QCA Research Notes

Charles C. Ragin
(cragin@uci.edu)

Version: April 2024

1. Stepwise Consistency Analysis

Charles C. Ragin (cragin@uci.edu)

Background

Consistency scores are fundamental to the practice of Qualitative Comparative Analysis (Ragin 2006). They are used most often to gauge the degree to which a condition or combination of conditions constitutes a subset of an outcome. If a rough subset relation can be established, the researcher may interpret this connection as suggestive of a sufficiency relationship between the condition or combination of conditions and the outcome. In other words, the subset relation indicates that the cases sharing a specified condition or combination of conditions also share an outcome (Ragin 2008). For example, a researcher might assess whether having a strong left party is sufficient for the formation of a generous welfare state. If countries with strong left parties consistently established generous welfare states, the evidence is supportive of a sufficiency relationship.

With crisp sets, the calculation of subset consistency is straightforward. In the example just presented, it's simply the number of cases combining strong left parties and generous welfare states divided by the number of cases with strong left parties. With fuzzy sets (Ragin 2000; 2008), the calculation is:

consistency of $X \le Y = \text{sum}(\min(X_i,Y_i)) / \text{sum}(X_i)$

where X_i is degree of membership in the condition or combination of conditions, and Y_i is degree of membership in the outcome. The numerator assesses the intersection of sets X and Y; the denominator allows expression of the intersection relative to the sum of the memberships in set X. The resulting consistency score indicates the proportion of set X that is contained within set Y. Consistency scores close to 1.0 provide clear evidence of a subset relationship, while scores less than 0.70 indicate substantial inconsistency. In general, if there are many cases with high membership in X coupled with low membership in Y, subset consistency scores are diminished accordingly.

Subset consistency scores are central to truth table analysis. The standard practice is to calculate the degree to which each combination of conditions, as specified by truth table rows, is a consistent subset of the outcome. Most applications of QCA use a consistency *threshold* to identify combinations of conditions with strong ties to the outcome. For example, a researcher might code truth table rows with consistency scores of 0.80 or greater as having strong ties to the outcome. Such rows would be coded "true" (1) on the outcome; rows with less than 0.80 consistency would be coded "false" (0).[1]

To illustrate, consider a truth table using data from the National Longitudinal Survey of Youth (NLSY79). As shown in table 1, the three conditions are race (white = 1; black = 0), gender (male = 1; female = 0), and class background (advantaged versus not-advantaged). Race and gender are conventional binary sets. Class background is operationalized as a fuzzy set based on the intersection of two other fuzzy sets, *educated-parent* and *not-low-parental-income*. To receive a high membership score in *advantaged-class-background*, a respondent's parent must be both educated and not low-income. The outcome is also a fuzzy set, degree of membership

[1] In practice, the coding of row outcomes as true or false comes after the deletion of truth table rows that do not meet whatever row frequency threshold the researcher has specified.

in the set of respondents *avoiding-poverty*. I describe the construction and calibration of the fuzzy sets in an appendix (see also Ragin and Fiss 2017).

Table 1: Impact of Race, Gender, and Advantaged Class Background on Poverty Avoidance

Row	**Race: white = 1 black = 0**	**Gender: male = 1 female = 0**	**Class: Membership in Advantaged (1) or ~Advantaged (0)**	**Number of respondents**	**Consistency of poverty avoidance**
1	1	1	1	1045	0.884
2	1	0	1	986	0.857
3	0	1	1	235	0.794
4	1	1	0	315	0.786
5	1	0	0	327	0.750
6	0	0	1	212	0.745
7	0	1	0	497	0.587
8	0	0	0	562	0.421

Three conditions yield a truth table with eight rows ($2^3 = 8$). The degree of membership of each case in each row is determined by the minimum of its memberships in the sets that make up the row. For example, a black female with a 0.3 membership in 'advantaged' has a membership of 0.3 in the row that combines black, female, and advantaged. This same case has a membership of 0.7 in the row that combines black, female, and not-advantaged (membership in not-advantaged = 1 – membership in advantaged), and it has a membership of 0 in the other six rows because black females have 0 membership in white and also 0 membership in male. Table 1 also reports the number of respondents in each row, with advantaged class background dichotomized at 0.5 (the cross-over point for fuzzy sets) to simplify the presentation.

Consistency of poverty avoidance is a set-analytic measure gauging the degree to which respondents with the combination of attributes in each row constitute a subset of respondents avoiding poverty (see Ragin 2008: chapter 3). In other words, it is an assessment of the degree to which the respondents in each row share the outcome in question—avoiding poverty. The consistency of poverty avoidance scores shown in the last column of table 1 are sorted from high to low, and range from 0.884 for *white•male•advantaged* to 0.421 for *black•female•~advantaged*.[2] As mentioned previously, it is standard practice for QCA users to establish a consistency threshold in order to distinguish combinations of conditions that are strongly linked to an outcome versus those that are not. Inspection of table 1 reveals a substantial gap in consistency scores between rows 6 and 7. Thus, one plausible consistency threshold for coding the outcome "true" (1) is 0.745. Using this threshold, the first six rows of table 1 are coded 1 on the outcome, while the bottom two are coded 0. Logical minimization of the coded truth table yields the following result:

avoidance of poverty ≥ white + advantaged

[2] Note: "~" indicates negation ("not-"); "•" indicates combined conditions (set intersection—logical and); "+" indicates alternate conditions or alternate combinations of conditions (set union—logical or).

This truth table solution states that respondents who are white or from advantaged class background are able to avoid poverty with at least moderate consistency (i.e., with scores of 0.745 or better). This set-analytic statement can be "clarified" (Ragin 2023:59-61) and rewritten as:

avoidance of poverty ≥ white + black•advantaged

by assigning the logical overlap between the two sets to "*white*." This restatement of the results of the truth table analysis makes it clear that all whites are able to avoid poverty with a consistency of 0.745 or better, but only blacks from advantaged class backgrounds experience comparable levels of insulation from poverty.

Stepwise Consistency Analysis

This research note presents an alternate, "stepwise" approach to the analysis of truth table consistency scores. In simple terms, the stepwise approach examines the consequences of lowering the consistency threshold row-by row beginning with the row with the highest consistency score. Table 2 illustrates the procedure. The initial consistency threshold is 0.884, row 1's consistency score (*white•male•advantaged*), which is the only row that is coded 1 on the outcome on the first step. The second threshold is pegged to row 2's consistency score, which lowers the consistency threshold to 0.857 and codes the first two rows to 1 on the outcome. The third threshold is pegged to row 3's consistency score, which lowers the threshold to 0.794 and codes the top three rows to 1 on the outcome, and so on through row 6, the last row with an acceptable consistency threshold. The last column of table 2 reports set-theoretic statements summarizing the stepwise results. The first cell in this column summarizes the first row results. The second cell summarizes the results of the union of the first two rows. The third cell summarizes the union of the first three rows, and so on through the remainder of the truth table. For example, the summary statement reported for the second row, *white•advantaged*, is based on the union of *white•male•advantaged* and *white•female•advantaged*. By progressively including more rows in a top-down, stepwise fashion, the scope of the conditions linked to poverty avoidance is gradually and methodically expanded, based on the rows with acceptable consistency scores.

Table 2: Stepwise Truth Table Results

Step	Conditions	Consistency	Cumulative Results (minimized)
1	white•male•advantaged	0.884	white•male•advantaged
2	white•female•advantaged	0.857	white•advantaged
3	black•male•advantaged	0.794	white•advantaged + male•advantaged
4	white•male•~advantaged	0.786	white•advantaged + male•advantaged + white•male
5	white•female•~advantaged	0.750	white + male•advantaged
6	black•female•advantaged	0.745	white + advantaged
7	black•male•~advantaged	0.587	white + advantaged + male
8	black•female•~advantaged	0.421	(all eight combinations are coded 1)

Stepwise consistency analysis provides a basis for interpreting the entire truth table, thereby avoiding the limitations that accompany focusing on a single outcome coding. For example, the first step confirms that the intersection of three favorable conditions (gender = male, race = white, and class = advantaged) offers the best insulation from poverty (see table 2). The second step indicates that for advantaged whites, gender is not a major factor when it comes to avoiding poverty. The third step extends the second step to embrace advantaged black males, yielding *male•advantaged* as a key combination of conditions. (The combination *black•female•advantaged* is not incorporated until step 6.) Step four offers the first appearance of not-advantaged as a condition, which yields the combination *white•male* as a recipe for avoiding poverty. It is noteworthy that white males are the first subpopulation to avoid poverty despite a lack of class advantage. Step 5 adds not-advantaged white females to the mix which in turn sets the stage for the debut of a single-ingredient recipe, *white*. Finally, step six incorporates the fourth combination that includes advantaged class background (*black•female•advantaged)*, which in turn permits the derivation of the second single-condition recipe, *advantaged*. Overall, the stepwise results (1) reinforce the importance of race and class, especially the benefits enjoyed by white males, and (2) highlight the disadvantages endured by black females.

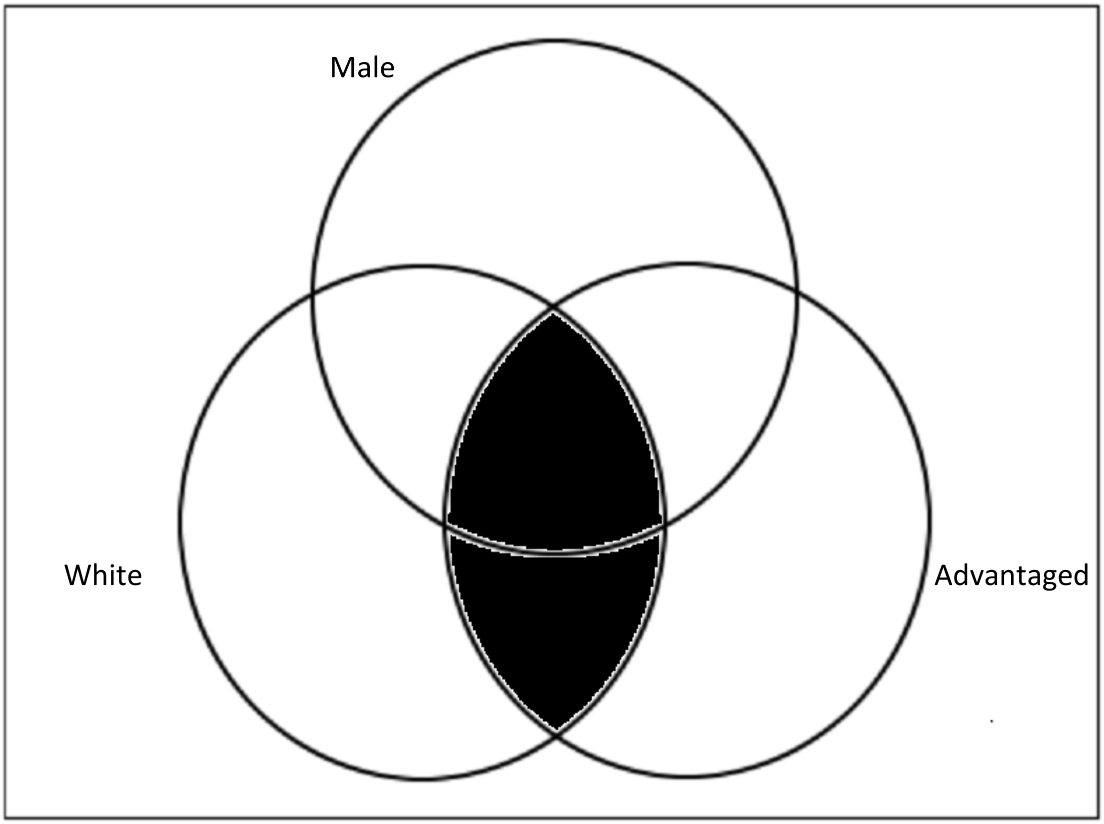
It is important to recognize that the truth table solutions listed in the last column of table 2 exhibit subset/superset relationships. The step 1 solution is a subset of the step 2 solution, which in turn is a subset of the step 3 solution, and so on. The subset relations are most visible when represented using Venn diagrams. Compare figures 1, 2 and 3. Figure 1 models step 2's solution—the intersection of white and advantaged. Figure 2 models step 4's solution and illustrates the three two-way set intersections that are linked to the avoidance of poverty. Figure 3 models step 6's solution—the union of white and advantaged.

This nuanced analysis of consistency scores is especially appropriate in situations where there is no obvious threshold value. For example, sometimes there is an array of consistency scores in the 0.75 to 0.90 range, offering several plausible threshold values. Rather than focusing on a single, possibly arbitrary threshold value, it is preferable to conduct a stepwise examination of the truth table, using multiple, rank-ordered consistency thresholds. Stepwise analysis also protects researchers from the charge that they have selected a threshold value in an opportunistic manner, in order to favor a specific interpretation of the evidence.

Limitations

Many, if not most, truth tables are straightforward and can be deciphered adequately using a single threshold value. Some truth tables, for example, have only a single row with a high consistency score. Stepwise consistency analysis is useful only when there are at least two rows with high consistency scores. Furthermore, the stepwise approach is likely to be less useful when Ns are small or when most truth table rows have few or no cases. The ideal context for the stepwise procedure is (1) a moderate-to-large number of cases, (2) a relatively well-populated truth table, and (3) several plausible consistency threshold values.

