From Fuzzy Sets to Crisp Truth Tables[1]

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

Version: April 2005

1. Overview

One limitation of the truth table approach is that it is designed for causal conditions are simple presence/absence dichotomies (i.e., Boolean or "crisp" sets). 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 are many 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. Section 2 of this paper 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. Ragin (2000) demonstrates that the subset relation is central to the analysis of multiple conjunctural causation, where several different combinations of conditions are sufficient for the same outcome.

While fuzzy sets solve the problem of trying to force-fit cases into one of two categories (membership versus nonmembership in a set), they are not well suited for conventional truth table analysis. 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 tables 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 and virtues 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 the use of different subsets of "remainders" to reduce complexity (see Ragin 1987; Ragin and Sonnett 2004). Analyses of this type are difficult without using a truth table as the starting point. Ragin and Sonnett (2004), for example, show how to use QCA to aid counterfactual analysis and link the analysis of counterfactual conditions to core practices in case-oriented research. Truth tables are central to the analysis of counterfactuals, and the techniques described in Ragin and Sonnett (2004) cannot be implemented without the aid of truth tables.

[1] I thank Gary Goertz and Benoit Rihoux for their many useful comments on this paper. I also thank Sean Davey for his work on the software implementing the ideas described here.

Section 3 of this paper 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 data on class voting in the advanced industrial societies, compiled by Paul Nieuwbeerta (see, e.g., Nieuwbeerta and de Graaf 1999; Nieuwbeerta and Ultee 1999; Nieuwbeerta, de Graaf and Ultee 2000). It is important to point out that the approach sketched in this paper 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 paper 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.

2. 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.

2.1 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 Security Council).

Categorical-level measurement, the foundation of crisp sets, is considered by many social scientists to be inferior to interval-level measurement. However, crisp sets are often better anchored in substantive and theoretical knowledge than conventional interval-scale measures. For example, it is a relatively simple matter to identify an interval-scale indicator of country wealth, which in turn provides a simple tool for evaluating the *relative* positions of countries on this dimension. By contrast, it is more challenging to define the set of "rich countries" in a qualitative manner and then specify which countries are fully in this set and which are not. The key difference is that qualitative distinctions are *explicit* and must be grounded in substantive and theoretical knowledge, while the *relative* rankings of an interval scale can be pegged simply to scores on a crude indicator of the underlying construct (e.g., GNP per capita as an indicator of country wealth).

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 single 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 point of maximum ambiguity regarding membership (.5).

TABLE 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 1, instead of using only two scores, 0 and 1, 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 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 1. The four-value scheme uses the numerical values 0, 0.25, 0.75, 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 evidence is not systematic or strictly comparable from case to case. A more fine-grained fuzzy set uses six values, as shown in column 4 of Table 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, 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 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 might translate to 0.3 rather than 0.1. 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 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 *not* relevant to 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 2 ABOUT HERE

In a fuzzy-set analysis both the outcome and the causal conditions are represented using fuzzy sets.[2] Table 2 shows a simple data matrix containing fuzzy membership scores. The data are relevant to class voting in the advanced industrial societies. In this example, the outcome of interest is the degree of membership in the set of countries with *weak* class voting (labeled *W* for *weak class voting*). This fuzzy set was constructed from survey evidence, compiled by Paul Nieuwbeerta, covering the post-World War II era. While levels of class voting have generally declined across the advanced industrial countries, the rank order of these countries with regards to levels of class voting has remained remarkably stable over time. This analysis focuses on the conditions linked to persistently low levels of class voting. The causal conditions used in this example are: (1) degree of membership in the set of countries with strong unions (*U*), (2) degree of membership in the set of countries with a high percentage of workers employed in manufacturing (*M*), (3) degree of membership in the set of highly affluent countries (*A*), and (4) degree of membership in the set of countries with substantial levels of income inequality (*I*). All fuzzy sets used in this analysis are six-value sets and are based on general characterizations of these countries over the post-World War II period. While finer gradations are possible with these data (as in column 5 of Table 1), the intent here is to demonstrate fuzzy-set procedures with a data set typical of those used in comparative research, where precise quantitative assessment of cases is often difficult.[3]

2.2 Operations on fuzzy sets

[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 persistently low levels of class voting. Instead, the focus is on practical procedures.

Before presenting the bridge between fuzzy sets and truth tables, 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, for example, is the crisp set of not-democracies. This simple mathematical principle holds in fuzzy algebra 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

$a = 1 - A$

(Lower-case letters are used to indicate negation.) Thus, for example, if the U.S. has a membership score of 0.9 in the set of "democratic countries," it has a score of 0.1 in the set of "not-democratic countries." That is, the U.S. would be mostly but not fully out of the set of not-democratic countries. The first four columns of Table 3 list all the original and all the negated membership scores of the four causal conditions listed in Table 2.

