Chapter 5: Qualitative Comparative Analysis Using Fuzzy Sets (fsQCA)

Forthcoming in Benoit Rihoux and Charles Ragin (editors), *Configurational Comparative Analysis*, Sage Publications, 2007

Charles C. Ragin
Department of Sociology
University of Arizona
Tucson, AZ 85721 USA
cragin@email.arizona.edu

One apparent limitation of the truth table approach is that it is designed for causal conditions that are simple presence/absence dichotomies (i.e., Boolean or "crisp" sets--chapter 3) or multichotomies (MVQCA--chapter 4). Many of the causal conditions that interest social scientists, however, vary by level or degree. For example, while it is clear that some countries are democracies and some are not, there is a broad range of in-between cases. These countries are not fully in the set of democracies, nor are they fully excluded from this set. Fortunately, there is a well-developed mathematical system for addressing partial membership in sets, fuzzy-set theory (Zadeh 1965). Section 1 of this chapter provides a brief introduction to the fuzzy-set approach, building on Ragin (2000). Fuzzy sets are especially powerful because they allow researchers to calibrate partial membership in sets using values in the interval between 0 (nonmembership) and 1 (full membership) without abandoning core set theoretic principles, for example, the subset relation. As Ragin (2000) demonstrates, the subset relation is central to the analysis of causal complexity.

While fuzzy sets solve the problem of trying to force-fit cases into one of two categories (membership versus nonmembership in a set) or into one of three or more categories (mvQCA), they are not well suited for conventional truth table analysis.[1] With fuzzy sets, there is no simple way to sort cases according to the combinations of causal conditions they display because each case's array of membership scores may be unique. Ragin (2000) circumvents this limitation by developing an algorithm for analyzing configurations of fuzzy-set memberships that bypasses truth table analysis altogether. While this algorithm remains true to fuzzy-set theory through its use of the containment (or inclusion) rule, it forfeits many of the analytic strengths that follow from analyzing evidence in terms of truth tables. For example, truth tables are very useful for investigating "limited diversity" and the consequences of different "simplifying assumptions" that follow from using different subsets of "remainders" to reduce

[1] The typical mvQCA application involves a preponderance of dichotomous causal conditions and one or two trichotomous conditions.

complexity (see Ragin 1987; Ragin and Sonnett 2004). Analyses of this type are difficult when not using truth tables as the starting point.

Section 2 of this chapter builds a bridge between fuzzy sets and truth tables, demonstrating how to construct a conventional Boolean truth table from fuzzy-set data. It is important to point out that this new technique takes full advantage of the gradations in set membership central to the constitution of fuzzy sets and is not predicated upon a dichotomization of fuzzy membership scores. To illustrate these procedures I use the same data set used in the previous chapters; however, I convert the original interval-scale data into fuzzy membership scores (which range from 0 to 1), and thereby avoid dichotomizing or trichotomizing the data (i.e., sorting the cases into crude categories). It is important to point out that the approach sketched in this chapter offers a new way to conduct fuzzy-set analysis of social data. This new analytic strategy is superior in several respects to the one sketched in *Fuzzy-Set Social Science* (Ragin, 2000). While both approaches have strengths and weaknesses, the one presented here uses the truth table as the key analytic device. A further advantage of the fuzzy-set truth-table approach presented in this chapter is that it is more transparent. Thus, the researcher has more direct control over the process of data analysis. This type of control is central to the practice of case-oriented research.

1. Fuzzy sets

In many respects fuzzy sets are simultaneously qualitative and quantitative, for they incorporate both kinds of distinctions in the calibration of degree of set membership. Thus, fuzzy sets have many of the virtues of conventional interval-scale variables, but at the same time they permit set theoretic operations. Such operations are outside the scope of conventional variable-oriented analysis.

1.1 Fuzzy sets defined

QCA was developed originally for the analysis of configurations of crisp set memberships (i.e., conventional Boolean sets). With crisp sets, each case is assigned one of two possible membership scores in each set included in a study: 1 (membership in the set) or 0 (nonmembership in the set). In other words, an object or element (e.g., a country) within a domain (e.g., members of the United Nations) is either in or out of the various sets within this domain (e.g., membership in the U.N. Security Council). Crisp sets establish distinctions among cases that are wholly qualitative in nature (e.g., membership versus nonmembership in the U.N. Security Council).

Fuzzy sets extend crisp sets by permitting membership scores in the interval between 0 and 1. For example, a country (e.g., the U.S.) might receive a membership score of 1 in the set of rich countries but a score of only 0.9 in the set of democratic countries. The basic idea behind fuzzy sets is to permit the scaling of membership scores and thus allow partial or fuzzy membership. A membership score of 1 indicates full membership in a set; scores close to 1 (e.g., 0.8 or 0.9) indicate strong but not quite full membership in a set; scores less than 0.5 but greater than 0 (e.g., 0.2 and 0.3) indicate that objects are more "out" than "in" a set, but still weak members of the set; a score of 0 indicates full nonmembership in the set. Thus, fuzzy sets combine qualitative and quantitative assessment: 1 and 0 are qualitative assignments ("fully in" and "fully out," respectively); values between 0 and 1 indicate partial membership. The 0.5 score is also qualitatively anchored, for it indicates the point of maximum ambiguity (fuzziness) in the assessment of whether a case is more "in" or "out" of a set.

Fuzzy membership scores address the varying degree to which different cases belong to a set (including the two qualitative states, full membership and full nonmembership), not how cases rank relative to each other on a dimension of open-ended variation. Thus, fuzzy sets pinpoint qualitative states while at the same time assessing varying degrees of membership between full inclusion and full exclusion. In this sense, a fuzzy set can be seen as a continuous variable that has been purposefully calibrated to indicate degree of membership in a well defined set. Such calibration is possible only through the use of theoretical and substantive knowledge, which is essential to the specification of the three qualitative breakpoints: full membership (1), full nonmembership (0), and the cross-over point, where there is maximum ambiguity regarding whether a case is more "in" or more "out" of a set (.5).

[Table 5.1 about here]

For illustration of the general idea of fuzzy sets, consider a simple three-value set that allows cases to be in the grey zone between "in" and "out" of a set. As shown in Table 5.1, instead of using only two scores, 0 and 1, this three-value logic adds a third value 0.5 indicating objects that are neither fully in nor fully out of the set in question (compare columns 1 and 2 of Table 5.1). This three-value set is a rudimentary fuzzy set. A more elegant but still simple fuzzy set uses four numerical values, as shown in column 3 of Table 5.1. The four-value scheme uses the numerical values 0, 0.33, 0.67, and 1.0 to indicate "fully out," "more out than in," "more in than out," and "fully in," respectively. The four-value scheme is especially useful in situations where researchers have a substantial amount of information about cases, but the nature of the evidence is not identical across cases. A more fine-grained fuzzy set uses six values, as shown in column 4 of Table 5.1. Like the four-value fuzzy set, the six-value fuzzy set utilizes two qualitative states ("fully out" and "fully in"). The six-value fuzzy set inserts two intermediate levels between "fully out" and the cross-over point ("mostly out" and "more or less out") and two intermediate levels between the cross-over point and "fully in" ("more or less in" and "mostly in").

