Field Methods
24(2) 155-174

DOI: 10.1177/1525822X11433998
http://fm.sagepub.com

Structuring the Calibration of Qualitative Data as Sets for Qualitative Comparative Analysis (QCA)

Xavier Basurto[1] and Johanna Speer[2]

Abstract

Most studies that apply qualitative comparative analysis (QCA) rely on macro-level data, but an increasing number of studies focus on units of analysis at the micro or meso level (i.e., households, firms, protected areas, communities, or local governments). For such studies, qualitative interview data are often the primary source of information. Yet, so far no procedure is available describing how to calibrate qualitative data as fuzzy sets. The authors propose a technique to do so and illustrate it using examples from a study of Guatemalan local governments. By spelling out the details of this important analytic step, the authors aim at contributing to the growing literature on best practice in QCA.

[1] Duke Marine Lab, Nicholas School of the Environment, Duke University, Beaufort, NC, USA
[2] Division of Cooperative Science, Humboldt University, Berlin, Germany

Corresponding Author:
Xavier Basurto, Duke Marine Lab, Nicholas School of the Environment, Duke University, 135 Duke Marine Lab Rd., Beaufort, NC 28516, USA
Email: xavier.basurto@duke.edu

Keywords
qualitative comparative analysis (QCA), qualitative data, fuzzy sets, calibration, intermediate *n* study, best practice

Background and Goals

Qualitative comparative analysis (QCA) is a novel analytical tool that offers the possibility to compare intermediate numbers of cases and to assess the necessity and sufficiency of conditions in relation to an outcome. It is based on set theory and Boolean algebra, and its fuzzy-set version (fsQCA) draws on fuzzy logic (Zadeh 1965). The overall aim of any QCA is "to allow systematic cross-case comparisons, while at the same time giving justice to within-case complexity" (Rihoux and Ragin 2009:xviii).[1]

QCA has been welcomed by many social scientists because it promises to maintain a constant dialog between theory and evidence throughout the analytical process. The goal of this article is to advance this dialog in QCA studies that draw on qualitative data. To do so, we introduce a systematic and transparent procedure that allows scholars to transform qualitative data from interviews or secondary sources (e.g., texts from archives, web sites, company profiles, nongovernmental organizations [NGO] leaflets) into fuzzy sets.[2] We aim to contribute to best practice in research that uses QCA (Ragin 2000, 2008; Rihoux and Ragin 2009; Schneider and Wagemann 2010).

The number of empirical studies that rely on QCA as an analytical tool to systematically compare intermediate numbers of cases continues to increase and has surpassed several hundred in the last years (www.compasss.org). QCA first emerged in comparative sociology, a subdiscipline within sociology dominated by qualitative case-oriented scholars (Yamasaki and Rihoux 2009). Therefore, the majority of QCA applications focus on macro social phenomena, such as peasant revolts or regime change, the traditional areas of interest to comparative sociologists or comparative political scientists (e.g., Grofman and Schneider 2009). But in recent years, we have also observed an increasing number of QCA applications in other areas, such as natural resource management (Heikkila 2003; Lam and Ostrom 2010; Rudel 2005; Wade and Goldstein 2003), firms and markets (Skoko et al. 2006), and decentralization (Linder 2010). Studies in these areas usually cover units of analysis at the micro or the meso level and are often based on qualitative interview data.

A nonexhaustive review of the main repository of QCA literature (www.compasss.org) shows that studies that rely on data collected at the

meso or micro level (e.g., by conducting interviews with individuals) do not provide much detail on how they transformed their qualitative data into sets (Haworth-Hoeppner 2000; Lam and Ostrom 2010; Larsen 2009; Linder 2010; Marx 2008; Schneider and Sadowski 2010; Skoko et al. 2006). So far, the only information that the authors of such studies provide is the final table of the process of transformation that shows the qualitative classification that corresponds to each crisp or fuzzy-set value. How each author arrives at these classifications remains unclear, which makes it difficult to assess the validity, reliability, and replicability of the results of these studies.[3]

The transformation of qualitative or quantitative data into crisp or fuzzy sets is an important analytic step that has a strong influence on the results of the QCA. We argue that until now, the transformation of qualitative data has not received enough attention from QCA scholars and that a calibration technique is needed to make QCA studies more transparent and replicable.

The technique we propose complements available calibration techniques for quantitative data and contributes to the development of standards of good practice in QCA. Existing direct and indirect calibration techniques for quantitative data (Ragin 2008) cannot be applied to qualitative data as such data need to be coded and summarized before fuzzy-set values can be determined. Besides, qualitative data provide much more information for the dialog between theory and evidence than do quantitative data. For example, they allow the researcher to adjust the theoretically guided definitions of set anchor points by providing more detailed information on the empirical context in which the conditions and the outcome are assessed.

Our approach also differs from other techniques for turning qualitative data into scaled categories (i.e., numbers). Common techniques, such as counting the occurrence of codes in a document or coding interviewees' answers according to a predetermined scale that can be quantified afterward, cannot be used for set calibration. Calibrating qualitative data as sets requires the development of definitions of fuzzy-set values based on theoretical and substantive knowledge and not the generation of an arbitrary numeric representation of the data.

