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.

https://doi.org/10.1017/9781009006781.005 Published online by Cambridge University Press

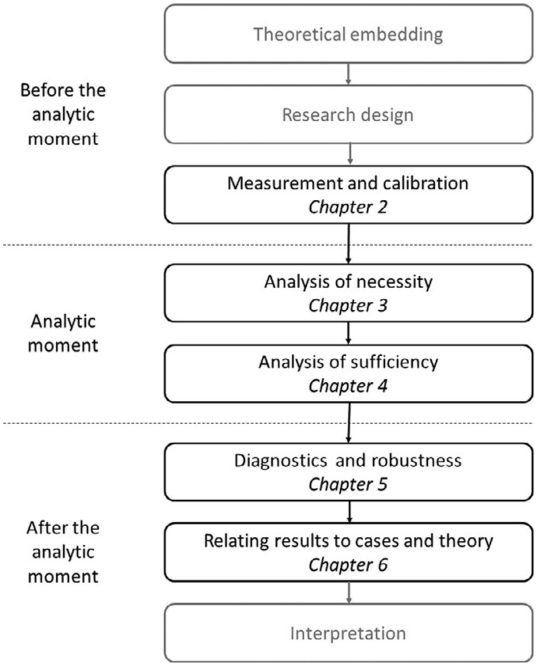

Figure 1.1 Steps of QCA and relevant book chapters

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.

https://doi.org/10.1017/9781009006781.005 Published online by Cambridge University Press

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.

https://doi.org/10.1017/9781009006781.005 Published online by Cambridge University Press

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:

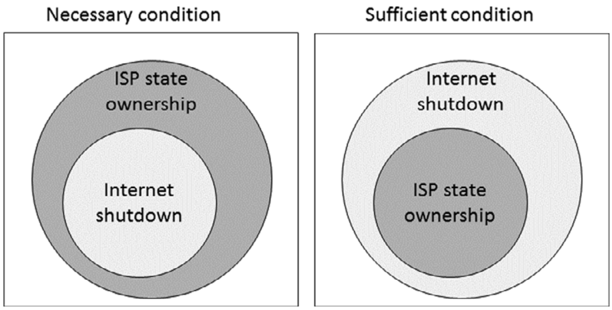

Figure 1.2 Necessary and sufficient conditions as set relations

> 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).

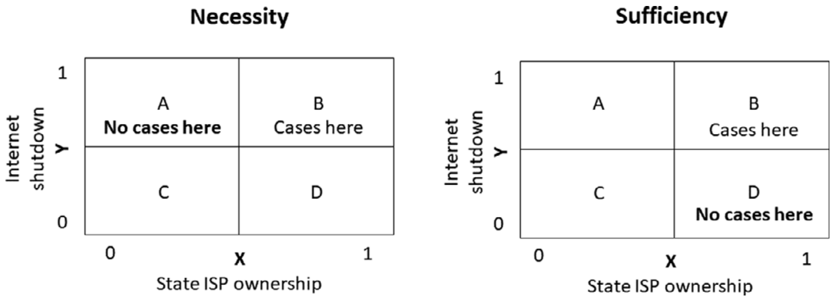

Figure 1.3 Necessity and sufficiency for crisp sets

1.2.3 During the Analytic Moment

After calibrating the raw data into set membership scores, we are ready to proceed to the data analysis in QCA. During the analytic moment, we want to identify whether there are necessary and/or sufficient combinations of conditions for the outcome of interest.

