

From Niche to Mainstream Method? A Comprehensive Mapping of QCA Applications in Journal Articles from 1984 to 2011

Benoît Rihoux¹, Priscilla Álamos-Concha¹, Damien Bol²,
Axel Marx³, and Ilona Rezsöhazy

DOI: 10.1177/1065912912468269c

Abstract

This article provides a first systematic mapping of QCA applications, building upon a database of 313 peer-reviewed journal articles. We find out that the number of QCA applications has dramatically increased during the past few years. The mapping also reveals that csQCA remains the most frequently used technique, that political science, sociology, and management are the core disciplines of application, that macrolevel analyses, medium-*N* designs, and a mono-method use of QCA remain predominant. A particular focus is also laid on the ratio between the number of cases and number of conditions and the compliance to benchmarks in this respect.

Introduction

Over the past few years especially, it seems that different variants of QCA have been increasingly discussed and applied, in a broad range of disciplines and around different substantial topics. So far, no systematic mapping of the QCA literature has been produced, apart from a few recent articles with a more focused topic in the fields of management studies (Chanson et al. 2005; Marx et al. 2012) or public policy analysis (Rihoux, Rezsöhazy, and Bol 2011). One exception is a recent article by Jordan et al. (2011), which provides a first attempt toward an inventory of QCA-related journal articles. Performing a very inclusive search through Web of Science, they identified 338 published journal articles referencing Ragin's (1987) seminal book *The Comparative Method*. According to their own coding, and expectedly, they established that the majority of these articles were in the fields of sociology (33%) and political science (government and public administration: 30%). They also found out that the number of articles per year increased quite significantly from 2004 onward. Arguably, though, a significant proportion of those articles do not contain a full-fledged QCA application.

The ambition of this contribution is to move beyond this broad inventory and to provide a first comprehensive mapping¹ and analysis of QCA applications focusing specifically on published peer-reviewed journal articles. To

this end, in a collective effort, we have updated the COMPASSS international bibliographical database.² First, conducting an extensive search both through bibliographical databases and requests to the core group of scholars in the COMPASSS network, we have gathered all peer-reviewed articles containing one QCA application.

For each publication, apart from the generic elements of the bibliographical references, we have also compiled the abstract, and systematically coded some core features on the basis of the full-text file: the type of QCA technique(s) used, the scientific discipline(s), the level(s) of analysis, the number of cases used for the QCA, the number of conditions, and the combination with other methods, if any. This specific set of features has been selected on two main grounds. On one hand, some of these have regularly been used in previous methodological mapping exercises (e.g., Bennett, Barth, and Rutherford 2003, Franchino 2005). On the other hand, QCA was initially designed (Ragin 1987) to handle small- and intermediate-*N* research situations, macrolevel cases (countries), in the fields of historical sociology and comparative politics, requesting a relatively low number of conditions, and not explicitly to be used in combination with other methods. Our overall hypothesis is that, given

¹Université catholique de Louvain, Louvain-La-Neuve, Belgium

²University of Montreal, Quebec, Canada

³University of Leuven, Belgium

the growth of the field and diversification of QCA users, QCA applications have moved away from Ragin's initial design. By concentrating on these features,³ we will be able to assess this hypothesis.

Arguably, the COMPASSS international bibliographical database has by now achieved a quite exhaustive inventory of the field in terms of publications, especially with regard to journal articles. At the time of writing, putting a cap on December 31, 2011, we have gathered 313 peer-reviewed journal articles with QCA applications (as defined in the next section)—we can estimate that our coverage for these is close to 100 percent. We have also identified and compiled 35 other journal articles containing a discussion about QCA (without a full-fledged application), 56 full books and 75 chapters in edited books discussing and/or applying QCA, and 279 publications or manuscripts in other formats⁴. Our coverage for these various other documents apart from our 313 core articles is difficult to estimate—probably between 70 and 80 percent, and diminishing during the past few years given the disciplinary diversification of QCA-related work (see Section 2). Overall, we have thus identified 750 items, out of a total of probably 900 to 1,000 as of late-2011. Considering these various formats of publications, our main observation is that the broader use of QCA started to pick up quite late after the publication of Charles Ragin's seminal volume (1987)—only from the late-1990s onward on the whole, and only from 2003 to 2004 onward for journal articles.

In the remaining part of this contribution, we have chosen to focus on the 313 full-size articles published in double-blind peer-reviewed journals, which contain at least one QCA *application*, and that are available in full-text format for coding and verification purposes. We chose to adopt a rather inclusive definition of an "application": any form of data processing with a QCA technique (crisp-set, multivalued, and/or fuzzy-set QCA), either on real-life data or with constructed data or replicated analyses. The group of articles that simply contains discussions of various aspects of QCA, without a real-life or constructed application, is a much broader and less precise group that lies beyond the scope of our contribution.

The vast majority of those 313 articles are in English, a small remaining minority being published mainly in French, German, and Japanese. In the next sections, we systematically survey this *corpus* of articles with regard to some of their core features outlined above.

Which Technique? Crisp-Set, Multivalued, or Fuzzy-Set QCA?

Figure 1 indicates, on the whole, that the overall evolution of published QCA applications as journal articles has followed five periods.