The procedure described here is based on the experience of the authors developing fieldwork-based projects to study decentralization issues in Central America using fsQCA (Basurto 2007, 2009; Speer 2010b). These projects relied mainly on qualitative data sources. Therefore, we needed to develop a technique for calibrating these data to fuzzy sets. The explanation of this technique is illustrated with examples from a study that we conducted in Guatemala.

The technique we propose consists of six steps, beginning with the preparation of data collection and ending when the fuzzy-set values of cases are assigned. Note that this is not an article about how to run an fsQCA. We assume that the reader has some basic understanding of fsQCA concepts and general procedures and is looking for guidance about how to prepare qualitative interview data for fsQCA analysis.

Background of the Guatemalan Study

Aim and Research Question

The aim of the study in Guatemala was to evaluate which accountability mechanisms contribute to good governance of public services. Large parts of the poor population in rural Guatemala lack access to basic public services. In light of the importance of these public services for improving the livelihoods of the rural poor and the low budget of the Guatemalan state, it is crucial that the state provides these services efficiently and that public spending responds to the needs of the poor. The research question of the study is therefore: Under what conditions are local governments responsive to their mainly poor electorate?

Theory and Hypothesis

The provision of public services to rural areas in developing countries has been found to be primarily affected by information asymmetries and conflicts of interest between the mayor and the population. These two problems were first conceptualized in political agency models that show how retrospective voting in elections can act as a disciplining device for politicians who shirk by either exercising reduced effort in carrying out their task or by diverting resources to private ends (Barro 1973; Ferejohn 1986). The objective from the point of view of the electorate, then, is to devise institutions that provide incentives for a self-interested mayor to refrain from opportunistic behavior and to respond to the needs of the population (Moe 1984).

The baseline political agency model shows how competitive elections can increase local government discipline. An extension of the model predicts that good access to information (GAI) for citizens (e.g., through active local media) can have a positive effect on local government responsiveness by making elections more effective as an accountability mechanism (Besley 2007:108–11, 128–32). Finally, another extension of the model shows that effective participatory governance (EPG) can increase local government's

discipline, since participatory governance forums serve both as a source of information and as a sanctioning mechanism (Speer 2010a). These arguments are summarized in the following research hypothesis: Competitive and free elections (CFE) and GAI combined and EPG alone are each sufficient for responsive local governance (RLG). In formal QCA notation, the hypothesis reads as: CFE × GAI + EPG –> RLG.[4]

Empirical Strategy

The study evaluated this hypothesis in an fsQCA of ten rural local governments (Speer 2010b). The unit of analysis of this study is the municipality (i.e., a local government's constituency). The qualitative data were mainly collected through semistructured interviews. Overall, we completed eighty-eight stakeholder interviews and eleven expert interviews. To obtain quantitative data (e.g., on the frequency of [MDC][5] meetings), we collected complementary secondary data, such as minutes of MDC meetings and municipal budgets as well as economic, political, and social information. The fuzzy-set QCA in the study is based on quantitative and qualitative measures, but for purposes of illustrating our technique, we will only refer to examples of qualitative measures.

The Procedure: From Qualitative Interview Data to Fuzzy-Set Values

In the following sections, we outline the procedure we developed to calibrate qualitative interview data to obtain fuzzy-set values ready to be loaded, for instance, into the fsQCA (2.0) software for analysis. We elaborate on the six steps of our procedure, describing them in chronological order. Step 1 describes the operationalization of the conditions and the outcome. Step 2 develops the anchor points and the elaboration of the qualitative interview guideline. Step 3 applies a content analysis to the raw interview data obtained in the field, and Step 4 explains how to summarize the code output. Step 5 determined the fuzzy-set scale and defined the fuzzy-set values. Finally, Step 6 assigned and revised the fuzzy-set values of the conditions and the outcome for each case. Before we started our procedure, we identified the theoretically relevant explanatory factors and the outcome to observe and derived a testable hypothesis (see Background of the Guatemalan Study). In QCA terminology, the factors are knowns as "conditions." We recommend stating the hypotheses in formal QCA notation, preferably identifying which conditions are expected to be necessary

and/or sufficient for the outcome. Yet, a deductive research design is not required for implementing the procedure we propose. It is also possible to carry out the fsQCA using a more open, inductive research strategy.[6]

Step 1: Identifying Measures of the Conditions and the Outcome

To develop a preliminary list of measures of the conditions and the outcome, start by operationalizing the theoretical concepts of the conditions and the outcome based on standard-scientific practice and/or your knowledge of the empirical context you are going to investigate (Ragin 2000). Measures may be added or dropped from this preliminary list during the research process based on the substantial information gained while studying your cases. Gathering qualitative data is likely to be a source of important case and contextual knowledge that will inform the operationalization of the theoretical concepts for your cases. Box 1 contains an example of the preliminary list of measures that we devised for the Guatemala study and the adjustments we made after returning from the field.

