Part I

Getting Started

1 Introduction
QCA in a Nutshell

1.1 Introduction and Learning Goals

This book offers a hands-on introduction and teaching resource for students, users, and teachers of Qualitative Comparative Analysis (QCA; Ragin, 1987, 2000, 2008b). Given its superior ability to model certain aspects of complexity, QCA has made inroads into virtually every social science discipline and beyond. Software solutions for QCA have also been developing at a fast pace. This book seeks to reduce the time and effort required when we first encounter the logic of not just a new method but also new software. It offers a genuinely simple, intuitive, and hands-on resource for implementing the state-of-the-art protocol of QCA using R, the most advanced software environment for QCA. Our book has an applied and practical focus.

In this introductory chapter, we use an empirical example to explain what QCA is and how it works. In the subsequent chapters, we will treat these features and steps in more depth. Using simple language and illustrations, our aim is to familiarize the reader with the basic analytic goals and steps of QCA and illustrate what kind of results this method produces. We then sketch the empirical spread of QCA and related software. Finally, we explain how this book is structured and how the reader can best use it.

Box 1.1 Learning Goals – QCA in a Nutshell

- Familiarity with the general analytic goals and motivations underlying the use of QCA.
- Basic understanding of the main analytic steps involved in doing a QCA.
- Basic understanding and interpretation of QCA results.

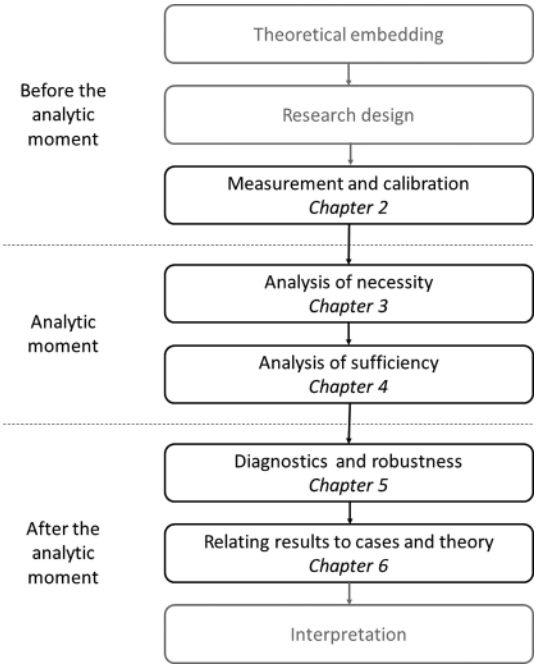

Figure 1.1 Steps of QCA and relevant book chapters
Note: Gray indicates an aspect not specific to QCA.

1.2 QCA in a Nutshell

We start by explaining what kind of research question and motivation would lead us to use QCA in the first place. Then, we introduce the example study that helps us to illustrate the basics of QCA – a study by Freyburg and Garbe (2018), who seek to explain internet shutdowns during elections in sub-Saharan Africa. Based on this example, we guide the reader through the different steps of QCA. We divide this process in the stages *before*, *during*, and *after* the analytic moment – a distinction that we also use for structuring our book into different parts. We conclude by summarizing these steps, and point to the subsequent chapters that cover their technical details (see Figure 1.1).

As Figure 1.1 highlights, some of the analytic steps, though essential, are not specific to QCA, but rather generic aspects of research design and interpretation. We do not cover these aspects in depth in this book, but illustrate them with the specific study by Freyburg and Garbe described in Box 1.2.

Box 1.2 Empirical example: Explaining internet shutdowns in sub-Saharan Africa (Freyburg and Garbe, 2018)

Research question: Why do even Africa's most stable democracies intentionally restrict public internet access during election periods?

Outcome: The occurrence or non-occurrence of internet shutdowns during election periods.

Cases: 33 presidential and parliamentary elections in sub-Saharan African (SSA) countries from 2014 to 2016.

Conditions: State ownership of the internet service providers (ISP); government is an autocracy (AUTOCRACY); occurrence of electoral violence (VIOLENCE).

Sets: Crisp.

Source: Freyburg, Tina, and Garbe, Lisa. 2018. Blocking the bottleneck: Internet shutdowns and ownership at election times in sub-Saharan Africa. *International Journal of Communication,* 12, 3896–3916.