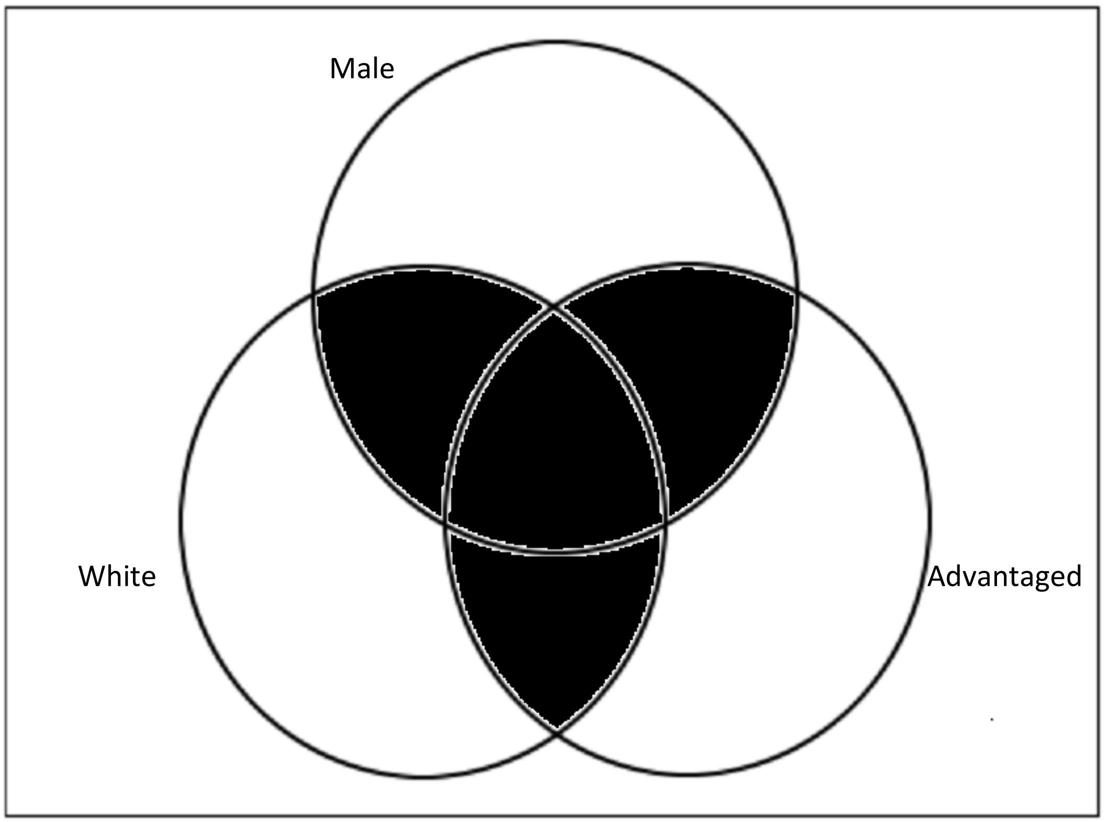
Figure 1: Venn Diagram for Truth Table Solution at Step 2.



avoidance of poverty ≥ white•advantaged

consistency = 0.857

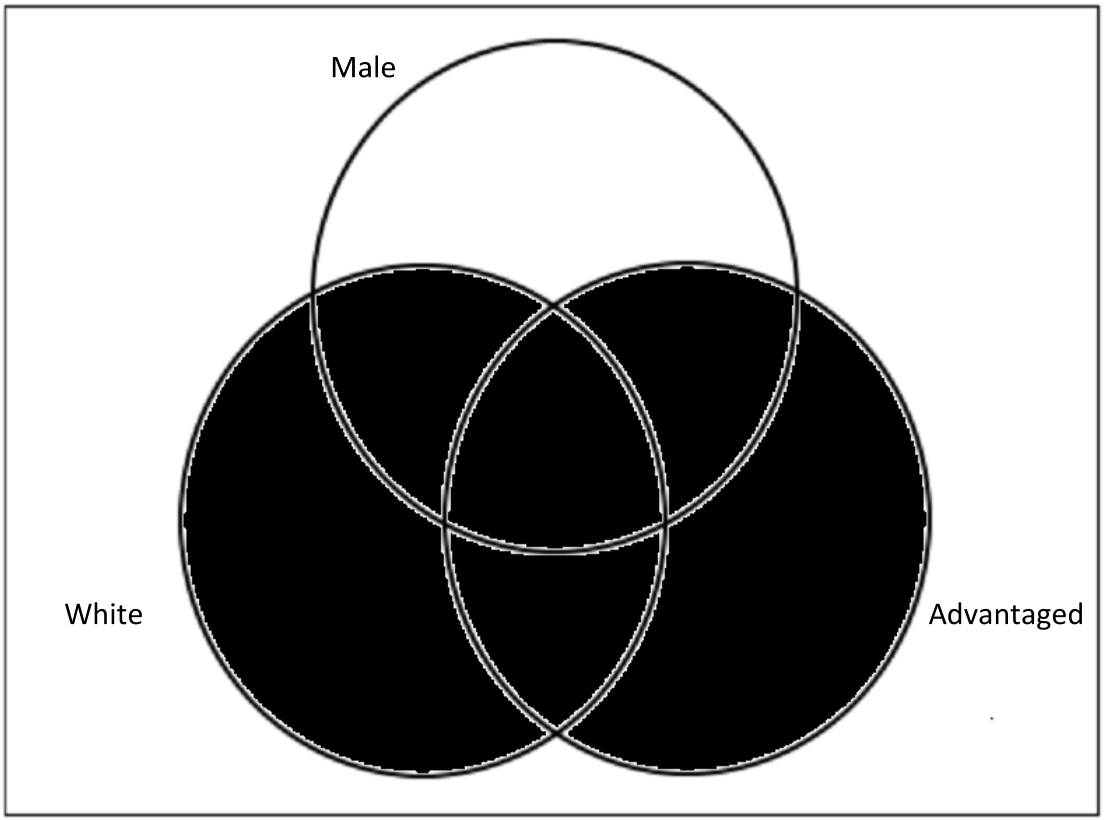
Figure 2: Venn Diagram for Truth Table Solution at Step 4.



avoidance of poverty ≥ white•advantaged + male•advantaged + white•male

consistency = 0.786

Figure 3: Venn Diagram for Truth Table Solution at Step 6.



avoidance of poverty ≥ white + advantaged

consistency = 0.745

Appendix: Calculating Fuzzy Set Membership Scores

Avoiding Poverty. To construct the fuzzy set of individuals avoiding poverty, I first calibrate respondents' degree of membership in poverty. I base the measurement on the official poverty threshold adjusted for household size and composition. The official poverty threshold is an absolute threshold, meaning it was fixed at one point in time and is updated solely for price changes.

The measure of poverty is based on the ratio of household income to the poverty level for that household. Using the direct method for calibrating fuzzy sets (see Ragin 2008, chapter 5), the threshold for full membership in the set of households *in-poverty* (fuzzy membership score = 0.95) is a ratio of 1.0 (household income is the same as the poverty level); the cross-over point (fuzzy membership score = 0.5) is a ratio of 2.0 (household income is double the poverty level); and the threshold for full exclusion from the set of households *in-poverty* (fuzzy membership score = 0.05) is a ratio of 3.0 (household income is three times the poverty level for that household).

The fuzzy set of households in poverty is a symmetric set; that is, it is truncated at both ends and the crossover point is set exactly at the halfway mark between the thresholds for membership and non-membership. Thus, the set of respondents *avoiding-poverty* is based on a straightforward negation of the set of respondents *in-poverty*. With fuzzy sets, negation is accomplished simply by subtracting membership scores from 1.0 (see Ragin 2008, chapter 2). That is, (membership in *avoiding-poverty*) = 1 − (membership *in-poverty*). For example, a case that is mostly but not fully *in* the set of respondents *in-poverty*, with a score of 0.90, is mostly but not fully *out* of the set of respondents *avoiding-poverty*, with a score of 0.10.

The use of a ratio of three times the poverty level for full membership in the set of cases avoiding poverty is a conservative cutoff value, but also one that is anchored in substantive knowledge regarding what it means to be out of poverty. For example, in 1989, the weighted average poverty threshold for a family of two adults and two children was about $12,500 (Social Security Administration 1998, table 3.E). Three times this poverty level corresponds to $37,500 for a family of four, a value that lies just slightly above the median family income of $35,353 in 1990 (U.S. Census Bureau, Historical Income Tables—Families, table F-7).

Advantaged Class Background. To assess degree of membership in *advantaged-class-background*, I combine two fuzzy sets, *not-low-income parents* and *educated-parents*. To obtain greater than 0.5 membership (i.e., more in than out) in *advantaged-class-background*, respondents must have greater than 0.5 membership in both *not-low-income-parents* and *educated-parents*. I describe the calibration of *not-low-income-parents* first.

I assess parental income by first computing the ratio of parental income to the household-adjusted poverty level for the parents' household. The numerator of this measure is based on the average of the reported 1978 and 1979 total net family income in 1990 dollars. The denominator is the household-adjusted poverty level for that household. The fuzzy set of respondents with not-low-income parents is similar in its construction to the fuzzy set of households avoiding poverty, described previously. That is, I first calculate the ratio of parents' household income to the poverty level, using NLSY data on the official poverty threshold in 1979, adjusted for household size and composition. Using the direct method of calibration, the threshold for full membership in the set with not-low parental income (.95) is a ratio of 5.5

(parents' income was five and a half times the poverty level). Respondents with ratios greater than 5.5 times receive fuzzy scores between 0.95 and 1.0. Conversely, the threshold for full exclusion (.05) from the set with not-low parental income is a ratio of 2 (parents' household income was only double the poverty level). Respondents with ratios less than 2 received fuzzy scores between 0.05 and 0. The cross-over point (membership = 0.50) is pegged at three times the household-adjusted poverty level.

I use the greater of mother's and father's years of education to assess parental education. In order to convert this variable to a fuzzy set, I use the following benchmarks: Parents with 12 or more years of schooling are more in than out the set of *educated-parents* (fuzzy score > 0.5). Those with fewer than 9 years of education are treated as fully out of the set (fuzzy score of 0), and those with 16 or more years of education are treated as fully in the set. The specific translation scheme is 0-8 years = 0; 9 years = 0.1; 10 years = 0.2; 11 years = 0.4; 12 years = 0.6; 13 years = 0.7; 14 years = 0.8; 15 years = 0.9; 16+ years = 1.0.

These two fuzzy sets, *not-low-income-parents* and *educated-parents* are joined by logical *and* (set intersection) to create *advantaged-class-background*. To implement logical *and* with fuzzy sets, it is necessary simply to take the lower of the two membership scores.

2. Dual Calibration

Charles C. Ragin (cragin@uci.edu)

Background

The first step in any set-analytic investigation, using either crisp or fuzzy sets, is the specification and construction of the relevant sets, including both the causal conditions and the outcome. Three main phases in this process are: (1) the identification of relevant conditions and outcomes, (2) the conceptualization of these conditions and outcomes as sets, and (3) the assignment of membership scores. With crisp sets, membership scores are either “1” (in the set) or “0” (out of the set). With fuzzy sets, scores range from 0 to 1, indicating the degree of membership of each case in a given set. When interval-scale variables are used as the basis for constructing fuzzy sets, the third phase involves specifying empirical benchmarks to structure the translation of variable scores to fuzzy membership scores (see Ragin 2008, chapter 5).

The translation of variables to sets involves an important reconceptualization of the underlying constructs. For example, “income” is easy to understand as a variable. However, in its “raw” form, it makes no sense as a set. The key difference between variables and fuzzy sets parallels the distinction between abstract nouns and adjectives. For example, “income” is an abstract noun, and abstract nouns make good variable names. Adding an adjective to create “high income,” by contrast, distinguishes a specific category or range of values and thus is a good starting point for constructing a fuzzy set. Another example: “education,” an abstract noun, describes a variable; “college-educated,” an adjective, describes a set.

The “variable/abstract noun” versus “set/adjective” distinction offers a good basis for understanding the second phase of calibration—the conceptualization of conditions and outcomes as sets. The set labels selected by the investigator should describe some case aspect that can be used as a basis for distinguishing cases in a qualitative manner (e.g., “*high* income”). With crisp sets, the assignment of cases to sets is all or nothing—in or out. Fuzzy sets, by contrast, allow partial membership in qualitative states and are thus simultaneously qualitative and quantitative (Ragin 2000: 153-155).

To calibrate interval-scale variables as fuzzy sets, most QCA researchers use the “direct method” of calibration discussed in chapter 5 of Ragin (2008). This method is based on the specification of selected interval variable values as indicating: (1) the threshold for full membership in the target set, which is translated to a fuzzy-set membership score of 0.95; (2) the threshold for full non-membership in the target set, which is translated to a fuzzy-set membership score of 0.05; and (3) the cross-over point, which is the dividing line between being “more in” versus “more out” of the target set, which is a fuzzy-set membership score of 0.50. The end result is typically an S-shaped curve, with low scores on the interval-scale variable approaching a score of 0 on the fuzzy set and high scores on the interval-scale variable approaching 1.0. The calibration procedure is automated in the software package fuzzy-set Qualitative Comparative Analysis (fsQCA) and is based on the user’s specification of the three benchmark values just described.

