Methodological Annex to "Explaining Government Preferences for Institutional Change in EU Foreign and Security Policy"

Mathias Koenig-Archibugi
London School of Economics and Political Science
July 2003

http://personal.lse.ac.uk/koenigar

The purpose of this note is to introduce fuzzy-set Qualitative Comparative Analysis, a novel method for social research that has been developed by Charles Ragin. The method has been used in my article "Explaining Government Preferences for Institutional Change in EU Foreign and Security Policy".[1]

At the basis of the method is the insight that patterns of causal necessity and sufficiency can be expressed in set-theoretic terms. For instance, to say that a having market economy is a *necessary* condition for a country to be democratic is equivalent to saying that, if a country is not in the set of countries with a market economy, then it will not be in the set of democratic countries. To say that being a democratic state is a *sufficient* condition for never being at war with another democratic state is equivalent to saying that, if a state is in the set of democratic states, then it will be in the set of states that are never at war with democratic states. This reasoning applies not only to single causes but also to combinations of causes, and this is especially important because many social-scientific theories consist of statements about the impact of combinations of causes. Consider for instance the thesis that sharing a democratic form of government *and* being economically interdependent is a sufficient condition for country A and country B (the dyad A-B) never to be at war with each other. This statement can be expressed in the following form: if the dyad A-B is in the set of

[1] For a full account cf. Ragin 2000.

democratic dyads *and* in the set of interdependent dyads (i.e. in the intersection of the two sets), then it will be in the set of peaceful dyads.

The set-theoretic conceptualization of causal necessity and sufficiency can be expressed in a more general form. Regarding necessity, whenever a causal condition is necessary (but not sufficient) for an outcome, instances of the outcome form a subset of instances of the causal condition. In the first example given above, the set of democratic countries is a subset of the set of countries with market economies. Regarding sufficiency, whenever a casual condition is sufficient (but not necessary) to an outcome, instances of the causal condition form a subset of instances of the outcome. In the second example given above, the set of democratic states is a subset of the set of states that are never at war against democracies.

There are two versions of QCA. The first is based on “crisp” sets, whereby a case is either in or out any given set. In other words, the outcome and the causal conditions are either present or absent (in the example above, a country is either in or out the set of democratic countries). The second version is based on “fuzzy” sets, whereby a case can still be fully in or fully out a set, but it can be also partially in. Membership in a set becomes a matter of degree. In the following I will first consider the tests of sufficiency and necessity in crisp-set QCA, and then discuss the additional contribution of fuzzy-set analysis.

The assessment of the necessity and sufficiency of a single cause in set-theoretic terms is a relatively simple task. It involves the cross-tabulation of the presence/absence of an effect against the presence/absence of the cause. If there are empirical instances where the effect (e.g. democracy) is present but the cause (e.g. market economy) is absent, then the test of necessity fails. Similarly, when there are empirical instances where the cause (e.g. democracy) is present but the effect (e.g. peaceful relations with other democracies) is absent, then the test of sufficiency fails.

However, the examination of a single cause is rarely the best analytical strategy in social research. This is because causality is generally complex, i.e. multiple (different causes can generate the same outcome) and conjunctural (the impact of a factor on the outcome depends on its interaction with other factors, i.e. what matters are combinations of causes). Most research questions therefore take into consideration several causes at the same time. Ragin has provided tools for assessing sufficiency and necessity under conditions of causal complexity.

Regarding the analysis of necessity, the possibility of several necessary conditions does not require a substantial departure from the procedure just mentioned: if an effect is present when a cause is absent, then the cause is never necessary regardless of whether other causes are present or absent. In other words, if a combination of causes is necessary for an outcome, then each of them is also necessary, and therefore the test of necessity focuses on the causes one at a time. This is different in the analysis of sufficiency. A cause may not be sufficient by itself, but it can be sufficient when other causes are present. A combination of causes might be sufficient (in the sense that whenever it is present, the effect is present) but each component cause might fail the test of sufficiency if examined one at a time. When analyzing sufficiency, therefore, the key is to examine all logically possible combinations (configurations) of the causes that the researcher wants to include in the analysis. These configurations are the fundamental analytic unit in QCA.

Using Ragin's Qualitative Comparative Analysis involves the following steps (an actual application is discussed below). As a preliminary but crucial step, the researcher selects the cases to be examined and defines the causal conditions that he or she considers especially relevant for the research question. Then the researcher assesses the necessity of each cause, by looking for cases where the outcome is present but the cause is absent. If there are such cases, then for that particular cause the test of necessity fails.

The analysis of sufficiency is somewhat more complicated. First, the researcher lists all logically possible combinations of the presence/absence of the causal conditions, as well as additional groupings that derive from merging configurations that share one or more attributes. (The number of possible groupings is $3^k - 1$, where k is the number of causal conditions examined).