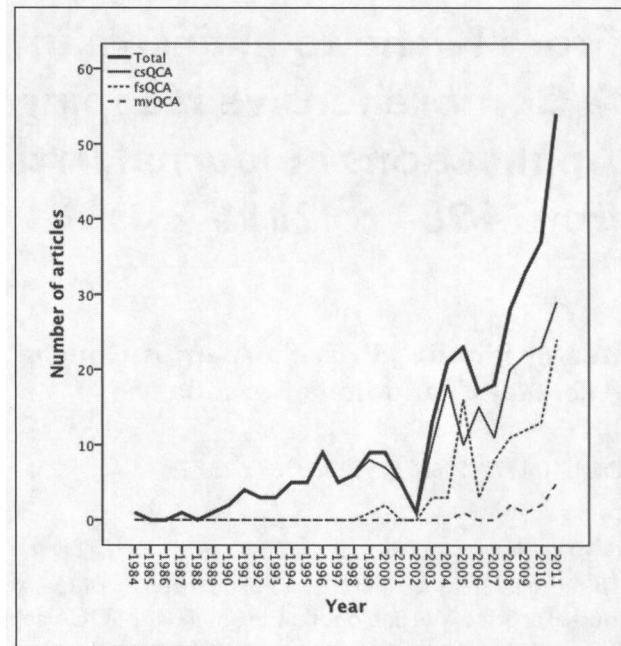


Figure 1. Number of articles by QCA technique, 1984–2011.⁵

The first period (1984–2000) is one of modest growth, almost exclusively corresponding to csQCA applications. The second period (2001 and 2002) is one of decline. Apparently, one cannot explain this decline by a long-lasting decline of interest for csQCA, as the use of this technique grew again in the next periods (see below). Two hypotheses can be put forward. On one hand, this could have been the result of Charles Ragin's (2000) second agenda-setting book on fuzzy sets: in the wake of the publication, researchers might have wanted to explore the possibilities of fuzzy sets and put the use of csQCA on hold. On the other hand, and in conjunction, the 2000 book was not accompanied by the quick availability of a complete software operating under windows; one had to wait until 2004 to 2005 to have full versions of the FS/QCA software.

The third period (2003–2005) is one of expansion, first mostly due to csQCA applications (2004) and then to fsQCA applications (2005). A closer look at 2005 articles shows that this first wave of fsQCA applications corresponds to a growing diversification in disciplinary terms (see also Figure 2 below).

The fourth period (the year 2006 specifically) is one of a second, less sharp decline, mostly not only due to a sharp decrease of fsQCA applications but also due to a stagnation in csQCA applications between 2005 and 2007. One explanation could be that the first more complete versions of both FS/QCA and TOSMANA (the two most often used software so far) only became available in

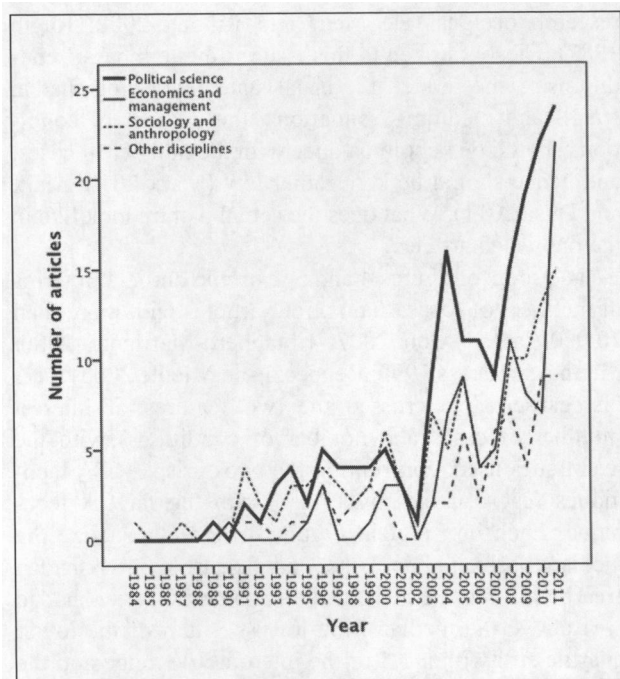


Figure 2. Number of articles by discipline, 1984–2011.⁶

2004 to 2005, producing a second “on hold” effect. In addition, no textbook that would have made QCA accessible to a broader public of researchers was yet on the market.⁷

Finally, the recent period (2007–2011) is one of sharp increase of both csQCA and fsQCA, as well as a first upward trend for mvQCA. By then, many elements were in place: a first English-language textbook (Rihoux and Ragin 2009) covering the different techniques,⁸ various software options (including gradual developments in R and STATA), and a broader training offer through a broader pool of qualified instructors.

On the whole, the most frequently used technique is still csQCA, that is, the technique that was first developed. It still amounts to 72 percent of applications, compared with 30 percent for fsQCA and only 3 percent for mvQCA that remains rather marginal at this stage. It is worth noting that, at this stage, the rougher binary csQCA has not been taken over by the finer-grained fsQCA. One reason may lie in the fact that quite some comparative researchers see distinct value in using dichotomous variables for analytic purposes (Collier and Adcock 1999; De Meur and Rihoux 2002; Grofman and Schneider 2009)—indeed, the resort to dichotomies provides the simplest operationalization of *relevant difference* and therefore provides strong analytic leverage in terms of parsimony (De Meur and Rihoux 2002, 151–52).

Which Discipline?

Over the whole period, the three core disciplinary fields of QCA applications are, respectively, political science (51%; more specifically: comparative politics and comparative policy analysis), sociology and anthropology (34%; specifically macrohistorical sociology, welfare-state studies, and sociology of labor and organizations), and economics and management studies (26%; mostly management studies in fact). Note that the total is higher than 100 percent as some articles are coded into more than one discipline. All other disciplines are still rather marginal at this stage, as they each account for less than 5 percent of the applications, though they together represent as many as 18 percent of all applications. These cover many fields, such as (in decreasing order of frequency): psychology and education studies, legal studies and criminology, health sciences, applied sciences, demography and development studies, history, geography, and philosophy.

The two predominant disciplinary fields mentioned above, that is, political science and sociology, correspond as it were to the disciplinary domains of Charles Ragin himself, and hence, one may assume that the spread of QCA occurred, at least until the late-1990s, along those specific disciplinary networks. This assumption can be tested by looking at Figure 2, which displays the evolution on a yearly basis.