Both crisp and fuzzy sets can be negated simply by subtracting set membership scores from 1. For example, suppose democratic countries are coded 1 and not-democratic countries are coded 0. Subtracting these two values from 1 reverses their scores. Likewise for fuzzy sets, a case with a membership score of 0.8 in “high income” has a membership score of 1 – 0.8 = 0.2 in “not-high income." It is important to be consistent in the labelling of the negated set; it should be labelled in a way that makes it clear that it is the negation of another set. For example, the negation of degree of membership in "high income" is *not* degree of membership in "low income." Degree of membership in "low income" should be conceptualized separately from the conceptualization of "high income." Too often, QCA researchers fail to recognize the importance of consistent conceptualization and labelling of negated sets. While it might seem appropriate for the negation of "high income" to be labelled "low income," there is an important difference between "not-high" and "low." A middle-income person with low membership in "high income" can also have low membership in "low income."

This research note introduces the method of *dual calibration*, along with a demonstration of the use of fsQCA software to implement the procedure. The method allows simultaneous calibration of "high-X versus not-high-X" and "low-X versus not-low-X" using four benchmark values instead of three. Using three, which is the usual practice, yields a single fuzzy set ranging from high-X to not-high-X. Using four benchmarks yields two fuzzy sets, one ranging from high-X to not-high-X and the other ranging from low-X to not-low-X.

Dual Calibration

It is often necessary to create more than one fuzzy set from a single source variable. Consider, for example, the causal conditions linked to poverty versus the causal conditions linked to having a well-paying job. Years of education has an impact on both outcomes. When thinking in terms of variables, and not sets, it would be routine to examine the correlation between years of education and these two outcomes. Education has a negative correlation with poverty and a positive correlation with having a well-paying job. When thinking in terms of sets, however, it is necessary to abandon abstract nouns like "education" and think in terms of sets and the adjectives that describe them. It is the "set of low-education respondents" who tend to end up in poverty and the set of "high-education respondents" who tend to end up with well-paying jobs. To examine these connections with set-analytic methods, it is necessary to calibrate education in two different ways, focusing on different ends of the distribution.

Consider the translation of years of education to the fuzzy set of respondents with low education and its negation, the set of respondents with not-low education:

	Membership in low education	**Membership in not-low education**
4 years	0.99	0.01
6 years	0.95	0.05
10 years	0.75	0.25
12 years	0.50	0.50
14 years	0.25	0.75
16 years	0.05	0.95
18 years	0.02	0.98
20 years	0.01	0.99

Using fsQCA's calibration procedure, to translate years of education into the fuzzy set of respondents with low education, the following command is executed:
compute lowed = calibrate(education,6,12,16)
where *education* is a variable in the dataset, measured in years, and *lowed* is the newly created fuzzy set. The first listed number (6) is the benchmark value for a fuzzy membership of 0.95; the second listed number (12) corresponds to the cross-over fuzzy score (0.5); the third listed number (16) corresponds to a fuzzy membership score of 0.05. The calculation of membership in the set of respondents with not-low education is simply:
compute notlowed = 1 – lowed

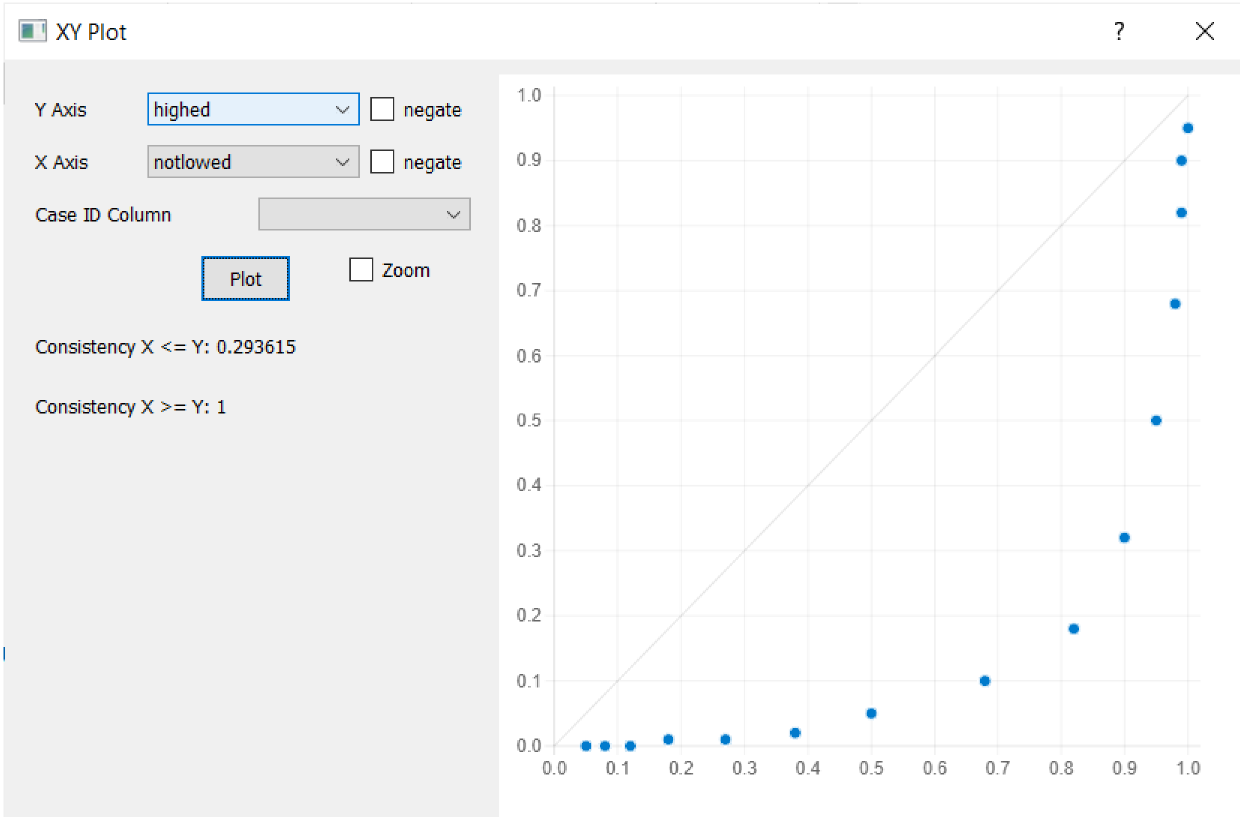
Next, consider the translation of years of education into the fuzzy set of respondents with high education and its negation, the set of respondents with not-high education

	Membership in high education	**Membership in not-high education**
4 years	0.01	0.99
6 years	0.01	0.99
10 years	0.02	0.98
12 years	0.05	0.95
14 years	0.25	0.75
16 years	0.50	0.50
18 years	0.75	0.25
20 years	0.95	0.05

To translate years of education to the fuzzy set of respondents with high education, the following command is executed in fsQCA:
compute highed = calibrate(education,20,16,12)
where education is a variable in the dataset and *highed* is the newly created fuzzy set. The first listed number (20) is the benchmark value for a fuzzy membership score of 0.95; the second listed number (16) corresponds to the cross-over fuzzy score (0.5); the third listed number (12) corresponds to a fuzzy membership score of 0.05. The calculation of membership in the set of respondents with not-high education is simply:
compute nothighed = 1 – highed

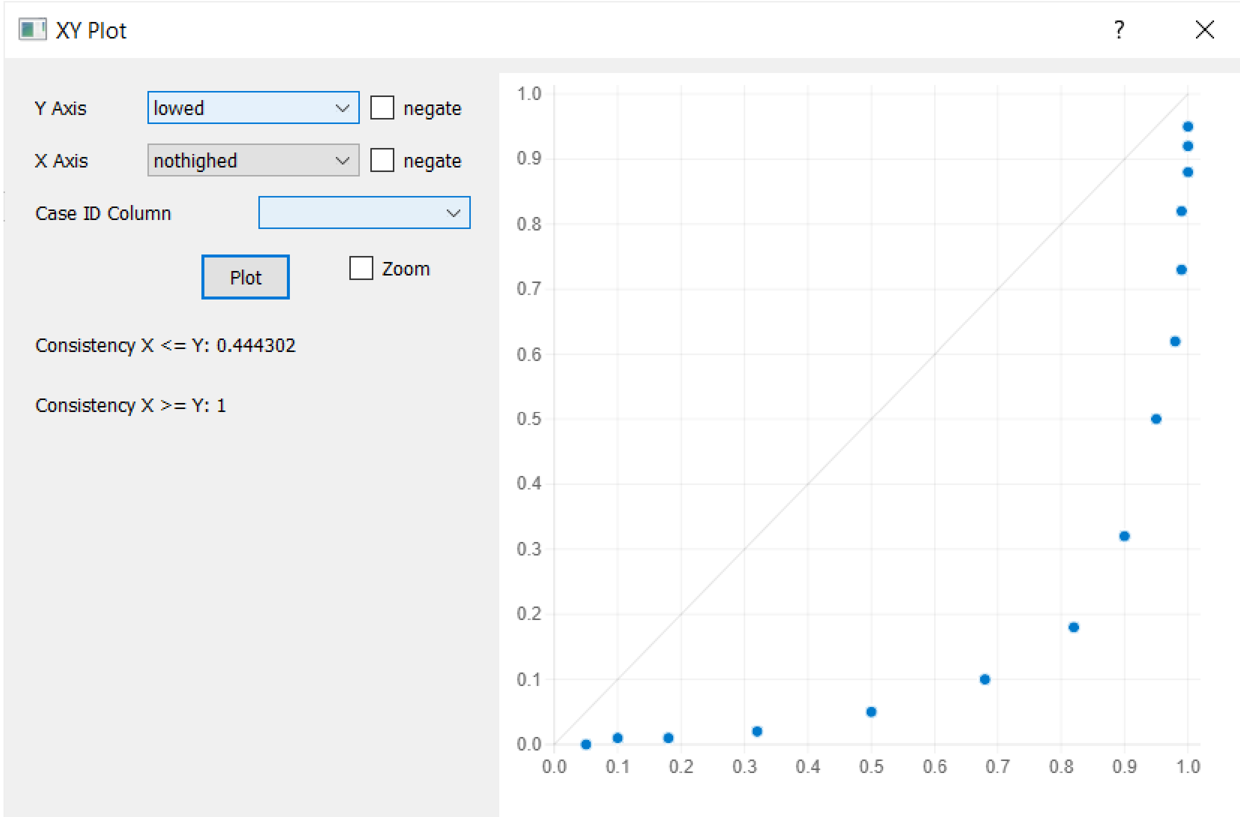
The difference between the two calibrations can be clearly seen in the plot of *highed* against *notlowed* and in the plot of *lowed* against *nothighed*. First, consider *highed* against *notlowed*, shown in Figure 1. The difference is striking. Moderately high membership scores in *notlowed* (e.g., in the 0.7 to 0.9 range) are linked to low membership scores in *highed* (e.g., scores in the 0.1 to 0.3 range). The two calibrations converge only when both membership scores are very close to 0 or to 1.

Figure 1: Plot of *highed* against *notlowed.*



Next, consider the plot of *lowed* against *nothighed*, shown in Figure 2. The pattern is virtually identical to the first plot. Again, scores in the 0.7 to 0.9 range of *nothighed* are paired with scores in the 0.1 to 0.3 range of *lowed,* and the two calibrations converge only when both membership scores are very close to 1 or to 0. Notice also that in both plots the superset consistency calculation ($X_i \geq Y_i$) is equal to 1, indicating that the Y axis values are uniformly less than or equal to the X axis values. Thus, *highed* is a subset of *notlowed*, and *lowed* is a subset of *nothighed*, which is consistent with common sense. For example, it is easier to achieve not-low education than it is to achieve high education.

Figure 2: Plot of *lowed* against *nothighed*.