For each of the conditions and the outcomewe developed a preliminary list of measures to be recorded in the field. Below we provide a list of four preliminary measuresthat we considered for measuring the degree of access to information (GAI).

Condition	Preliminary Measures
Good Access to Information (GAI)	Amount of coverage of local political issues in local media
	Circulation of local media
	Accessibility of official documentation in public information offices
	Frequency of NGO/donor information campaigns

In the field the researcher often learns new things about the measure first selected and adjustments need to be made accordingly. In the Guatemalan project, we learned that official documents were accessible in all municipal public offices and that NGOs do not provide information on local government decisions and plans. In this context, the accessibility of official documentation turned out not to display any variance and the frequency of NGO information provision was zero in all municipalities. Hence, we dropped these two measures and only measured access to information as a composite measure that reflected the amount of coverage of local politics in local media and their circulation. The final version of the list of measures for GAI looked like this:

Condition	Final Measures
Good Access to Information (GAI)	Availability of news on local politics for majority of population on TV
	Availability of news on local politics for majority of population on the radio
	Availability of news on local politics for majority of population in newspapers

Box 1. Developing a preliminary list of measures of conditions.

Step 2: Developing Anchor Points and the Interview Guideline

For the Guatemala study, we collected data on each measure using mainly semistructured qualitative interviews.[7] To prepare the interview guidelines for this type of interviews, first develop a list of anchor points of each fuzzy set. Anchor points are the three main thresholds that structure a fuzzy set: 1 (threshold for full membership), 0.5 (cross-over point), and 0 (threshold for nonmembership; Ragin 2000:160). Anchor points help a researcher clarify how to distinguish a case that is more in the set from a case that is less in the set.

The initial anchor points are based on researchers' knowledge of the theoretical concepts they aim to measure and their knowledge of the context of his or her cases.[8] Later on, after coming back from the field, the initial anchor points can be revised and adapted when necessary. Yet, even though they may change during the research process, developing anchor points a priori is essential for judging during an interview whether an interviewee's answer is detailed enough for measuring the fuzzy-set values of the cases and for elaborating the specifying questions in the interview guideline. Thinking about the anchor points also helps improve the definitions of the theoretical concepts. See Box 2 for an illustration of the development of anchor points.

We developed a list with anchor points for all measures of the conditions and the outcome. The list below shows the preliminary anchor points that we determined for three measures of participatory governance. Later in the research process we replaced them with the final fuzzy-set value definitions (see Step 5). We only present a few measures of participatory governance used in the Guatemalan study to illustrate the development of anchor points.

Condition: Participatory governance	
Measure	**Anchor points**
Frequency of Municipal Development Council (MDC) meetings	**0**: MDC has not met in the last year
	0,5: MDC has met six times in the last year
	1: MDC has met 12 times or more in the last year
Provision of information from village heads to Municipal Corporation	**0**: The Municipal Corporation has not received any information on community preferences in the last year
	0.5: The Municipal Corporation received the preference lists from half of the communities in the last year
	1: The Municipal Corporation received the preference lists from all community representatives in the last year
Provision of information from Municipal Corporation to MDC about allocation of funds: oral or printed information on municipal budget revenues and expenses and on allocation of central government funds	**0**: The Municipal Corporation has not informed on any of these issues in the last year
	0,5: The Municipal Corporation has informed about selected topics at irregular intervals in the last year
	1: The Municipal Corporation has informed about both issues in the legally mandated intervals in the last year

Box 2. Developing anchor points.

To design the interview guidelines, we suggest first creating one section for each condition and the outcome in the interview guideline (Step 2.1). For example, since we examined three conditions and one outcome in the Guatemalan study, the interview guideline initially had four sections. Second, define an introductory eliciting question for each section and within the sections include a subquestion on each measure (Step 2.2). Starting with an open initial eliciting question leads the interviewees into the topic and allows them to talk about the ideas that first came to their minds and thus inform us about their relevance. The subquestions elicit more targeted information about the measures. Including additional subquestions that explore new dimensions of the theoretical concept or additional measures is also recommended. Finally, add specifying questions for following up on subquestions (Step 2.3). This is useful when a respondent would not answer in enough detail to determine the fuzzy-set value of a measure of a condition or outcome. We provide examples of Steps 2.1–2.3 in Box 3.

Step 3: Interview Coding

After completing the data collection, perform a content analysis of the raw interview data using qualitative data analysis software (e.g., Atlas.ti). The coding procedure we describe here does not differ substantively from other iterative coding procedures that draw on theoretical concepts and empirical data to define codes. See, for instance, Bernard and Ryan (2010:84–85). To code our Guatemalan interview data, we developed an initial list of codes based on the preliminary list of measures of the conditions and the outcome that we developed in Step 1. When interviewees had pointed out an additional dimension of one of our theoretical concepts that we had not captured in our preliminary list of measures (and thus for which there was no code in the initial list of codes for that dimension), we added it in the course of the content analysis using open and in vivo coding.[9] For instance, the content analysis of the interviews from Guatemala revealed large differences in access to local media between rural and urban areas. Hence, we added the codes "mediause_rural" and "mediause_urban" and subsequently took them into account in measuring access to information through local media.