Analyses of Necessity and Sufficiency

As we have seen in Section 1.2.1, to analyze necessity (see Chapter 3), we usually focus on whether each condition or a specific combination of conditions is a superset of the outcome. In other words, we check whether the outcome only occurs in the presence of that condition or that combination of conditions. For us to declare that a condition is necessary, there should not be instances where the outcome occurs but the necessary condition is missing. For example, ISP state majority ownership is necessary for internet shutdowns, if there is no election without ISP state majority where the internet is shut down. Figure 1.3 visualizes this: we do not want to observe cases in quadrant A in the upper-left side. These are the cases that would contradict the statement of necessity, because here the internet was shut down even though the state was not the majority ISP holder. More generally, we write necessity relations as $X \leftarrow Y$. We can read this as 'outcome Y implies condition X', because outcome Y cannot occur in the absence of condition X.[8]

For the analysis of sufficiency (see Chapter 4), we look at conditions that always lead to the outcome. Earlier, we saw in Section 1.2.1 that a sufficient condition is a subset of the outcome. In other words, we want to identify (combinations of) conditions which do not occur together with the absence of

[8] Note that Freyburg and Garbe (2018) did not find any necessary conditions for internet shutdowns in their study. We use this example for illustrative purposes.

https://doi.org/10.1017/9781009006781.005 Published online by Cambridge University Press

the outcome. In the Freyburg and Garbe (2018) example, if ISP state ownership is sufficient for internet shutdown, this means that there must not be cases with ISP state ownership that do not result in internet shutdowns. In Figure 1.3, we thus do not want to see cases in the lower-right quadrant D. We generally write sufficiency relations as $X \rightarrow Y$. We can read this as 'condition X implies outcome Y', because there are no cases with condition X but without outcome Y. To analyze sufficiency, we start by looking at all possible combinations of conditions. We then evaluate which of these combinations are subsets of – in other words, sufficient for – the outcome. To reveal all logically possible configurations of conditions, in QCA we have the *truth table* as a useful analytic tool.

Table 1.1 shows the truth table for the occurrence of internet shutdowns in the study by Freyburg and Garbe (2018). Each row of the truth table represents a specific combination of conditions. And all truth table rows together display all the possible logical AND combinations of conditions. The conditions are either present (1) or absent (0). We observe these configurations in one or several cases. We then determine sufficiency by checking the membership of these cases in the outcome set. If all cases with the same configuration also display the outcome, so-called consistency takes on the value 1. In this case, the configuration is considered sufficient for the outcome (OUTPUT 1). In Chapter 4, we expand on how to perform an analysis of sufficiency by constructing and analyzing truth tables.

To put it in very simple terms, the goal of the truth table analysis is to find the shortest possible expressions of those combinations of conditions that are sufficient for the outcome by eliminating irrelevant (redundant) conditions. For example, if we compare the first two sufficient rows in Table 1.1, we can see that they are identical except the condition AUTOCRACY once being absent and once being present. From this we can conclude that the combination of state ISP ownership in a setting of electoral violence is sufficient for internet shutdowns, irrespective of the presence or absence of an autocratic regime. Similarly, the third and fourth row only differ in the condition VIOLENCE; we can hence 'minimize away' this condition as irrelevant through a pairwise comparison of two otherwise identical sufficient truth table rows.

Performing a truth table analysis in this manner, Freyburg and Garbe (2018) find the following sufficient solution for the occurrence of internet shutdowns (S1). In this expression, the negation of a condition is denoted with a $\sim$. The multiplication sign * reads as logical 'AND', and the plus sign + reads as logical 'OR'.

$$S1: ISP * AUTOCRACY + {\sim}ISP * VIOLENCE \rightarrow SHUTDOWN \quad (1.1)$$

In plain words, this means that there are two scenarios that have typically resulted in internet shutdowns in recent elections in SSA countries. In the first scenario, an autocratic state that holds the majority in ISP ownership shuts

https://doi.org/10.1017/9781009006781.005 Published online by Cambridge University Press

https://doi.org/10.1017/9781009006781.005 Published online by Cambridge University Press

Table 1.1 Truth table, outcome 'Internet Shutdown'