At first glance, the four-value and six-value fuzzy sets might seem equivalent to ordinal scales. In fact, however, they are qualitatively different from such scales. An ordinal scale is a mere ranking of categories, usually without reference to such criteria as set membership. When constructing ordinal scales, researchers do not peg categories to degree of membership in sets; rather, the categories are simply arrayed relative to each other, yielding a rank order. For example, a researcher might develop a six-level ordinal scheme of country wealth, using categories that range from destitute to super rich. It is unlikely that this scheme would translate automatically to a six-value fuzzy set, with the lowest rank set to 0, the next rank to 0.1, and so on (see column 4 of Table 5.1). Assume the relevant fuzzy set is the set of *rich countries*. The lower two ranks of the ordinal variable might both translate to "fully out" of the set of rich countries (fuzzy score = 0). The next rank up in the ordinal scheme might translate to 0.4 rather than 0.2 in the fuzzy set scheme. The top two ranks might translate to "fully in" (fuzzy score = 1), and so on. In short, the specific translation of ordinal ranks to fuzzy membership scores depends on the fit between the content of the ordinal categories and the researcher's conceptualization of the fuzzy set. This point underscores the fact that researchers must calibrate memberships scores using substantive and theoretical knowledge when developing fuzzy sets. Such calibration should not be mechanical.

Finally, a continuous fuzzy set permits cases to take values anywhere in the interval from 0 to 1, as shown in the last column of Table 5.1. The continuous fuzzy set, like all fuzzy sets, utilizes the two qualitative states (fully out and fully in) and also uses the cross-over point to distinguish between cases that are more out from those that are more in. As an example of a continuous fuzzy set, consider membership in the set of rich countries, based on GNP per capita. The translation of this variable to fuzzy membership scores is neither automatic nor mechanical. It would be a serious mistake, for instance, to score the poorest country 0, the richest country 1, and then to array all the other countries between 0 and 1, depending on their positions in the range of GNP per capita values. Instead, the first task in this translation would be to specify three important qualitative anchors: the point on the GNP per capita distribution at which full membership is reached (i.e., definitely a rich country, membership score = 1), the point at which full nonmembership is reached (i.e., definitely not a rich country, membership score = 0), and the point of maximum ambiguity in whether a country is "more in" or "more out" of the set of rich countries (a membership score of 0.5, the cross-over point). When specifying these qualitative anchors, the investigator should present a rationale for each breakpoint.

Qualitative anchors make it possible to distinguish between relevant and irrelevant variation. Variation in GNP per capita among the unambiguously rich countries is *not* relevant to membership in the set of rich countries, at least from the perspective of fuzzy sets. If a country is unambiguously rich, then it is accorded full membership, a score of 1. Similarly, variation in GNP per capita among the unambiguously not-rich countries is also irrelevant to degree of membership in the set of rich countries. Thus, in research using fuzzy sets it is not enough simply to develop scales that show the relative positions of cases on distributions (e.g., a conventional index of wealth such as GNP per capita). It is also necessary to use qualitative anchors to map the links between specific scores on continuous variables (e.g., an index of wealth) and fuzzy set membership (e.g., degree of membership in the set of rich

countries).
[Table 5.2 about here]

In a fuzzy-set analysis both the outcome and the causal conditions are represented using fuzzy sets.[2] Table 5.2 shows a simple data matrix containing fuzzy membership scores. The data are the same used in the two previous chapters and show causal conditions relevant to the breakdown/survival of democracy in interwar Europe. In this example, the outcome of interest is the degree of membership in the set of countries with democracies that survived the many political upheavals of this period (SURVIVED). Degree of membership in the set of countries experiencing democratic breakdown (BREAKDOWN) is simply the negation of degree of membership in SURVIVED (see discussion of negation below). The causal conditions are degree of membership in the set of developed countries (DEVELOPED), degree of membership in the set of urbanized countries (URBAN), degree of membership in the set of industrialized countries (INDUSTRIAL), degree of membership in the set of literate countries (LITERATE), and degree of membership in the set of countries experiencing political instability during this period (UNSTABLE). The table shows both the original data (interval-scale values or ratings) and the corresponding fuzzy membership scores (denoted with "FZ" suffixes). The fuzzy membership scores were calibrated using a procedure detailed in Ragin (2007).[3] This procedure is based on the researcher's qualitative classification of cases according to the six-value scheme shown in Table 5.1. The original interval-scale data are then rescaled to fit the metric indicated by these qualitative codings.

1.2 Operations on fuzzy sets

Before presenting the bridge between fuzzy sets and truth table analysis, I discuss three common operations on fuzzy sets: negation, logical *and*, and logical *or*. These three operations provide important background knowledge for understanding how to work with fuzzy sets.

Negation. Like conventional crisp sets, fuzzy sets can be negated. With crisp sets, negation switches membership scores from "1" to "0" and from "0" to "1." The negation of the crisp set of democracies that survived, for example, is the crisp set of democracies that collapsed. This simple mathematical principle holds in fuzzy algebra

---

[2] Crisp-set causal conditions can be included along with fuzzy-set causal conditions in a fuzzy-set analysis.

[3] The primary goal of this paper is to illustrate a method for creating crisp truth tables from fuzzy-set data. Accordingly, this presentation does not focus on how these fuzzy sets were calibrated or even on the issue of which causal conditions might provide the best possible specification of the social structural circumstances linked to the survival of democracy in Europe during this period. Instead, the focus is on practical procedures.

as well, but the relevant numerical values are not restricted to the Boolean values 0 and 1, but extend to values between 0 and 1. To calculate the membership of a case in the *negation* of fuzzy set *A* (i.e., *not-A*), simply subtract its membership in set *A* from 1, as follows:

(membership in set *not-A*) = 1 - (membership in set *A*)

or

$\sim A = 1 - A$

(The tilde sign "~" is used to indicate negation.) Thus, for example, Finland has a membership score of .64 in SURVIVED; therefore, its degree of membership in BREAKDOWN is .36. That is, Finland is more out than in the set of democracies that collapsed.

Logical and. Compound sets are formed when two or more sets are combined, an operation commonly known as set intersection. A researcher interested in the fate of democratic institutions in relatively inhospitable settings might want to draw up a list of countries that combine being "democratic" with being "poor." Conventionally, these countries would be identified using crisp sets by crosstabulating the two dichotomies, poor versus not-poor and democratic versus not-democratic, and seeing which countries are in the democratic/poor cell of this 2 X 2 table. This cell, in effect, shows the cases that exist in the intersection of the two crisp sets. With fuzzy sets, logical *and* is accomplished by taking the minimum membership score of each case in the sets that are combined. The minimum membership score, in effect, indicates degree of membership of a case in a combination of sets. Its use follows "weakest link" reasoning. For example, if a country's membership in the set of poor countries is 0.7 and its membership in the set of democratic countries is 0.9, its membership in the set of countries that are both poor and democratic is the smaller of these two scores, 0.7. A score of 0.7 indicates that this case is more in than out of the intersection.

[Table 5.3 about here]

For further illustration of this principle, consider Table 5.3. The last two columns demonstrate the operation of logical *and*. The penultimate column shows the intersection of DEVELOPED and URBAN, yielding membership in the set of countries that combine these two traits. Notice that some countries (e.g., France and Sweden) with high in DEVELOPED but low membership in URBAN have low scores in the intersection of these two traits. The last column shows the intersection of DEVELOPED, URBANIZED, and UNSTABLE. Only one country in interwar Europe had a high score in this combination, Germany. In general, as more sets are added to a combination of conditions, membership scores either stay the same or decrease. For each intersection, the *lowest* membership score provides the degree of membership in the combination.