Step 4: Summarizing the Interview Data to Qualitative Classifications

Next, carry out a systematic analysis of the coded qualitative data (code output) and summarize all quotations within one case for each code. We

The main instrument of data collection that was used in Guatemala were semistructured interviews with the mayor, village representatives, local civil society representatives, key informants, and local journalists. In the following, we illustrate each step in the elaboration of the interview guideline.

Step 2.1: In line with the research hypothesis, the interview guideline contains four sections on "elections," "access to information," "participatory governance," and "local government responsiveness." The first three constitute conditions and the last one is the outcome. We use a few measures of the condition "participatory governance" as an example.

Step 2.2: Each section starts with an open introductory eliciting question. After that, we included open sub-questions for all measures to gather information on them. This is illustrated in the table below for three measures of the effectiveness of participatory governance.

Section: Participatory governance (Condition)	
Eliciting question: How does the Municipal Development Council (MDC) in this municipality work?	
Measure	Sub-question
Frequency of meetings	How many times did the MDC meet in the last 12 months?
Provision of information from village heads to Municipal Corporation	How did the village heads inform the municipal corporation about the priorities of their villages in the last 12 months?
Provision of information from Municipal Corporation to MDC	What information about the activities of the municipal corporation did you receive in the MDC in the last 12 months?

Step 2.3: Each open sub-question was complemented by a closed specifying question that was used to gather further information from interviewees who did not answer the sub-question in detail. These specifying questions were devised based on the anchor points. The final structure of the interview guideline including the specifying questions is shown in the following table:

Section: Participatory governance		
Eliciting question: How does the Municipal Development Council (MDC) in this municipality work?		
Measure	Sub-question	Specifying question
Frequency of meetings	How many times did the MDC meet in the last 12 months?	If it has met: have you met every month or every few months?
Provision of information from village heads to Municipal Corporation	How did the village heads inform the municipal corporation about the priorities of their villages in the last 12 months?	How many of the village heads have submitted a list of projects to the mayor? Did you submit a list?
Provision of information from Municipal Corporation to MDC	What information about the activities of the municipal corporation did you receive in the MDC in the last 12 months?	In what form? How often?

Box 3. Elaborating the interview guideline.

relied again on Atlas.ti for this step, but it can also be done without data analysis software.

For our study, we extracted interview quotations in three different ways (see summary in Box 4). First, we examined all quotations with the same code from all cases and all interviewees (Step 4.1). For example, we reviewed all quotations that had been coded with "mediause_rural." This first review

The table below summarizes the most important aspects of reviewing the interview data sorted in three different ways to arrive at the final table of measured values for all cases and all codes.

Logic of review	Tasks	Output
Step 4.1: For each code review data from all cases and all interviewees	Ask yourself for each code: ✓ Is there sufficient reliable information for all cases? ✓ Is there sufficient variation across cases? ✓ Do I need to add this code as a measure to the list?	Final list of measures of conditions and outcome
Step 4.2: For each code review data from all cases sorted by interviewee group	Check for each code and interviewee group whether there are biases due to: ✓ relation with other actors ✓ social position ✓ level of education ✓ other factors	List of systematic biases in responses of interviewee groups for each code
Step 4.3: For each code review data from all interviewees sorted by case	Summarize quotations for each code within each case. If there are contradictions, resolve them using information on: ✓ systematic biases in responses (see Step 4.2) ✓ interview situation, consistency of answers ✓ common interview problems ✓ secondary data	Table with measured values for all cases (see Box 6 for an example)

Box 4. Summarizing qualitative data for each case.

of all data on each code allowed us to check whether there were sufficient reliable data for all cases on the respective measure. Furthermore, we could assess whether there was enough variation between cases in each measure and, if so, which range of values of the measure we observed. Measures for which we could not collect sufficient data in the field or that turned out to display no variation were dropped from the list. The result of this first general revision is a final list of measures for the conditions and the outcome.

Second, we extracted the quotations for each code sorted by type of interviewee (Step 4.2). For instance, we looked at all answers from village representatives that had been coded as "downward information flow from the local government." This allowed us to detect biases in the responses of certain types of interviewees and to take into account the particular characteristics of different types of interviewees in evaluating their answers. For example, Guatemalan village representatives were often economically or politically dependent on the mayor and therefore avoided saying that the mayor did not inform them about his spending decisions.

Being aware of systematic biases in responses of interviewees is crucial for Step 4.3, in which we summarized all interview quotations with the same code for each case in a qualitative classification.[10] For this, we reviewed for example the interview quotations from all ten interviewees of the Guatemalan municipality of San Bueno[11] that we had coded with "downward information flow from the local government." Then, we

summarized all these quotations from the mayor, the village representatives, NGO members, and so on, to create the qualitative classification of San Bueno for the measure "provision of information from the Municipal Corporation to the Municipal Development Council." In our example, the qualitative classification was "the local government of San Bueno provides yearly information on total revenues and expenses."