TABLE 3 ABOUT HERE

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. 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. For further illustration of this principle, consider the right-hand side of Table 3. The last two columns demonstrate the operation of logical *and*. The fifth column shows the intersection of the first two columns, yielding membership in the set of countries that combine being affluent with income inequality. The algebraic expression for this intersection is $A*I$; the asterisk is used to indicate set intersection. The next column shows the membership scores that result when this combination is further elaborated with information on degree of membership in the set of countries with not-strong unions (u). The algebraic expression for this combination is $A*I*u$: highly affluent countries with substantial income inequality and weak unions. One case (the U.S., in the last row) has very strong membership in this combination (0.9). As more sets are added to a combination of conditions, membership scores either stay the same or decrease. In 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 4 ABOUT HERE

For illustration of the use of logical *or*, consider Table 4 which reproduces the first four columns of Table 3. The right-hand side of Table 4 shows the operation of logical *or*. The fifth column shows countries that have a high percentage of workers in manufacturing (*M*) *or* strong unions (*U*): $M + U$. (Addition is used to indicate logical *or*.) The next column shows a relatively complex function, combining logical *or* and logical *and*: the set of countries that combine low income inequality (*i*) with either strong unions (*U*) or a high percentage of workers employed in manufacturing (*M*): $i*(M + U)$. To ascertain each case's membership in $i*(M + U)$, first determine the maximum of *M* and *U*, and then find the minimum of *i* and the maximum just derived.

2.3 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 social structural conditions linked to persistently low levels of class voting).

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.

TABLE 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 6. This table shows the membership of countries in three of the causal conditions used in this example and in the eight causal combinations that can be formed from these three fuzzy sets. 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 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 corner can be calculated using fuzzy algebra, as shown in Table 6. (This link between fuzzy-set vector spaces and crisp truth tables is explored in more detail below.)

TABLE 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 the *hypothetical* data listed in Table 7 and plotted in Figure 1. Table 7 shows membership scores in two fuzzy sets, the set of religiously heterogeneous countries and the set of countries with weak class voting. Observe that the class voting scores are consistently greater than or equal to the religious heterogeneity scores. This pattern is consistent with the fuzzy subset relation. If membership in the causal combination is high, then membership in the outcome also must be high. Note, however, that 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 combinations of causal conditions capable of generating high membership in the outcome (i.e., there are causal conditions other than religious heterogeneity linked to strong membership in the set of countries with persistently low levels of class voting). 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.

TABLE 7 ABOUT HERE

Figure 1 illustrates the fuzzy subset relation. 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. As just noted, the points in the upper left-hand region of the plot are not "errors," as they would be in a linear regression analysis. Rather, these points have strong membership in the outcome due to the operation of *other* combinations of causal conditions. The vacant lower triangle in this plot corresponds roughly to empty cell #4 of Table 5. Just as cases in cell #4 of Table 5 would be inconsistent with the crisp subset relation, cases in the lower-right triangle of Figure 1 would be inconsistent with the fuzzy subset relation.

FIGURE 1 ABOUT HERE

Table 7 and Figure 1 illustrate the fuzzy subset relation using a single causal condition. Note, however, that this same assessment could be conducted using degree of membership in a *combination* of causal conditions. Degree of membership in a causal combination (such as those shown in Table 6) can be used in an assessment of fuzzy subset relations by comparing these scores with membership scores in the outcome, as in Table 7 and Figure 1. This examination establishes whether degree of membership in a combination of causal conditions is a fuzzy subset of degree of membership in the outcome, a pattern of results consistent with an argument of sufficient causation (Ragin 2000).

3. Constructing truth tables from fuzzy sets

The bridge from fuzzy-set data to crisp 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 different 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 6 for a three-dimensional vector space. Some corners of the vector space may have many cases with strong membership; other corners may have only a few cases with any degree of membership at all. When constructing a crisp truth table from fuzzy sets, it is important to take these differences into account. The third pillar is the 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 from fuzzy-set data.

3.1 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). 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 may have varying degrees of membership in the different corners of the vector space.

Still, truth tables are relevant to fuzzy-set analysis. The researcher can use truth table rows as specifications of the corners of the vector space and uses the truth table to summarize statements about the characteristics of corners. In essence, truth table rows become the predicates in statements about vector space corners. For example, the researcher might determine if degree of membership in a corner of the vector space is a subset of degree of membership in the outcome. (This assessment would involve evidence on all cases included in the study, as illustrated in Table 7 and Figure 1.) The researcher could then append information about the results of this assessment to the corresponding truth table row.