	Conditions					
ISP ownership	Electoral violence	Autocracy	Output	Number of cases	Consistency	Cases
0	1	0	1	1	1	Gabon_16
0	1	1	1	3	1	Republic of Congo_16, Sudan, North_15, Uganda_16
1	0	1	1	4	1	Equatorial Guinea_16, Ethiopia_15, Gambia_16, Togo_15
1	1	1	1	2	1	Burundi_15, Chad_16
0	0	0	0	7	0	Burkina Faso_15, CAR_15, Ghana_16, Guinea_15, Ivory Coast_16,Lesotho_15, Malawi_14
0	0	1	0	2	0	CAR_15, Mauritania_14
1	0	0	0	11	0	Benin_15, Benin_16, Botswana_14, Djibouti_16, Guinea_Bissau_16, Ivory Coast_16, Namibia_14, Niger_16, Nigeria_15, Zambia_15, South Africa_14
1	1	0	0	3	0	Mozambique_14, Tanzania_15, Zambia_16

down the internet during elections. In the second configuration, ISP ownership is not public and there were significant incidents of electoral violence. The authors then discuss the cases that display these specific configurations. This helps them to illustrate the underlying mechanisms behind why private ISP owners may be willing to temporarily shut down their services.

We are not only interested in why the outcome occurs, but we also usually want to analyze why the outcome does *not* occur. For example, it is not only interesting to know under what conditions governments in SSA countries shut the internet down, we also want to learn about the conditions under which they do *not* shut the internet down. We call this the ‘negation’ of the outcome. Accordingly, the QCA often involves four distinct analyses: necessary conditions for the outcome, necessary conditions for the non-occurrence of the outcome, sufficient conditions for the outcome, and sufficient conditions for the negated outcome. For example, Freyburg and Garbe (2018) find that different configurations of conditions are sufficient for the non-occurrence of the outcome (S2) ‘No internet shutdown’ (see S2):

$$S2 : \sim ISP * \sim VIOLENCE + ISP * \sim AUTOCRACY \rightarrow \sim SHUTDOWN \tag{1.2}$$

They find two paths that do not result in an internet shutdown: private ISP ownership when there is no electoral violence, or state ISP ownership in non-authoritarian governments. These QCA results reflect all the different aspects of causal complexity. Firstly, we have evidence of conjunctural causality: the single conditions do not produce internet shutdowns (or their absence) on their own, but only in combination with others. For example, autocratic government is only a factor leading to shutting the internet down if the government also has the majority ownership of ISP. Secondly, the results also refect equifinality: there is more than just one way the outcome can occur. For example, internet shutdowns during elections can happen when there is state ownership of an ISP in an autocracy; but shutdowns also happen when there is violence during the election combined with no state ownership of an ISP. Thirdly, these results reflect asymmetric causality: the sufficient conditions for no internet shutdowns (S2) are not just the ‘opposite’ of the sufficient conditions for internet shutdowns (S1).

Less-than-Perfect Set Relations

We should note that social science data tend to be considerably more noisy than the examples we presented so far in Figure 1.3, which depicts perfect subset relations. In practice, we often allow for small deviations from perfect patterns of necessity and sufficiency. In other words, we want these deviations not to

https://doi.org/10.1017/9781009006781.005 Published online by Cambridge University Press

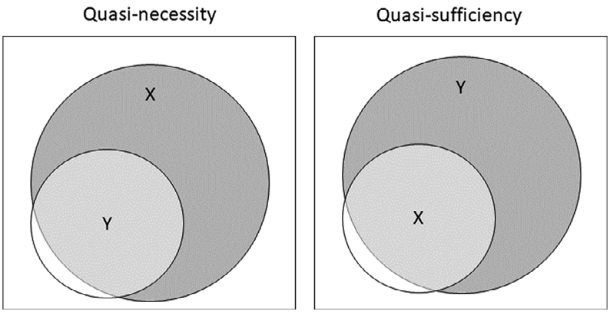

Figure 1.4 Less-than-perfect set relations

exceed a certain, still acceptable, degree. Figure 1.4 illustrates this. For this purpose, with QCA several parameters of fit for necessity and sufficiency allow us to quantify the extent to which set relations deviate from being 'perfect'. In Figure 1.4, these parameters would tell us the relative size of the area where the condition set and the outcome set do not overlap. We introduce the parameters of fit for necessity and sufficiency in Chapters 3 and 4, respectively.