1.2.1 General Goal and Motivation for Using QCA

We can think of empirical research methods as tools in a toolbox. Deciding on which tool to use depends on how suitable it is for performing a given task. For instance, we use a screwdriver to tighten a screw, but not to cut a board. Similarly, the choice of an empirical research method depends on its suitability to answer a given research question. In this section, we introduce four characteristics of QCA: its orientation toward explaining outcomes, case orientation, its set-theoretic foundation, and its approach to modeling causal complexity.[1]

Causes-of-Effects Research Questions

QCA helps us address so-called causes-of-effects type of research questions that ask for the reasons why certain phenomena occur (Mahoney and Goertz, 2006). For example, the study by Freyburg and Garbe (2018) starts with a

[1] In this chapter, we use causal language in line with the extant literature, such as 'causes-of-effects' and 'causal complexity'. However, we advise caution in simply taking the results of applied QCA as indicating causality. Whether or under which circumstances the solutions generated with QCA can be interpreted in causal terms is subject to a debate that goes beyond the scope of this book. We hold the position that with QCA one can come closer to a causal interpretation if the cross-case evidence generated with QCA is combined with within-case analyses on the causal mechanisms; see Rohlfing and Schneider (2018). In Chapter 6 on set-theoretic multi-method research (SMMR), we spell out the principles of how to combine QCA with case studies.

particular puzzle: why do even Africa's most stable democracies intentionally restrict public internet access during election periods? Accordingly, their study asks for the conditions that explain why internet shutdowns at election times occur in sub-Saharan Africa. Conversely, we would not use QCA to identify, say, how much a change in leadership affects internet shutdowns. That would be a so-called effects-of-causes question asking for the effect of a specific variable. QCA does not help us identify the magnitude of the effect of a single factor in isolation. Instead, its core motivation is to account for the complex interplay of different factors in bringing about the outcome of interest.

Formalized Comparative Case Studies

QCA is often presented as a method specifically designed for comparing small numbers of cases. However, the number of cases itself is not a good reason for choosing QCA as a method. When Charles Ragin initially introduced QCA (1987), he intended it to be a method for researchers who wish to combine the best features of both qualitative and quantitative methods. QCA is particularly suitable for addressing causes-of-effects questions because it combines formalized comparison with a strong focus on the complexity and individuality of cases. Thus, QCA enables systematic comparisons of relatively small to large numbers of cases (as a rule of thumb, $N \geq 10$). For example, Freyburg and Garbe (2018) use QCA to compare the occurrence or non-occurrence of internet shutdowns in 33 elections in sub-Saharan African (SSA) countries. The use of QCA can only make sense if the phenomenon of interest has two features: it is plausible to frame it in terms of set relations and of causal complexity.

A Set-Analytic Approach

QCA is a set-theoretic method that has its foundations in Boolean algebra and its fuzzy sets extensions (see Chapter 2). This means, first, that we analyze social phenomena as sets. For example, Freyburg and Garbe examine the set of elections during which the state shut down the internet. In their analysis, 10 elections are members of this set, while 23 elections did not resort to internet shutdown. Second, we analyze how different phenomena relate to each other in terms of set relations. Essentially, we want to know whether specific sets of cases are subsets of other sets of cases. For example, Freyburg and Garbe (2018, p. 3901) assert that the set of elections in SSA countries with internet shutdown is a subset of elections in SSA countries where the state has a majority ownership of the internet service provider:

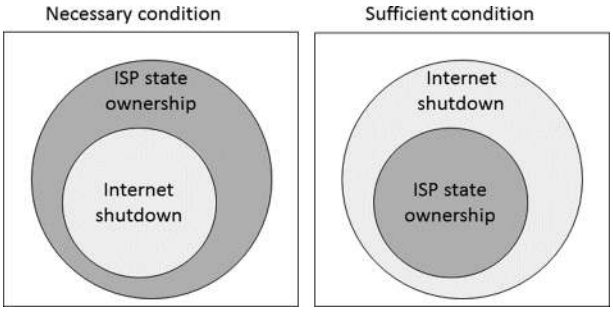

Figure 1.2 Necessary and sufficient conditions as set relations

> We expect a company's commitment to comply with a government's request to shut down its services to be decisively determined by its ownership, notably, facilitated by state majority ownership and hindered otherwise.

This, in fact, is just another way of saying that state majority ownership is a *necessary condition* for internet shutdowns. The left-hand side of the Euler diagram in Figure 1.2 illustrates this.[2] A condition (here: state majority ownership) is a superset of an outcome (here: internet shutdowns) if the outcome is 'hindered' when the condition is not present. This is why on the left-hand side of Figure 1.2, the set of cases with internet shutdowns is fully inside the set of cases with state majority ownership.