Logical or. Two or more sets also can be joined through logical *or*--the union of sets. For example, a researcher might be interested in countries that are "developed" *or* "democratic" based on the conjecture that these two conditions might offer equivalent bases for some outcome (e.g., bureaucracy-laden government). When using fuzzy sets, logical *or* directs the researcher's attention to the *maximum* of each case's memberships in the component sets. That is, a case's membership in the set formed from the *union* of two or more fuzzy sets is the maximum value of its memberships in the component sets. Thus, if a country has a score of 0.3 in the set of democratic countries and a score of 0.9 in the set of developed countries, it has a score of 0.9 in the set of countries that are "democratic *or* developed."

[Table 5.4 about here]

For illustration of the use of logical *or*, consider Table 5.4. The last two columns of Table 5.4 show the operation of logical *or*. The penultimate column shows countries that are DEVELOPED or URBAN. Notice that the only countries that have low membership in this union of sets are those that have low scores in both component sets (e.g., Estonia, Greece, Portugal, and Romania). The last column shows degree of membership in the union of three sets, DEVELOPED, URBAN, or UNSTABLE. Only Estonia and Romania have low scores in this union.

1.3 Fuzzy subsets

The key set theoretic relation in the study of causal complexity is the subset relation. As discussed in Ragin (2000), if cases sharing several causally relevant conditions uniformly exhibit the same outcome, then these cases constitute a subset of instances of the outcome. The subset relation just described signals that a specific combination of causally relevant conditions may be interpreted as *sufficient* for the outcome. If there are other sets of cases sharing other causally relevant conditions and these cases also agree in displaying the outcome in question, then these combinations of conditions also may be interpreted as sufficient for the outcome. The interpretation of *sufficiency*, of course, must be grounded in the researcher's substantive and theoretical knowledge; it does not follow automatically from the demonstration of the subset relation. Regardless of whether the concept of sufficiency is invoked, the subset relation is the key device for pinpointing the different combinations of conditions linked in some way to an outcome (e.g., the combinations of conditions linked to democratic survival or breakdown in interwar Europe).

With crisp sets it is a simple matter to determine whether the cases sharing a specific combination of conditions constitute a subset of the outcome. The researcher simply examines cases sharing each combination of conditions and assesses whether or not they agree in displaying the outcome. In crisp-set analyses, researchers use truth tables to sort cases according to the causal conditions they share, and the investigator assesses whether or not the cases in each row of the truth table agree on the outcome. The assessment specific to each row can be conceived as a 2x2 crosstabulation of the presence/absence of the outcome against the presence/absence of the combination of causal conditions specified in the row. The subset relation is indicated when the cell

corresponding to the presence of the causal combination and the absence of the outcome is empty, and the cell corresponding to the presence of the causal combination and the presence of the outcome is populated with cases, as shown in Table 5.5.
[Table 5.5 about here]

Obviously, these procedures cannot be duplicated with fuzzy sets. There is no simple way to isolate the cases sharing a specific combination of causal conditions because each case's array of membership scores may be unique. Cases also have different degrees of membership in the outcome, complicating the assessment of whether they "agree" on the outcome. Finally, with fuzzy sets cases can have partial membership in every logically possible combination of causal conditions, as illustrated in Table 5.6. This table shows the membership of countries in three of the causal conditions used in this example (DEVELOPED, URBAN, and LITERATE) and in the eight causal combinations that can be generated using these three fuzzy sets. These eight causal combinations also can be seen as eight logically possible causal arguments. As explained in *Fuzzy-Set Social Science*, fuzzy sets representing causal conditions can be understood as a multidimensional vector space with 2^k corners, where k is the number of causal conditions. The number of corners in this vector space is the same as the number of rows in a crisp truth table with k causal conditions. Empirical cases can be plotted within this multi-dimensional space, and the membership of each case in each of the eight corners can be calculated using fuzzy algebra, as shown in Table 5.6. For example, the membership of Austria in the corner of the vector space corresponding to developed, urban, and literate (D*U*L, the last column of Table 5.6) is the minimum of its memberships in developed (0.74), urban (.14) and literate (.98), which is .14. Austria's membership in the not-developed, not-urban, and not-literate (~D*~U*~L) corner is the minimum of its membership in not-industrial (1 - 0.74 = 0.26), not-urban (1 - 0.14 = 0.86), and not-literate (1 - 0.98 = 0.02), which is 0.02. The link between fuzzy-set vector spaces and crisp truth tables is explored in greater depth below.
[Table 5.6 about here]

While these properties of fuzzy sets make it difficult to duplicate crisp-set procedures for assessing subset relationships, the fuzzy subset relation can be assessed using fuzzy algebra. With fuzzy sets a subset relation is indicated when membership scores in one set (e.g., a causal condition or combination of causal conditions) are consistently less than or equal to membership scores in another set (e.g., the outcome). For illustration, consider Figure 5.1, the plot of degree of membership in BREAKDOWN (the negation of SURVIVED) against degree of membership in the ~D*~U*~L (not developed, not urban, not literate) corner of the three-dimensional vector space. (The negation of the fuzzy membership scores for SURVIVE in Table 5.2 provides the BREAKDOWN membership scores.) This plot shows that almost all countries' membership scores in this corner of the vector space (~D*~U*~L) are less than or equal to their corresponding scores in BREAKDOWN. The characteristic upper-left triangular plot indicates that the set plotted on the horizontal axis is a subset of the set plotted on the vertical axis. The (almost) vacant lower triangle in this plot corresponds to empty cell #4 of Table 5.5. Just as cases in cell #4 of Table 5.5 are inconsistent with the crisp subset relation, cases in the lower-right triangle of Figure 5.1 are inconsistent with the fuzzy subset relation. Thus, the evidence in Figure 5.1 supports the argument that membership in ~D*~U*~L is a subset of membership in BREAKDOWN, which in turn provides supports for the argument that this combination of conditions (not developed, not urban, and not literate) is sufficient for democratic breakdown.

[Figure 5.1 about here]

Note that when membership in the causal combination is high, membership in the outcome also must be high. However, the reverse does not have to be true. That is, the fact that there are cases with relatively low membership in the causal combination but substantial membership in the outcome is not problematic from the viewpoint of set theory because the expectation is that there may be several different causal conditions or combinations of causal conditions capable of generating high membership in the outcome. Cases with low scores in the causal condition or combination of conditions but high scores in the outcome indicate the operation of alternate causal conditions or alternate combinations of causal conditions.

Figure 5.1 illustrates the fuzzy subset relation using only one corner of the three-dimensional vector space shown in Table 5.6. As shown below, this same assessment could be conducted using degree of membership in the other seven corners (causal combinations) shown in the table. These eight assessments would establish which causal combinations formed from these three causal conditions are subsets of the outcome (BREAKDOWN), which in turn would signal which combinations of conditions might be considered sufficient for the outcome.