Set Relations with Fuzzy Sets

It is easy to grasp set relations with crisp sets where we only deal with the occurrence or non-occurrence of conditions and the outcome. In the case of fuzzy sets, cases are members in sets to different degrees. In principle, the analyses of necessity and sufficiency for fuzzy sets work in a very similar manner as with crisp sets: we identify supersets (necessity) or subsets (sufficiency) of the outcome. The main difference is that we also take into account the more fine-grained differences 'in degree' in set membership, not only binary differences 'in kind'. In practice, the way this plays out is that we analyze whether the condition set is greater than (necessity) or smaller than (sufficiency) the outcome set. For example, a condition is a superset of the outcome (necessary) if the membership scores of cases in that condition are greater than or equal to their membership in the outcome set.

In the left part of Figure 1.5, we illustrate how this looks when we plot the condition and the outcome. The cases that cluster below the diagonal conform to a pattern of necessity because their membership in the condition set is always greater than or equal to their membership in the outcome set. Conversely, the cases above the diagonal are inconsistent with the statement of necessity. For sufficiency, it is the other way around: as the right part of Figure 1.5 illustrates, the set of cases that are a subset of the outcome (sufficiency) is the set of cases whose membership in the condition set is smaller than or equals their membership in the outcome set. For sufficiency, the cases that weaken the sufficiency statement are hence those that cluster below the diagonal.

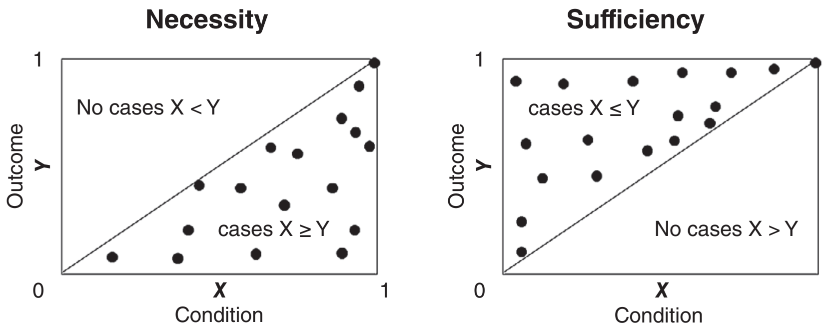

Figure 1.5 Necessity and sufficiency for fuzzy sets

1.2.4 After the Analytic Moment

Once we have identified necessary and sufficient (combinations of) conditions for the outcome, we can apply several procedures for increasing the confidence in our results, for clarifying the extent to which our findings can 'travel' to other contexts, and for deriving more abstract conclusions. Figure 1.6 illustrates how we usually draw inferences at three distinct levels with QCA: the actual set of cases we analyzed, the universe of cases, and the body of theory or abstract knowledge to which we seek to contribute with our research (Thomann and Maggetti, 2020).

On the one hand, we will perform a variety of diagnostics and check the robustness of the results within our sample. Readers will learn more details about this in Chapter 5. Moreover, after the QCA analysis, we often bring individual cases back to the forefront. For instance, we would perform case studies in order to analyze the mechanisms that underlie the different sufficient paths toward the outcome. The rules governing the choice of cases based on a QCA result are spelled out in the framework of set-theoretic multi-method research (SMMR) and we discuss them in more detail in Section 6.3. Freyburg and Garbe (2018), for example, discuss the cases of the Republic of Congo and Uganda in depth, arguing that they are especially relevant because they involve internet shutdowns by private companies. We will also want to learn from cases that deviate from the perfect necessity or sufficiency pattern. This can help us identify potential limitations of our research.