The challenge of Step 4.3 is to summarize the information of several interviewees in one statement that best reflects the case.[12] As in all qualitative data, it is possible that two or more interviewees contradict each other. Based on our knowledge of the cases, the context, and the data, we were in a position to solve such contradictions in the replies of interviewees and to decide how to weigh the different answers of interviewees for the same measure of interest. For triangulating interviewees' answers, we drew on information about the Guatemalan context, about each case, about potential sources of biases in answers from interviewees, and about secondary data, such as minutes of the meetings of a council and municipal budgets. The decisions on contradictions in the data and the information based on which they were made need to be transparent in presenting the analysis results.

Step 5: Determining the Precision of Fuzzy Sets and Defining their Values

Before you can match the qualitative classifications to fuzzy-set values, choose the degree of precision of the fuzzy sets to use and define each of their values. For the Guatemalan study, we determined the degree of precision of our fuzzy sets based on the level of detail in our qualitative data. Our data lent themselves to a four-value fuzzy set (i.e., a fuzzy set with the values: "Fully out [0]," "more out than in [0.33]," "more in than out [0.67]," and "fully in [1]"). It was not suitable for a more finely scaled fuzzy set[13] as it became increasingly difficult to assign different fuzzy-set values to two cases if these values were very close to each other.

Following Ragin (2000, 2008), we based the definition of the fuzzy-set values on the theoretical concept of interest and on our in-depth knowledge of the cases and their context. Using the theoretical concept for the definition of fuzzy-set values is important for deriving implications for theory development from the findings of the fsQCA (Goertz 2006). At the same time, it is necessary to take into account the sociocultural context of the cases because the cases are not compared to an absolute ideal case in a QCA. Instead, the context-based cases are compared with each other.

In our study, we first consulted the theoretical debate surrounding the definition of the concept to determine the main elements of "effective participatory governance" or "competitive elections." The discourse on the definition of effective participatory governance reveals, for example, that one of its essential elements is "inclusiveness" (i.e., the effectiveness of participatory governance depends on who participates).

Second, we adjusted the theoretical ideal to the sociocultural context of our cases. We defined full membership in a fuzzy set (fuzzy-set value = 1) by constructing an imaginary ideal case in the context of the universe of our cases (i.e., rural Guatemalan municipalities). The definition of this ideal case might not coincide with the qualitative classification of the best empirical case. Rather, it is the best imaginable case in the context of the study that is logically and socially possible (Ragin 2000:165–71). For example, for the measure "provision of information from the Municipal Corporation to MDC about municipal revenues and expenses," the value of 1 was adjusted to the context of the Guatemalan law, which foresees that the municipal government informs the MDC four times a year about its revenues and expenses. The provision of information to the MDC three or four times a year is therefore defined to correspond to a fuzzy-set value of 1, even though in other contexts this frequency of information provision could be quite low.

We used the same approach to define the lowest value of a fuzzy set, that is, nonmembership in a set (fuzzy-set value = 0). However, for constructing the imaginary case that is fully out of a set, it is more important to draw on the theoretical concept than on the socioeconomic context. The definition of the theoretical concept one uses usually contains one or several key characteristics. For example, in elections, at least two serious candidates need to be running if they are to be regarded as competitive. Cases that do not display at least one of the defining characteristics of the concept to a low degree need to be assigned a 0, which stands for nonmembership in the set of municipalities with competitive elections. Again, this value might not necessarily coincide with the lowest measured value among the sampled cases.

Box 5 provides an example of the definitions of the four fuzzy-set values of two measures of participatory governance for the Guatemalan study.

An alternative to adjusting the definition of the fuzzy-set values to the sociocultural context of the cases is to relabel the concept of the condition or the outcome. For example, we could have relabeled the measure "provision of information from Municipal Corporation to MDC about municipal revenues and expenses" to "compliance with mandated budget information."

The table below shows how the four fuzzy-set values were defined for two of the measures of effective participator governance.

Condition	Measures	Fuzzy-Set Value Definitions
Effective Participatory Governance (EPG)	Participation of all required groups of actors	**0:** None of the required groups participates **0.33:** Less than half the organizations participates **0.67**: Half or more of the organizations participate **1**: All the required groups participate
	Provision of information from Municipal Corporation to Municipal Development Council about municipal revenues and expenses	**0**: No revenues and expenses are communicated **0.33**: All revenues and expenses are communicated once a year **0.67**: All revenues and expenses are communicated twice a year **1**: All revenues and expenses are communicated three or four times a year

Box 5. Defining fuzzy-set values.

Step 6: Assigning and Revising Fuzzy-Set Values

After defining the fuzzy sets, assign values within the fuzzy sets to each case in the data set by matching the qualitative classifications you derived in Step 4 with the fuzzy-set values you defined in Step 5. We illustrate how we conducted this matching exercise in Box 6.

