Ragin, “Chapter 5: Qualitative comparative Analysis using fuzzy sets (fsQCA),” on-line (but, not exactly the same) version of Chapter 5 that is in Rihoux and Ragin (2009). http://docs.google.com/viewer?a=v&q=cache:pogzM__72B4J:www.u.arizona.edu/~cragin/fsQCA/download/Chapter5_CCA.pdf+Charles+RAgin+Chapter+5&hl=en&gl=us&pid=bl&srcid=ADGEESgG4VG28nJS_fwzPeHE3rgGkhvyEBFALuv5WBTedTZsJEZg2qhatJ_4YZ04FZ26fwZ9upCP2xQgnCZMfZUoPOD24AHQ1E9UbDYjBGdWY8ol3cOoKdrOT0_hEaVs38xvgDxy-D0P&sig=AHIEtbQ_KAJOfNYNHKkvEjve3xKyKnHczg, 2007.

C. C. Ragin, *Redesigning Social Inquiry: Fuzzy Sets and Beyond*, Univ. of Chicago Press, Chicago, IL, 2008.

C. C. Ragin and J. Sonnett, “Between complexity and parsimony: limited diversity, counterfactual cases and comparative analysis,” in *Vergleichen in der Politikwissenschaft*, ed. S. Kropp and M. Minkenberg, Wiesbaden: VS Verlag for Sozialwissencahften.

B. Rihoux and C. C. Ragin (eds.), *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques*, SAGE, Los Angeles, CA 2009.

C. Wagemann and C. Q. Schneider, “Standards of good practice in qualitative comparative analysis (QCA) and fuzzy sets,” COMPASSS Working Paper WP2007-51, posted on November 2007, available on-line at: http://www.compasss.org/pages/resources/wpfull.html

L.-X. Wang and J. M. Mendel, “Generating fuzzy rules by learning from examples,” *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 22, no. 2, pp. 1414–1427, 1992.

D. Wu and J. M. Mendel, “Linguistic summarization using IF-THEN rules,” *Proc. FUZZ-IEEE 2010*, pp. 2166-2173, Barcelona, Spain, July 2010.

D. Wu and J. M. Mendel, “Linguistic summarizations using IF-THEN rules and interval type-2 fuzzy sets,” to appear in *IEEE Trans. on Fuzzy Systems*, 2011.

R. R. Yager, “A new approach to the summarization of data,” *Information Sciences*, vol. 28, pp. 69–86, 1982.

R. R. Yager, “On linguistic summaries of data,” in Piatetsky-Shapiro, G. and Frawley, B. (Eds.): *Knowledge Discovery in Databases*, MIT Press, Cambridge, USA, pp. 347-363, 1991.

R. R. Yager, “Linguistic summaries as a tool for database discovery,” in *Proc. FUZZ-IEEE'95/IFES'95*, Yokohama, Japan, pp. 79–82, 1995.

R. R. Yager, “Database discovery using fuzzy sets,” *In'l. J. of Intelligent Systems*, vol. 11, pp. 691-712, 1996.

L. A. Zadeh, “Outline of a new approach to the analysis of complex systems and decision processes,” *IEEE Trans. on Systems, Man, and Cybernetics*, vol. SMC-3, pp. 28-44, 1973.

L. A. Zadeh, “Fuzzy logic = computing with words,” *IEEE Trans. on Fuzzy Systems*, vol. 4, pp. 103-111, 1996.

L. A. Zadeh, “From computing with numbers to computing with words—from manipulation of measurements to manipulation of perceptions,” *IEEE Trans. on Circuits and Systems–1: Fundamental Theory and Applications*, vol. 45, pp. 105-119, Jan. 1999.

L. A. Zadeh, “A new direction in AI,” *AI Magazine*, pp. 73-84, Spring 2001.

L. A. Zadeh, “Toward a generalized theory of uncertainty (GTU)—an outline,” *Information Sciences*, vol. 172, pp. 1–40, 2005.

L. A. Zadeh, “Toward human level machine intelligence—is it achievable? The need for a paradigm shift,” *IEEE Computational Intelligence Magazine*, pp. 11-22, August 2008.

S. Zadrożny, and Kacprzyk J., “Fuzzy querying using the 'query-by example' option in a Windows-based DBMS,” in *Proc. Third European Congress on Intelligent Techniques and Soft Computing - EUFIT'95* (Aachen, Germany), vol. 2, pp. 733-736, 1995.