On the other hand, there will also be an element of abstraction 'upwards' (see Figure 1.6), where we relate the results to existing theories and concepts or draw other abstract conclusions on policy recommendations. The detailed procedure for how this can be done in QCA is explained in Section 6.2 and Chapter 7. For instance, Freyburg and Garbe (2018) conclude from their analysis that ISP ownership is critical to understanding authoritarian practices violating citizens'

https://doi.org/10.1017/9781009006781.005 Published online by Cambridge University Press

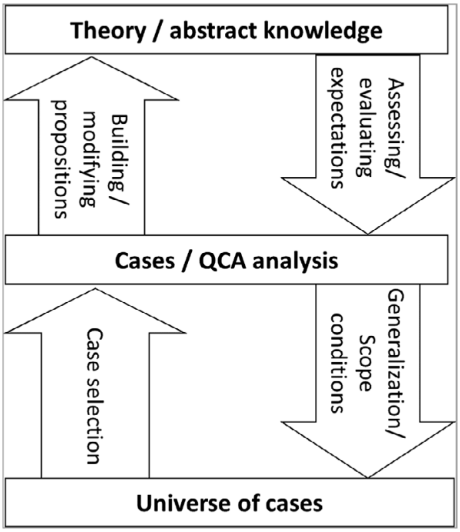

Figure 1.6 Levels of inference in QCA research
Note: Own illustration inspired by Thomann and Maggetti (2020).

freedom of expression in the digital sphere. Finally, we draw inferences 'downwards' toward the relevant universe of cases (see Figure 1.6). This process often involves defining *scope conditions*. Scope conditions represent the conceptual and empirical boundaries within which the theory we use applies or where we expect the phenomenon of interest to take place. They are directly related to how we have earlier defined the relevant universe of cases. When discussing scope conditions, we reflect on the extent to which the results can 'travel' in spatial and temporal terms, but also in terms of the types of cases we might be able to generalize to (Goertz and Mahoney, 2009). In this vein, Freyburg and Garbe (2018) call for more research that accounts for temporally and spatially varying levels of internet penetration and the effects of different ownership structures.

1.3 Spread of QCA and Related Software

While still not considered a 'mainstream' method, QCA has established itself as a methodological tool in the social sciences and beyond. Between 1994 and 2019, in the Web of Science Core Collection database, we find 611 publications describing or using QCA (see Figure 1.7).[9]

[9] Search terms were QCA AND 'Qualitative Comparative Analysis' in the title, abstract, author keywords, and Keywords Plus. The Web of Science Core Collection includes sciences, social

https://doi.org/10.1017/9781009006781.005 Published online by Cambridge University Press