Finally, revise and adjust the assigned fuzzy-set values (i.e., the values in the last column of Box 6) for all cases and all measures. This revision is a crucial part of the dialog between theory and evidence. Going through one measure across all cases, the scholar can evaluate whether the fuzzy-set value differences between cases reflect real differences between the cases according to case knowledge and whether the interview data are well captured by the fuzzy-set values. If there are discrepancies, the scholar needs to go back to Step 4 to revise the interview data summary for overlooked clues or biases in the data affecting the resulting qualitative classifications. If not, the scholar returns to Step 5 to check whether the definitions of the fuzzy-set values reflect all relevant dimensions of the theoretical concept and have been adjusted appropriately to the context.

This process of figuring out why a fuzzy-set value of a case does not seem to fit given the researcher's case knowledge can make the researcher aware of important aspects that have so far been neglected in her or his definition of the fuzzy-set values. It is critical to ensure that the revision process is not used as a way to adjust the data to display a nice pattern of causality. Instead, it is a process in which the researcher verifies that the data are well aligned with the theoretical concepts she or he is interested

In the table below we illustrate how we matched the verbal measure values for the provision of information by the Municipal Corporation to the Municipal Council on revenues and expenses (derived in Step 4) with the definitions of the fuzzy-set values (developed in Step 5).

For each verbal measure value (column 2) we chose the closest value from the four-value fuzzy set (column 3). In Case A of the example below, we decided that the verbal measure value "yearly information on total revenues and expenses" best matched with 0.33 = "all revenues and expenses are communicated once a year." Hence, we assigned Case A the fuzzy-set value of 0.33.

Measure: Provision of information from Municipal Corporation to Municipal Development Council about municipal revenues and expenses			
Case	Verbal measure value	Fuzzy-set value definitions	Assigned fuzzy-set value
Case A	Yearly information on total revenues and expenses	1.0 = All revenues and expenses are communicated three or four times a year 0.67= All revenues and expenses are communicated twice a year. 0.33 = All revenues and expenses are communicated once a year 0.0 = No revenues and expenses are communicated	0.33
Case B	No information given		0.00
Case C	Three or four reports given at a public hearing in written form		1.00

Box 6. Assigning fuzzy-set values.

in, the conditions and outcome are well represented by their measures, and the case evidence is adequately summarized in the fuzzy-set values.

After the revision, there may still be fuzzy sets with a skewed distribution of cases (e.g., all cases could be concentrated in the lower half of a fuzzy set). This happens when the researcher has not encountered empirical instances of some logically possible configurations of conditions and consequently one corner of the property space of conditions remains empty. Such limited diversity may be in itself an issue of interest for future inquiry (Ragin 2000:168–69, 198–201). For example, in Guatemala none of the selected municipalities has local media that critically cover local government decisions and thus provide independent information to voters. Local media was intended to be a measure of "good access to information," but the fuzzy-set values of this measure were low for all cases. Why we did not observe rural municipalities with independent local media might be an issue that merits future inquiry.

At this point, the researcher is ready to aggregate the fuzzy-set values of all measures into the condition to which they belong and create a summary table as show in Table 1. This table contains the fuzzy-set values of the conditions and the outcome for all cases in the Guatemalan study. Aggregating measures can be done in different ways depending on the theoretical concept and the particular research question (Goertz 2006). In our example, we have taken the maximum of the three measures of GAI (see Box 1),

Table 1. Final Fuzzy-Set Values of Cases

Case Municipality	Conditions			Outcome
	GAI	CFE	EPG	RLG
Case A	0.33	0.67	0.22	0
Case B	1	0.33	0.67	0
Case C	0.67	0	0.33	1
Case E	0	0.67	0.33	0
Case F	0	0.67	0.67	0.67
Case G	0.67	0	1	0.67
Case H	0.33	1	1	1
Case I	0	1	1	1
Case J	0.33	1	1	0.67
Case K	0.33	0	0.67	0.33

Note. CFE = competitive and free elections; EPG = effective participatory governance; GAI = good access to information; RLG = responsive local governance.

since it does not matter through which medium people are informed and, hence, the three measures are substitutable.[14]

Contribution to QCA Best Practice

We agree with seasoned QCA users that regardless of whether one uses quantitative data, qualitative interview data, historical documents, or secondary text data, the determination of set values should be based on a researcher's theoretical and substantive knowledge and not on internal criteria such as the mean or the mode (Ragin 2008:30). Yet, we argue that this commonly accepted proposition about how to determine set values has not been sufficiently developed for qualitative interview data. While two techniques to calibrate quantitative data have been proposed (Ragin 2008), to our knowledge no equivalent calibrating procedure is available for qualitative interview data.[15] So far, studies that use qualitative data for a fsQCA do not provide any details on how they have transformed their data to fuzzy sets (Metelits 2009; Schneider and Sadowski 2010). The details of this analytical step remain opaque to the audience of such research. We regard the availability of a well-developed calibration technique as an essential step toward increasing the reliability and replicability of any study using QCA. In the technique we propose, the definition of set values used for QCA is made transparent and thus open to criticism and improvement from other scholars.