Dual Calibration Using fsQCA

While it is perfectly acceptable to calibrate high and low using separate commands, as in the example just provided, dual calibration is automated in fsQCA. A single command generates two calibrated sets, one with the suffix HI and another with the suffix LO. For example, suppose a researcher's goal is to compare the impact of high education with the impact of not-low education, a more inclusive set, on an outcome. She could use the "dualcal" command to generate *educHI* and *educLO* from the source variable, years of education. The researcher supplies the source variable name, four benchmarks values, and the stem name for the two fuzzy sets (*educ*, in this example).[3]

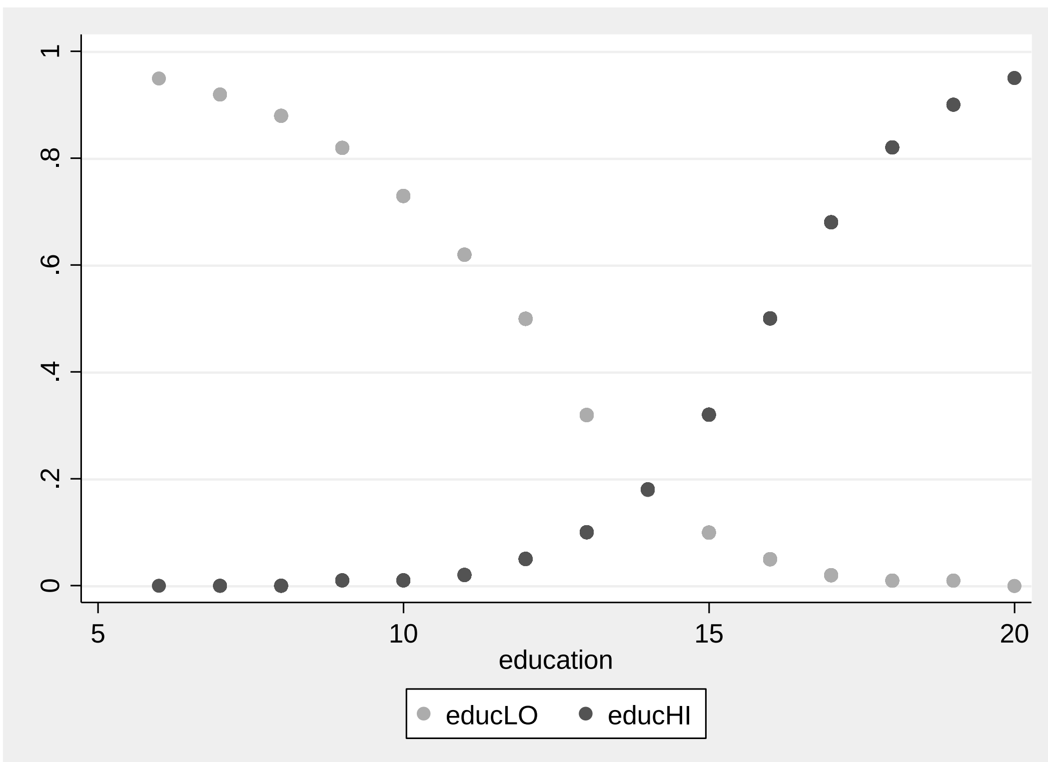
A key focus in using the dualcal command is the specification of the four benchmark values. Notice that in the detailed example presented above the calibration of *highed* uses 12 years as the benchmark for 0.05 membership and 16 years as the benchmark for the cross-over point (0.5). Observe that *lowed* uses these same two benchmark values, only reversed: 12 years is the benchmark value for the cross-over point (0.5), and 16 years is the benchmark value for 0.05 membership. Thus, the same benchmark values can be used in the calibration of two fuzzy sets. The following chart maps the connections between the four benchmark values and fuzzy membership scores in *educHI* and *educLO*:

[3] Of course, it would be necessary to negate *educLO* in order to generate degree of membership in the set of cases with not-low education.

Benchmark	educHI calibration	educLO calibration
20 years	0.95 (high)	not used
16 years	0.5 (cross-over)	0.05 (not low)
12 years	0.05 (not high)	0.5 (cross-over)
6 years	not used	0.95 (low)

The dualcal command would be executed as follows:
compute educ = dualcal(education,20,16,12,6)
The fuzzy set *educHI* utilizes the first three benchmark values, while *educLO* utilizes the last three benchmark values, in reverse order. The plot of *educLO* and *educHI* against their source variable (*education*) is shown in Figure 3. The fuzzy sets created using this procedure are identical to sets generated using separate commands to produce *lowed* and *highed*, shown previously.

Figure 3: Plot of educLO and educHI Against Years of Education



Discussion

My goal in developing dualcal was to discourage what is probably the most common way of calibrating interval/ratio variables as fuzzy sets. The usual practice is to (1) select a benchmark for high-X, (2) select a benchmark for low-X, and then (3) define the cross-over point in the vicinity of the median or in the vicinity of the midpoint between the high and low benchmarks. The end result is that the not-high benchmark is too low because it has been pegged to the low-X benchmark. Also, the cross-over point for X/not-X, by implication, is also too low. The purpose of dualcal is to encourage users to consider the calibration of high-X and low-X at the same time. In effect, the dualcal procedure forces users to pay attention to the disjuncture of the calibration of high-X versus not-high-X from low-X versus not-low-X.

The dualcal procedure has wide applications; however, it is not always appropriate. For example, the benchmark values for “rich” and “poor” are unlikely to align such that the “not-poor” benchmark coincides with the “rich/not-rich” cross-over point, and the “not-rich” benchmark coincides with the “poor/not-poor” cross-over point. In situations where the benchmarks do not align, the two sets, high-X and low-X, should be calibrated separately.[4]

To specify benchmark values simultaneously for high and low membership in a fuzzy set requires extra effort on the part of researchers. As noted previously, researchers too often rely on a single calibration using three benchmarks and then misconstrue and mislabel the negated set (e.g., as “low” X instead of “not-high” X). When selecting benchmark values using the dualcal procedure, researchers should keep in mind that they are specifying four benchmark values. Ideally, each value selected should have some empirical referent or rationale (Ragin 2008; Ragin and Fiss 2017; see also Research Note #6).

It is best to start with the specification of the benchmark for high membership using relatively strict bounds. That is, the specification of “high” should capture only cases that clearly have full or very close to full membership in the set in question and are clearly well above the benchmark value for not-low. Likewise, the specification of “low” should capture cases that have full or very close to full non-membership in the set in question and are clearly below the benchmark for not-high. When specifying the benchmark value for the cross-over point for membership in “high/not-high,” keep in mind that it is also the benchmark for non-membership in “low.” When specifying the benchmark for “not-high,” keep in mind that it is also the cross-over point for “low/not-low.” These restraints on specifying benchmarks have the potential to make calibrations more precise as well as more grounded in evidence.

[4] Thanks to Claude Rubinson for suggesting this example.

3. Constructing Macro-Conditions

Charles C. Ragin (cragin@uci.edu)

Background

One of the difficulties routinely faced by researchers using configurational methods is the large number of logically possible combinations (i.e., configurations) that can be generated by a relatively small number of causal conditions. The following table illustrates the relationship between the number of conditions specified in a truth table analysis and the number of truth table rows (i.e., configurations) generated by QCA software:

Conditions	Combinations
3	8
4	16
5	32
6	64
7	128
8	256
9	512
10	1,024
12	4,096

Consequently, most applications of QCA use a relatively small number of conditions—usually in the range of three to eight. Not only does having more conditions exponentially increase the number of combinations, but also the truth table solutions generated in such situations are often complex and may resist straightforward interpretation. It is important to understand that each configuration (i.e., each corner of the property space defined by the causal conditions) constitutes, in effect, a qualitatively distinct state. Configurational logic dictates treating each combination as (potentially) a different "whole." Of course, when there are 4,096 combinations (or even as few as 32), it is difficult to visualize each one of these distinct states.

This research note describes procedures for combining two or more conditions to create a macro-condition. Combining conditions reduces the dimensionality of the property space implemented in the truth table. For example, suppose a researcher initially specifies six conditions (64 truth table rows). If two related conditions can be combined to form a single macro-condition, the dimensionality of the property space is reduced to five (32 truth table rows). If an additional pair of conditions can be combined, the property space is further reduced to four dimensions (16 truth table rows).

When conditions are combined, the resulting macro-condition is often pitched at a higher level of conceptual abstraction than the component conditions. For example, Ragin and Fiss (2017) combine two fuzzy sets, *not-low parental income* and *educated parent*, to create the macro-condition *favorable family background*. They joined the two component conditions using *Logical AND* which entails taking the minimum of the two component membership scores. Using *Logical AND*, a respondent with a fuzzy membership score of 0.8 in *not-low parental* *income* and a fuzzy membership score of 0.3 in *educated parent* receives a score of 0.3 in *favorable family background*. Using the minimum, only respondents with high membership scores on both component conditions receive high membership scores in the macro-condition.

There are several different ways to combine conditions. The method selected depends on the researcher's substantive knowledge, research goals, and conceptual agility. In the discussion that follows, I describe five different methods for combining conditions: *Logical AND*, *Logical OR*, *Exclusive OR*, *Ranking*, and *Averaging*. Four of the five methods assume that the component conditions are fuzzy sets, which means that the components must be meaningfully calibrated before they can be combined to form macro-conditions.

1. Joining Conditions Using *Logical AND*

When combining conditions using *Logical AND*, the "weakest link," that is, the lowest component membership score, "rules" the degree of membership in the combination. For example, suppose a researcher wants to address degree of membership in the set of *less-developed*, *debtor* countries. The formula for combining these conditions using "weakest link" reasoning is:

macro-condition = fuzzyand(less-developed, debtor)

(The *fuzzyand* operation in fsQCA returns the "minimum" of a set of values.) For example, a case with a score of 0.6 in *less-developed* and 0.9 in *debtor* would be awarded a score of 0.6 in the macro-condition combining these two characteristics. The researcher could then use the derived macro-condition as a single causal condition, in place of the two component conditions.

The weakest link approach also can be used to assess the degree to which cases conform to an ideal type. Suppose, for example, a researcher wants to determine the degree to which each organization in a study conforms to the ideal typic bureaucratic form. The first step would be to define the essential features of bureaucracy and to measure the degree of membership of each organization in each of the essential features, constructed as fuzzy sets. Next, the researcher would compute the degree of membership of each case in the combination of these elements by taking the minimum of the component membership scores. To achieve strong membership in the condition *bureaucratic organization* a case would have to register strong membership (i.e., greater than 0.5) in each of the component conditions.

2. Joining Conditions Using *Logical OR*

The second approach to joining conditions focuses on conditions that separately satisfy a macro-condition. For example, consider the macro-condition *credit worthiness*. The relevant component conditions are *ownership of significant assets* and *high income*. It's not necessary to have high membership in both conditions to register high membership in credit worthiness because the satisfaction of either condition establishes credit worthiness. *Logical OR* returns the maximum membership score, consistent with an emphasis on putting "the best foot forward." For example, if a person has a membership score of 0.8 in *ownership of significant assets* and a membership score of 0.10 in *high income*, *Logical OR* returns the maximum of the two scores, 0.80 membership in *credit worthiness*. In fsQCA *Logical OR* is implemented using *fuzzyor*:

macro-condition = fuzzyor(ownership-of-significant-assets, high-income)

Logical OR can be utilized whenever two or more conditions are considered to be "substitutable" as causal conditions for an outcome. The maximum "rules" the combination. Combining multiple conditions using *Logical OR* tends to result in macro-conditions with predominantly high membership scores, especially if the component conditions do not strongly overlap. The opposite is true of combining multiple conditions using *Logical AND*—the resulting macro-conditions tend to have low scores, again, if the component conditions do not strongly overlap.

3. Joining Conditions Using *Exclusive OR*

An *Exclusive OR* relationship exists between two conditions when both conditions contribute to an outcome when present, but only if the other condition is absent. In short, condition 1 contributes to the outcome only if condition 2 is absent, and condition 2 contributes to the outcome only if condition 1 is absent. Both conditions are linked to the outcome, but not when they are both present. The *Exclusive OR* operation is best understood as a way to model incompatible strategies. For example, business organizations frequently face tradeoffs between incompatible strategies, such as between cost leadership strategies focused on pushing down costs and looking for savings throughout the whole organization, on the one hand, and differentiation strategies focused on improving quality or developing new product features, on the other. Pursuing both at the same time is typically not feasible, as an organization cannot typically offer the lowest cost products and the products with the highest quality and features at the same time; rather, they present mutually exclusive market positions. Either one is feasible, but both simultaneously are not.[5]

The implementation of the *Exclusive OR* relationship involves using both fuzzyand and fuzzyor:

macro-condition = fuzzyor(fuzzyand(cost-cutting, ~quality-enhancing),
fuzzyand(~cost-cutting, quality-enhancing))

First, the minimum of condition 1 and the negation of condition 2 is computed, along with the minimum of condition 2 and the negation of condition 1. Next, the maximum of these two minimums is computed. The resulting macro-condition score shows the degree to which cases are employing one of the two component conditions to the exclusion of the other.

4. Joining Conditions by Averaging Component Scores

[5] Thanks to Peer Fiss for developing this example.

The fourth method for joining conditions is to average the component membership scores. For example, suppose an employer reviewing job candidates believes that extensive job experience can compensate for weak educational credentials, and vice versa. One simple way to allow compensation is to average the relevant component scores. When evaluating job candidates the employer would compute the average of degree of membership in the set with excellent educational credentials and degree of membership in the set with extensive job experience, based on the belief that high scores in one aspect can compensate for low scores in the other. A candidate with 0.9 membership in excellent educational credentials and 0.3 membership in extensive job experience would receive a score (0.9 + 0.3)/2 = 0.6 in the macro-condition—overall attractiveness as a job candidate. In this example, the candidate was able to compensate for a low score on extensive job experience with a high score on excellent educational credentials. In this approach, a case with a high score in one set and a low score in the other is awarded a middle-range score in the macro-condition. By contrast, using *Logical AND*, this case would end up with a low score (0.3, the minimum); using *Logical OR* the applicant would be awarded a high score (0.9, the maximum).