Another subset relation that we want to explore with QCA is when a condition is sufficient for an outcome. For instance, state majority ownership would be a *sufficient condition* for internet shutdowns if, whenever the company in charge has state majority ownership, the internet is being shut down during elections. As the right-hand side of Figure 1.2 shows, this is another way of saying that the set of cases with state majority ownership (the condition) is a subset of those cases where the internet was shut down (the outcome). Usually, however, we are less interested in the necessity or sufficiency of single

[2] A Euler diagram allows us to visualize the relationship between various sets by displaying them as overlapping circles (or other shapes) surrounded by a box. Each circle in such a diagram denotes one of the sets included in the analysis. Cases that are situated within the circle are members of that particular set, while cases situated outside the circle are non-members. The box around the circles represents the set of all possible cases that are situated within the scope conditions of a particular research or, in other words, the 'universal' set.

conditions. Instead, we are interested in how social phenomena are often the result of combinations of different conditions. QCA helps us explore just that.

Causal Complexity

Indeed, a core feature underlying QCA that distinguishes it from many other methods is that it acknowledges that we can rarely understand social phenomena by focusing on the role of a single factor on its own. Instead, usually complex combinations of conditions bring about a specific outcome. Thus, when we use QCA we can model the presence of three core elements of *causal complexity*, where different sets combine with the logical operations AND, OR, and NOT (Schneider and Wagemann, 2012, pp. 76–90). For example, internet shutdowns might only occur when there is both state ownership of ISPs AND the government in power is authoritarian (Freyburg and Garbe, 2018). This means that we assume *conjunctural causation*: a given factor might only perform its causal role together with another condition. Second, there might be more than one scenario in which internet shutdowns occur. For instance, internet shutdowns might either occur under conditions of state ownership and authoritarian government OR in order to prevent escalation when there is a high level of electoral violence. In other words, many roads may lead to Rome. The assumption of *equifinality* captures that a given event may have several mutually non-exclusive explanations.

Finally, as we shall see later, the occurrence of internet shutdowns in elections in SSA countries has reasons that do not simply mirror the factors that explain its non-occurrence (Freyburg and Garbe, 2018). Instead, the occurrence of an event – such as internet shutdowns – may have different explanations than its non-occurrence – such as when the internet was NOT shut down during an election. In QCA, we call this phenomenon *asymmetric causation*. In real life, there are many examples of this: for instance, while money alone may not make you happy, its absence can be enough to make you unhappy (Thomann et al., 2018).

In summary, we use QCA when research questions ask for the causes of a given effect, when we are interested in the prevalence of set relations and when we assume that empirical relations are complex. If this is the case, then QCA can serve a variety of uses, some of them more theory driven, others more exploratory (see Berg-Schlosser et al., 2008; Schneider and Wagemann, 2010a; Thomann and Maggetti, 2020). We discuss more assumptions underlying QCA and its implementation within a variety of research approaches in the concluding Chapter 7.

Box 1.3 Core points – Goal and motivation for QCA

- The rationale for using QCA should be based on the affinities between the method and characteristics of the research question and phenomena at hand, rather than number of cases alone.
- QCA is suited for addressing 'causes-of-effects' types of questions that ask for the reasons why a certain phenomenon occurs, rather than 'effects-of-causes' questions that ask for the effect of a specific variable on the outcome.
- QCA should be used when the phenomenon under study is best understood in terms of set relation of necessity and sufficiency.
- QCA should be used when the phenomenon under study is assumed to be causally complex in terms of conjunctural causation, equifinality, and asymmetry.

1.2.2 Before the Analytic Moment

Before we actually analyze data with QCA, that is, before the 'analytic moment', we have to make several decisions related to research design[3] and then assign set membership scores to our cases, the so-called process of calibration.

Research Design

Based on the research question that we have formulated, we can define what the outcome is that we want to explain. For instance, Freyburg and Garbe (2018) seek to explain the occurrence of internet shutdowns during elections in sub-Saharan Africa. As a next step, we will need to select and define a set of conditions that should be relevant in explaining this outcome. This step is called *model specification*. To avoid complications during the analysis, most QCA studies include between three and seven conditions for explaining an outcome of interest.[4] Just as with any other quantitative or qualitative method, we will choose the conditions that we include in our study based on the existing body of relevant theory and empirical findings related to the given research question.

[3] Further useful literature on research design includes Brady and Collier (2010), Gerring (2011) and Goertz and Mahoney (2012).

[4] This is due to problems of theoretical interpretation, on the one hand, and problems related to limited diversity, on the other (i.e., combinations of conditions for which there is not enough empirical evidence; see Chapter 4). The number of conditions can be larger in two-step QCA; see Section 5.4.1, and Schneider (2019).