Second, the researcher examines the cases conforming to each of these combinations and ascertains whether they agree in displaying the outcome.[2] If a grouping has no negative instances of the outcome (i.e. cases where the outcome is absent) and a certain number of positive instances of the outcome, the researcher may judge that it is sufficient for the outcome.

[2] This summary considers the "veristic" version of the sufficiency test, in which one single disconfirming case falsifies the claim of sufficiency. This is for simplicity and because this article conducts a veristic test.

However, some of the groupings that pass the sufficiency test are contained within other groupings and thus are logically redundant. Therefore, as the third step, the researcher uses minimization algorithms or a containment rule to simplify these groupings into a logical equation for the outcome of interest. This summary equation describes in a parsimonious way the causes or combination of causes that are sufficient for the outcome.

Finally, the researcher examines the equation and evaluates the simplifying assumptions that went into it. Simplifying assumptions are statements about the hypothetical outcome of combinations of causal conditions that do not occur in the population studied. They are a reflection of the limited diversity of naturally occurring social phenomena. On the basis of theoretical and empirical knowledge about the object of study, the researcher may assume that a given combination of causal conditions, if it had occurred, would have been sufficient for the outcome. Alternatively, the researcher might consider that a certain simplifying assumption is not plausible and decide not to include it in the solution.

To illustrate, I summarize how Alexander Hicks, Joya Misra, and Tang Nah Ng used QCA to explore the emergence of the welfare state.[3] They asked why by the 1920s some developed countries had consolidated income security programs and others did not. Of the fifteen countries they considered, six had such programs[4]. On the basis of the existing literature on the origins of the welfare state, they hypothesized that this difference is due principally to five crucial factors: (i) the mobilization of the working class through unions and socialist/labor parties, (ii) the presence of a patriarchal state with a traditional authoritarian ideology, (iii) the presence of a unitary (i.e. non-federal) democracy, (iv) government by Liberal parties, and (v) government by Catholic parties.

In applying QCA, Hicks *et al.* first determined for each country if those causal conditions were present or absent. They obtained a table whose rows display the presence/absence of the causal conditions and the presence/absence of the outcome for the fifteen countries. Then they eliminated redundancies by applying minimization algorithms to the table and obtained a parsimonious logical expression for the

---

[3]Hicks, Misra and Ng 1995.

[4]For the sake of brevity I report the analysis of the 1929 data only.

outcome.[5] This expression told the researchers that two combinations of causal conditions are sufficient for program consolidation: (1) the combination of working-class mobilization, patriarchal statism, and Catholic government in the absence of unitary democracy; (2) the combination of working-class mobilization, Liberal government, and unitary democracy. The first combination occurred in Germany and Austria, the second in Denmark, Sweden, United Kingdom, and New Zealand. There was no case of program consolidation in the absence of working-class mobilization, which therefore can be considered a necessary condition.

The analysis confirmed to the authors that there had been two historical paths to the formation of the welfare state: the first is the "Bismarckian" path of authoritarian traditionalism and paternalist social policies; the second is the "Lib-Lab" path where Liberal parties in democratic countries were able to pursue social reforms effectively thanks to the unitary character of their polity. The analysis also confirmed that both paths were taken in response to the increasing pressure of the organized working class. Overall, through their use of QCA the authors were able to identify one necessary condition and two sufficient combinations of conditions for the emergence of the welfare state within the specified domain (industrialized democratic countries before World War II).

In his recent work, Ragin has improved QCA by integrating it with fuzzy-set logic. As showed above, crisp-set QCA works with dichotomic variables (presence and absence of outcomes and causal conditions). The advantage of fuzzy sets is that they allow the use of QCA when the variables take more than two values. But fuzzy sets are considerably different from conventional ordinal or interval measurements. Fuzzy sets embody both qualitative states (full membership and full nonmembership in a set) and variation by level (degrees of membership between 0 and 1). In other words, fuzzy-set scores can be 1 (full membership), 0 (full nonmembership), or any number between 0 and 1 (partial membership). By combining qualitative and quantitative assessments in a single instrument, fuzzy sets help to deal analytically with the fact that social researchers, when dealing with an analytical category (e.g.

---

[5]The expression is: working-class mobilization • patriarchal statism • Catholic government • ~unitary democracy + working-class mobilization • Liberal government • unitary democracy → consolidation. The expression embodies some simplifying assumption on empirically unobserved combinations, which the authors considered plausible.