The computation of an average is simply the sum of the component membership scores divided by the number of conditions:

macro-condition = (excellent-educational-credentials + extensive-job-experience)/2

It is important to note that averaging component scores may lead to an abundance of middle-range scores and thus yield a clustering of scores near the cross-over point, the zone of maximum fuzziness. Also, there is likely to be considerable heterogeneity in the middle range. For example, the following three cases would each receive a macro-condition score of 0.5: (0, 1), (1, 0), and (0.5, 0.5).

5. Joining Conditions by Calibrating Status Combinations

The fifth and final technique for creating macro-conditions uses crisp-set conditions such as race, gender, family compositions, and marital status and assigns fuzzy membership scores to different combinations of statuses. For example, Ragin and Fiss (2017) use marital status (married versus not married) and household composition (with children versus no children) to create four categories: married/no children, married/with children, unmarried/no children, and unmarried/with children. These four combinations can be ranked with respect to the degree to which they offer protection from poverty. Ragin and Fiss (2017) label the macro-condition formed from these two dichotomies *favorable domestic situation* and assign fuzzy membership scores as follows:

Status Combination	Fuzzy Membership in Favorable Domestic Situation
married/no children	1.0
married/with children	0.6
unmarried/no children	0.4
unmarried/with children	0

The assigned scores are based on the research literature on poverty which treats marriage as offering a modest degree of insulation from poverty and the presence of children in the household as a liability when it comes to avoiding poverty. Of course, other fuzzy membership scores could be assigned, as long as the rank order is preserved. For example, the two middle combinations could be assigned scores of 0.7 and 0.3 instead of 0.6 and 0.4. Researchers should strive for transparency in their assignment of membership scores, and they should establish a firm grounding in relevant research literature and substantive knowledge.

4. Superset Relations and Necessary Conditions

Charles Ragin (cragin@uci.edu)

Background

When X, a causally relevant antecedent condition, is a consistent superset of Y, an outcome, a researcher may cite this evidence as support for the argument that X is a necessary-but-not-sufficient condition for Y. With crisp sets, a superset relation is indicated when there are very few or no cases displaying the outcome but not the causal condition. With fuzzy sets, a superset relation is indicated when there are very few or no cases with membership scores in the outcome that exceed their membership scores in the causal condition. Once a consistent superset relation has been established, it is possible (1) to assess whether the superset *makes sense* as a necessary condition and then (2) to determine whether it is *theoretically relevant*. For example, the set of former milk drinkers is a superset of the set of heroin users. However, drinking milk does not make sense as a gateway substance for heroin addiction. Even if a condition does make sense as a shared antecedent condition, it still must pass a second test, theoretical relevance. An example: air to breathe is truly a necessary condition for social revolution; however, this condition is not relevant to theory. Thus, it is important to recognize the contingent character of the link between the demonstration of a superset relation and necessity. The demonstration of necessity relies not only on $X_i \geq Y_i$ (where "≥" indicates "is a superset of"), but also upon corroborating empirical evidence, preferably at the case level, as well as supportive theory.[6]

Unfortunately, many QCA researchers treat the finding of a consistent superset relation as proof of necessity, when in fact such a finding is only a first step. For this reason, researchers should reserve the language of necessity for studies that go beyond the demonstration of a consistent superset relation. Studies should cite supportive empirical evidence and relevant theory, as well. Furthermore, when describing set-analytic results, it is often prudent to use generic terms to describe superset relations. For example, instead of describing a superset relation as indicating a necessary condition, researchers could state simply that instances of an outcome (e.g., social revolution) consistently share one or more antecedent conditions (e.g., peasant insurrection). If the antecedent condition is shared by most, but not all instances of an outcome, it could be described as a "widely shared antecedent condition."

This research note explores the superset relation ($X_i \geq Y_i$) in applications of QCA. This topic is not adequately addressed in the QCA literature because truth table analysis—the core analytic device in QCA—focuses almost exclusively on subset relations ($X_i \leq Y_i$). The key assessment is the degree to which combinations of conditions, specified in truth tables, form consistent subsets of the outcome. Thus, as conventionally practiced, QCA is not well-equipped for the analysis of necessary conditions.

The remainder of this research note is seeks to clarify several issues regarding superset relations and necessary conditions. I first discuss practical issues in the assessment of superset

---

[6] In most QCA applications, the understanding of necessity is neither strict nor absolute. Not only do researchers make allowances for inconsistencies due to data and calibration errors, but they also employ a broad conceptualization of necessity, as in "X is *usually* necessary for Y" as opposed to "if no X, then no Y."

relations, focusing on five common shortcomings. I then suggest an analytic protocol for the assessment of superset relations.

Some Common Shortcomings

1. Not assessing the consistency of superset relations. It is advisable to evaluate the set relation between every causal condition and the outcome prior to conducting a truth table analysis. The simplest way to do this is to use the XY Plot procedure in fsQCA, which reports the consistency of $X_i \geq Y_i$ and $X_i \leq Y_i$ for each bivariate plot. It is important to identify superset relations ($X_i \geq Y_i$) at the outset of an analysis because they may have important implications for theory as necessary conditions. Also, truth table analysis focuses on subset relations, which means that conditions that are supersets of the outcome may be overlooked. In fact, parsimonious solutions often drop consistent supersets of the outcome from recipes identified via truth table analysis.

2. Excluding consistent superset conditions from truth table analysis. If the researcher identifies a consistent superset relation that makes sense as a necessary condition, then it should be included as a condition in the truth table analysis. Even though truth table analyses focus on subset relations, it is important to examine all relevant conditions when conducting an analysis. A necessary condition may be required in a recipe for a causal combination to pass the sufficiency test.

3. Concluding that a condition is necessary simply because it appears in all the recipes in a truth table solution. It is tempting to view a causal condition that appears in each recipe identified via truth table analysis as a necessary condition. After all, its presence seems to be required. However, recall that solution coverage is usually much less than 100%, meaning that there are additional causal pathways to the outcome, not captured by the current analysis. These additional pathways may not require the presence of the ingredient shared by the recipes identified in the current analysis. The only effective test of the superset relation is the assessment of the consistency of $X_i \geq Y_i$.

4. Failing to notice that a necessary condition has been dropped from a parsimonious solution. As noted above, parsimonious solutions often drop conditions that are supersets of the outcome. To understand why this occurs it is necessary to consider the coding of superset conditions across the three kinds of truth table rows: rows coded 1 on the outcome, rows coded 0 on the outcome, and remainder rows. Superset conditions are likely to be skewed toward a coding of “present” in rows coded 1 on the outcome and, to a lesser degree, also in rows coded 0 on the outcome. (Cases that display strong membership in a superset condition but not in the outcome do not undermine the consistency of the superset relation.) However, in the *remainder* rows, superset conditions are likely to be skewed toward a coding of “absent.” Recall that truth table minimization uses a process of “incremental elimination,” such that rows with the same outcome can be paired if they differ on only one condition. The condition that differs across the two rows can be eliminated. All remainder rows are up for grabs in the derivation of the parsimonious solution and thus are well-positioned to combine with rows coded “1” on the outcome. This arrangement sets the stage for the pair-wise elimination of the superset condition. Parsimonious solutions utilize any remainder combination that yields a simpler truth table solution, regardless of whether the remainder combination makes sense from the viewpoint of the researcher’s substantive knowledge or theory.

5. Compounding necessary conditions (i.e., using logical “and” to model jointly necessary conditions) without explicitly testing the necessity of their combination. It is important to remember that most of the time we are able (in social science) to identify conditions that are “usually” necessary or “almost always” necessary. The usual consistency with the superset relation is around 0.90 when we claim we have found a pattern suggestive of necessity. This fact (i.e., that consistency is not a perfect 1.0) complicates the assessment of jointly necessary conditions. For example, if the consistency of $X_i \geq Y_i$ is 1.0 and the consistency of $Z_i \geq Y_i$ is 1.0, then we can state with confidence that the consistency of $X_i \bullet Z_i \geq Y_i$ is 1.0 (i.e., the consistency of min(X_i, Z_i) as a superset of Y_i is 1.0). If there is any inconsistency, however, the data set will include cases where $X_i < Y_i$ and cases where $Z_i < Y_i$. If the superset consistency of $X_i \geq Y_i$ is 0.90, there will be a modest number of cases where $X_i < Y_i$. If the superset consistency of $Z_i \geq Y_i$ is also 0.90, then the consistency of $X_i \bullet Z_i \geq Y_i$ will be less than 0.90, to the extent that the cases where $X_i < Y_i$ are not the same as the cases where $Z_i < Y_i$ (i.e., the inconsistent cases do not overlap). The important point is that when you combine “usually” necessary conditions, their intersection (i.e., their joint necessity) may NOT meet the “usually” necessary threshold (0.9 in this example).

Recommended protocol for necessary conditions

1. It is always a good idea to examine scatterplots of the calibrated outcome (Y axis) against calibrated causal conditions (X axis), one at a time, at the outset of an investigation. Check both X and ~X and both Y and ~Y (all four assessments) to explore the possible triangularity of the plot. Pay attention to the consistency scores for both $X_i \leq Y_i$ (X is a subset of Y) and $X_i \geq Y_i$ (X is a superset of Y)

2. Include conditions that are consistent supersets of the outcome ($X_i \geq Y_i$) in the truth table analysis (which searches for sufficient combinations of conditions). Be sure to code them correctly in the intermediate solution dialogue box.

3. Focus on the complex or the intermediate solution; parsimonious solutions tend to eliminate conditions that are supersets of the outcome.

4. If any condition appears in all the recipes for the outcome, then it is a common ingredient. It is also a necessary condition if it has high consistency, on its own, as a superset of the outcome and makes sense as a necessary condition.

5. Conditions that are supersets of the outcome but do not appear in all the recipes in the solution (intermediate or complex) should not be treated as “necessary” conditions.

Discussion

Superset consistency assesses the degree to which membership in a condition is greater than or equal to membership in an outcome. Thus, a superset condition functions as a ceiling, setting an upper limit on the expression of an outcome. All breaches of that ceiling undermine the consistency of the superset relation. When interpreted as necessary conditions, superset conditions are especially relevant to theory, for they reveal possible empirical constraints on the outcome. Thus, testing for superset relations should be a routine part of conventional applications of QCA, especially in the early phases of an analysis.

5. Clarifying Causal Recipes

Charles C. Ragin (cragin@uci.edu)

Background

Applications of QCA use truth tables to derive logically minimal expressions of the combinations of conditions linked to an outcome. Results take the form of logical statements, which join relevant conditions ("ingredients") together in "causal recipes" using logical *and*. In turn, recipes are joined together using logical *or*, indicating alternate combinations of conditions for an outcome. The resulting logical statements are known as "sums of products" expressions because the use of logical *and* is analogous to multiplication and the use of logical *or* is analogous to addition. For example, the expression A•B + C•D → Y states that there are two combinations of conditions, A combined with B and C combined with D, that are consistently linked to the expression of outcome Y. If a case has membership in either recipe, it is likely to exhibit outcome Y.

QCA's emphasis on the derivation of *logically minimal* statements is an important feature of the approach. For example, if cases of A•B are linked to outcome Y and cases of A•B•C are also linked to Y, the second, more complex expression is absorbed by the first, yielding A•B as the logically minimal expression. Likewise, if cases of A•D•E and A•D•~E are both linked to outcome Y (the tilde "~" indicates negation), the two recipes can be combined to generate A•D. These and other principles of Boolean algebra make QCA solutions logically minimal.

A given truth table row may be incorporated into more than one causal recipe. For example, both recipes in the truth table solution described above (A•B + C•D → Y) incorporate the truth table row A•B•C•D. Whenever truth table rows are incorporated into multiple recipes, the resulting recipes partially overlap in their coverage. Again, the goal is to derive a logically minimal solution, which in simple terms means that each recipe has as few ingredients as possible and that there are as few recipes as possible. In applications of QCA, having overlapping causal recipes is the rule, not the exception, and follows from the goal of deriving logically minimal truth table solutions.

To illustrate overlapping coverage, consider a simple truth table using data from the National Longitudinal Survey of Youth (NLSY79). As shown in table 1, the three conditions are race (white = 1; black = 0), gender (male = 1; female = 0), and class background (advantaged versus not-advantaged). Race and gender are conventional binary sets. Advantaged class background is operationalized as a fuzzy set based on the intersection of two other fuzzy sets, *educated-parent* and *not-low-parental-income*. The truth table's outcome is also a fuzzy set, degree of membership in the set of respondents *avoiding poverty*. For further information on the construction and calibration of these fuzzy sets, see Ragin and Fiss (2017).

Inspection of table 1 reveals a substantial gap in consistency scores between rows 6 and 7. Thus, one plausible consistency threshold for coding the outcome "true" (i.e., 1) is 0.745 (see table 1). Using this threshold, the first six rows of table 1 are coded 1 on the outcome, while the remaining two rows are coded 0. Logical minimization of the coded truth table yields the following result:

white + advantaged → avoidance of poverty

This truth table solution states that respondents who are white or from advantaged class background are able to avoid poverty with at least moderate consistency (i.e., with consistency scores of 0.745 or better). The two causal recipes overlap, which is apparent from the fact that their intersection is not a null set. Instead, intersecting the two recipes yields their overlap: *white•advantaged* (rows 1 and 2 of the truth table).