Thus, in the translation of fuzzy sets to crisp truth tables, the truth table represents *statements* about the corners of the vector space formed by the fuzzy sets. 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 is a subset of degree of membership in the outcome.

3.2 Assessing the distribution of cases across causal combinations

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. 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 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. In other words, if most cases have very low or zero membership in a combination, then it is pointless to assess that combination's link to the outcome. The empirical base for such an assessment would be too weak.

TABLE 8 ABOUT HERE

Table 8 shows the distribution of the membership scores of the 12 countries across the 16 logically possible combinations of the four causal conditions. In essence, the table lists the 16 corners of the four dimensional vector space that is formed by the four 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.[4] A membership score greater than 0.5 in a

[4] Note that if a case has 0.5 membership in a causal condition, then its maximum membership in any

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 last row of Table 8 shows the number of cases with greater than 0.5 membership in each corner.

The key task in this phase of the analysis is to establish a number-of-cases threshold, that is, to develop a rule for classifying some combinations of conditions as "relevant" and others as "remainders" based on the number of cases with greater than 0.5 membership in each. The rule established by the investigator must reflect the nature of the evidence and the character of the study. Important considerations include the total number of cases, 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 demonstration is comprised of 12 cases and 16 logically possible combinations of conditions. In this situation, a reasonable frequency threshold is one case. Thus, the eight 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 strong empirical instances of any of them. The causal combinations with at least one case with greater than 0.5 membership are retained for further examination.

When the number of cases is large (e.g., 100s), it is important to establish a *frequency threshold* for the relevance or viability of causal combinations. In such analyses, some corners may have several cases with greater than 0.5 membership due to measurement or coding errors. It would be prudent, therefore, to treat low-frequency causal combinations the same as those lacking strong empirical instances altogether (frequency = 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 the subset relation with the outcome (e.g., at least five cases with greater than 0.5 membership). 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 high threshold.

3.3 Assessing the consistency of fuzzy subset relations

Once the empirically *relevant* causal combinations have been identified using the procedures just described (section 3.2), 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? What are the different combinations of conditions linked to strong membership in the outcome? Notice that the hypothetical data in Table 7 are perfect from a set theoretic viewpoint. All the membership scores in the causal condition are less than or equal to their corresponding membership scores in the outcome. Based on this evidence, a researcher could claim that this causal condition (religious heterogeneity) is a subset of the outcome (weak class voting). Thus, religious heterogeneity could be interpreted (hypothetically) as a sufficient condition for weak class voting.

Social science data are rarely this uniform, however. When there are cases that are inconsistent with the subset relation, it is important to be able to assess the *degree* to which the empirical evidence is consistent with the set theoretic relation in question. For example, suppose the value of religious

---

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.

homogeneity in the first row of Table 7 was 1.0 instead of 0.7. It would be inconsistent with the set theoretic relation because this value exceeds the corresponding outcome membership score, 0.9. While the set theoretic relation would no longer hold consistently across the cases listed in Table 7, it would still be very close to perfect, with 11 out of the 12 cases consistent (92%).

Ragin (2003) proposes a measure of set theoretic consistency based on fuzzy membership scores. It is simply the sum of the *consistent* membership scores in a causal condition or combination of causal conditions (e.g., in column 1 of Table 7) divided by the sum of *all* the membership scores in a cause or causal combination. In Table 7, as presented, the value of this measure is 1.0 (4.7/4.7) because all the membership scores in column 1 are consistent. If the value of religious homogeneity in the first row of Table 7 is changed to 1.0, however, consistency drops to 0.8 (4/5). The numerator is 1.0 fuzzy units lower than the denominator because of the one inconsistent score of 1.0. The reduction of consistency to 0.8 (from perfect consistency, 1.0) is substantial because 1.0 (the value substituted for 0.7 in the first row) is a large membership score.

The consistency measure described in Ragin (2003) can be refined further so that it gives credit for near misses and penalties for causal membership scores that exceed their mark, the outcome membership score, by a wide margin. This adjustment can be accomplished by adding to the numerator in the formula just sketched the portion of each *inconsistent* causal membership score that is consistent with the outcome. For example, if the value of religious homogeneity in the first row of Table 7 is changed to 1.0, then most of it is consistent, up to the value of the outcome membership score, 0.9. This portion is added to the numerator of the consistency measure. Using this more refined measure of consistency yields an overall consistency score of 0.98 (4.9/5). This refined consistency score is more compatible with the evidence. After all, all but one of the scores are consistent, and the one inconsistent score is a very near miss. Thus, a consistency score very close to 1.0 would be expected.