2. Using crisp truth tables to aid fuzzy set analysis

The bridge from fuzzy set analysis to truth tables has three main pillars. The first pillar is the *direct correspondence* that exists between the rows of a crisp truth table and the corners of the vector space defined by fuzzy-set causal conditions (Ragin 2000). The second pillar is the assessment of the *distribution of cases* across the logically possible combinations of causal conditions (i.e., the distribution of cases within the vector space defined by the causal conditions). The cases included in a study have varying degrees of membership in each corner of the vector space, as shown in Table 5.6 for a three-dimensional vector space. Some corners of the vector space may have many cases with strong membership; other corners may have no cases with strong membership. When using a crisp truth table to analyze the results of multiple fuzzy set assessments, it is important to take these differences into account. The third pillar is the fuzzy set assessment of the *consistency of the evidence* for each causal combination with the argument that it is a subset of the outcome. The subset relation is important because it signals that there is an explicit connection between a combination of causal conditions and an outcome. Once these three pillars are in place, it is possible to construct a crisp truth table summarizing the results of multiple fuzzy set assessments and then to analyze this truth table using Boolean algebra.

2.1 The correspondence between vector space corners and truth table rows

A multidimensional vector space constructed from fuzzy sets has 2^k corners, just as a crisp truth table has 2^k rows (where k is the number of causal conditions). There is a one-to-one correspondence between causal combinations, truth table rows, and vector space corners (Ragin 2000). The first four columns of Table 5.7 show the correspondence between truth table rows and corners of the vector space. In crisp-set analyses cases are sorted into truth table rows according to their specific combinations of presence/absence scores on the causal conditions. Thus, each case is assigned to a unique row, and each row embraces a unique subset of the cases included in the study. With fuzzy sets, however, each case has varying degrees of membership in the different corners of the vector space and thus varying degrees of membership in each truth table row (as illustrated in Table 5.6).

[Table 5.7 about here]

When using a truth table to analyze the results of fuzzy set assessments, the truth table rows do not represent subsets of cases, as they do in crisp set analyses. Rather, they represent the 2^k causal arguments that can be constructed from a given set of causal conditions. In this light, the first row of the crisp truth table is the causal argument that ~D*~U*~L is a subset of the outcome (democratic BREAKDOWN in this example); the outcome for this row is whether the argument is supported by the fuzzy-set evidence. The second row addresses the ~D*~U*L causal combination, and so on. If both arguments (~D*~U*~L and ~D*~U*L) are supported, then they can be logically simplified to ~D*~U, using Boolean algebra. Thus, in the translation of fuzzy set analyses to crisp truth tables, the rows of the truth table specify the different causal arguments based on the logically possible combinations of causal conditions as represented in the corners of the vector space of causal conditions. Two pieces of information about these corners are especially important: (1) the *number* of cases with strong membership in each corner (i.e., in each combination of causal conditions), and (2) the *consistency* of the empirical evidence for each corner with the argument that degree of membership in the corner (i.e., causal combination) is a subset of degree of membership in the outcome.

2.2 Specifying frequency thresholds for fuzzy-set assessments

The distribution of cases across causal combinations is easy to assess when causal conditions are represented with crisp sets, for it is a simple matter to construct a truth table from such data and to examine the number of cases crisply sorted into each row. Rows without cases are treated as "remainders." When causal conditions are fuzzy sets, however, this analysis is less straightforward because each case may have partial membership in every truth table row (i.e., in every corner of the vector space), as Table 5.6 demonstrates with three causal conditions. Still, it is important to assess the distribution of cases' membership scores across causal combinations in fuzzy-set analyses because some combinations may be empirically trivial. If all cases have very low membership in a combination, then it is pointless to conduct a fuzzy set assessment of that combination's link to the outcome.[4]

Table 5.6 shows the distribution of the membership scores of the 18 countries across the eight logically possible combinations of the three causal conditions. In essence, the table lists the eight corners of the three-dimensional vector space that is formed by the three fuzzy sets and shows the degree of membership of each case in each corner. This table demonstrates an important property of combinations of fuzzy sets, namely, that each case can have only a single membership score greater than 0.5 in the logically possible combinations formed from a given set of causal conditions (shown in bold type).[5] A membership score greater than 0.5 in a causal combination signals that a case is more in than out of the causal combination in question. A score greater than 0.5 also indicates which corner of the multidimensional vector space formed by causal conditions a given case is closest to. This property of fuzzy sets makes it possible for investigators to sort cases according to corners of the vector space, based on their degree of membership. The penultimate column of Table 5.7 shows the number of cases with greater than 0.5 membership in each corner, based on the evidence presented in Table 5.6. For example, Table 5.6 shows that five countries have greater than 0.5 membership in ~D*~U*~L (not developed, not urban, and not literate) and thus are good instances of this combination.

The key task in this phase of the analysis is to establish a number-of-cases threshold for assessing fuzzy subset relations. That is, the investigator must formulate a rule for determining which combinations of conditions are relevant, based on the number of cases with greater than 0.5 membership in each combination. If a combination has enough cases with membership scores greater than 0.5, then it is

---

[4] If the membership scores in a causal combination are all very low, then it is very easy for that combination to satisfy the subset relation signaling sufficiency (where scores in the causal combination must be less than or equal to scores in the outcome). However, the consistency with the subset relation in such instances is meaningless, for the researcher lacks good instances of the combination (i.e., cases with greater than .5 membership in the causal combination).

[5] Note that if a case has 0.5 membership in any causal condition, then its maximum membership in a causal combination that includes that condition is only 0.5. Thus, any case coded 0.5 will not be "closest" to any single corner of the vector space defined by the causal conditions.

reasonable to assess the fuzzy subset relation, as in Figure 5.1. If a combination has too few cases with membership scores greater than .5, then there is no point in conducting this assessment.

The number-of-cases threshold chosen by the investigator must reflect the nature of the evidence and the character of the study. Important considerations include the total number of cases included in the study, the number of causal conditions, the degree of familiarity of the researcher with each case, the degree of precision that is possible in the calibration of fuzzy sets, the extent of measurement and assignment error, whether the researcher is interested in coarse versus fine-grained patterns in the results, and so on. The data set used in this simple demonstration is comprised of only 18 cases and eight logically possible combinations of conditions. In this situation, a reasonable frequency threshold is at least one case with greater than 0.5 membership in a combination. Thus, the three combinations of conditions lacking a single case with greater than 0.5 membership are treated as "remainders" in the analysis that follows, for there are no solid empirical instances of any of them.

When the number of cases is large (e.g., hundreds of cases), it is important to establish a higher frequency threshold. In such analyses, some corners may have several cases with greater than 0.5 membership due to measurement or coding errors. It is prudent in these situations to treat low-frequency causal combinations the same as those lacking strong empirical instances altogether (number of cases with greater than 0.5 membership = 0). When the total number of cases in a study is large, the issue is not which combinations have instances (i.e., at least one case with greater than 0.5 membership), but which combinations have enough instances to warrant conducting an assessment of its possible subset relation with the outcome. For example, a researcher's rule might be that there must be at least five or at least ten cases with greater than 0.5 membership in a causal combination in order to proceed with the assessment of the fuzzy subset relation. By contrast, when the total number of cases is small, it is possible for the researcher to gain familiarity with each case, which in turn mitigates the measurement and coding errors that motivate use of a higher threshold.