Table 1: Impact of Race, Gender, and Advantaged Class Background on Poverty Avoidance

Row	Race: white = 1 black = 0	Gender: male = 1 female = 0	Class: Membership in Advantaged (1) or ~Advantaged (0)	Number of respondents	Consistency of poverty avoidance	Outcome code
1	1	1	1	1045	0.884	1
2	1	0	1	986	0.857	1
3	0	1	1	235	0.794	1
4	1	1	0	315	0.786	1
5	1	0	0	327	0.750	1
6	0	0	1	212	0.745	1
7	0	1	0	497	0.587	0
8	0	0	0	562	0.421	0

Clarifying Causal Recipes

This research note argues that deriving logically minimal solutions, while clearly desirable, may undermine the interpretability of truth table solutions. Interpretability can often be enhanced by "clarifying" truth table solutions specifying multiple recipes. To clarify recipes it is necessary to assign overlapping coverage exclusively to one of the causal recipes exhibiting the overlap. The resulting statement is no longer logically minimal, but may be easier to interpret due to the separation of previously overlapping recipes. It is appropriate to call the procedure *clarification* for two reasons: (1) The purpose of clarification is to separate the combinations of conditions linked to an outcome in a way that highlights their differences. (2) In the field of chemistry, *clarification* is a term that is used to describe the process of separating mixtures into their component substances.

Clarification is a four-step process (Ragin 2023:59-61). It is easiest to grasp when there are only two recipes:

#1 #2

white + advantaged → avoidance of poverty

Step 1: select the recipe that is to receive the overlap (recipe #1). Step 2: negate recipe #1. Step 3: intersect negated recipe #1 with recipe #2. Step 4: rewrite the solution with the modified recipe #2. Consider, for example, the clarification of the solution to table 1:

logically minimal solution:	*white + advantaged*
recipe to receive overlap (recipe #1):	*white*
negated recipe #1:	*~white*
intersection of negated recipe w/recipe #2:	*~white•advantaged*
clarified solution:	*white + ~white•advantaged*

This restatement of the results of the truth table analysis makes it clear that all whites are able to avoid poverty with a consistency of 0.745 or better, but only blacks from advantaged class backgrounds experience comparable levels of insulation from poverty. Clarifying the truth table findings in this way shifts the focus to *white* versus ~*white* and to the fact that nonwhites must come from advantaged class backgrounds in order to avoid poverty to a degree that is comparable to whites.

The selection of the recipe assigned to receive the overlap depends on the researcher's interest and focus. The key issue is which assignment offers superior interpretive gain. Consider the consequences of switching the assignments of recipes #1 and #2 in the example just presented:

logically minimal solution:	*white + advantaged*
recipe to receive overlap (recipe #2):	*advantaged*
negated recipe #2:	*~advantaged*
intersection of negated recipe w/recipe #1:	*~advantaged•white*
clarified solution:	*advantaged + ~advantaged•white*

In this version of the results, the focus is on *advantaged* versus ~*advantaged* and the fact that not-advantaged whites are able to avoid poverty. The issue of which clarified solution to stress is a question of the researcher's focus, specifically, in this example, whether the focus is on *white* versus *not-white* or on *advantaged* versus *not-advantaged*.

A More Complex Example

The simple example presented in table 1 is not characteristic of QCA applications; causal recipes with only one ingredient are relatively rare. Consider the following, more complex example: Olav Stokke's (2004) truth table for successful shaming of violators of international fishing agreements (see table 2). Using the three-letter condition labels, as shown in the table, the QCA solution to the table can be stated as follows:

#1 #2

adv•~inc + adv•shd•rev → successful shaming

The two causal recipes for successful shaming are: (#1) supportive scientific advice (*adv*) in situations where it is not inconvenient for the target of shaming to alter its behavior (~*inc*), and (#2) supportive scientific advice (*adv*) in situations where there is both domestic reverberations for being shamed (*rev*) and a need to strike future deals (*shd*).

Table 2: Stokke's truth table for successful shaming

Advice (adv)	Commitment (com)	Shadow (shd)	Inconvenience (inc)	Reverberation (rev)	Success
1	0	1	1	1	1
1	1	1	1	1	1
1	1	1	0	0	1
1	0	0	0	0	1
1	0	0	1	0	0
1	0	0	1	1	0
0	0	0	1	0	0
1	1	1	1	0	0

1. *Advice* (*adv*): whether the shamers can substantiate their criticism with reference to explicit recommendations of the regime's scientific advisory body;
2. *Commitment* (*com*): whether the target behavior explicitly violates a conservation measure adopted by the regime's decision-making body;
3. *Shadow of the future (shd*): perceived need of the target of shaming to strike new deals under the regime--such beneficial deals are likely to be jeopardized if criticism is ignored;
4. *Inconvenience* (*inc*): the inconvenience (to the target of shaming) of the behavioral change that the shamers are trying to prompt;
5. *Reverberation* (*rev*): whe domestic political costs to the target of shaming for not complying (i.e., for being scandalized as a culprit).

Notice that there is logical overlap between the two recipes: instances of *adv•~inc•shd•rev*, formed from the intersection of the two recipes, conform to both. It is possible to assign this overlap to recipe *adv•~inc*, and thereby clarify and separate the two causal recipes. Following the procedure used in the first example, I award the overlapping coverage to *adv•~inc*.

logically minimal solution:	*adv•~inc + adv•shd•rev*
recipe to receive overlap (#1):	*adv•~inc*
negated recipe #1:	*~adv + inc*
negation intersected with recipe #2:	*(~adv + inc)•adv•shd•rev*
results of intersection	*adv•inc•shd•rev*
clarified solution	*adv•~inc + adv•inc•shd•rev*

The clarified solution focuses on whether the behavioral change is inconvenient to the targets of shaming. If it is not inconvenient (~*inc*), then the conditions for successful shaming are simple: supportive scientific advice (*adv*). However, if the behavioral change is inconvenient (*inc*), then two additional conditions for successful shaming must be met. In addition to supportive advice (*adv*), there must be a need to strike future deals (*shd*) and there must be domestic reverberations (*rev*) for being shamed.

Awarding the overlap to recipe #2 (i.e., to *adv•shd•rev*) is possible, as in the first example, but the clarified solution that results is not as interpretable as the one just presented. Awarding the overlap to *adv•shd•rev*, the clarified solution is:

(~shd + ~rev)•adv•~inc + adv•shd•rev → successful shaming

As a general rule, interpretive gains are greater when the recipe selected to receive the overlap has fewer conditions.

Clarifying Solutions with Three Recipes

The greater the number of recipes in a truth table solution, the greater the challenge of clarifying the recipes. Also, as is evident in the previous example, the selection of the recipe to receive the overlap can have a substantial impact on the interpretability of the clarified solution. Consider a truth table solution with three recipes. For this demonstration, I use data on social movement organizations ("challengers") published by William Gamson (1990) in *The Strategy of Social Protest.* Gamson developed a sampling frame for social movement organizations (SMOs) in the United States from 1800 to 1945. He focused on SMOs that gained new advantages for their constituents within fifteen years of their period of activism. The relevant presence/absence conditions for the outcome ("new advantages") are presented in Gamson (1990) as follows:

bur: the challenging group developed a bureaucratic organizational structure;
low: the challenging group's constituency was low status (e.g., workers, minorities);
dis: the challenging group's goal was to displace a person in a position of power;
hlp: the challenging group received help from outsiders (e.g., another SMO);
acp: the challenging group won acceptance as a representative of its constituents.

Truth table analysis yields a solution with the following three recipes for new advantages:

#1 #2 #3
~low•~dis + bur•~dis•acp + ~dis•hlp•acp → new advantages

With three recipes there are three possible overlaps: recipe #1 with recipe #2, recipe #1 with recipe #3, and recipe #2 with recipe #3. All three pairings intersect without yielding a null set, indicating meaningful overlap. Clarification begins with recipe #1 receiving its overlapping coverage from recipe #2:

recipe to receive overlap (#1):	*~low•~dis*
negated recipe #1:	*low + dis*
negation intersected with recipe #2:	*(low + dis)• bur•~dis•acp*
results of intersection	*low•bur•~dis•acp*
clarified recipes (#1 and #2)	*~low•~dis + low• bur•~dis•acp*

The focus of the clarified recipes is on SMOs representing not-low status constituents (~*low*) versus SMOs representing low-status constituents (*low*). Essentially, the clarified recipes indicate that recipes for new advantages differ dramatically depending on the status of an

SMO's constituents. In order to achieve new advantages, an SMO representing low-status constituents (e.g., minorities or workers) must have an organizational structure (*bur*), non-displacement goals (~*dis*), and win acceptance as a representative of its constituents (*acp*). If, however, an SMO's constituents are not low status, the key to achieving new advantages is simple—having non-displacement goals.

Next, recipe #1 receives its overlapping coverage from recipe #3:

recipe to receive overlap (#1):	*~low•~dis*
negated recipe #1:	*low + dis*
negation intersected with recipe #3:	*(low + dis)• ~dis•hlp•acp*
results of intersection	*low•~dis•hlp•acp*
clarified recipes (#1 and #3)	*~low•~dis + low•~dis•hlp•acp*

Once again, the focus of the clarified recipes is on SMOs representing not-low-status constituents (~*low*) versus those representing low-status constituents (*low*). The recipe for SMOs representing not-low-status constituents (#1) is simple and unchanged. After all, recipe #1 receives the overlapping coverage from recipes #2 and #3. The recipe for SMOs representing low status constituents (#3) differs only slightly from clarified recipe #2. Instead of establishing a bureaucratic organizational structure (*bur*), clarified recipe #3 has help from outsiders (*hlp*).

Finally, recipe #2 receives its overlapping coverage from recipe #3:

recipe to receive overlap (#2):	*low• bur•~dis•acp*
negated recipe:	*~low + ~bur + dis + ~acp*
negation intersected with recipe #3:	*(~low+ ~bur +dis + ~acp)• low~dis•hlp•acp*
results of intersection	*low•~bur•~dis•hlp•acp*
clarified recipes (#2 and #3)	*low•bur•~dis•acp +low•~bur•~dis•hlp•acp*

Thus, the fully clarified solution with all three recipes is:

#1 #2 #3

~low•~dis + low• bur•~dis•acp + low•~bur•~dis•hlp•acp → new advantages

The primary interpretive gain is the low-status/not-low-status distinction and the associated contrast regarding the conditions required for the achievement of new advantages. SMOs representing low-status constituents are at a clear disadvantage when it comes to winning new advantages.

Because recipes #2 and #3 are so similar, there is more to be gained from combining them than from clarifying and separating them. Consider again the two recipes (#2 and #3), clarified with respect to their overlapping coverage with recipe #1:

low•bur•~dis•acp + low•~dis•hlp•acp

Factoring the two recipes yields:

low•~dis•acp•(bur + hlp)

The factored expression indicates that SMOs representing low-status constituents (*low*) can achieve new advantages if they have non-displacement goals (~*dis*), win acceptance as a representative of their constituents (*acp*), and have either a bureaucratic organization (*bur*) or help from outsiders (*hlp*). Thus, merging recipes #2 and #3 yields the following (final) truth table solution:

~low•~dis + low•~dis•acp•(bur + hlp) → new advantages

Conclusion

There are several ways to enhance the interpretability of truth table solutions. This research note focuses primarily on the interpretive gains that follow from clarifying recipes, which separates recipes' overlapping coverage and highlights their differences. Several issues are addressed. First, it matters which recipe is chosen to receive the overlapping coverage. It is important for the clarified recipes to resonate with the substantive issues that motivate the research. Second, one tentative conclusion is that clarification works best when the recipe selected to receive the overlap has the fewest conditions. Third, recipes that are very similar, such as #2 and #3 in the last example, may not be good candidates for clarification. Instead, researchers should consider merging them into a single recipe. Fourth, researchers may have to go through a process of trial and error in order to derive the most interpretable versions of their truth table solutions.

The larger point is that researchers should not hesitate to manipulate recipes algebraically. Truth table results are formulated as logical statements and thus are open to fine tuning. The application of a few simple rules to truth table results can yield substantial interpretive gains. In all three examples presented in this research note, the clarified solutions highlighted important contrasts: (1) poverty avoidance by whites versus nonwhites; (2) responses to shaming that varied depending on whether the behavioral change was inconvenient; and (3) the impact of the social status of an SMO's constituency on the conditions required for its success.

6. Documenting Calibration

Charles C. Ragin (cragin@uci.edu)

Background

A central focus of set analytic research is the evaluation of subset and superset relationships. A fuzzy subset relationship exists between a causal condition and an outcome when set membership scores in a causal condition (e.g., the fuzzy set of individuals with high academic test scores) are consistently *less than or equal to* set membership scores in an outcome (e.g., the fuzzy set of individuals avoiding poverty). A fuzzy superset relation exists when membership scores in a causal condition (e.g., set of countries experiencing state breakdown) are consistently *greater than or equal to* membership scores in an outcome set (e.g., the set of countries undergoing social revolution). Thus, it matters a great deal how fuzzy sets are constructed and how membership scores are assigned to cases. Serious miscalibrations can easily distort or undermine the analysis of set-analytic relationships.