Notice that the revised measure of consistency just sketched prescribes a substantial penalty for *large* inconsistencies. Suppose again that the value of religious homogeneity in the first row of Table 7 is 1.0, but this time assume that the corresponding value of the outcome, weak class voting, is only 0.3. The consistent portion of the 1.0 membership score is 0.3, yielding an overall addition of only 0.3 to the numerator. The resulting consistency score in this instance would be 0.86 (4.3/5). This lower score reflects the fact that the one inconsistent score exceeds its target by a very wide margin.

TABLE 9 ABOUT HERE

Table 9 shows the degree to which the eight causal combinations (those with at least one case with greater than 0.5 membership) are consistent subsets of the outcome, weak class voting, using data on all 12 countries for each calculation. The consistency scores are based on the formula just described, which gives credit for near misses. (For ease of interpretation the combinations have been sorted in descending order according to their consistency scores.) It is important to remember that Table 9 presents summary *statements* about the corners of the vector space defined by the four fuzzy-set causal conditions used in this analysis. Each row, in effect, answers the question: Is degree of membership in this corner of the vector space a subset of degree of membership in the outcome? The analysis of the evidence in this table is thus an analysis of statements about vector space corners.

3.4 Constructing and analyzing the truth table

It is a short step from tables like Table 9 to 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 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.[5] 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 have weak class voting). Simple inspection of the consistency values in Table 9 reveals that there is a substantial gap in consistency scores between the second and third causal combinations; degree of consistency with the subset relation drops from 1.00 (perfect consistency) to 0.87. This gap provides an easy basis for differentiating consistent causal combinations from inconsistent combinations, as shown in the last column of Table 9, which is labeled *Outcome*. In most analyses of this type, however, the consistency cut-off value will be less than 1.0, for perfect set-theoretic consistency is not common with fuzzy-set data.

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., 0.75 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, as in Table 9, and observe whether a substantial gap occurs in the upper ranges of consistency scores. With the large gap between rows two and three of Table 9, it is clear that using probabilistic criteria to aid the selection of a cut-off value would simply obfuscate the obvious. In general, the cut-off value should not be less than 0.75; a cut-off value ≥ 0.85 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.

The crisp truth table resulting from these procedures is contained within Table 9. The first four columns show the codings of the causal conditions; the last column shows the crisp-set outcome (consistent versus not-consistent) attached to each truth table row (vector space corner). The eight combinations of causal conditions not listed in Table 9 are "remainders." The results of the analysis of this truth table with remainders defined as false (i.e., no simplifying assumptions used) show:

A*m*u ---> W

where: *W* = membership in the set of countries with weak class voting; *A* = affluent country; *M* = high manufacturing employment; *U* = strong unions; upper-case letters indicate the original fuzzy sets; lower-case letters indicate negated fuzzy sets; and "--->" signals the subset relation. The equation indicates that weak class voting (*W*) occurs in countries where affluence is combined with both weak membership in the set of countries with high manufacturing employment and weak membership in the set of countries with strong unions. Inspection of Table 8 reveals that the best instances of this combination of conditions are the U.S., France, Germany, and the Netherlands.

When simplifying assumptions are drawn from the pool of eight remainders, a more parsimonious solution results:

u ---> W

According to this equation, weak class voting has a single source, weak unions (*u*). As Table 2 shows, the countries with the weakest unions are the U.S. and France. This solution of the truth table is dependent on six simplifying assumptions (see Ragin 1987; 2000) describing unobserved combinations of causal conditions. The six assumptions are drawn from the eight combinations of causal conditions lacking strong empirical instances (i.e., those with no cases with membership scores exceeding 0.5).[6]

---

[5] Any rows not meeting the frequency threshold selected by the investigator (based on the number of cases with greater than .5 membership) should be 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.

[6] Ragin and Sonnett (2004) demonstrate how to use these two solutions to conduct counterfactual

At this juncture it is important to point out a property of fuzzy sets that sharply 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 both an outcome (e.g., weak class voting) and the *negation* of the outcome (e.g., strong class voting). 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 for the original outcome. Thus, the fuzzy-set analysis of the negation of the outcome (e.g., strong class voting in this demonstration) must be conducted separately from the analysis of the outcome (e.g., weak class voting). This property of fuzzy sets, in effect, allows for *asymmetry* between the results of the analysis of the causes of an outcome and 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. Consider the example presented in this paper. The analysis shows that persistently low levels of class voting are found in countries with weak unions. In effect, this analysis indicates the main impediment to class voting. The question of which conditions are *impediments* to class voting is not the same as the question of which conditions are *productive* of class voting (see Lieberson 1985 on the asymmetry of social causation). Thus, the asymmetry of fuzzy-set analysis dovetails with theoretical expectations of asymmetric causation.