2.3 Assessing the consistency of fuzzy subset relations

Once the empirically *relevant* causal combinations have been identified using the procedures just described, the next step is to evaluate each combination's *consistency* with the set theoretic relation in question. Which causal combinations are subsets of the outcome? Social science data are rarely perfect, so it is important to assess the *degree* to which the empirical evidence is consistent with the set theoretic relation in question. Ragin (2006) describes a measure of set theoretic consistency based on fuzzy membership scores (see also Kosko 1993; Smithson and Verkuilen 2006). The formula is:

$$\text{Consistency } (\mathbf{X_i} \leq \mathbf{Y_i}) = \Sigma(\min(\mathbf{X_i},\mathbf{Y_i}))/\Sigma(\mathbf{X_i})$$

where "min" indicates the selection of the lower of the two values, $\mathbf{X_i}$ represents membership scores in a combination of conditions, and $\mathbf{Y_i}$ represents membership scores in the outcome. When all of the $\mathbf{X_i}$ values are less than or equal to their corresponding $\mathbf{Y_i}$ values, the consistency score is 1.00; when there are only a few near misses, the score is slightly less than 1.00; when there are many inconsistent scores, with some $\mathbf{X_i}$ values greatly exceeding their corresponding $\mathbf{Y_i}$ values, consistency drops below 0.5.[6] This measure of consistency prescribes substantial penalties for *large* inconsistencies, but small penalties for near misses (e.g., an $\mathbf{X_i}$ score of .85 and a $\mathbf{Y_i}$ score of .80).

The last column of Table 5.7 reports fuzzy subset consistency scores, using the formula just described. The assessment is conducted for the five combinations that meet the frequency threshold--the combination must have at least one case with greater than 0.5 membership (see Table 5.6). All 18 cases were included in each subset assessment, following the pattern shown in Figure 5.1. In essence, the consistency scores assess the degree to which the evidence for each combination conforms to the upper triangular pattern shown in Figure 5.1. Note that the consistency of the evidence in Figure 5.1 with the subset relation is 0.98, indicating a very high degree of consistency.

2.4 Constructing the truth table

It is a short step from tables like Table 5.7 to crisp set truth tables appropriate for the Quine procedure of QCA. The key determination that must be made is the consistency score to be used as a cut-off value for determining which causal combinations pass fuzzy set theoretic consistency and which do not. Causal combinations with consistency scores at or above the cut-off value are designated fuzzy subsets of the outcome and are coded 1; those below the cut-off value are not fuzzy subsets and are coded 0.[7] In effect, the causal combinations that are fuzzy subsets of the outcome delineate the kinds of cases in which the outcome is found (e.g., the kinds of countries that experienced democratic breakdown). Simple inspection of the consistency values in Table 5.7 reveals that there is a substantial gap in consistency scores between the first and second causal combinations; degree of consistency with the subset relation drops from 0.98 (close to perfect consistency) to 0.83. This gap

---

[6] It is important to point out that when the formula for the calculation of fuzzy set-theoretic consistency is applied to crisp-set data, it returns the simple proportion of consistent cases. Thus, the formula can be applied to crisp and fuzzy data alike.

[7] Rows not meeting the frequency threshold selected by the investigator (based on the number of cases with greater than 0.5 membership) are treated as remainder rows. Designating such rows as remainders is justified on the grounds that the evidence relevant to these combinations is not substantial enough to permit an evaluation of set-theoretic consistency.

provides an easy basis for differentiating consistent causal combinations from inconsistent combinations, as shown in the last column of Table 5.9, which shows the coding of the outcome for truth table analysis. For purposes of comparison, it would be reasonable also to use 0.80 as the cut-off value and conduct an alternate analysis with the first two rows coded as "1" (true). In most analyses of this type, the consistency cut-off value will be substantially lower than perfect consistency, for perfect set-theoretic consistency is not common with fuzzy-set data.[8] Together, the first three columns plus the last column of Table 5.7 form a simple truth table appropriate for standard (crisp set) truth table analysis using the Quine algorithm of QCA. The results of this truth table analysis are not presented here. I present instead an analysis of a more fully specified truth table, using all five causal conditions.

2.5. Application of the procedure

To facilitate comparison of the fuzzy-set analysis with the analyses presented in chapters 3 (crisp-set QCA) and 4 (multi-value QCA), the analysis presented in this section uses all five causal conditions shown in Table 5.2: DEVELOPED, URBAN, INDUSTRIAL, LITERATE, and UNSTABLE. I first show the results using BREAKDOWN as the outcome and then the results using SURVIVED as the outcome.

With five causal conditions, there are 32 (i.e., 2^5) corners to the vector space formed by the fuzzy set causal conditions. These 32 corners correspond to the 32 rows of the crisp truth table formed from the dichotomous versions of these conditions (see chapter 3) and also to the 32 logically possible arguments that can be constructed using five causal conditions. While the eighteen cases all have some degree of membership in the 32 causal combinations, they are, of course unevenly distributed within the five-dimensional vector space. Table 5.8 shows the distribution of cases across the causal combinations (which also constitute corners of the vector space). Specifically, the penultimate column of this table shows the number of cases with greater than 0.5 membership in each combination. (Causal combinations that fail this frequency threshold of at least one case are not shown.) Altogether, there are good instances (i.e., countries with greater than .5 membership) of ten of the 32 logically possible combinations of conditions. The remaining 22 are "remainders" and thus are available

---

[8] Ragin (2000) demonstrates how to incorporate probabilistic criteria into the assessment of the consistency of subset relations, and these same criteria can be modified for use here. The probabilistic test requires a benchmark value (e.g., 0.80 consistency) and an alpha (e.g., 0.05 significance). In the interest of staying close to the evidence, it is often useful simply to sort the consistency scores in descending order and observe whether a substantial gap occurs in the upper ranges of consistency scores. In general, the cut-off value should not be less than 0.75; a cut-off value of 0.85 or higher is recommended. While the measure of consistency used here can range from 0.0 to 1.0, scores between 0 and 0.75 indicate the existence of substantial inconsistency.

as potential counterfactual cases for further logical simplification of the truth table (see Ragin and Sonnett 2004).

The last column of Table 5.8 show the degree of consistency of each causal combination with the argument that it is a subset of the outcome BREAKDOWN. In short, this column shows the truth value of the statement: Membership in the combination of conditions in this row is a subset of membership in the outcome. The rows have been sorted to show the distribution of consistency scores, which range from 0.99 to 0.24. In order to prepare this evidence for conventional truth table analysis it is necessary simply to select a cut-off value for consistency and recode it as a dichotomy. Following the rough guidelines sketched in the previous sections, a cut-off value of 0.80 was selected, which results in six rows coded "1" (true) for the truth table outcome, and four rows coded "0" (false). The reduction of this simple truth table with remainders (i.e., rows without cases) set to "0" (false) shows:

BREAKDOWN ≥ developed*urban*industrial +
DEVELOPED*LITERATE*INDUSTRIAL*UNSTABLE

The set-theoretic consistency of this result is 0.87; the coverage of BREAKDOWN by the two causal combinations is 0.79. (For an explanation of these two measures see Ragin 2006.) The results indicate two paths to democratic breakdown. The first path combines three conditions: low level of development, low urbanization, and low industrialization. In short, this paths reveals that democratic breakdown in the interwar period occurred in some of the least advanced areas of Europe. Countries with strong membership in this combination include Estonia, Hungary, Poland, Portugal, and Romania. The second path is quite different; it combines four conditions: high level of development, high literacy, high industrialization, and political instability. Countries with strong membership in this combination include Austria and Germany. These results are not altogether surprising. The conditions used in this illustration are very general and not based on detailed case-oriented study. Still, it is important to point out that the analysis reveals that there were two very different paths, thus demonstrating the utility of the method for the investigation of causal complexity.