"great powers"), are often prepared to consider unambiguously some cases as members and some other as nonmembers of the category in question (e.g. the US is fully within the set of great powers while Luxembourg is fully out), but at the same time they might consider other cases as being neither fully in nor fully out (e.g. Italy).[6]

As noted above, causal necessity and sufficiency are best conceptualized in set-theoretic terms, and this makes fuzzy sets particularly useful for this kind of analysis. Recall that, whenever a causal condition is *necessary* (but not sufficient) for an outcome, instances of the outcome form a subset of instances of the causal condition. The basic idea behind the fuzzy-set analysis of necessity is: when fuzzy membership scores in the outcome are less than or equal to fuzzy membership scores in the causal condition, then instances of the outcome can be considered a subset of instances of the causal condition. Therefore, when a researcher finds that in all cases fuzzy membership scores in the outcome are less than or equal to fuzzy membership in the cause, then he or she can cite this as evidence that the cause is necessary for the outcome.

Recall also that, whenever a casual condition is *sufficient* (but not necessary) to an outcome, instances of the causal condition form a subset of instances of the outcome. The basic idea behind the fuzzy-set analysis of sufficiency is: when fuzzy membership scores in a causal condition or combination of causal conditions are less than or equal to fuzzy membership scores in the outcome, then instances of the causal condition or combination of causal conditions are a subset of instances of the outcome. Therefore, when a researcher finds that in all cases fuzzy membership scores in the cause or causal combination are less than or equal to fuzzy membership scores in the outcome, then he or she can cite this as evidence that the cause or casual combination is sufficient for the outcome.

The basic steps of fuzzy-set analysis are similar to the application of crisp-set QCA as it has been summarized above. The main differences are the following:

1. The researcher does not determine the presence or absence of the outcome and the causes for each case, but assigns to each case a fuzzy membership score in the outcome and in the causal conditions. The translation of raw data into fuzzy membership scores is a crucial aspect of the method and depends ultimately on the

[6] See Cioffi-Revilla 1981 on fuzzy-set logic applied to international relations theory.

substantive knowledge and judgment of the researcher. The allocation of scores to cases must be transparent and allow other researchers to form an opinion about its plausibility.

2. In analyzing necessity, the researcher looks for causal conditions with membership scores that are consistently greater than outcome membership scores. If there is a causal condition where this happens in all cases, then this condition passes the test of necessity. Suppose that the researcher hypothesizes that a market economy is a necessary condition for democracy. She constructs a population of countries and plots each individual causal condition against the outcome. If she finds a pattern such as in Figure 2, the thesis of necessity is supported. In this hypothetical situation, all cases are below the diagonal, indicating that the membership scores of the causal condition are consistently greater that the membership score of the outcome. The outcome is therefore a subset of the causal condition, which is the set-theoretic way to express necessity.

Figure 2 about here

3. In analyzing sufficiency, the researcher compares the membership scores of the outcome not only with the score of each individual cause, as in the analysis of necessity, but also with the scores of all possible causal expressions.[7] The membership of each case on each causal expression is computed using the fuzzy-set operations of intersection and negation. Intersection means that the score of a combination of causal conditions equals the lowest among the scores of the causal conditions in the expression.[8] Negation means that the researcher examines not only the membership of a case in set A, but also its membership in set non-A,

[7]The causal expressions correspond to all logically possible combinations of the causal conditions. Their number is $3^k - 1$, where k is the number of causal conditions included in the analysis.

[8] For instance, if a country has a membership score of 0.80 in the set of democratic countries and a membership score of 0.60 in the set of countries open to international trade, then it has score of 0.60 in the set of countries that are democratic *and* open.

where “fuzzy membership in set not-A” equals “1 - fuzzy membership in set A”.[9] Suppose that the researcher hypothesizes that the combination of democracy and openness to international trade is a sufficient condition for a country to be pacific, i.e. not willing to wage a war of aggression. The researcher plots each causal expression against the outcome and hopes to find a pattern such as in Figure 3. All cases are above the diagonal, indicating that the membership scores of the outcome (peacefulness in the example) are consistently greater that the membership scores of the causal expression consisting of the intersection of the causal conditions (democracy and openness). The causal expression is therefore a subset of the outcome, which is the set-theoretic way to express sufficiency.

Figure 3 about here

Like in crisp-set QCA, the researcher eliminates those expressions that pass the test of sufficiency but are logically redundant and obtains a logically parsimonious statement of causal sufficiency. This statement is then evaluated in terms of any simplifying assumption it incorporates and in light of the theoretical concerns.

References

Cioffi-Revilla, Claudio A. 1981. Fuzzy Sets and Models of International Relations. *American Journal of Political Science* 25 (1): 129-159.

Hicks, Alexander, Joy Misra, and Nah Ng Tang. 1995. The Programmatic Emergence of the Social Security State. *American Sociological Review* 60 (3): 329-350.

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

[9] For instance, if a country has a membership score of 0.80 in the set of democratic countries, it has a score of 0.20 in the set of “not democratic countries”.