4. Why Not Dichotomize Fuzzy Sets?

The construction of a crisp truth table from fuzzy-set data, as just sketched, is somewhat laborious. It involves two relatively difficult procedures: (1) the assessment of the distribution of cases across causal combinations, and (2) the assessment of the degree of consistency of each causal combination with the subset relation vis-a-vis the outcome. Further, both assessments involve the selection of cut-off values, which may seem arbitrary.[7] Why not simply recode fuzzy sets to crisp sets and conduct a conventional crisp-set analysis using the dichotomized membership scores? After all, a fuzzy score of 0.5 differentiates cases that are "more in" versus "more out" of a fuzzy set. The use of this cross-over value to create crisp sets from fuzzy sets appears at first glance to be a straightforward extension of the approach.

TABLE 10 ABOUT HERE

The best way to evaluate the viability of this option is to re-analyze the fuzzy-set data presented in Table 2, first converting the fuzzy sets to crisp sets. Table 10 shows the crisp-set data that result from the application of the cross-over rule (which dichotomizes fuzzy data at the 0.5 membership score) to the fuzzy-set data presented in Table 2. There is clearly a gain in simplicity, comparing Tables 2 and 10. The cases in Table 10 are either fully "in" (1) or fully "out" (0) of the relevant sets; in Table 2, the

---

analyses. The first solution maximizes complexity; the second maximizes parsimony. They show how theoretical and substantive knowledge can be used to specify various middle paths between these two solutions.

[7] Actually, the range of plausible cut-off values is relatively narrow, and the range can be narrowed further when researchers are familiar with their cases and with the relevant theoretical and substantive literatures.

memberships are fuzzy. However, there are some costs. Notice, for example, that the data set now includes a contradiction. Belgium and Denmark have the same scores on the four crisp causal conditions, yet they have different scores on the outcome. In a conventional crisp-set analysis, it is necessary to address this contradiction in some way before the analysis can proceed.

With one contradiction and eight remainders, there are several ways to analyze the evidence in Table 10. In an effort to match the second fuzzy-set solution as closely as possible in the analysis that follows, the contradiction is set to "false" and the eight remainders are used as "don't care" combinations, which makes them available for use as simplifying assumptions. (A "don't care" row may be assigned either 1 or 0 on the outcome by fs/QCA, depending on which assignment yields the most parsimonious solution.) The results of the crisp-set analysis show:

u + a + I*m ---> W

where: *W* = membership in the set of countries with weak class voting; *U* = strong unions; *A* = highly affluent country; *I* = substantial income inequality; *M* = high manufacturing employment; upper-case letters indicate the original fuzzy sets; lower-case letters indicate negated fuzzy sets; "*" indicates logical *and* (combinations of conditions); "+" indicates logical *or* (alternate causal conditions or alternate causal combinations); and "--->" indicates the subset relation. The equation states that there are three alternate bases for weak class voting (*W*): weak unions (*u*), a relative lack of affluence (*a*), or the combination of income inequality and lower levels of employment in manufacturing (*I*m*). This solution is dependent on a number of simplifying assumptions (not examined here) because eight of the sixteen truth table rows lack cases. The key difference between the crisp-set solution and the second fuzzy-set solution is that the crisp-set solution adds two new terms: *a* (not highly affluent) and *I*m* (substantial income inequality combined with lower levels of employment in manufacturing). Thus, the crisp-set solution is both more complex and more inclusive than the second fuzzy-set solution.

The two new terms, *a* and *I*M*, appear in the crisp-set solution because of its *lower standard of set-theoretic consistency*. As illustrated in Figure 1, perfect set-theoretic consistency is achieved with fuzzy sets when the cases all reside above the main diagonal of the scatterplot. With crisp sets, however, perfect set-theoretic consistency is much easier to achieve. As long as there are no cases in the fourth quadrant of the fuzzy plot (the lower right quadrant), then the set plotted on the horizontal axis can be described as a subset of the set plotted on the vertical axis. This lower standard defines more cases and thus more causal combinations as consistent. For example, both Norway and Denmark score 0.4 on *a* (not highly affluent, one of the crisp-set solution terms) and 0.2 on *W* (low levels of class voting). In the fuzzy-set analysis, these cases both are *inconsistent* because their score on the causal condition exceed their scores on the outcome. These fuzzy-set inconsistencies directly undermine the argument that *a* is a subset of *W*. From a crisp-set perspective, however, these cases are *consistent* because they display both an absence of affluence (their 0.4 scores are recoded to 0s) and an absence of weak class voting (their 0.2 scores on the outcome are also recoded to 0s). A similar pattern emerges for *I*m*: some of the cases defined as consistent in the crisp-set analysis are inconsistent in the fuzzy-set analysis. Thus, the additional causal terms appearing in the crisp-set solution are due to its lower consistency standard. Given these results, it appears that the practice of dichotomizing fuzzy sets to create crisp sets for crisp-set analysis is not an attractive option.