Indeed, from 1984 (the very first application, in *American Sociological Review*, by Ragin, Mayer, and Drass) to 1995, applications correspond almost exclusively to political science and sociology, the only exceptions being early articles in the field of criminology, the discipline of the first programmer of QCA software, the late Kris Drass.

By contrast, the whole period after 2002 until now is one of disciplinary diversification. In particular, management studies have been the fastest-growing field of published applications in the most recent period—not surprisingly because QCA appears particularly well suited for many core management studies topics (Marx, Cambré, and Rihoux 2012). The frequency of articles in other diverse disciplines is also rising gradually over the past few years. Note, however, that the predominant field is still political science, also during the past few years—a closer look at the articles demonstrates that this corresponds first and foremost to a sharp increase of applications in the subdiscipline of policy analysis (also highlighted by Rihoux, Rezsöházy, and Bol 2011).

Which Level of Analysis?

As mentioned above, Ragin’s seminal work (1987) chiefly dealt with macrolevel cases such as political

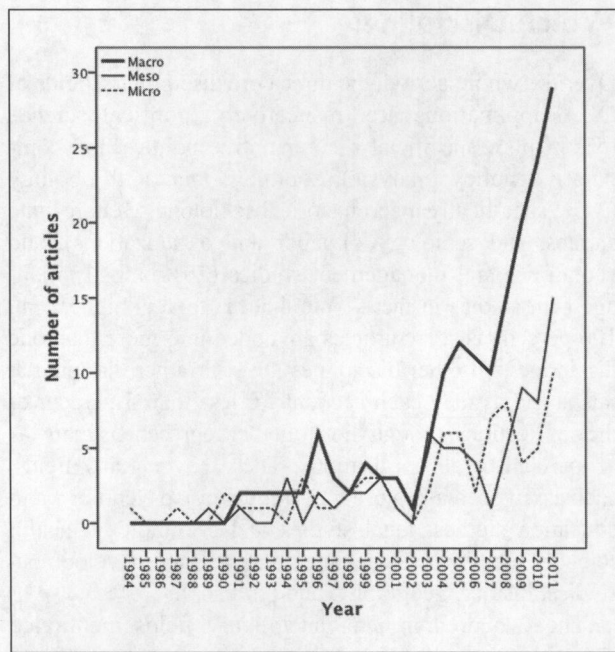


Figure 3. Number of articles by level of analysis, 1984–2011.⁹

systems or welfare states. If we consider the whole period, however, while the proportion of macrolevel applications (49%) is indeed quite high, there are also 21 percent mesolevel and 21 percent microlevel applications. The picture over the whole period is in fact somewhat more complex.

Figure 3 shows first that, until 2004, there were no systematic differences in the respective proportions of micro-, meso- and macrolevel applications. In fact, there were some microlevel applications from quite early on (e.g., Abell 1990; Drass and Spencer 1987; Ragin, Mayer, and Drass 1984; Williams and Farrell 1990). Second, the real differences are observed from 2004 onward, with the sharp rise of macrolevel applications, which have thus become predominant over the past few years. Again, this is strongly linked with the sharp increase of macrolevel—that is, mainly cross-country—comparisons in the field of policy analysis (Rihoux, Rezsöházy, and Bol 2011). Third, and last but not the least, meso- as well as microlevel applications have also been on the rise in the most recent period (2010–2011). In other words, the recent disciplinary diversification, in particular in the field of management and organization studies, also goes along with nonnegligible proportions of micro- and mesolevel analyses.

How Many Cases, How Many Conditions?

QCA has been often presented, initially at least, as an approach and set of techniques geared toward medium-*N*

research designs (De Meur and Rihoux 2002; Ragin 1987). Closely linked to this issue is the number of conditions in the model, the usual statement being that in small- and medium-*N* situations, the number of conditions should be kept in balance with the number of cases, and hence should be kept rather low (Marx 2010; Marx and Dusa 2011). What does the actual picture look like in the published articles?

To produce Figures 4 and 5, extreme outliers in terms of numbers of cases, that is, four articles with more than 20,000 cases (Alon 2007; Chiappero-Martinetti 2000; Miethe and Drass 1999; Regoeczi and Miethe 2003) were disregarded. We crossed the two features of interest (number of cases and number of conditions) with the techniques used, confronting the two “crisp” QCA techniques (csQCA and mvQCA¹⁰) with the fsQCA technique. Each box plot in Figures 4 and 5 visualize the median (horizontal line), the second quartile (gray-shaded area below the median), the third quartile (gray-shaded area above the median), the lowest datum of the lower quartile still within 1.5 of the interquartile range and the highest datum of the upper quartile still within 1.5 of the interquartile range (the two extremities of the whiskers), and the “outlier” values¹¹ if applicable.

With regard to the number of cases, our expectation was that fsQCA applications would tend to be larger-*N* than crisp (csQCA or mvQCA) applications. The two box plots in Figure 4 show that this is not the case: the median is almost equal (22 and 23 cases for fsQCA and cs/mvQCA, respectively), and the quartiles do not differ so much, apart from a slightly higher proportion of fsQCA applications toward medium-to-large-*N* cases. The two distributions do not differ significantly in statistical terms.

To obtain a more dynamic view, with regard to the number of cases, we have established a simple distinction between “small-*N*” (less than 10 cases), “medium-*N*” (between 10 and 50 cases) and “large-*N*” (more than 50 cases) designs, and examined trends over time (Figure 6). Over the whole period, the respective proportions of small-, medium-, and large-*N* applications amount to 12, 60, and 28 percent. Thus, medium-*N* applications are on the whole much more frequent. If we consider the larger-*N* spectrum of applications, there is quite a spread of designs: for instance, 6 percent of applications between 100 and 199 cases, 3 percent between 200 and 499 cases, and 8 percent above 500 cases. In other words, roughly one-tenth of all QCA applications so far are using numbers of cases that would usually be associated with statistical analyses.