Having said this, we are aware that our technique is not immune to criticism QCA has received from researchers using mainly statistical tools. Such researchers have stated that the definition of the crisp and fuzzy-set values is arbitrary or can be adapted by the researcher to get the desired results (Wade and Goldstein 2003). The possibility that this might happen cannot be completely ruled out in the procedure we propose as it relies heavily on case and context knowledge of a researcher. However, we consider that reliance on such knowledge is not a weakness but a tremendous strength in the interpretation of qualitative and quantitative data. The potential of the technique we propose is its ability to maintain a constant dialog between theory and evidence, which is important for the validity of any systematic approach to qualitative data analysis (Bernard and Ryan 2010:110–11). As we have illustrated, until the fuzzy-set values are defined, the researcher must continuously think thoroughly about the definition of the theoretical concepts she or he uses and their main elements or subdimensions. In addition, QCA researchers can show in sensitivity tests whether slight changes in the definitions of the set values affect their results or not. Another possibility for demonstrating the robustness of the results of the QCA is to use alternative measures for a given concept.

Finally, our procedure is only one example of different ways calibration techniques could be devised. We welcome discussion among scholars about the advantages and disadvantages of this procedure. An open discussion on how to improve the use of qualitative interview data in QCA increases the credibility of the results that are produced using this method and facilitates its adoption among scholars who are accustomed to using other analytical approaches.

Acknowledgments

The authors thank Gustavo Garcia-López, Morgan Gopnik, Rebecca Gruby, Charles Ragin, Peter Schneider, and two anonymous referees for their thoughtful contributions. Author names in alphabetical order express equal contribution to this piece.

Notes

1. QCA can compare factors systematically and rigorously across intermediate numbers of cases. It increases the external validity of the results compared to single case studies without ignoring within case complexity, as happens in standard quantitative analyses (Ragin et al. 2003).
2. Our technique can also be used for crisp sets, which require less fine-grained data.

3. While journal space limitations will often make the disclosure of all details of the calibration process within a journal article challenging, at a minimum authors should mention what calibration procedure was used. We encourage reviewers to request all details of the calibration procedure during the peer-review process. To authors, we suggest making such description available (e.g., in the journal's online materials, on www.compasss.org, or on the author's homepage) and noting in the methods section whether reviewers had access to the details of the calibration procedure and where they are available.
4. The Boolean algebra notation of the hypothesis in the box uses: "+" to represent logical "OR," "*" for logical "AND." Lower-case letters refer to logical negation/absence of a condition. The connecting arrow means that the formula on the left is expected to be causally linked to the outcome on the right.
5. The MDCs are participatory governance forums for the planning, monitoring, and evaluation of municipal projects for building or repairing social infrastructure. In the MDC, representatives from the communities and local civil society as well as officials of the municipal and central government meet once a month.
6. We recognize that there are many traditions in the social sciences that follow other strategies for designing their research. We are confident that our calibration technique can be equally useful to scholars who use more open research designs. For a discussion of several ways on how to select the conditions for a QCA see Berg-Schlosser and de Meur (2009:25–32).
7. See Bernard (2006) for information on interviewing techniques and data collection in developing countries.
8. The challenge to assign a numerical value such as an anchor point to a verbal label is not unique to fsQCA but is frequently encountered, for example, in the design of questionnaires for quantitative research. The assignment of anchor points is easier for some measures for which there are already agreed-on thresholds, such as the cut-off point for poor countries in international gross domestic product rankings, and more difficult for others where no consensus has been established so far.
9. See Haworth-Hoeppner (2000) for an example of the use of open coding for crisp-set QCA.
10. We use the term "qualitative classification" to refer to the verbal statement that indicates the expression of a qualitative measure for a case. We follow the terminology used by Adcock and Collier (2001). The quantitative equivalent of a qualitative classification is the numerical score of a case on a measure or indicator.
11. We use a fictional name to illustrate this example.
12. This data-reduction step necessarily reduces the complexity of the qualitative data. Qualitative classifications cannot reflect different interpretations of a

concept among several actors in one case. Such interesting additional findings could be made available through a narrative of the cases that could complement any QCA.

13. For an overview of differently scaled fuzzy sets and their advantages and disadvantages, see Ragin (2008:30–33). More precise fuzzy sets can be used when more detailed qualitative data are available or when quantitative data are calibrated to fuzzy-set values.
14. For a discussion of the criteria that can be used to decide on whether measures should be aggregated taking the average, the maximum, or the minimum, see Ragin (2000:321–28) or Goertz (2006:135–42).
15. Not all studies using qualitative interview data for fsQCA, however, need to calibrate their data following the technique described in this article. Sometimes it might be adequate for the researcher to offer the interviewee a predetermined Likert scale of answers that correspond directly to fuzzy-set values.

References

Adcock, R., and D. Collier. 2001. Measurement validity: A shared standard for qualitative and quantitative research. *The American Political Science Review* 95:529–46.

Barro, R. J. 1973. The control of politicians: An economic model. *Public Choice* 14:19–42.

Basurto, X. 2007. Policy, governance and local institutions for biodiversity conservation in Costa Rica. Unpublished Ph.D. diss., University of Arizona.