In the language of Ragin and Sonnett (2004), the results just presented constitute the "complex" (or detailed) solution. A "parsimonious" solution can be generated by re-analyzing the truth table with the "remainder" rows (combinations lacking good instances) set to "don't care." (This coding of truth table rows is explained in chapter 3.) This re-analysis of the truth table results in a very simple solution:

BREAKDOWN ≥ developed*urban + UNSTABLE

Again, there are two paths, but this time the paths are quite simple. Following the logic developed in Ragin and Sonnett (2004), however, this solution is "too parsimonious," because the simplifying assumptions that it incorporates via counterfactual analysis are untenable. Therefore, the first solution is the preferred solution; no intermediate solution can be generated without incorporating "difficult" counterfactuals.

Table 5.9 shows the results of the analysis of the same five causal conditions with SURVIVED as the outcome. Because the five causal conditions are the same, the vector space of causal conditions is unchanged, and the distribution of cases within the vector space is unchanged. Once again, there are ten causal combinations with "good instances" (i.e., at least one case with greater than 0.5 membership) and 22 causal combinations lacking good empirical instances. The key difference between Tables 5.9 and 5.8 is the last column, which in Table 5.9 shows the degree of consistency of each causal combination with the statement: Membership in the combination of conditions in this row is a subset of membership in the outcome (SURVIVED). Again the rows have been sorted to show the distribution of the consistency scores. Applying the same cut-off criterion than was applied to Table 5.8 (at least 0.80 consistency) yields only the first row coded "1" (true) and other nine rows coded "0" (false).

Once again, to derive the complex (or detailed) solution, the remainder rows (causal combinations lacking good empirical instances) are set to "0" (false). The results are:

SURVIVED ≥ DEVELOPED*URBAN*LITERATE*INDUSTRIAL*unstable

The set-theoretic consistency of this result is 0.89; the coverage of SURVIVED by this single combination is 0.44 (see Ragin 2006). The one path to survival combines a high level of development, high urbanization, high literacy, high industrialization, and political stability. Countries with high scores in this combination include Belgium, the Netherlands, and the United Kingdom. In essence, the countries with democracies that survived were in advanced area of Europe and avoided political instability. In short, they avoided the two paths to BREAKDOWN shown previously.

The parsimonious solution (which allows the incorporation of remainders into the solution) is as follows:

SURVIVED ≥ DEVELOPED*URBAN*unstable

In essence, the parsimonious solution is a streamlined version of the complex solution. However, this reduction in complexity requires the incorporation of simplifying assumptions that entail "difficult" counterfactuals, as does the possible "intermediate" solutions for this analysis (Ragin and Sonnett 2004). Thus, once again, the complex solution is the preferred solution. More generally, these five causal conditions do a better job of accounting for membership in BREAKDOWN than they do of accounting for membership in SURVIVED. The coverage calculation for BREAKDOWN was 0.79, while it was only 0.44 for SURVIVED. This asymmetry suggests that important causal conditions linked to democratic survival are not represented in the truth table. For example, France, Ireland, and Sweden all have very high membership in SURVIVED, but low membership in the causal combination linked to SURVIVED in the complex solution. Close examination of these cases would provide important clues for specifying additional paths to democratic survival in interwar Europe.

At this juncture it is important to point out a property of fuzzy sets that distinguishes them from crisp sets. Briefly stated, with fuzzy sets it is mathematically possible for a causal condition or causal combination to be a subset of an outcome (e.g., democratic survival) and a subset of the *negation* of that outcome (e.g., democratic breakdown). This result is mathematically possible because degree of membership in a causal condition or combination (e.g., a score of 0.3) can be less than the outcome (e.g., 0.6) and less than the negation of the outcome (1 - 0.6 = 0.4). It is also possible for a causal condition or combination to be *inconsistent* with both the outcome and its negation by exceeding both (e.g., causal combination score = 0.8, outcome membership score = 0.7; negation of the outcome membership score = 0.3). The important point is that there is no mathematical reason, with fuzzy sets, to expect consistency scores calculated for the *negation* of an outcome to be perfectly negatively correlated with consistency scores calculated using the original outcome. Thus, the fuzzy-set analysis of the negation of the outcome (e.g., democratic breakdown) must be conducted separately from the analysis of the outcome (e.g., democratic survival).

This property of fuzzy sets, in effect, allows for *asymmetry* between the results of the analysis of the causes of an outcome and the results of the analysis of the causes of its negation. From the viewpoint of correlational methods, this property of fuzzy sets is perplexing. From the viewpoint of theory, however, it is not. The question of which causal factors produce or generate an outcome is different from the question of which causal factors impede or prevent an outcome from occurring (see Lieberson 1985 on the asymmetry of social causation). Thus, the asymmetry of fuzzy-set analysis dovetails with theoretical expectations of asymmetric causation.

Conclusion

The various procedures sketched in this chapter should not be viewed as "inferential," at least not in the way this term is typically used in quantitative research. QCA does not seek to infer population properties from a sample, nor does it seek to make causal inferences, *per se*. Rather the goal is to aid causal interpretation, in concert with knowledge of cases. The practical goal of the techniques presented in this chapter, and of QCA more generally, is to explore evidence descriptively and configurationally, with an eye toward the different ways conditions may combine to produce a given outcome. Unlike conventional quantitative methods such as regression analysis and related multivariate procedures, there is no "single correct answer" to draw from the analysis of the data. Rather, different results follow from different decisions regarding frequency and consistency thresholds and the like. While these different results are likely to show a strong family resemblance, the choice as to which is "best" may be decided, in the end, only by the cases. The ultimate goal of this chapter is to provide researchers interested in complex causation a variety of strategies and tools for uncovering and analyzing it, while at the same time bringing researchers closer to their evidence.

References

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

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

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

Ragin, Charles C. 2000. *Fuzzy-Set Social Science*. Chicago: University of Chicago Press.

Ragin, Charles C. 2006. "Set Relations in Social Research: Evaluating Their Consistency and Coverage." Political Analysis 14(3):291-310.

Ragin, Charles C. 2007. "Calibration Versus Measurement." Forthcoming in David Collier, Henry Brady, and Janet Box-Steffensmeier (eds.), *Methodology* volume of *Oxford Handbooks of Political Science*.

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

Smithson, Michael and Jay Verkuilen. 2006. *Fuzzy Set Theory*. Thousand Oaks, CA: Sage.

Zadeh, Lotfi. 1965. "Fuzzy sets." *Information and Control* 8: 338-353.

Appendix: A summary of the procedure

The central focus of this chapter is the process of analyzing crisp truth tables constructed from multiple fuzzy set analyses. The basic steps are:

1. Create a data set with fuzzy-set membership scores. (Crisp sets may be included among the causal conditions.) The fuzzy sets must be carefully defined (e.g., degree of membership in the set of "countries with high levels of literacy"). Pay close attention to the calibration of fuzzy membership scores, especially with respect to the three qualitative anchors: full membership (1.0), full nonmembership (0.0), and the cross-over point (0.5). In general, calibration requires good grounding in theoretical and substantive knowledge, as well as in-depth understanding of cases. The procedures described in this chapter work best when the 0.5 membership score and membership scores close to 0.5 are used sparingly when coding the causal conditions.