5. A summary of the procedure

The primary focus of this paper is the process of constructing crisp truth tables from fuzzy-set data. 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 persistently low levels of class voting"). 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). Fuzzy sets are often unimodal at 1.0 or 0.0, or bimodal at both 0.0 and 1.0. In general, calibration requires good grounding in theoretical and substantive knowledge, as well as in-depth understanding of cases. The procedures described in this paper 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 fs/QCA or into a program that can save data files in a format compatible with fs/QCA (e.g., Excel: comma delimited files; SPSS: tab delimited files; variable names must be on the first row). The data set should include both the outcome and as many of the possibly relevant causal conditions as possible. Open the data file using fs/QCA version 1.4. (Click *Help* on the start-up screen to identify fs/QCA 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 fs/QCA, 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 fs/QCA 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 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*). To the left of *number* is *weight*, an indictor of the amount of empirical evidence for each corner, based on the sum of the memberships in the corner. *Weight* correlates very strongly with *number*. Two columns to the right of *number* is *yconsist*, the consistency measure assessing the degree to which membership in that corner is a subset of membership in the outcome. Also reported in the truth table spreadsheet is *nconsist*, a measure of the degree to which each causal combination is a subset of the *negation* of the outcome (1 - each outcome scores). As previously discussed, in fuzzy-set analysis it is possible for a causal combination to be a subset of both the outcome and its negation.

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 *yconsist* 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 *yconsist* 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. Next input 1s and 0s into empty outcome column, which is labeled with the name of the outcome and listed to the left of the *yconsist* 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.[8] If the truth table spreadsheet has many rows, code the outcome column using the *Delete and code* function in the *Edit* menu.

8. After deleting rows that do not meet the selected frequency threshold (step 5) and coding the outcome according to *yconsist* (steps 6 and 7), click *Continue* and a *Truth Table Analysis* window will appear. (In some versions of Microsoft Windows, this window will appear underneath the other windows and must be clicked to the surface.) For this analysis, select the same causal conditions and outcome used to create the truth table (in step 4). In the *Specify* panel set *Positive Cases* (1) to *True* and all the others to *False*. As explained in Ragin and Sonnet (2004), this specification of the analysis yields the most complex solution. To derive its counterpart, the most parsimonious solution, repeat the procedure just described, only this time around set *Remainders* to *Don't Cares* in the *Specify* panel of the procedure. Conceive of the most complex and the most parsimonious solutions as the two endpoints of a single complexity/parsimony continuum. 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 intermediate solution.

The various procedures sketched in this paper 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 paper, 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. The choice as to which is "best" may be decided, in the end, only by the cases. The ultimate goal of this paper 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.

---

[8] When the values of *yconsist* and *nconsist* are both close to 1.0, it is possible to code the outcome with a dash (-) to indicate either "1 or 0." These rows are known as *Don't Cares* and may be manipulated independently by the analyst in the *Specify* panel of the truth table procedure.

Table 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": $.5 < x_i < 1$
		.75 = more in than out		
			.7 = more or less in	
	.5 = neither fully in nor fully out			.5 = cross-over: neither in nor out
			.3 = more or less out	
		.25 = more out than in		
			.1 = mostly but not fully out	Degree of membership is more "out" than "in": $0 < x_i < .5$
0 = fully out		0 = fully out	0 = fully out	
	0 = fully out			0 = fully out

Table 2: Fuzzy-set data on class voting in the advanced industrial societies

Country	Weak Class voting (W)	Affluent (A)	Income Inequality (I)	Manufacturing (M)	Strong Unions (U)
Australia	0.7	0.9	0.7	0.3	0.7
Belgium	0.7	0.7	0.1	0.1	0.9
Denmark	0.1	0.7	0.3	0.1	0.9
France	0.9	0.7	0.9	0.1	0.1
Germany	0.6	0.7	0.9	0.3	0.3
Ireland	0.9	0.1	0.7	0.9	0.7
Italy	0.6	0.3	0.9	0.1	0.7
Netherlands	0.9	0.7	0.3	0.1	0.3
Norway	0.1	0.7	0.3	0.7	0.9
Sweden	0.0	0.9	0.3	0.9	1.0
United Kingdom	0.3	0.7	0.7	0.9	0.7
United States	1.0	1.0	0.9	0.3	0.1