_____. 2009. The role of cross-scale linkages for maintaining local autonomy: The case of biodiversity conservation in Costa Rica (working paper, Bloomington: Indiana University, workshop in political theory and policy analysis).

Berg-Schlosser, D., and G. de Meur. 2009. Comparative research design. In *Configurational comparative methods. Qualitative comparative analysis (QCA) and related techniques,* eds. B. Rihoux and C. C. Ragin, 19–32. Thousand Oaks, CA: SAGE.

Bernard, H. R. 2006. *Research methods in anthropology: Qualitative and quantitative approaches*. Walnut Creek, CA: AltaMira.

Bernard, H. R., and G. W. Ryan. 2010. *Analyzing qualitative data. Systematic approaches*. Thousand Oaks, CA: SAGE.

Besley, T. 2007. *Principled agents? The political economy of good government.* New York: Oxford University Press.

Ferejohn, J. 1986. Incumbent performance and electoral control. *Public Choice* 50:5–25.

Goertz, G. 2006. *Social science concepts. A user's guide*. Princeton, NJ: Princeton University Press.

Grofman, B., and C. Q. Schneider. 2009. An introduction to crisp set QCA, with a comparison to binary logistic regression. *Political Research Quarterly* 62:662–72.

Haworth-Hoeppner, S. 2000. The critical shapes of body image: The role of culture and family in the production of eating disorders. *Journal of Marriage and the Family* 62:212–27.

Heikkila, T. 2003. Institutional boundaries and common-pool resource management: A comparative analysis of water management programs in California. *Journal of Policy Analysis and Management* 23:97–117.

Lam, W., and E. Ostrom. 2010. Analyzing the dynamic complexity of development interventions: Lessons from an irrigation experiment in Nepal. *Policy Sciences* 43:1–25.

Larsen, M. 2009. Vulnerable daughters in times of change: A set-theoretic analysis of the "missing girls" problem in India. *COMPASSS Working Paper 2009-55*, http://www.compasss.org (accessed October 10, 2011).

Linder, W. 2010. On the merits of decentralization in young democracies. *Publius: The Journal of Federalism* 40:1–30.

Marx, A. 2008. Limits to non-state market regulation: A qualitative comparative analysis of the international sport footwear industry and the Fair Labor Association. *Regulation and Governance* 2:253–73.

Metelits, C. M. 2009. The consequences of rivalry: Explaining insurgent violence using fuzzy sets. *Political Research Quarterly* 62:673–84.

Moe, T. M. 1984. The new economics of organization. *American Journal of Political Science* 28:739–77.

Ragin, C. C. 2000. *Fuzzy-set social science*. Chicago: University of Chicago Press.

_____. 2008. *Redesigning social inquiry: Fuzzy sets and beyond.* Chicago: University of Chicago Press.

Ragin, C. C., D. Shulman, A. Weinberg, and B. Gran. 2003. Complexity, generality, and qualitative comparative analysis. *Field Methods* 15:323–40.

Rihoux, B., and C. C. Ragin, eds. 2009. *Qualitative comparative analysis (QCA) and related techniques*. Thousand Oaks, CA: SAGE.

Rudel, T. K. 2005. *Tropical forests—Regional paths of destruction and regeneration in the late twentieth century*. New York: Columbia University Press.

Schneider, C. Q., and C. Wagemann. 2010. Standards of good practice in qualitative comparative analysis (QCA) and fuzzy-sets. *Comparative Sociology* 9:397–418.

Schneider, P., and D. Sadowski. 2010. Governance configurations and academic outcomes: The example of Ph.D. education. Discussion Papers 201001, Institute of Labor Law and Industrial Relations in the European Community (IAAEG), http://www.iaaeg.de/images/documents/dp/dp%20012010.pdf (accessed October 10, 2011).

Skoko, H., B. Krivokapic-Skoko, M. Skare, and A. Ceric. 2006. ICT adoption policy of Australian and Croatian SMEs. *Managing Global Transitions* (University of

Primorska, Slovenia), http://www.fm-kp.si/zalozba/ISSN/1581-6311/4_025-040.pdf (accessed October 10, 2011).

Speer, J. 2010a. A political agency model of the impact of municipal development councils on local government performance. Paper presented at the 2010 Annual Meeting of the European Society for Public Choice, Izmir, Turkey, April 8–11.

_____. 2010b. The joint effect of participatory governance, elections, and access to information on local government responsiveness. Paper presented at the 14th Annual Conference of The International Society for New Institutional Economics, Stirling, Scotland, June 17–19.

Wade, R. H., and M. Goldstein. 2003. The causes of village cooperation: Comment on Ragin, Shulman, Weinberg, and Gran. *Field Methods* 15:341–50.

Yamasaki, S., and B. Rihoux. 2009. A commented review of applications. In *Configurational comparative methods. Qualitative comparative analysis (QCA) and related techniques*, eds. B. Rihoux and C. C. Ragin, 123–46. Thousand Oaks, CA: SAGE.

Zadeh, L. A. 1965. Fuzzy sets. *Information Control* 8:338–53.