2. Input the fuzzy-set data directly into fsQCA or into a program that can save data files in a format compatible with fsQCA (e.g., Excel: comma delimited files; SPSS: tab delimited files; simple, SPSS-type variable names should appear on the first row of the data file). The data set should include both the outcome and as many of the possibly relevant causal conditions as feasible. Open the data file using fsQCA version 2.0 dated June 2006 or later. (Click *Help* on the start-up screen to identify fsQCA version and date; the most up-to-date version can be downloaded from www.fsqca.com.)

3. Select a preliminary list of causal conditions. In general, the number of causal conditions should be modest, in the range of three to eight. Often causal conditions can be combined in some way to create "macrovariables" using the procedures described in Ragin (2000). These macrovariables can be used in place of their components to reduce the dimensionality of the vector space. For example, a single macrovariable might be used to replace three substitutable causal conditions joined together by logical *or*, which dictates using their maximum membership score. (In the *Data Sheet* window of fsQCA, click *Variables*, then *Compute*, and then use the *fuzzyor* function to create this type of macrovariable.)

4. Create a truth table by specifying the outcome and the causal conditions. In fsQCA this function is accessed by clicking *Analyze*, *Fuzzy Sets*, and *Truth Table Algorithm*. The resulting truth table will have 2^k rows, reflecting the different corners of the vector space. (The 1s and 0s for the causal conditions in this spreadsheet identify the different corners of the vector space.) For each row, the program reports the number of cases with greater than 0.5 membership in the vector space corner (in the column labeled *number*). Two columns to the right of *number* is *consistency*, the measure assessing the

degree to which membership in that corner is a subset of membership in the outcome.

5. The researcher must select a frequency threshold to apply to the data listed in the *number* column. When the total number of cases included in a study is relatively small, the frequency threshold should be 1 or 2. When the total *N* is large, however, a more substantial threshold should be selected. It is very important to inspect the distribution of the cases when deciding upon a frequency threshold. This can be accomplished simply by clicking on any case in the *number* column and then clicking the *Sort* menu and then *Descending*. The resulting ordered list of the number of cases with greater than 0.5 membership in each corner will provide a snapshot of the distribution and also may reveal important discontinuities or gaps. After selecting a threshold, delete all rows that do not meet it. This can be accomplished (for tables that have been sorted according to *number*) by clicking on the first case that falls below the threshold (in the *number* column), clicking the *Edit* menu, and then clicking *Delete current row to last*. The truth table will now list only the rows (corners of the vector space) that meet the frequency threshold.

6. Next is the selection of a consistency threshold for distinguishing causal combinations that are subsets of the outcome from those that are not. This determination is made using the measure of set-theoretic consistency reported in the *consistency* column. In general, values below 0.75 in this column indicate substantial inconsistency. It is always useful to sort the consistency scores in descending order so that it is possible to evaluate their distribution. This should be done *after* rows that fall below the frequency threshold have been deleted from the table (step 5). Click on any value in the *consistency* column; click the *Sort* menu; and then click *Descending*. Identify any gaps in the upper range of consistency that might be useful for establishing a threshold, keeping in mind that it is always possible to examine several different thresholds and assess the consequences of lowering and raising the consistency cut-off.

7. Input 1s and 0s into the empty outcome column, which is labeled with the name of the outcome and listed to the left of the *consistency* column. Using the threshold value selected in the previous step, enter a value of 1 when the consistency value meets or exceeds the consistency threshold and 0 otherwise. If the truth table spreadsheet has many rows, you may want to code the outcome column using the *Delete and code* function in the *Edit* menu.

8. Once the outcome column is completely filled in, click the *Standard Analysis* button at the bottom of the truth table spreadsheet. Clicking this button will give you two solutions, the complex solution (with remainders set to "false") and the parsimonious solution (with remainders set to "don't care"). Conceive of the complex and

parsimonious solutions as the two endpoints of a single complexity/parsimony continuum (see Ragin and Sonnett 2004). Any solution that is a subset of the most parsimonious solution and a superset of the most complex solution is a valid solution of the truth table. These intermediate solutions use only a subset of the simplifying assumptions that are used in the most parsimonious solution. Ragin and Sonnett (2004) explain how to use theoretical and substantive knowledge to derive an optimal solution. They link these procedures to counterfactual analysis, a technique that is central to case-oriented research.

Table 5.1: Crisp versus fuzzy sets

Crisp set	Three-value fuzzy set	Four-value fuzzy set	Six-value fuzzy set	"Continuous" fuzzy set
1 = fully in	1 = fully in	1 = fully in	1 = fully in	1 = fully in
			.9 = mostly but not fully in	Degree of membership is more "in" than "out": $.5 < X_i < 1$
		.67 = more in than out	.6 = more or less in	
	.5 = neither fully in nor fully out			.5 = cross-over: neither in nor out
			.4 = more or less out	
		.33 = more out than in		Degree of membership is more "out" than "in": $0 < X_i < .5$
			.1 = mostly but not fully out	
0 = fully out	0 = fully out	0 = fully out	0 = fully out	0 = fully out

Table 5.2: Data matrix showing original variables and fuzzy-set membership scores

Country	Survived	Survived-FZ	Developed	Developed-FZ	Urban	Urban-FZ	Literate	Literate-FZ	Industrial	Industrial-FZ	Unstable	Unstable-FZ
Austria	-9.00	0.01	720	0.74	33.4	0.14	98.0	0.98	33.4	0.76	10.00	0.65
Belgium	10.00	0.98	1,098	0.99	60.5	0.89	94.4	0.96	48.9	0.98	4.00	0.04
Czechoslovakia	7.00	0.85	586	0.42	69.0	0.96	95.9	0.97	37.4	0.91	6.00	0.13
Estonia	-6.00	0.12	468	0.15	28.5	0.07	95.0	0.96	14.0	0.02	6.00	0.13
Finland	4.00	0.64	590	0.43	22.0	0.03	99.1	0.98	22.0	0.09	9.00	0.49
France	10.00	0.98	983	0.97	21.2	0.02	96.2	0.97	34.8	0.83	5.00	0.07
Germany	-9.00	0.01	795	0.85	56.5	0.83	98.0	0.98	40.4	0.96	11.00	0.77
Greece	-8.00	0.03	390	0.05	31.1	0.10	59.2	0.11	28.1	0.38	10.00	0.65
Humgary	-1.00	0.41	424	0.08	36.3	0.20	85.0	0.81	21.6	0.08	13.00	0.91
Ireland	8.00	0.91	662	0.62	25.0	0.04	95.0	0.96	14.5	0.02	5.00	0.07
Italy	-9.00	0.01	517	0.25	31.4	0.11	72.1	0.38	29.6	0.49	9.00	0.49
Netherlands	10.00	0.98	1,008	0.97	78.8	0.99	99.9	0.99	39.3	0.94	2.00	0.01
Poland	-6.00	0.12	350	0.03	37.0	0.22	76.9	0.55	11.2	0.02	21.00	0.98
Portugal	-9.00	0.01	320	0.02	15.3	0.01	38.0	0.02	23.1	0.12	19.00	0.98
Romania	-4.00	0.25	331	0.02	21.9	0.03	61.8	0.15	12.2	0.02	7.00	0.22
Spain	-8.00	0.03	367	0.04	43.0	0.41	55.6	0.08	25.5	0.22	12.00	0.86
Sweden	10.00	0.98	897	0.93	34.0	0.15	99.9	0.99	32.3	0.70	6.00	0.13
United Kingdom	10.00	0.98	1,038	0.98	74.0	0.98	99.9	0.99	49.9	0.98	4.00	0.04