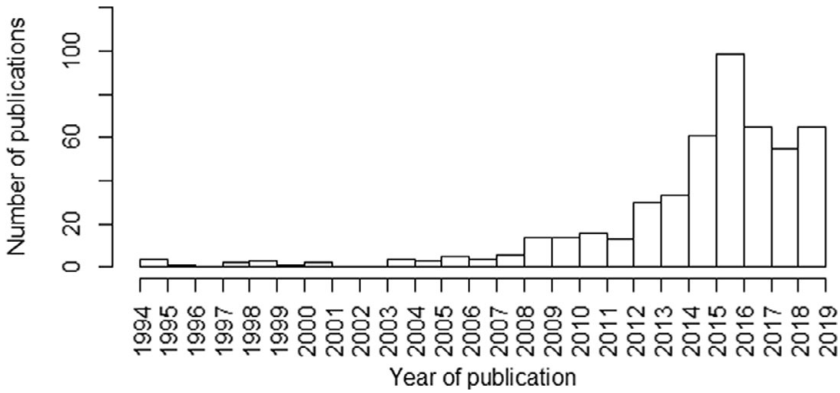

Figure 1.7 Development of QCA publications, 1994–2019

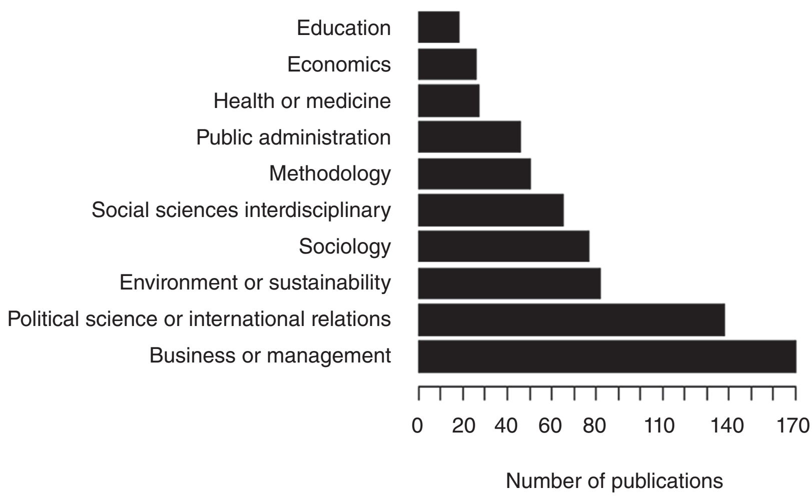

Figure 1.8 QCA-related publications per discipline, 1994–2019

Source: Own illustration based on Web of Science including the 10 most frequent disciplines. Publications can be classified under more than one discipline.[10]

As Figure 1.8 shows, QCA is used prominently in four fields: political science and international relations, sociology, environment and sustainability, and particularly business and management (Wagemann et al., 2016).

sciences, and arts and humanities. The figures include publications between 1987 and 2019, and no publications before 1994 are recorded in the database. Date of search: November 11, 2019.

[10] Combined categories as follows (double counts are possible): Environment or Sustainability = ENVIRONMENTAL STUDIES, ENVIRONMENTAL SCIENCES, PUBLIC

https://doi.org/10.1017/9781009006781.005 Published online by Cambridge University Press

Table 1.2 Top 10 journals publishing QCA-related research, 1987–2019

Journal name	N
Journal of Business Research	61
Sociological Methods & Research	21
Quality & Quantity	17
International Journal of Social Research Methodology	11
European Journal of Political Research	7
Journal of European Public Policy	7
Political Analysis	7
Political Research Quarterly	6
Comparative Political Studies	5
Field Methods	5

Source: Web of Science.

Table 1.2 identifies the top 10 academic outlets with a strong record of publishing QCA-related research. They cover the disciplines of sociology and business research and include some of the most prestigious political science journals. It is notable that five out of the 10 journals are methodological outlets: *Sociological Methods and Research, Quality and Quantity,* the *International Journal of Social Research Methodology, Political Analysis,* and *Field Methods*. These figures reflect that QCA has become an increasingly robust, sophisticated, and attractive tool for empirical social research.

In parallel, Figure 1.9 shows how the use of `R` packages has increased exponentially in recent years (see also Thiem and Dusa, 2013; Verweij and Trell, 2019) and are catching up in numbers with more widely used 'traditional' software, fs/QCA (Ragin and Davey, 2016), and Tosmana (Cronqvist, 2019).[11] The latter rely on 'point-and-click' style graphical user interfaces, whereas `R` is a command-line software environment. Once users master `R` to a sufficient degree – something our book is meant to help with – it provides several advantages. First, analytic flexibility is much higher, as one can customize very precise analytic steps, rather than relying on a few standardized options. The `R` packages `SetMethods` (Oana and Schneider, 2018) and `QCA` (Dusa, 2018) offer superior functionality compared to other QCA-specific software (see Table 7.4 for details). Second, and related, because `R` is open-source, methodological innovations continue to quickly find their way to QCA practitioners via the

ENVIRONMENTAL OCCUPATIONAL HEALTH and GREEN SUSTAINABLE SCIENCE TECHNOLOGY; Methodology = SOCIAL SCIENCES MATHEMATICAL METHODS and STATISTICS PROBABILITY; Health or Medicine= HEALTH CARE SCIENCES SERVICES, HEALTH POLICY SERVICES and MEDICINE GENERAL INTERNAL.