In our example, Freyburg and Garbe (2018) chose three conditions to analyze internet shutdowns during elections in SSA. The first condition is state ownership of the ISP because strategies of repression, more generally, are particularly effective if the government has control over that particular resource or infrastructure. The second condition is whether the government is an autocracy. Previous studies claim the manipulation of internet access is more prevalent in autocracies. Finally, electoral violence is an important condition because it is thought to trigger protests by opposition forces. Internet shutdown would make it harder for these forces to organize and communicate. After selecting the conditions and the outcome, we will carefully conceptualize these as sets, and think about how we can observe (measure) them in our analysis.[5]

Another step in designing the research is *case selection*. Case selection involves a set of decisions about defining cases (Ragin, 1987), the universe of cases, scope conditions, and the set of cases we include in the analysis. First, we need to define what constitutes a case, and hence what our unit of analysis is (Ragin and Becker, 1992). The *unit of analysis* is the entity of interest which we study as a whole, at the level of which we draw inferences. For instance, Freyburg and Garbe (2018) look at elections as the unit of analysis for studying whether internet shutdowns occur during elections.[6] Choosing an appropriate unit of analysis is a theoretical question: we need to determine at what level we expect the phenomenon of interest – here, internet shutdowns – to take place. Next, we will think about the scope of our research. Freyburg and Garbe (2018) define the scope of their research to involve elections after decolonialization and since the introduction of internet and social media in Africa. The scope conditions help us define the entire universe of possible cases which would in principle be relevant to analyze the research question.

Finally, we always choose cases within the boundaries of the scope we defined. We can either work with the entire universe of cases (or the population), or select a specific set of cases (or a sample) from it. Freyburg and Garbe (2018) apply several selection criteria to choose cases from this universe, both in order to ensure comparability and due to considerations of data availability. Applying temporal criteria, they focus on the period 2014–2016. In spatial terms, they include only SSA countries. Conceptually, they focus on national elections only. This leads them to compare all the 33 presidential and parliamentary elections in SSA between 2014 and 2016, with the exception of

[5] Further literature on conceptualization and measurement includes Adcock and Collier (2001) and Goertz (2012).

[6] The unit of analysis is different from the unit of observation, which is the unit at the level of which we collect data. For example, one can collect individual-level data on public opinion (unit of observation) to obtain a measure of public opinion in a country (unit of analysis) for explaining a country-level phenomenon, for instance, party change.

the ones in Sao Tome and Mauritius, for which data were not available. We will return to the question of case selection in Section 7.2.[7]

Measurement and Calibration

Before we can proceed to the analytic moment, we need to prepare the empirical material – the 'data' – that we can use to compare the cases in the QCA. We have already seen that, in QCA, we think of conditions and outcomes as sets to which cases belong or not. We now need to determine, for each case, the extent to which it belongs to these different sets. The first step in doing so is *measurement*: we need to think about how we can observe the concept that this set stands for in the real world (Adcock and Collier, 2001). For example, to determine the set of elections with ISP state ownership, Freyburg and Garbe (2018) use the percentage of outstanding shares that the state has in ISP in the country. Once we have collected the qualitative and/or quantitative empirical information to measure the conditions and the outcome, we have obtained the 'raw data' for our QCA. In a next step, we need to transform the available data on the cases so that they reflect the sets we are interested in. We call the process of transforming raw data into set membership scores, in order to determine whether and to what extent cases belong to a particular set, *calibration*. For example, Freyburg and Garbe (2018) consider elections in countries where the state has more than 51 per cent of the shares in at least one ISP in the country as members of the set of ISP state ownership, whereas the rest of the elections are not considered members of this set.

In QCA, there are different types of sets. *Crisp sets* are binary: they only distinguish between cases that are members of a set (membership score of 1) and cases that are not members of a set (membership score of 0). Freyburg and Garbe (2018) calibrate crisp-set data on three conditions (ISP state ownership, autocracy, and electoral violence) and the outcome (internet shutdowns) so that each election in SSA countries has a membership score of 1 or 0 in all these sets. However, sometimes we are also interested in the different degrees to which cases belong to a set. For example, we could be interested in different intensities of electoral violence. To this end, we can use *fuzzy sets*, where cases can also belong or not belong to a set to various degrees. Fuzzy-set membership scores vary from 0 to 1 (see Section 2.2.1). For the moment, we stick with so-called crisp sets, that is, sets that only allow membership scores of 1 and 0, to explain the QCA.

[7] Issues of case selection, condition selection and research design with QCA more broadly are also discussed in Berg-Schlosser and De Meur (2009), Rihoux and Ragin (2009, chapter 2) and Mello (2021, chapter 3).