Table 5.3: Illustration of logical *and*

Country	Developed	Urban	Unstable	Developed and Urban	Developed, Urban and Unstable
Austria	.74	.14	.65	.14	.14
Belgium	.99	.89	.04	.89	.04
Czechoslovakia	.42	.96	.13	.42	.13
Estonia	.15	.07	.13	.07	.07
Finland	.43	.03	.49	.03	.03
France	.97	.02	.07	.02	.02
Germany	.85	.83	.77	.83	.77
Greece	.05	.10	.65	.05	.05
Hungary	.08	.20	.91	.08	.08
Ireland	.62	.04	.07	.04	.04
Italy	.25	.11	.49	.11	.11
Netherlands	.97	.99	.01	.97	.01
Poland	.03	.22	.98	.03	.03
Portugal	.02	.01	.98	.01	.01
Romania	.02	.03	.22	.02	.02
Spain	.04	.41	.86	.04	.04
Sweden	.93	.15	.13	.15	.13
United Kingdom	.98	.98	.04	.98	.04

Table 5.4: Illustration of logical *or*

Country	Developed	Urban	Unstable	Developed or Urban	Developed or Urban or Unstable
Austria	.74	.14	.65	.74	.74
Belgium	.99	.89	.04	.99	.99
Czechoslovakia	.42	.96	.13	.96	.96
Estonia	.15	.07	.13	.15	.15
Finland	.43	.03	.49	.43	.49
France	.97	.02	.07	.97	.97
Germany	.85	.83	.77	.85	.85
Greece	.05	.10	.65	.10	.65
Hungary	.08	.20	.91	.20	.91
Ireland	.62	.04	.07	.62	.62
Italy	.25	.11	.49	.25	.49
Netherlans	.97	.99	.01	.99	.99
Poland	.03	.22	.98	.22	.98
Portugal	.02	.01	.98	.02	.98
Romania	.02	.03	.22	.03	.22
Spain	.04	.41	.86	.41	.86
Sweden	.93	.15	.13	.93	.93
United Kingdom	.98	.98	.04	.98	.98

Table 5.5: Crosstabulation of outcome against presence/absence of a causal combination

	Causal combination absent	Causal combination present
Outcome present	1. not directly relevant	2. cases here
Outcome absent	3. not directly relevant	4. no cases here

Table 5.6: Fuzzy set membership of cases in causal combinations

Country	Membership in causal conditions			Membership in corners of vector space formed by causal conditions							
	DEVELOPED (D)	URBAN (U)	LITERATE (L)	~D*~U*~L	~D*~U*L	~D*U*~L	~D*U*L	D*~U*~L	D*~U*L	D*U*~L	D*U*L
Austria	0.74	0.14	0.98	0.02	0.26	0.02	0.14	0.02	0.74	0.02	0.14
Belgium	0.99	0.89	0.96	0.01	0.01	0.01	0.01	0.04	0.11	0.04	0.89
Czechoslovakia	0.42	0.96	0.97	0.03	0.04	0.03	0.58	0.03	0.04	0.03	0.42
Estonia	0.15	0.07	0.96	0.04	0.85	0.04	0.07	0.04	0.15	0.04	0.07
Finland	0.43	0.03	0.98	0.02	0.57	0.02	0.03	0.02	0.43	0.02	0.03
France	0.97	0.02	0.97	0.03	0.03	0.02	0.02	0.03	0.97	0.02	0.02
Germany	0.85	0.83	0.98	0.02	0.15	0.02	0.15	0.02	0.17	0.02	0.83
Greece	0.05	0.10	0.11	0.89	0.11	0.10	0.10	0.05	0.05	0.05	0.05
Hungary	0.08	0.20	0.81	0.19	0.80	0.19	0.20	0.08	0.08	0.08	0.08
Ireland	0.62	0.04	0.96	0.04	0.38	0.04	0.04	0.04	0.62	0.04	0.04
Italy	0.25	0.11	0.38	0.62	0.38	0.11	0.11	0.25	0.25	0.11	0.11
Netherlands	0.97	0.99	0.99	0.01	0.01	0.01	0.03	0.01	0.01	0.01	0.97
Poland	0.03	0.22	0.55	0.45	0.55	0.22	0.22	0.03	0.03	0.03	0.03
Portugal	0.02	0.01	0.02	0.98	0.02	0.01	0.01	0.02	0.02	0.01	0.01
Romania	0.02	0.03	0.15	0.85	0.15	0.03	0.03	0.02	0.02	0.02	0.02
Spain	0.04	0.41	0.08	0.59	0.08	0.41	0.08	0.04	0.04	0.04	0.04
Sweden	0.93	0.15	0.99	0.01	0.07	0.01	0.07	0.01	0.85	0.01	0.15
United Kingdom	0.98	0.98	0.99	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.98

Table 5.7: The correspondence between truth table rows and vector space corners

Developed	Urban	Literate	Corresponding Vector Space Corner (Table 5.6)	N of cases with membership in causal combination > .5	Consistency with subset relation vis-a-vis the outcome (N = 18 in each assessment)	Outcome code (based on consistency score)
0	0	0	~D*~U*~L	5	0.98	1
0	0	1	~D*~U*L	4	0.83	0
0	1	0	~D*U*~L	0	(too few cases with scores > .5)	remainder
0	1	1	~D*U*L	1	0.74	0
1	0	0	D*~U*~L	0	(too few cases with scores > .5)	remainder
1	0	1	D*~U*L	4	0.46	0
1	1	0	D*U*~L	0	(too few cases with scores >.5)	remainder
1	1	1	D*U*L	4	0.34	0

Table 5.8: Distribution of cases across causal combinations and set-theoretic consistency of causal combinations as subsets of BREAKDOWN

DEVELOPED	URBAN	LITERATE	INDUSTRIAL	UNSTABLE	N of cases with > .5 membership	Consistency as a subset of BREAKDOWN
0	0	0	0	1	3	0.99
0	0	0	0	0	2	0.98
1	1	1	1	1	1	0.91
1	0	1	1	1	1	0.89
0	0	1	0	1	2	0.88
0	0	1	0	0	2	0.83
0	1	1	1	0	1	0.67
1	0	1	0	0	1	0.58
1	0	1	1	0	2	0.44
1	1	1	1	0	3	0.24

Table 5.9: Distribution of cases across causal combinations and set-theoretic consistency of causal combinations as subsets of SURVIVED

DEVELOPED	URBAN	LITERATE	INDUSTRIAL	UNSTABLE	N of cases with > .5 membership	Consistency as a subset of SURVIVED
1	1	1	1	0	3	0.89
1	0	1	0	0	1	0.79
1	0	1	1	0	2	0.74
0	1	1	1	0	1	0.69
0	0	1	0	0	2	0.51
0	0	1	0	1	2	0.51
1	1	1	1	1	1	0.40
1	0	1	1	1	1	0.40
0	0	0	0	0	2	0.32
0	0	0	0	1	3	0.23

Figure 5.1: Plot of degree of membership in BREAKDOWN against degree of membership in ~D*~U*~L