The translation of interval-scale variables into fuzzy sets—the primary focus of this research note—should never be mechanical. Instead, it should be based as much as possible on substantive case knowledge, with full attention to the conceptualization and labeling of the set in question. Basing set membership scores entirely on such criteria as the range, the mean, the median, the rank order, or the standard deviation typically compromises the utility of fuzzy sets. The specification of membership scores must be rooted as much as possible in substantive knowledge. This understanding of set membership scores justifies the conceptualization of membership assignment as a process of calibration, for the procedure involves using external standards and criteria to interpret raw scores (Ragin, 2008:72ff).

The dependence of set-analytic research on careful calibration of set membership scores stands in sharp contrast to conventional quantitative analysis with respect to its use of variables. For the conventional variable to be useful in a multivariate procedure such as multiple regression, it needs only to vary in a meaningful way. Often, the specific metric of a conventional variable is ignored by researchers because it is arbitrary or difficult to interpret. After all, metrics disappear when correlations are computed, and matrices of bivariate correlations, or their mathematical equivalents, provide the foundation for the most commonly used statistical procedures.

The central purpose of this research note is to underscore the need for transparency in the conceptualization of fuzzy sets and in the calibration of set membership scores. Too often, articles submitted to research journals provide little or no detail regarding the derivation and assignment of fuzzy membership scores. But this information is crucial to the evaluation of an application of QCA. For example, if a researcher specified a too-low threshold for full membership in a condition, s/he might *increase* both the condition's mean fuzzy membership score and the proportion of cases with full or close-to-full membership in the condition. These distortions, in turn, might spuriously inflate consistency with the *superset* relation between the causal condition and the outcome. Likewise, specifying a too-high threshold for full membership in a condition might *depress* both the mean fuzzy membership score and the proportion of cases with full or close-to-full membership in the condition. These distortions, in turn, might spuriously inflate consistency with the *subset* relation between the causal condition and the outcome. In short, set calibration can have a direct impact on the assessment of the consistency of set analytic relationships. Full disclosure of calibration practices and procedures provides the primary safeguard against spuriously inflated or deflated consistency scores.

My focus in this research note is specifically on the translation of conventional interval-scale variables to fuzzy sets using the calibration function that is included in fsQCA and other versions of QCA software. To calibrate interval-scale variables as fuzzy sets, researchers use the "direct calibration" method discussed in Ragin (2008, chapter 5). This method is based on the specification of selected interval-variable values corresponding to (1) the threshold for full membership in the target set, which is translated to a fuzzy membership score of 0.95; (2) the threshold for full non-membership in the target set, which is translated to a fuzzy membership score of 0.05; and (3) the cross-over point, which is the dividing line between being "more in" versus "more out" of the target set, which translates to a fuzzy membership score of 0.50. The end result is typically an S-shaped XY plot of the fuzzy scores (Y axis) against the raw values (X axis), with low values on the interval-scale variable approaching a fuzzy membership score of 0 and high values on the interval-scale variable approaching a fuzzy membership score of 1.[7]

Figure 1 provides an example illustrating the calibration of degree of membership in the fuzzy set of respondents avoiding poverty. The source variable is the ratio of the respondent's income to the poverty level for that household. The three interval-scale values are as follows: The threshold for full non-membership in the outcome (avoiding poverty) is an income to poverty ratio of 1.0 (household income is at the poverty level); the cross-over point is an income to poverty ratio of 2.0 (household income is twice the poverty level: and the threshold for full membership in avoiding poverty is an income to poverty ratio of 3.0 (household income is three times the poverty level).

Documenting Calibration

To illustrate my concerns regarding the documentation of calibration, I use data from the National Longitudinal Survey of Youth, as analyzed by Ragin and Fiss (2017) in *Intersectional Inequality*. I first discuss the minimal requirements necessary for documenting calibration, which is a table reporting the source variable and the values used to structure the translation of raw values to fuzzy membership scores. Second, I recommend the use of an online appendix to provide substantive rationales—whenever possible—for the selection of the three values used to translate raw values to fuzzy membership scores.

[7] The "dual calibration" procedure described in detail in research note #2, utilizes four benchmark values. The top three are used to calibrate membership in high X; the bottom three are used in reverse order to calibrate membership in low X. This refinement of the calibration procedure prevents the conflation of not-high X and low X by making their separate calibration explicit.

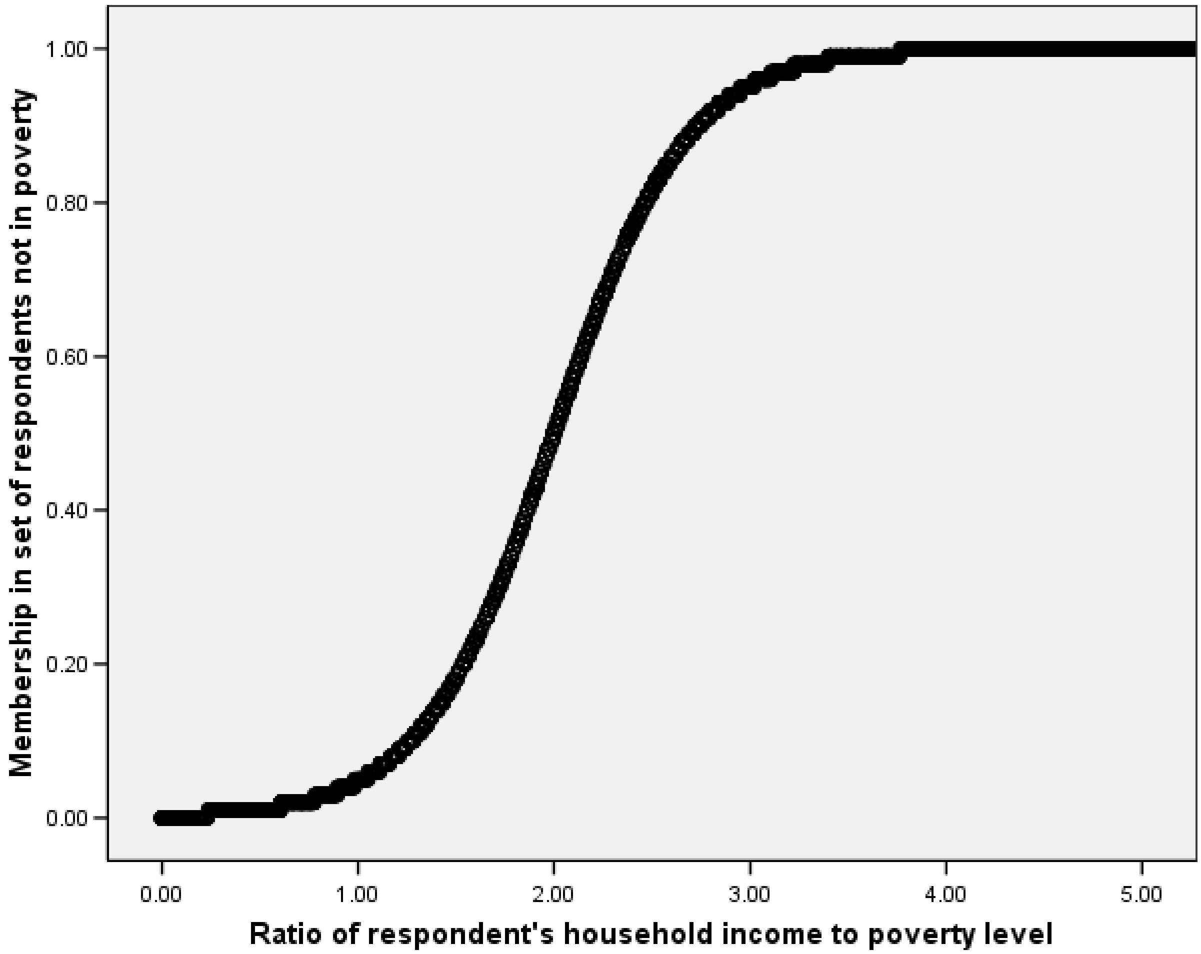


Table 1 presents examples of fuzzy set calibration drawn from *Intersectional Inequality* (Ragin and Fiss 2017). Applications of fsQCA that employ fuzzy sets should use table 1 as a model. The first column names the target fuzzy set; the second column identifies the source variable. Note that the same source variable may be used to generate more than one fuzzy set. For example, the ratio of parents' income in 1978/9 to the official poverty level for the parents' household (based on the number of adults, the number of children, and so on) is the source variable for two fuzzy sets, the fuzzy set of respondents with low-income parents and the fuzzy set of respondents with high-income parents. The third and fourth columns report the means and standard deviations of the source variables. These statistics are not used in the calibration process; rather, they are reported so that readers have access to baseline descriptive statistics on the distributions of the source variables. Columns 5-7 provide the key information utilized in the calibration process: the values of the source variables that correspond to fuzzy set membership scores: the threshold for non-membership (.05), the cross-over point (.5), and the threshold for full membership (.95).

Table 1: Documenting Fuzzy Set Calibration*

Target Fuzzy Set	Source Variable	Source Variable Mean	Source Variable Standard Deviation	Threshold for full nonmembership (0.05)	Cross-over point (0.50)	Threshold for full membership (0.95)
In Poverty	ratio of household income to poverty level for household	3.49	2.75	3.0	2.0	1.0
Not in Poverty	ratio of household income to poverty level for household	3.49	2.75	1.0	2.0	3.0
Low-income Parents	ratio of parent's income in 1978/9 to poverty level for household	5.61	4.09	5.5	3.0	2.0
High-income Parents	ratio of parent's income in 1978/9 to poverty level for household	5.61	4.09	3.0	5.5	8
Low AFQT Score	AFQT score expressed as a percentile	43.74	29.26	30th	20th	10th
High AFQT Score	AFQT score expressed as a percentile	43.74	29.26	65th	80th	93rd

* N = 4,185

Parental income and AFQT test scores (from the Armed Forces Qualifying Test) are both calibrated in two different ways in order to permit examination of key differences in causal processes. For example, is it having low-income parents that is linked to being in poverty or is it having not-high-income parents? By calibrating parental income to reflect membership in these two different target sets, it is possible to address this question using fuzzy sets. Likewise, is it having *high* AFQT scores that is linked to avoiding poverty or is it having *not-low* AFQT scores? Again, by calibrating AFQT scores to reflect membership in these two different target sets, it is possible to address this question.

Including a table documenting the calibrations used in a study is essential to the evaluation of the study. Additionally, it is important to document the reasoning behind the selection of the three values used to translate raw values to fuzzy membership scores. Because the rationale for each set of benchmarks can be lengthy, it is best to provide this information in an on-line appendix. The following example describes the calibration of low and high AFQT scores. It demonstrates a situation where the external criteria used to guide calibration are explicit and readily available. The entire write-up that follows would be included in the online appendix.

> **Source**. To construct the two fuzzy-set measures based on the AFQT, degree of membership in the set of respondents with high AFQT scores and degree of membership in the set of respondents with low AFQT scores, I utilize categories specified by the Department of Defense to place enlistees. The military divides the AFQT scale into five categories based on percentiles. These five categories have substantive importance in that they determine eligibility for military service as well as assignment into different qualification groups. Persons in categories I (93rd to 99th percentile) and II (65th to 92nd percentile) are considered to be above average in trainability; those in category III (31st to 64th percentile) are considered about average; those in category IV (10th to 30th percentile) are designated as below average in trainability; and those in category V (1st to 9th percentile) are designated as substantially below average.
>
> **High AFQT Score**. The threshold for full membership (0.95) in the set of respondents with *high* AFQT scores was placed at the 93rd percentile, in line with the military's designation of the lower boundary of their highest category; the cross-over point (0.5) was set at the 80th percentile; and the threshold for full non-membership (0.05) in the set of respondents with high AFQT scores was placed at the 65th percentile, the bottom of the military's second highest AFQT category.
>
> **Low AFQT Score**. The threshold for full membership (0.95) in the set of respondents with *low* AFQT scores was placed at the 10th percentile, in line with its usage by the military; respondents who scored lower than the 10th percentile received fuzzy membership scores between 0.95 and 1.0. The cross-over point (0.5) was set at the 20th percentile, and the threshold for non-membership was set at the 30th percentile, again reflecting the practical application of AFQT scores

> by the military. Respondents who scored better than the 30th percentile received fuzzy scores between 0.05 and 0 in degree of membership in the set of respondents with low AFQT score.

Unfortunately, detailed calibration guidance, such as that provided in the above example, is uncommon, and researchers must use alternate calibration strategies. When the number of cases is small to moderate, case knowledge can be brought to bear. For example, when calibrating degree of membership in the set of democracies, knowledge of the political history of specific countries can be used to aid the selection of benchmark values. Often, it is useful to sort cases in descending order on the source variable and then examine the column of case names in the data spreadsheet to guide the identification of benchmark values, using case knowledge.