Table 3: Illustration of logical *and* using data from Table 2

Country	Affluent (A)		Income Inequality (I)		Manufacturing (M)		Strong Unions (U)		Affluent* Income Inequality	Affluent* Income Inequality* Weak Unions
	A	a	I	i	M	m	U	u	A*I	A*I*u
Australia	0.9	0.1	0.7	0.3	0.3	0.7	0.7	0.3	0.7	0.3
Belgium	0.7	0.3	0.1	0.8	0.1	0.9	0.9	0.1	0.1	0.1
Denmark	0.7	0.3	0.3	0.7	0.1	0.9	0.9	0.1	0.3	0.1
France	0.7	0.3	0.9	0.1	0.1	0.9	0.1	0.9	0.7	0.7
Germany	0.7	0.3	0.9	0.1	0.3	0.7	0.3	0.7	0.7	0.7
Ireland	0.1	0.9	0.7	0.3	0.9	0.1	0.7	0.3	0.1	0.1
Italy	0.3	0.7	0.9	0.1	0.1	0.9	0.7	0.3	0.3	0.3
Netherlands	0.7	0.3	0.3	0.7	0.1	0.9	0.3	0.7	0.3	0.3
Norway	0.7	0.3	0.3	0.7	0.7	0.3	0.9	0.1	0.3	0.1
Sweden	0.9	0.1	0.3	0.7	0.9	0.1	1.0	0.0	0.3	0.0
United Kingdom	0.7	0.3	0.7	0.3	0.9	0.1	0.7	0.3	0.7	0.3
United States	1.0	0.0	0.9	0.1	0.3	0.7	0.1	0.9	0.9	0.9

Table 4: Illustration of logical *or* using data from Table 2

Country	Affluent (A)		Income Inequality (I)		Manufacturing (M)		Strong Unions (U)		Manufacturing + Strong Unions	Low Inequality * (Manufacturing + Strong Unions)
	A	a	I	i	M	m	U	u	M+U	i*(M+U)
Australia	0.9	0.1	0.7	0.3	0.3	0.7	0.7	0.3	0.7	0.3
Belgium	0.7	0.3	0.1	0.9	0.1	0.9	0.9	0.1	0.9	0.9
Denmark	0.7	0.3	0.3	0.7	0.1	0.9	0.9	0.1	0.9	0.7
France	0.7	0.3	0.9	0.1	0.1	0.9	0.1	0.9	0.1	0.1
Germany	0.7	0.3	0.9	0.1	0.3	0.7	0.3	0.7	0.3	0.1
Ireland	0.1	0.9	0.7	0.3	0.9	0.1	0.7	0.3	0.9	0.3
Italy	0.3	0.7	0.9	0.1	0.1	0.9	0.7	0.3	0.7	0.1
Netherlands	0.7	0.3	0.3	0.7	0.1	0.9	0.3	0.7	0.3	0.3
Norway	0.7	0.3	0.3	0.7	0.7	0.3	0.9	0.1	0.9	0.7
Sweden	0.9	0.1	0.3	0.7	0.9	0.1	1.0	0.0	1.0	0.7
United Kingdom	0.7	0.3	0.7	0.3	0.9	0.1	0.7	0.3	0.9	0.3
United States	1.0	0.0	0.9	0.1	0.3	0.7	0.1	0.9	0.3	0.1

Table 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 6: Memberships of countries in causal combinations, illustrated with three conditions

Country	Income Inequality		Manufacturing		Strong Unions									
	I	i	M	m	U	u	i*m*u	i*m*U	i*M*u	i*M*U	I*m*u	I*m*U	I*M*u	I*M*U
Australia	0.7	0.3	0.3	0.7	0.7	0.3	0.3	0.3	0.3	0.3	0.3	0.7	0.3	0.3
Belgium	0.1	0.9	0.1	0.9	0.9	0.1	0.1	0.9	0.1	0.1	0.1	0.1	0.1	0.1
Denmark	0.3	0.7	0.1	0.9	0.9	0.1	0.1	0.7	0.1	0.1	0.1	0.3	0.1	0.1
France	0.9	0.1	0.1	0.9	0.1	0.9	0.1	0.1	0.1	0.1	0.9	0.1	0.1	0.1
Germany	0.9	0.1	0.3	0.7	0.3	0.7	0.1	0.1	0.1	0.1	0.7	0.3	0.3	0.3
Ireland	0.7	0.3	0.9	0.1	0.7	0.3	0.1	0.1	0.3	0.3	0.1	0.1	0.3	0.7
Italy	0.9	0.1	0.1	0.9	0.7	0.3	0.1	0.1	0.1	0.1	0.3	0.7	0.1	0.1
Netherlands	0.3	0.7	0.1	0.9	0.3	0.7	0.7	0.3	0.1	0.1	0.3	0.3	0.1	0.1
Norway	0.3	0.7	0.7	0.3	0.9	0.1	0.1	0.3	0.1	0.7	0.1	0.3	0.1	0.3
Sweden	0.3	0.7	0.9	0.1	1.0	0.0	0.0	0.1	0.0	0.7	0.0	0.1	0.0	0.3
United Kingdom	0.7	0.3	0.9	0.1	0.7	0.3	0.1	0.1	0.3	0.3	0.1	0.1	0.3	0.7
United States	0.9	0.1	0.3	0.7	0.1	0.9	0.1	0.1	0.1	0.1	0.7	0.1	0.3	0.1