Figure 6 shows that, until 2006, there was no systematic predominance of one type of research design, with sizable shares of both medium-*N* and large-*N* applications, while small-*N* applications remained almost nonexistent. By contrast, from 2007 onward, medium-*N* applications

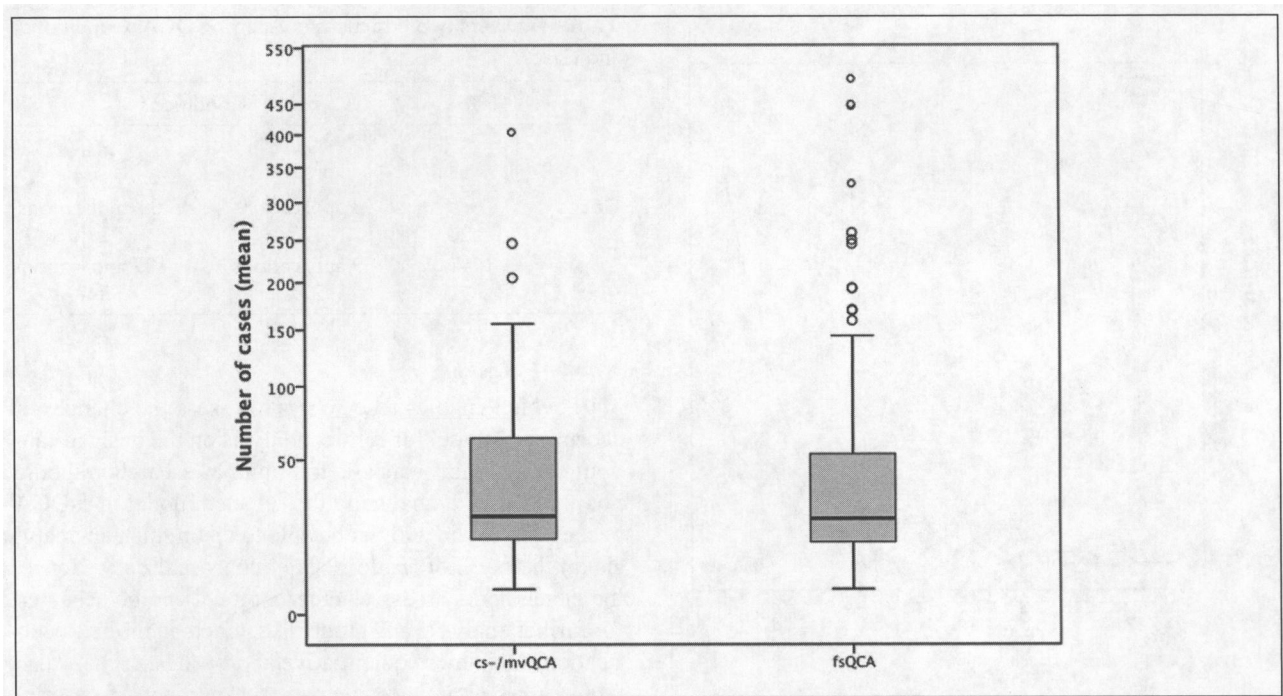


Figure 4. Number of cases by technique.¹²

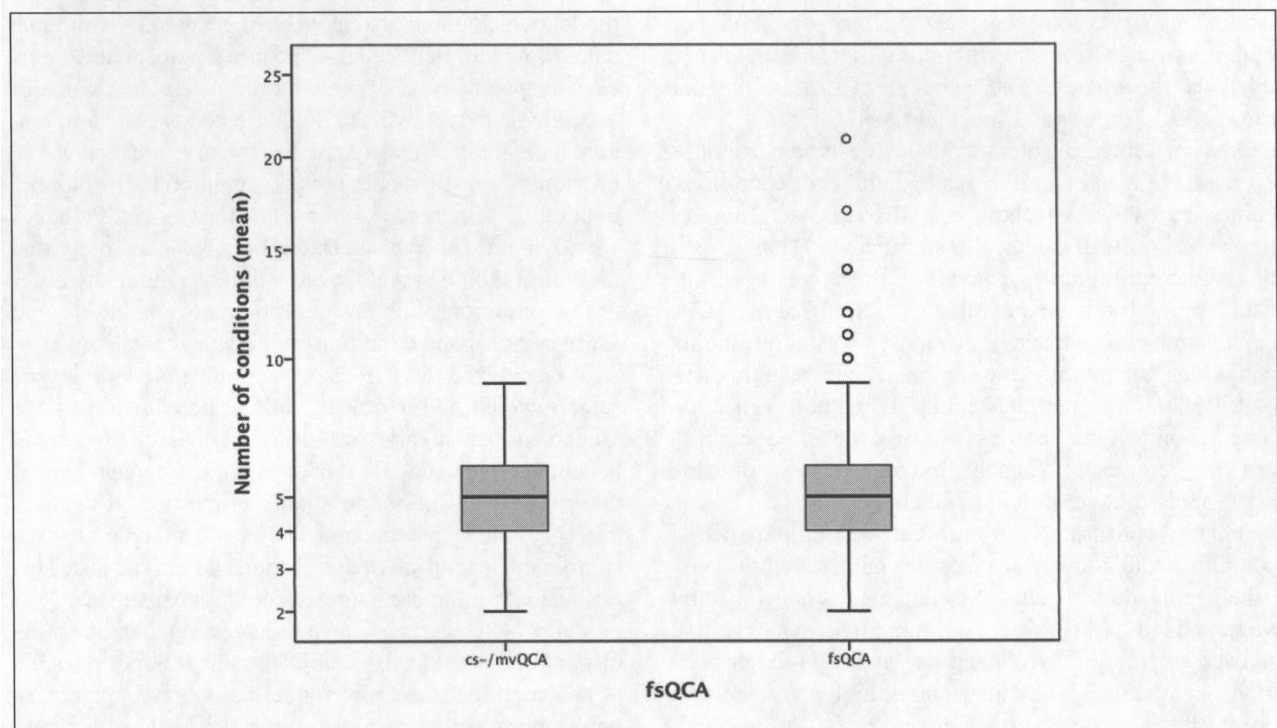


Figure 5. Number of conditions by technique.¹³

have clearly become predominant, in line with Ragin’s initially identified “niche” for QCA, even if two other “niches” of large-*N* and small-*N* applications have also been developing to some extent.