Distributional criteria should be used only as a last resort. Too often, researchers use mechanistic criteria to guide calibration. For example, a researcher might use the raw value at the 80th percentile as the threshold for full membership, the raw value at the 50th percentile as the crossover point, and the raw value at the 20th percentile as the threshold for full non-membership. This translation strategy is likely to yield equivocal results. Not only does it lack grounding in substantive and/or case knowledge, it may also contradict one of the essential features of fuzzy sets—namely, their asymmetric coding. As explained in “Dual Calibration” (research note #2), *not-high* X is not the same as *low* X, and *high* X is not the same as *not-low* X. *High* X and *low* X should be calibrated separately. Thus, when using mechanistic criteria as a last resort, researchers need to make allowance for the fact that the benchmark value for the threshold for non-membership in X corresponds to the value for *not-high* X which may be substantially higher than the value for *low* X.

Full transparency with respect to calibration is essential to both the evaluation and the advancement of fuzzy-set-analytic methods. Researchers who document their calibrations not only enhance the value of their work, but they also contribute to the foundation of set-analytic research, setting the stage for greater accumulation of knowledge. Furthermore, careful documentation of calibration strengthens the community of researchers using fuzzy sets, as researchers share their expertise and their experiences with each other.

7. Subset-SupersetAnalysis

Charles C. Ragin (cragin@uci.edu)

Background

In social research crisp and fuzzy sets are most commonly utilized in truth table analysis. The usual goal is to identify the combinations of causally relevant conditions linked to an outcome. The truth table provides a platform for the (mostly inductive) examination of all logically possible combinations of a given set of conditions. Combinations that consistently share the outcome in question, as indicated by their subset consistency scores, are identified via truth table analysis and then logically simplified. The end product is a logical statement describing the different "recipes" for an outcome. There are many other actual and potential uses of both crisp and fuzzy sets. Ragin and Fiss (2017), for example, examine to degree to which memberships in two or more sets "coincide." Set coincidence is high when the membership in one set overlaps strongly with the membership in another. Ragin and Fiss (2017: 91) document a strong degree of set coincidence between respondents with not-low parental income and respondents avoiding poverty. The degree of set coincidence is much stronger for whites than for blacks.

This research note describes an alternate analytic procedure utilizing crisp and fuzzy sets, subset-superset analysis. Like truth table analysis, subset-superset analysis examines combinations of conditions linked to an outcome. However, subset-superset analysis differs from truth table analysis in that it is more deductive than truth table analysis, which tends to be inductive.

The Logic of Subset-Superset Analysis

The starting point of subset-superset analysis is the user's specification of a causal recipe for an outcome. The specification can be based on relevant research literature, theory, substantive knowledge, case studies, or the results of a previous truth table analysis. Subset-superset analysis assesses the consistency and coverage of the user-specified causal recipe, as well as the consistency and coverage of all subsets of conditions specified in the recipe. The procedure is called "subset-superset analysis" because the examination of a subset of the conditions specified in a causal recipe constitutes a superset of that recipe. For example, a researcher might speculate, based on relevant theory, that causal recipe a•~b•c•~d is sufficient for outcome y ("•" indicates set intersection—combined conditions; "~" indicates set negation).[8] Using the subset-superset procedure, the researcher would test not only the initial recipe but also all subsets of the conditions appearing in the recipe, for instance, the combination a•~b. By definition, degree of membership in a•~b is greater than or equal to membership in a•~b•c•~d, just as the set of males is a superset of the set of red-headed males. In fact, the coverage of all subsets of conditions appearing in a•~b•c•~d are supersets of a•~b•c•~d. More generally, the fewer the conditions in a recipe, the larger the set of cases it embraces.

[8] Notice that the researcher specifies the directionality of each causal condition. That is, the researcher must specify whether it is the causal condition (as calibrated in the data spreadsheet) or its negation that is linked to the outcome.

The number of recipes assessed in a subset-superset analysis is $2^k - 1$, where k is the number of conditions in the initial recipe. For a four-condition initial recipe, the procedure assesses a total of 15 set formulations. For example, suppose a researcher proposes the following recipe: a•~b•c•~d (i.e., a single combination of four conditions). The subset-superset procedure calculates the consistency and coverage of the initial recipe as well as the consistency and coverage of all possible subsets of the four conditions, as follows:

initial four-condition recipe:	a•~b•c•~d
four three-condition recipes:	a•~b•c, a•~b•~d, a•c•~d, ~b•c•~d
six two-condition recipes:	a•~b, a•c, a•~d, ~b•c, ~b•~d, c•~d
four one-condition recipes:	a, ~b, c, ~d

For each recipe tested, the procedure reports the consistency and coverage of the recipe, with respect to its subset relation with the outcome. In other words, the procedure assesses the causal sufficiency of each recipe—do cases with this combination of conditions constitute a consistent subset of cases with the outcome? That is, do these cases share the outcome? If consistency is high (e.g., at least 0.80), it is reasonable to evaluate coverage. If consistency is low, then the calculation of coverage is not interpretable (Ragin 2008: 55).

The immediate goals of a subset-superset analysis is to determine: (1) whether the initially hypothesized recipe is consistently linked to the outcome in question and, if so, to what degree is it linked, and (2) whether there is a more parsimonious version of the initial recipe that achieves comparable levels of consistency and coverage. Generally speaking, the consistency of a more parsimonious recipe is likely to be inferior to the consistency of the initial recipe. However, the difference may be small. The more important consideration is determining which of the subset recipes exceed the user's benchmark consistency threshold (e.g., .80 consistent) and whether they offer coverage that is superior to the initial recipe.

Comparison with Truth Table Analysis

The starting point of truth table analysis is the specification of the logically possible combinations of conditions identified as relevant to an outcome by the researcher. For example, suppose a researcher specifies four conditions as relevant to outcome y: a, b, c, and d. The truth table procedure assesses all 16 combinations of these four conditions, as follows:

~a•~b•~c•~d	a•~b•~c•~d
~a•b•~c•~d	a•b•~c•~d
~a•~b•~c•d	a•~b•~c•d
~a•b•~c•d	a•b•~c•d
~a•~b•c•~d	***a•~b•c•~d***
~a•b•c•~d	a•b•c•~d
~a•~b•c•d	a•~b•c•d
~a•b•c•d	a•b•c•d

The subset-superset procedure, by contrast, starts with only one of these sixteen combinations—the one that is in bold italics—and then explores simpler versions of the recipe.

An important part of truth table analysis is the consideration of "remainders"—combinations of causally relevant conditions that lack empirical instances. The three different ways of treating remainders provide the basis for the derivation of the "complex," "parsimonious," and "intermediate" solutions to the truth table. The complex solution bars the use of remainders. The parsimonious solution uses any remainders that it can, regardless of whether they might be considered "easy" or "difficult" counterfactuals, from the perspective of the researcher's theoretical and substantive knowledge. The intermediate solution uses only those remainders that are considered "easy" counterfactuals, again, from the perspective of the researcher's theoretical and substantive knowledge (see Ragin 2008: 160-175).

Like the parsimonious truth table solution, subset-superset procedure uses remainders in a knowledge-neutral manner. That is, there is no check on the use of remainders that might be considered difficult counterfactuals. Suppose, for example, that a researcher using the subset-superset procedure specifies the recipe a•~b•c•~d as sufficient for outcome y. As part of the analysis, the subset-superset procedure will also test the sufficiency of recipe a•~b•c. But suppose there are no cases of a•~b•c•d (i.e., no cases with greater than 0.5 membership in the combination), and our substantive knowledge tells us that a•~b•c•d would be a difficult counterfactual. From the perspective of the intermediate truth table solution, the test of the sufficiency of a•~b•c is not warranted.[9]

An Example of the Subset-Superset Procedure

Between World War I and World War II many democracies in Europe suffered setbacks. The rise of Hitler in Germany is a well-known and dramatic instance of the unravelling of a democracy, but there were other cases of democratic breakdown among the 18 European countries (e.g., Italy and Spain). The following example of the subset-superset procedure uses the breakdown of democracies in Europe between WWI and WWII as the outcome, coded as a fuzzy set. The four fuzzy-set conditions thought to be linked to democratic breakdown are political instability, low level of development, urbanized, and low level of industrialization. The researcher's hunch is that:

lindusfz*instablfz*ldevelfz*urbanfz → breakdown

where lindusfz = low level of industrialization, instablfz = political instability, ldevelfz = low level of development, and urbanfz = urbanized. Here's the initial output:

[9] Perhaps, based on case knowledge, a•~b•c might be considered a better initial recipe than a•~b•c•~d, and the whole issue of the difficult counterfactual a•~b•c•d, from this perspective, might be seen as a useless distraction.

Table 1: Initial Results of Subset-Superset Procedure

Recipe	Consistency	Coverage
lindusfz*instablfz*ldevelfz*urbanfz	1	0.142405
lindusfz*instablfz*urbanfz	0.971429	0.14346
lindusfz*ldevelfz*urbanfz	1	0.144515
lindusfz*urbanfz	0.94	0.148734
instablfz*ldevelfz*urbanfz	1	0.152954
ldevelfz*urbanfz	0.826816	0.156118
instablfz*urbanfz	0.980861	0.216245
urbanfz	0.381513	0.239451
lindusfz*instablfz*ldevelfz	0.897163	0.533755
lindusfz*instablfz	0.889845	0.545359
instablfz*ldevelfz	0.894017	0.551688
instablfz	0.901592	0.657173
lindusfz*ldevelfz	0.86511	0.703586
lindusfz	0.708377	0.722574
ldevelfz	0.837472	0.7827

The top row shows the consistency and coverage of the initial recipe, which combines all four conditions. It is perfectly consistent, but has very low coverage. The remaining rows show the results for subsets of conditions. The procedure presents the results in the form of a spreadsheet which can be manipulated in ways that enhance interpretability. For example, it is important to focus on the recipes that pass the investigator's consistency threshold (0.80). The rows can be sorted on consistency (descending) by clicking on the "consistency" column heading (twice), and then clicking on the first row that falls below the consistency threshold. Next, click the "Edit" menu; then click "Delete current row to last row." Only the recipes that pass consistency will remain. One further manipulation is needed—the remaining rows can be sorted according to coverage (descending). Simply click the "coverage" column heading (twice). The best recipes are now positioned at the top of the spreadsheet, as follows:

Table 2: Consistent Recipes Sorted on Coverage

Recipe	Consistency	Coverage
ldevelfz	0.837472	0.7827
lindusfz*ldevelfz	0.86511	0.703586
instablfz	0.901592	0.657173
instablfz*ldevelfz	0.894017	0.551688
lindusfz*instablfz	0.889845	0.545359
lindusfz*instablfz*ldevelfz	0.897163	0.533755
instablfz*urbanfz	0.980861	0.216245
ldevelfz*urbanfz	0.826816	0.156118
instablfz*ldevelfz*urbanfz	1	0.152954
lindusfz*urbanfz	0.94	0.148734
lindusfz*ldevelfz*urbanfz	1	0.144515
lindusfz*instablfz*urbanfz	0.971429	0.14346
lindusfz*instablfz*ldevelfz*urbanfz	1	0.142405

Notice that the initial recipe is now shown to be the worst recipe, due to its very low coverage (0.142). In fact, the bottom seven recipes all have poor coverage, ranging from 0.142 to 0.216. These seven rows also share the condition urbanfz, which suggests that urbanfz is not a key ingredient and should be dropped from the analysis. The six top recipes all have vastly superior coverage than the bottom seven. Coverage is especially strong for the top three recipes: ldevelfz, lindusfz*ldevelfz, and instablfz. Of course, lindusfz*ldevelfz is included in ldevelfz and is therefore redundant which yields the following solution, based on the three recipes with the highest coverage:

instablfz + ldevelfz → breakdown

The solution has a coverage of 0.8882 and a consistency of 0.8488.[10] It indicates that democracy unraveled in less developed countries and in countries that were politically unstable. Both conditions are sufficient.

[10] Membership in the solution was calculated using the fuzzyor compute function; consistency and coverage were computed using XYPlot, with membership in the outcome on the Y axis and membership in the solution on the X axis.

Conclusion

The subset-superset procedure offers researchers important tools for exploring and interrogating causal recipes. It is not a substitute for truth table analysis, which assesses all logically possible combinations of conditions, but can be used as an aid to truth table analysis. For example, it can be used to interrogate a recipe derived using truth table analysis. Furthermore, its deductive orientation makes it ideal for testing hypotheses about causal recipes, regardless of the source of the hypothesis. In the application just presented, the procedure revealed that the hypothesized recipe was too narrowly formulated. Instead of a relatively complex combination of four INUS conditions, the subset-superset procedure identified two sufficient conditions.

References

Gamson, William. 1990. *The Strategy of Social Protest*, second edition. Belmont, CA: Wadsworth.

Ragin, Charles C. 1987. *The Comparative Method: Moving beyond Qualitative and Quantitative Strategies*. Berkeley, CA: 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. 2008. *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago: University of Chicago Press.

Ragin, Charles C., and Peer C. Fiss. 2017. *Intersectional Inequality: Race, Class, Test Scores, and Poverty*. Chicago, IL: University of Chicago Press.

Ragin, Charles C. 2023. *Analytic Induction for Social Research*. Berkeley, CA: University of California Press.

Stokke, Olav Schram. 2004. "Boolean Analysis, Mechanisms, and the Study of Regime Effectiveness". In *Regime Consequences: Methodological Challenges and Research Strategies*. Edited by Arild Underdal and Oran R. Young. Dordrecht: Kluwer Academic.