[11] See http://compasss.org/software/ for a more complete list of QCA-related software.

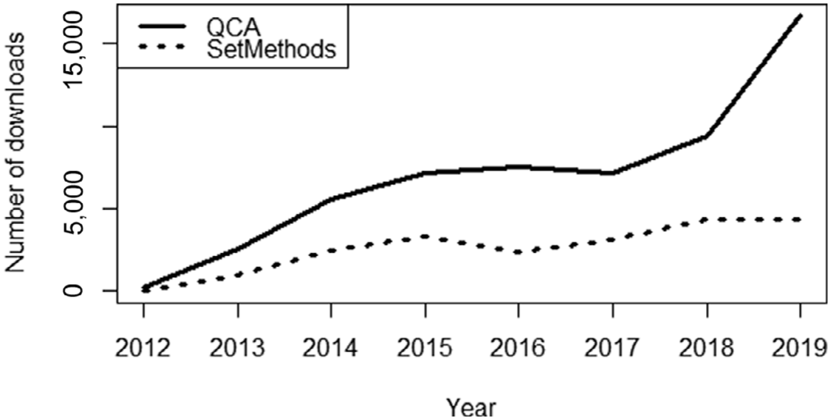

Figure 1.9 Downloads of R packages QCA and SetMethods, 2012–2019

Source: CRAN repository, retrieved November 27, 2019. The figure does not include downloads from github, therefore underrepresentation of the use of SetMethods is likely.

dedicated R packages. Third, QCA analyses in R are easier to replicate, as one can write down and store every analytic step that was performed in order to obtain the results. This is useful not only for others, but also for organizing one's own research project, especially bigger ones that last for longer periods of time and entail many analyses.

Increased computational sophistication may come with practical hurdles for QCA users of the R software. This book serves a simple purpose: to facilitate the efficient teaching, use, and independent learning of state-of-the-art QCA with the best available technology. As readers will see, learning QCA with R is actually a didactic advantage rather than a burden. The use of command lines helps to better understand the logic of the analytic steps performed during QCA. Additionally, in order to aid users in their R learning process, this book is complemented by two resources available in the online appendix, which can be found at https://doi.org/10.7910/DVN/S9QPM5. First, we provide readers with an 'Intro to R' chapter to kickstart the software induction. Second, we offer a series of datasets and template files containing all the basic R commands used through the different steps of a QCA which can be directly customized and copied into R scripts.

1.4 Summary and Logic of the Book

In this chapter, we have introduced QCA as a method that helps us identify necessary and sufficient conditions for an outcome of interest. QCA allows us to model three core aspects of causal complexity. First, conjunctural causation

https://doi.org/10.1017/9781009006781.005 Published online by Cambridge University Press

means that different conditions combine to produce an outcome. Second, equifinality means that one outcome may have several, mutually non-exclusive explanations. Finally, asymmetric causation prevails as the occurrence of the outcome has a different explanation than its non-occurrence. QCA is a set-theoretic method. This means that we attribute cases to sets that represent the outcome we want to explain and the conditions we assume to be relevant for this outcome, and we analyze necessary and sufficient conditions as set relations.

1.4.1 Structure of the Book

The book guides the reader through the different analytic steps of QCA before, during, and after the truth table analysis. In the phase before the analytic moment, we design our research, conceptualize cases and sets, and transform them into 'data'. The process of attributing cases to sets is called 'calibration'. Chapter 2 further develops the notions of sets and describes several available procedures for calibrating raw data into set-membership scores. This chapter also discusses in more depth the logical operators that we use to combine sets and their negation. Based on this, we introduce different ways of forming the concepts at the heart of our set-analytic procedure. While we return to the question of case selection in Section 7.2, our book does not aim to provide an extensive theoretical treatment of general research design issues. Issues of case selection, condition selection, and research design with QCA more broadly are discussed in Berg-Schlosser and De Meur (2009), Rihoux and Ragin (2009, chapter 2), and Mello (2021, chapter 3).