As for the number of conditions (Figure 5), the differences between crisp and fuzzy variants of QCA are almost nonexistent (apart from a few csQCA applications with very high number of conditions): the median is

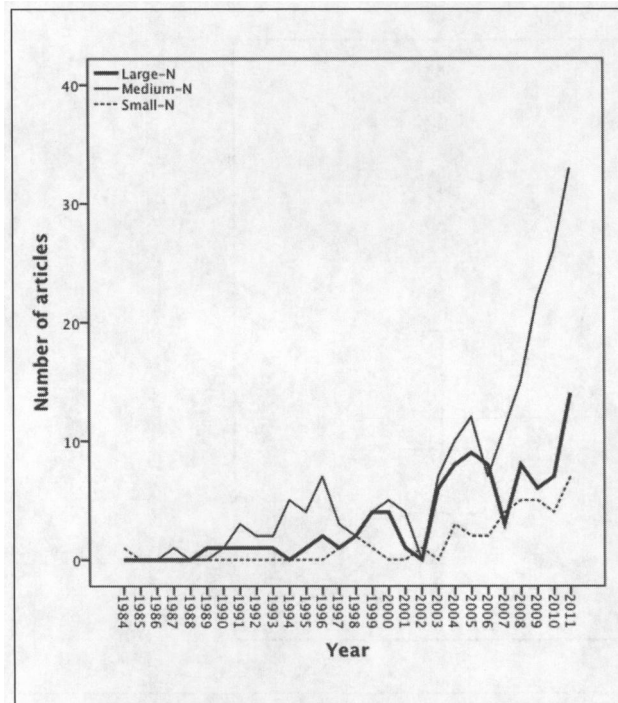


Figure 6. Number of articles by research designs, 1984–2011.¹⁴

equivalent (five conditions) as well as the 25th and 75th percentiles (four and six conditions, respectively). In other words, whatever the technique, the usual practice ranges between four and six conditions.

We performed different statistical tests, for example, correlating the number of cases by number of conditions, controlled by QCA technique, and over time. None of these tests proved statistically significant.¹⁵ Hence, over the whole period, and whatever the technique used, the usual practices have not changed significantly. One explanation could be that, if one moves way above the six conditions, the interpretation becomes increasingly difficult in theoretical and conceptual terms, not to mention other technical difficulties such as the exponentially growing number of “logical remainder” cases and the associated problem of limited diversity.

Finally, concerning the ratio between the number of conditions and number of cases, we have tested the published applications against the benchmarks developed by Marx and Dusa (2011; see also Marx 2010). Marx (2010) showed that csQCA generated explanatory models (i.e., models with no contradictions and/or high consistency¹⁶) on the basis of random data if one does not take into account a balance between the number of conditions and cases, with an upper limit to the number of conditions. This is troublesome because the researcher cannot distinguish whether a model has any explanatory power, as it might also be accepted on the basis of random data. Marx and Dusa (2011) developed a benchmark table,¹⁷ for each combination of conditions (up to 13) and cases (up to

Table 1. Diversity of Practices in Applying QCA: Conditions and Cases.¹⁹

		Conditions	
		Few ≤4	Many > 4
Cases	Few ≤20	53 applications 17%	84 applications 27%
	Many > 20	69 applications 22%	107 applications 34%

300), which enables researchers to assess the chances of accepting a model for further analysis on the basis of random data. If this chance is too high, a researcher should not proceed with analyzing the selected model in csQCA because he or she will not be able to distinguish an analysis on the basis of random data versus real data. Hence, the benchmarks assess whether a model can be accepted for further analysis, and guide researchers in model specification. As in more quantitative approaches, csQCA also requires researchers to increase the number of cases if one wants to take on board additional condition variables.

The threshold for accepting a model is set at 10 percent, that is, there is a 10 percent chance of accepting a model, which could also have been generated on random data. More stringent benchmarks are 5 percent and 1 percent, respectively, a 5 percent and 1 percent chance of accepting a model, which could also be generated on random data. Publications on the benchmarks are recent, and previous applications did not pay attention to the number of conditions included in an explanatory model. Practices varied significantly between applications as is presented in Table 1: some include few conditions with many cases or few conditions with few cases or many conditions with few cases or many conditions with many cases.¹⁸

We analyzed the 219 csQCA applications in our data set, of which 131 (60% of the applications) pass the 10 percent benchmark test, 115 (52%) pass the 5 percent benchmark test, and 100 (46%) pass the 1 percent benchmark test. Thus, although a large proportion of applications pass the 10 percent and 5 percent benchmark tests, a significant proportion of applications do fail the test. This means that researchers proceeded with analyzing a QCA model, which could also have been generated on the basis of random data. For these applications, it is not possible to ascertain that the QCA model holds any explanatory power because other models with completely different explanatory conditions would equally generate high consistency scores or would not show any contradictions.²⁰ As indicated above, most csQCA applications contain a small to medium number of cases. The significant change from the 10 percent to the 1 percent test indicates that several csQCA applications use a small to a medium number of cases, in which context, the inclusion or