Table 7: Illustration of fuzzy subset relation

Religious heterogeneity	Weak class voting
0.7	0.9
0.1	0.9
0.1	0.1
0.3	0.3
0.9	0.9
0.7	0.7
0.3	0.9
0.3	0.7
0.3	0.7
0.1	0.1
0.0	0.0
0.9	1.0

Table 8: Assessing the distribution of cases across combinations of causal conditions

	a.i.m.u	a.i.m.U	a.i.M.u	a.i.M.U	a.I.m.u	a.I.m.U	a.I.M.u	a.I.M.U	A.i.m.u	A.i.m.U	A.i.M.u	A.i.M.U	A.I.m.u	A.I.m.U	A.I.M.u	A.I.M.U
1.Australia	.100	.100	.100	.100	.100	.100	.100	.100	.300	.300	.300	.300	.300	.700	.300	.300
2.Belgium	.100	.300	.100	.100	.100	.100	.100	.100	.100	.700	.100	.100	.100	.100	.100	.100
3.Denmark	.100	.300	.100	.100	.100	.300	.100	.100	.100	.700	.100	.100	.100	.300	.100	.100
4.France	.100	.100	.100	.100	.300	.100	.100	.100	.100	.100	.100	.100	.700	.100	.100	.100
5.Germany	.100	.100	.100	.100	.300	.300	.300	.300	.100	.100	.100	.100	.700	.300	.300	.300
6.Ireland	.100	.100	.300	.300	.100	.100	.300	.700	.100	.100	.100	.100	.100	.100	.100	.100
7.Italy	.100	.100	.100	.100	.300	.700	.100	.100	.100	.100	.100	.100	.300	.300	.100	.100
8.Netherlands	.300	.300	.100	.100	.300	.300	.100	.100	.700	.300	.100	.100	.300	.300	.100	.100
9.Norway	.100	.300	.100	.300	.100	.300	.100	.300	.100	.300	.100	.700	.100	.300	.100	.300
10.Sweden	.000	.100	.000	.100	.000	.100	.000	.100	.000	.100	.000	.700	.000	.100	.000	.300
11.United Kingdom	.100	.100	.300	.300	.100	.100	.300	.300	.100	.100	.300	.300	.100	.100	.300	.700
12.United States	.000	.000	.000	.000	.000	.000	.000	.000	.100	.100	.100	.100	.700	.100	.300	.100
13. Number: score>.5	0	0	0	0	0	1	0	1	1	2	0	2	3	1	0	1

Table 9: Assessing the consistency of causal combinations with the fuzzy subset relation

Affluence	Inequality	Manufacturing	Unions	Consistency	Outcome
1	0	0	0	1.00	1
1	1	0	0	0.97	1
0	1	1	1	0.87	0
1	1	0	1	0.82	0
0	1	0	1	0.76	0
1	0	0	1	0.70	0
1	1	1	1	0.65	0
1	0	1	1	0.54	0

Table 10: Crisp-set data on class voting in the advanced industrial societies

Country	Weak Class voting (W)	Affluent (A)	Income Inequality (I)	Manufacturing (M)	Strong Unions (U)
Australia	1	1	1	0	1
Belgium	1	1	0	0	1
Denmark	0	1	0	0	1
France	1	1	1	0	0
Germany	1	1	1	0	0
Ireland	1	0	1	1	1
Italy	1	0	1	0	1
Netherlands	1	1	0	0	0
Norway	0	1	0	1	1
Sweden	0	1	0	1	1
United Kingdom	0	1	1	1	1
United States	1	1	1	0	0

Figure 1: Scatterplot Showing Subset Relationship Using Hypothetical Data