Q

Qualitative Comparative Analysis

Malcolm J. Beynon
Cardiff University, UK

INTRODUCTION

This article concerns itself with qualitative comparative analysis (QCA), introduced in Ragin (1987), it is a technique that attempts to identify the causal relationship between variables and some outcome. QCA is employed in comparative case-oriented research, for studying a small-to-moderate number of cases in which a specific outcome has occurred, compared with those where it has not. Unlike conventional statistical analysis, however, QCA does not ask about the independent effect of a variable on the likelihood of an outcome. Rather, it considers configurations of values on the independent variables as cases (Ragin, 1987; Kitchener, Beynon, & Harrington, 2002).

The central goal of QCA is to mimic some of the basic analytic procedures that comparative researchers use routinely when making sense of their cases. The key difference between QCA and traditional case-oriented methods is that with QCA it is possible to extend these basic analytic procedures to the examination of more than a handful of cases (Ragin & Rihoux, 2004).

Achen (2005) in a celebration of the surrounding attitude of Ragin's analysis demonstrated in QCA, recreates his argument, summarising it in three claims:

- **Claim 1:** Case-study analysis has a methodology of its own, and there is much to learn from it that can be learned in no other practical way;
- **Claim 2:** Quantitative research, as most researchers practice, does not match the character of serious case-study methods and, more importantly, does not match social reality; and
- **Claim 3:** In light of Claims 1 and 2, conventional quantitative methods should be de-emphasized in favor of a new methodology that Ragin has developed, QCA, which aims to replicate the logic of case study analysis in a mathematical framework different from statistical theory.

These claims set QCA apart from traditional decision support methodologies. Its peripheral role in decision making and support is exemplified in applications that have included social policy development (Hicks, 1994), labor management practices in textile mills (Coverdill & Finlay, 1995), and managing barriers to the diffusion of the home and community-based services waiver program (Kitchener et al., 2002), as well as more diverse applications such as regulating biomedicine in Europe and North America (Varone, 2006). In summary, in 2004, Ragin and Rihoux (2004) identified around 250 reference applications of QCA, for a comprehensive list of applications see, http://www.compasss.org/Bibli%20database.htm.

QCA uses Boolean algebra to implement a mode of logical comparison through which each case is represented as a combination of causal and outcome conditions (Ragin, Mayer, & Drass, 1984). Based on a conceptual model, the method identifies different logical combinations of variables, using AND (*) or OR (+) expressions, which might be necessary and/or sufficient to produce the outcome. One feature of QCA analysis is the option to include the use of 'remainders', combinations of causal condition variables that are not represented by specific cases.

Schneider and Wagemann (2006) offer an informative discourse on the relationship between causality and concomitant necessary and/or sufficiency, including the exposition; a cause is defined as necessary if it must be present for a certain outcome to occur, in contrast, sufficiency is present if, whenever we see the cause, then we also see the outcome. Following these and other definitions, necessity and sufficiency statements lead to the use of set theoretic relations as indicated by the 'if .. then ..' structure. It is thus possible to represent and think about necessity and sufficiency by making use of set theoretic approaches such as Boolean algebra and fuzzy sets.

The application of QCA in a fuzzy environment was developed in Ragin (2000), with fs-QCA, whereby there is continuous measurement in which cases are coded on a 0 to 1 scale, according to their degree of membership in a particular set (such as those cases associated with an outcome). Lieberman (2005) considers how

Ragin, through fs-QCA, attempts to steer a middle path between "quantitative" and "qualitative" research. A multi-valued version of QCA was presented within TOSMANA (Cronqvist, 2005). Caren and Panofsky (2005) furthered QCA in a temporal perspective, acknowledging the temporal order in which variables occur might provide as much information as does the interaction of these variables. Comparisons with regression based analyses have also been undertaken (Seawright, 2005).

BACKGROUND

The operational rudiments of QCA are based on set-theoretic relations, Ragin (2000) argues that set-theoretic relations are asymmetrical and concern explicit connections between theory and analysis (see also Ragin, 2006). The structure of QCA surrounds the ability to elucidate the causes (condition variable values) that associate cases (such as people or countries) to a considered outcome. Ragin and Rihoux (2004), under the guise of best QCA practices, exposits a 'technical and procedural pieces of advice' for its utilization, summarized in the next 13 points:

1. Select cases in a rigorous manner. The way cases are selected should be stated explicitly.
2. To the extent possible, develop an "intimacy" with each case.
3. Select the condition variables in a rigorous fashion, in a theoretically and empirically informed way. Do not select too many condition variables. It is best to focus on conditions that seem decisive from the perspective of either substantive or theoretical knowledge.
4. When the raw data is quantitative and the *N* is not too large, display the data in tables so colleagues can test other operationalizations of your conditions.
5. When using conventional (crisp) sets, explain clearly how each condition is dichotomized. Justify the placement of the 0/1 threshold on empirical and/or theoretical grounds.
6. If possible, display the truth table and indicate which observed cases correspond to each combination of conditions.
7. If the truth table contains contradictory configurations, resolve them.
8. Proceed systematically to four analyses: those for the configurations with a positive outcome (coded 1), first without and then with the inclusion of remainders; and then analyze the configurations with a negative outcome (coded 0), first without and then with the inclusion of remainders. In order to do so, quite naturally, cases with a "0" outcome and cases with a "1" outcome should be included in the research.
9. The analysis should be done with software using the Quine-McCluskey algorithm and not by hand.
10. Resolve any "contradictory simplifying assumptions" that may have been generated in the process of minimization with the inclusion of remainders.
11. Provide some information (even in a shorthand manner) about the main iterations of the research (back to cases, back to theories, fine-tuning of the model, etc.).
12. At the end of each truth table analysis, report all combinations of conditions linked to the outcome.
13. Proceed to a real "return to the cases" (and/or theory, depending on research goals) at the end of the analysis, using the truth table solution as a guide.

Needless to say, there is no *perfect* QCA analysis. In real life research, some of these best practices are not easy to implement, and indeed they could be time-consuming.

Further, expansion of some of the points relating to the summary processes involved with QCA is next presented. As mentioned, the utilization of QCA requires the choice of case variables, Amenta and Poulsen (1994) describe a number of different techniques for choosing such variables. This is a nontrivial act, since the researcher should keep the subsequent model as parsimonious as possible. A final step in QCA is to minimize the equation, in the sense of producing a final equation that is logically equivalent to the original but that has the fewest possible additive terms, each of which consists of the fewest possible multiplicative terms.

QCA uses the Quine-McCluskey method to minimize the equations formed from a truth table, as discussed in Ragin (1987). The Quine-McCluskey method has two components. First, additive terms that differ only by having opposite values of one multiplicative variable