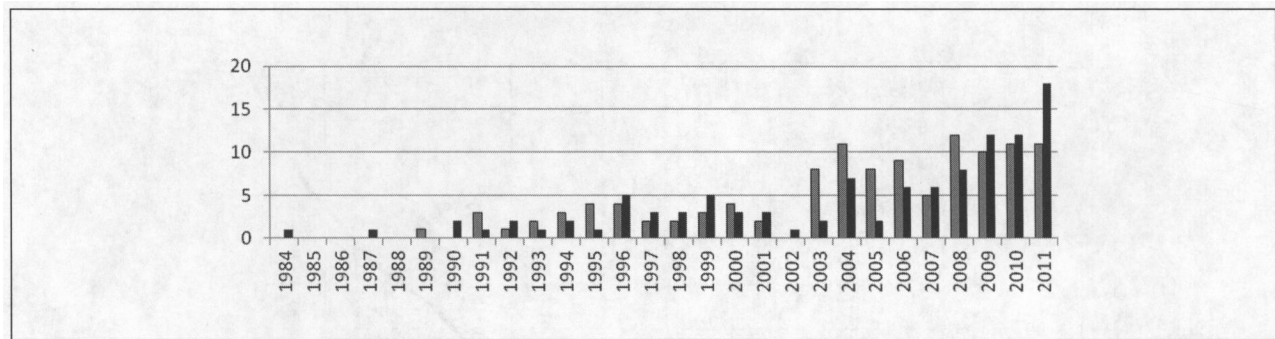


Figure 7. Evolution of number of csQCA applications that pass (gray) or fail (black) the 5% benchmark test.²¹

exclusion of only a few cases can make an important difference and can influence the analysis. Emmenegger, Kvist, and Skaaning, in their contribution on QCA applications in welfare-state research, identify and discuss several studies that failed the benchmark test.

An analysis of the evolution of applications that pass or fail the test shows that there is no significant change in the proportions over time. On the contrary, Figure 7 presents the evolution for the 5 percent benchmark test and shows that, during the past three years (2009–2011) more applications have failed than passed the test. By contrast, during the five previous years (2003–2008) with the exception of 2007, more applications had passed than failed the benchmark test.

Could those less favorable results in the most recent period (2009–2011) be the result of the quick spread of csQCA publications in journals, with perhaps less rigorous use of the csQCA techniques by a broader and more diverse group of researchers? Or could it be that the broadening array of journals (see the exponential diversity of journals having published QCA applications during the past few years, Figure 9) means that the pool of peer reviewers has expanded beyond the core of the more “purist” csQCA developers, and hence that some reviewers have become more lenient on issues such as the conditions:cases ratio? Whatever the explanation, future applications should take the benchmarks more seriously into account when performing a csQCA.

Combinations with Other Methods?

Finally, to what extent have QCA techniques been explicitly combined with other methods, be they quantitative (statistical) or qualitative? Indeed, some QCA developers tend to suggest that this should gradually become a “good practice” in the use of QCA (Rihoux et al. 2009, 170–72; Schneider and Wagemann 2012). The answer is very clear: The most frequent practice, by far, is still to use QCA as the single analytical tool (61.3% of the applications). Publications combining QCA with at least one quantitative/statistical

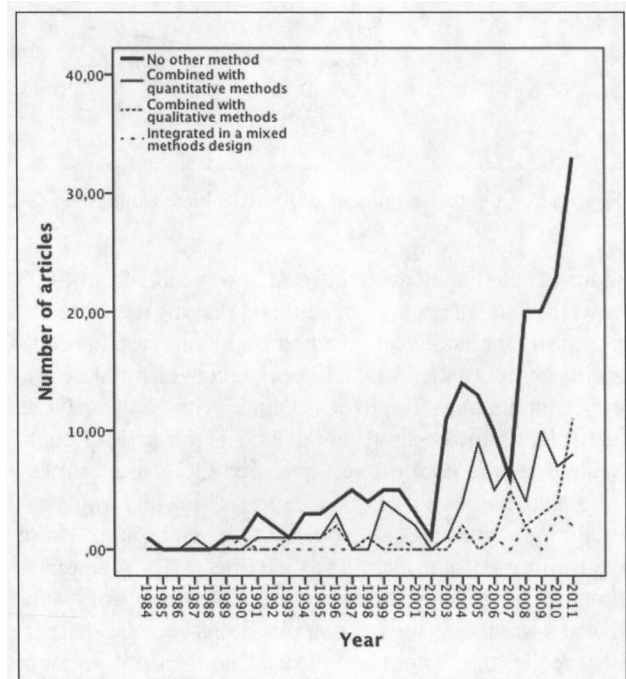


Figure 8. Number of articles and combined methods, 1984–2011.²³

technique are quite frequent (23.3%), whereas combinations with at least one qualitative technique (10.9%) and “mixed” methods designs combining QCA with both qualitative and quantitative techniques (4.5%) remain less frequent. Of course this could partly be due to the space limitations in journal article formats—combined methods are probably more frequently observed in larger-format publications such as monographs or research reports. Also noteworthy is that the combination between QCA and qualitative techniques has risen sharply over the past few years—a sign that an increasing number of researchers go beyond the “simple” exploitation of case studies and now exploit specific qualitative techniques (e.g., interviews, direct observation).

The overriding feature from Figure 8 is that the trend of “mono-method” exploitation of QCA has become

