

## Introduction to QCA

Wendy Olsen

Qualitative comparative analysis is a methodology whose acronym originated in the writings of Charles Ragin (1987, 2000). This school of analysis has earlier roots in substantive social research. Comparative social research has a long history, including Theda Skocpol's substantial studies of comparative historical trajectories. QCA also has roots in the work of JS Mill. Mill argued that the presence of a factor in a series of cases, all leading to an outcome of interest, suggests that that factor positively contributes to Y and may be a cause of Y. However it could be an accidental association and a spurious finding.

The meaning of Mill's own work in this area is contested, because both empiricists and substantivists can refer to Mill as a source of advice. Empiricists tend to argue that if the two factors occur together, and if X's absence is associated with Y's absence, and Y follows X temporally, then you have proof of an association. This school is known as the deductive-nomothetic (law-seeking) school. Other empiricists argue that all we can know relates to X and Y, not to any real underlying causes, since X and Y reflect all that is measurable. Such perfectionists are often also atomistic empiricists, and atomism relates to the assumption that you can break down every social and physical thing into its components parts in order to understand it. Atomists tend to ignore the multi-level, non-nested stratified nature of society, and instead to focus only on the data to hand. They dislike speculation and tend to say that X and Y are simply associated (not assuming a causal relation, which they may say is ultimately un-knowable).

To adduce that an X is necessary as a cause of Y, we need cases where X is present and Y is present, and at [www.compass.org](http://www.compass.org) you can obtain software which helps you to do that. The QCA urges you to use substantive comparative aspects of the objects you are studying – that is, you create your own data set. If there is a configuration where X is absent and Y is absent, and another where X is present and Y is present, then you can begin to explore whether X causes Y. But the multi-dimensional case is much more complicated. If we have X's 1 to 20, and one main outcome Y, then there are thousands of possible configurations of the X's and Y. Each has a frequency, and some would imply that X3 is necessary for Y whilst others would imply that X5 is sufficient for Y. 'Sufficiency' is defined as Y occurring if X5 occurs. 'Necessity' is defined as Y occurring only if X3 occurs. You can use software to find out whether, in cases of multiple trajectories that lead to a single outcome such as 'getting well', there are complex causes like:

If (X5 but NOT X3 ) OR (X7 AND X9) then Y

The if-then interpretation is a causal interpretation of an association. You must take the time element into account. You must carefully avoid spurious associations. You can use retroduction to avoid making interpretive mistakes. Retroduction means asking what must be the case in order for this configuration to have occurred. The configuration overall (i.e. the set of configuration which actually occurred) is THE CASE within which we have numerous 'cases' that are comparable. The whole is a complex social whole. Skocpol says we must see that macro-causal forces are at

work, which are unique and date-specific. Ragin argues that we can benefit from parsing out the multiple configurations leading to Y.

Ragin also argues that NOT-Y is interesting. These non-instances are cases where Y=0 or 'no'. We also code the cases which are instance as Y=1. The causal analysis of NOT-Y is symmetric in crisp-set QCA but becomes non-symmetric in fuzzy set social science (Ragin, 2000).

Fuzzy sets refer to planning a labelling scheme that is numeric but ordinal, such that:

0 refers to 'not in this set'  
low numbers below .5 refer to 'not really in this set'  
0.5 represents neither in nor out of the set;  
higher numbers represent 'not fully in but more in than out of this set'  
1.0 refers to 'fully in this set'.

Ragin gives the example of poor countries, which are only a subset of the low-income countries. One might use multiple indicators to generate a fuzzy set coding scheme. In SPSS we call the resulting mini-programme an 'algorithm'. It solves your problem of the fuzzy coding whilst also producing an ordinal variable.

Ragin argues that we use Boolean logic to multiply together fuzzy sets – not multiplicative logic or additive separability as in linear regression. Thus, configurations of fuzzy sets have more permutations than configurations of crisp sets.

The software for analysing fuzzy sets is FS/QCA. See [www.compass.org](http://www.compass.org) for working papers and see [www.fsqca.org](http://www.fsqca.org) to download the software and manuals. We strongly recommend Ragin (1987 or 1994) as a starter for QCA and Ragin (2000) as a starter for FS/QCA. The first half of Ragin (2000) re-iterates the QCA arguments.

## References

- Ragin Charles, C. (1987). *The comparative method : moving beyond qualitative and quantitative strategies*. Berkeley, University of California Press.
- Ragin Charles, C. (1991). *Issues and alternatives in comparative social research*. Leiden, Brill.
- Ragin, C. C. (1994). Constructing social research : the unity and diversity of method. Thousand Oaks, CA; London, Pine Forge Press Sage.  
--Ragin (1994) makes a very good textbook for dissertation writers.
- Ragin, C. (2000) *Fuzzy Set Social Science* Chicago: University of Chicago Press
- Skocpol, T. and A. Conference on Methods of Historical Social (1984). Vision and method in historical sociology. Cambridge [Cambridgeshire] ; New York, Cambridge University Press.

9/9/04

[www.durham.ac.uk/case.2004](http://www.durham.ac.uk/case.2004)