The 'analytic moment' refers to the actual analyses of necessity and sufficiency. Chapter 3 introduces the analysis of necessity, when the condition is a superset of the outcome and the respective parameters of fit. Chapter 4 is about analyzing sufficient conditions that are subsets of the outcome set with the help of truth tables and parameters of fit. In this chapter, we also expand on the problem of 'limited empirical diversity', when our cases do not populate all the rows in our truth table. We present various strategies to deal with this problem. For a more in-depth theoretical discussion on issues pertaining to the analytic moment, especially regarding pitfalls in the analysis of necessity and sufficiency, we point the reader to Schneider and Wagemann (2012, chapters 8 and 9).

After the analytic moment, we need to interpret the results and check how 'good' they are. In Chapters 5 and 6, we discuss advanced analytic tools that can – or even should – be used after a QCA result has been obtained. Finally, in Chapter 7 we discuss the steps and options involved in interpreting QCA results more generally, and outline standards of good practice. Throughout the book, we highlight how case knowledge informs QCA at the various stages of the analysis. We use a rich body of examples, some fictitious, some based on real data, to illustrate the different concepts and steps.

1.4.2 How to Use This Book

With this book, we seek to make the lives of students, teachers, researchers, and practitioners as easy as possible when performing the most complete QCA protocol with `R`. For this, we adopt an applied and practical focus. For recent comprehensive theoretical introductions to QCA, we point the reader to, among others, Rihoux and Ragin (2009) and Schneider and Wagemann (2012). These books provide in-depth theoretical discussions of many of the issues treated in this book, but without the focus on hands-on implementation using `R`. Dusa (2018) provides a more advanced guide to QCA with `R`, focused particularly on the `QCA` package, but less accessible for readers without prior knowledge of the method and software. Mello (2021) also provides a beginner-oriented introduction to QCA, including an appendix of the main `R` functions, but does not include the state-of-the-art techniques treated in the later chapters of this book, particularly those to be performed after the analytic moment.

All the chapters in the book follow the same structure. We first introduce the topic of the chapter and specify learning goals. This is followed by a very brief conceptual description of the analytic procedure treated in the chapter. Most often, we construct some hypothetical examples to illustrate this conceptual introduction. Next, we show how to implement this procedure in `R` using RStudio (RStudio Team, 2019) with packages `QCA` and `SetMethods`, sometimes using an empirical example from a published QCA study. We explain the structure of the `R` outputs and how to interpret them. Each chapter concludes with a summary of take-home points. Throughout the book, we use boxes to highlight learning goals, core points, empirical examples, and tips for good practices.

In addition, the freely available online appendix to this book provides a rich body of additional resources. First, it contains additional guidance on implementing specific parts of the analysis in `R`, including a hands-on introduction to using `R` and handling datasets in RStudio. Second, the online appendix features all datasets and `R` scripts to replicate all analyses performed in this book. Third, users can find ready-made template commands with explanations of the most important options online, structured along the analytic steps of QCA. Fourth, the online appendix contains test questions (and answers), as well as exercises (and solutions) which readers and instructors can use as learning resources.

We suggest that readers start the book from page 1 and follow thoroughly the material in Chapters 1–4, as well as Section 5.2 and Chapter 7. In later chapters, we assume that readers know the material introduced in preceding chapters.

In order to test whether or not readers have mastered the material covered in a chapter, we suggest trying our online test questions. Sections 5.3 and 5.4 and Chapter 6 are catered to more advanced users and contain analytic steps that can be used after the analytic moment in order to make more out of the QCA results obtained. We wish the reader an enjoyable reading and learning experience!