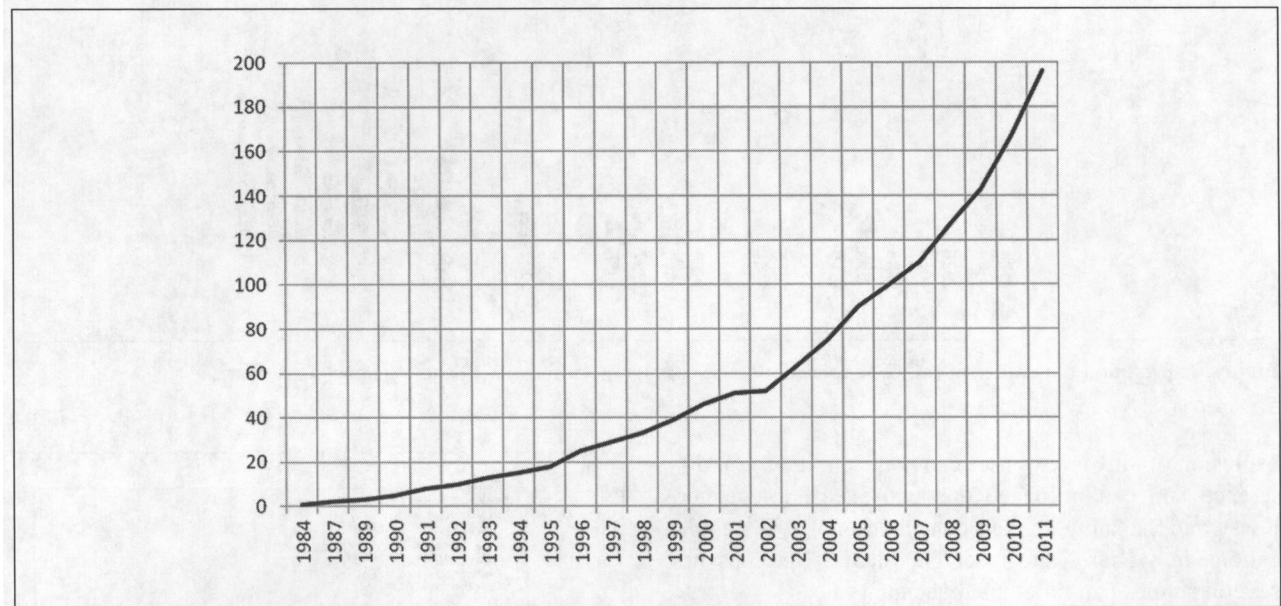


Figure 9. Cumulative number of journals having published QCA applications.²⁴

more predominant over the past few years, from 2008 onward. One interpretation could be that the mainstreaming of the method, accompanied by its increased acceptability by journal editors and peer reviewers, makes it easier for a manuscript to be accepted with “only” a QCA analysis, as the method has gained legitimacy as such. Second, as the pool of well-qualified QCA users broadens, also across disciplines, and as the techniques—including software—become more elaborate, more well-informed and technically sharper,²² QCA applications find their way through a growing variety of established journals. As the QCA applications become sharper, they request a greater level of detail in their presentation in an article-length format, which makes it difficult to combine with another method. A third interpretation is that the QCA techniques have gained in leverage, with more precise and qualified results (taking advantage of richer modes of calibration, more refined uses of “logical remainder” cases, more theory-informed arbitrations, more elaborate causal/configurational reasoning, richer models, the use of consistency and coverage coefficients, etc.), and therefore enable a researcher to successfully push articles relying solely on QCA through the peer review process.

Conclusion: Toward Mainstreaming?

The purpose of this article was to provide a first comprehensive mapping of QCA applications in peer-reviewed journal articles, from the very first application in 1984

until late-2011. Some quite clear trends are emerging: (a) The diffusion of QCA remained quite modest until 2002, but has increased significantly over the past few years especially; (b) the most frequently used technique is still csQCA, although fsQCA has also expanded significantly during the past few years; (c) the two predominant fields of application remain political science and sociology, but management studies has recently become the fastest-growing field; (d) macrolevel applications tend to be more numerous; (e) the use of different techniques (cs, mv, or fsQCA) has little impact on the number of conditions or cases; (f) QCA applications remain predominantly attached to medium-*N* designs; (g) the benchmarks in terms of number of cases/number of conditions ratio are not better abided by in the more recent period than in previous periods; and (h) QCA is most often used as a single technique, not combined with other quantitative or qualitative techniques. In other words, in spite of the development of fsQCA and a relative disciplinary diversification, and contrary to our overall hypothesis, QCA applications have not on the whole moved radically away from the crisp-set approach, which was developed by Charles Ragin (1987) in his seminal book.

To what extent have QCA techniques actually reached a stage at which they are mainstreamed in journals, that is, recognized as a valid and legitimate set of techniques, in their own right, in the respective disciplines? A first way to objectify this is to examine the evolution of the cumulative number of different peer-reviewed journals that have published QCA articles.

Table 2. Peer-Reviewed Journals with Highest Number of QCA Applications.²⁵

Journal title	No. of articles	Year of first QCA article
<i>American Journal of Sociology</i>	10	1992
<i>American Sociological Review</i>	10	1984
<i>European Journal of Political Research</i>	9	2003
<i>Journal of Business Research</i>	8	1995
<i>Sociological Methods and Research</i>	8	1994
<i>Revue internationale de politique comparée</i>	7	2004
<i>Methodological Innovations Online</i>	6	2010
<i>Journal of European Social Policy</i>	5	1996
<i>Social Forces</i>	5	1996

Figure 9 indicates that, until the year 2001, the range of journals publishing QCA applications grew quite regularly but modestly (up to about 50 different journals). By contrast, from 2004 onward, the pool of journals publishing QCA applications has increased exponentially, quadrupling over a five-year period. As a result, by the end of 2011, 196 different journals had published at least one QCA article. The spread has thus become quite significant.

A second way to answer this question is to look at *which journals* have published QCA applications. In particular, have highly ranked journals in the respective disciplines done so? If one looks at those journals that have published the largest number of QCA applications (Table 2), one notices quite a few top-tier journals in sociology (*American Journal of Sociology* [AJS] and *American Sociological Review* [ASR]) and in political science (*European Journal of Political Research*). The following are rather more specialized journals, either disciplinary or methodological, well-established in their respective fields.

Considering top-tier journals, there is a strong contrast between sociology, where the two highest-ranked journals (AJS and ASR) have been open to QCA from early on, and all the other disciplines, where access to the highest journals has, on the whole, been more recent. This being said, highly ranked journals have by now published QCA articles in many disciplines.²⁶

To sum up, 25 years after Charles Ragin's *The Comparative Method*, and contrary to our overall hypothesis, published QCA applications have remained quite proximate to Ragin's initial intentions in terms of design and use. QCA has also entered, during the past few years, a phase of mainstreaming in quite a few disciplines. To

what extent this mainstreaming will accelerate during the next few years remains to be seen. It will probably depend on the extent to which the QCA techniques will be further refined (some core ongoing refinements are discussed in other contributions in this mini-symposium). Further software development, along with a next generation of textbooks (e.g., Schneider and Wagemann 2012; Thiem and Dusa 2012), are also likely to make QCA accessible to a broader community of researchers, and to further expand the range of published applications.

Authors' Note

Ilona Rezsöhazy, previously at Université catholique de Louvain.

Notes

1. Another mapping is provided by Thiem and Dusa (2012), whose own data set has been merged with ours. They apply a more restrictive definition of "QCA applications" as articles that focus on substantive empirical research questions and include neither articles that focus on methodological aspects of QCA with the use of data for demonstration purposes nor articles that replicate another published analysis.
2. Accessible through www.compass.org. COMPASS is a broad network of scholars and practitioners involved in the development of systematic cross-case analysis, and QCA in particular. This full update has been conducted by (A-Z) Priscilla Álamos-Concha, Benoît Rihoux, and Alrik Thiem, as part of a concerted COMPASS effort. We wish to thank the following colleagues (A-Z) for their assistance in this: Dirk Berg-Schlosser, Damien Bol, Bart Cambré, Barry Cooper, Taisuke Fujita, Santi Furnari, Judith Glaesser, Martha Gross, Amy Javernick-Will, Axel Marx, Wendy Olsen, Ilona Rezsöhazy, Ingo Rohlfing, Svend-Erik Skaaning, Jadir Soares Junior, Maarten Vink and Barbara Vis.
3. Some further features could also have been surveyed, such as the implementation of various "good practices" (e.g., informed use of "logical remainders," transparency in the calibration, implementation of robustness checks, and formulation of hypotheses in set-theoretical form). This could be the topic of another contribution with a focus on "good practices."
4. Such as papers, working papers, conference proceedings, PhD dissertations, reports, and manuals.
5. Source: COMPASS bibliographical database, <http://www.compass.org/bibdata.htm>.
6. Source: COMPASS bibliographical database, <http://www.compass.org/bibdata.htm>.
7. A first textbook, focused on csQCA exclusively, was published in 2002 (De Meur and Rihoux), but it was in French and therefore not accessible to most.
8. As well as a first German-language one (Schneider and Wagemann 2007).

9. Source: COMPASSS bibliographical database, <http://www.compass.org/bibdata.htm>.
10. There is a technical justification to consider csQCA and mvQCA together: mvQCA is simply an extension of csQCA (Cronqvist and Berg-Schlusser 2009).
11. Following the usual practice, the outliers correspond to values that exceed 1.5 times the interquartile range, above the upper quartile in our case (there are no outliers below the lower quartile).
12. Source: COMPASSS bibliographical database, <http://www.compass.org/bibdata.htm>.
13. Source: COMPASSS bibliographical database, <http://www.compass.org/bibdata.htm>.
14. Source: COMPASSS bibliographical database, <http://www.compass.org/bibdata.htm>.
15. Results of these tests can be made available on request. The only statistically significant correlation is produced by the inclusion of the four extreme outlier applications with more than 20,000 cases.
16. Consistency is a measure that captures the degree to which a configuration of explanatory conditions explains an outcome. It is assumed that (a) consistency will be very low when the explanatory model is ill-specified or does not make theoretical sense and that (b) high consistency indicates the validity of the analyzed explanatory model.
17. The table is developed for three degrees of stringency. The threshold for accepting a model is set at 10 percent, that is, there is a 10 percent chance of accepting a model that could also have been generated on random data. More stringent benchmarks are 5 percent and 1 percent, respectively, a 5 percent and 1 percent chance of accepting a model that could also be generated on random data. For example, if one has four conditions and wants to pass the 10 percent benchmark test, one needs at least twelve cases. One will need seventeen cases if a researcher wants to pass the 1 percent threshold. If a researcher has eleven cases or less with four explanatory conditions he or she should not proceed with the analysis. Following this table, potential problems arise if one has four conditions with less than twelve cases, five conditions with less than fifteen cases, six conditions with less than twenty-six cases, seven conditions with less than thirty cases, and eight conditions with less than forty-six cases, and so on.
18. The operationalization of “few” and “many” is arbitrary but only serves the purpose of illustrating the diversity of practices. Finer-grained tests do not show a significant correlation between the number of conditions and number of cases; see above.
19. Source: COMPASSS bibliographical database, <http://www.compass.org/bibdata.htm>.
20. In csQCA, resolving contradictions is the mechanism by which researchers develop their explanatory model (see Kogut and Ragin 2006; Marx 2010; Ragin 1987).
21. “Failing the 5% benchmark test” means that there is at least a 5 percent probability that random data would have produced a QCA solution with high consistency or no contradictions, which then would be used in the subsequent analytic steps. Source: COMPASSS bibliographical database, <http://www.compass.org/bibdata.htm>.
22. This will be analyzed more in detail in a forthcoming paper, also based on the COMPASSS database.
23. Source: COMPASSS bibliographical database, <http://www.compass.org/bibdata.htm>.
24. Source: COMPASSS bibliographical database, <http://www.compass.org/bibdata.htm>.
25. Source: COMPASSS bibliographical database, <http://www.compass.org/bibdata.htm>.
26. Some examples (A-Z): *the Academy of Management Journal, Administrative Science Quarterly, the American Journal of Medical Genetics, Evaluation, Governance, Journal of European Public Policy, Legal Studies, Mobilization, Socio-Economic Review, West European Politics*.