Chapter 6

Measures of Fit

Mello, Patrick A. (2021) *Qualitative Comparative Analysis: An Introduction to Research Design and Application*, Washington, DC: Georgetown University Press, Chapter 6.

set relations are important … in the same way that assessments of significance and strength are important in the analysis of correlational connections.

Charles C. Ragin [1]

In the course of its development over the past 30 years, QCA has undergone substantial methodological sophistication. As a Boolean approach the method initially only worked with binary values and did not allow for contradictory truth table rows where some cases show the outcome and others do not – such configurations needed to be resolved through measures of *research design*, before one could proceed with the analysis (Ragin et al. 1984; Rihoux and De Meur 2009; Rihoux and Ragin 2009). However, perfect set relations can rarely be found in the social sciences. More often one identifies a relationship that comes close to necessity or sufficiency, but where some empirical cases in the data do not fit such patterns. How to move on under these circumstances? What proportion of cases will be *enough* to merit a set-theoretic relationship?

Hence, following the introduction of fuzzy sets (Ragin 2000), which allowed for differentiated degrees of set membership, Charles Ragin (2006a) put forth the measures of fit *consistency* and *coverage* to assess imperfect set-theoretic relationships in empirical data. These were a major step forward in the methodological development of QCA. While these metrics should never replace case knowledge and substantive interpretation, they are a valuable addition to the method because they gave researchers benchmarks to assess and compare their results. They also furthered transparent standards for applied research (Oana et al. 2021; Rihoux and Ragin 2009; Schneider and Wagemann 2010; 2012).

Consistency and coverage can be compared to the well-known statistical indicators of significance and strength, as highlighted in the above quote. Similar to statistical significance, *consistency* measures the degree to which an empirical relationship between a condition or combination of conditions and the outcome comes close to set-theoretic necessity and/or sufficiency. Similar to statistical strength, *coverage* measures the empirical importance or relevance of a condition or combination of conditions (Ragin 2008, 45). That said, set relations should *not* be equated with metrics of statistical association. They measure different things but serve similar purposes.

While consistency and coverage remain the standard indicators to formally assess set-theoretic relationships, over the years it has been pointed out that these measures could not detect issues that may occasionally arise in empirical data. Hence researchers developed several additional metrics. Among these, "proportional reduction in inconsistency" (PRI) addresses simultaneous subset relations (Ragin 2006b; Schneider and Wagemann 2012), and "relevance of necessity" (RoN) seeks to distinguish trivial from relevant necessary conditions (Schneider and Wagemann 2012). Both of these have been incorporated into the major software packages and it has become customary to report these metrics in publications.[2] As with other parts of the method, measures of fit are an area of continuous development and methodological refinement. In that vein, several other metrics have been proposed, including alternative measures of consistency and coverage (Haesebrouck 2015; Stoklasa et al. 2017; Veri 2018; 2019), to assess the distance of an empirical pattern from a set-theoretic relationship (Eliason and Stryker 2009), measures to evaluate the importance of single conditions in QCA solutions (Damonte 2018), distinct approaches to assess the degree of necessity (Dul et al. 2020; Vis and Dul 2018), and metrics for combinations of conditions that are jointly necessary (Bol and Luppi 2013).

Building on the introduction of set relations in Chapter 4, this chapter introduces the standard measures of fit that are essential for QCA, explains how they are calculated, and provides illustrative examples to demonstrate the issues at stake. This means that the discussion is focused on consistency, coverage, PRI, and RoN – whereas a broader treatment of metrics is beyond the scope of this chapter. Of course, for a standard application of QCA, users do not have to *manually* calculate the measures of fit – that is what the software does – but users will benefit from going through the calculations to understand how different scores are derived and how these can be interpreted.

Before proceeding, one caveat is in order. The measures of fit discussed in this chapter can help to identify set-theoretic relationships in empirical data, but any interpretation, particularly a *causal interpretation* of the identified patterns must always be grounded in theory and substantive knowledge (see the discussion in Chapter 4). Like the old adage, *correlation is not causation*, we should acknowledge that *set relation is not causation*. Whether identified patterns are meaningful always depends on the research design of a study (Chapter 2), the strength of the empirical evidence, and its theoretical and substantive interpretation by the researcher.

Set-Theoretic Consistency

The measure of *consistency* is used to assess "the degree to which the cases sharing a given combination of conditions […] agree in displaying the outcome in question" (Ragin 2008, 44). In other words, consistency helps to determine the "fit of the empirical evidence with an assumed set-theoretic relationship" (Mello 2017, 127). As such, the measure is applied to assess the consistency of necessary conditions and of sufficient conditions, and it can be used on combinations as well as individual conditions.

In formal terms, consistency is calculated to reflect the extent to which there is a set relation between instances of a condition and an outcome. When all values for the outcome Y are equal to or less than the respective values for X, then Y is a *subset* of X (vice versa, X is a superset of Y) and hence X is a *necessary condition* for Y. Likewise, if all values for Y are greater than or equal to the respective values for X, then Y is a *superset* of X (which also means that X is a subset of Y) and X is thus a *sufficient condition* for Y. This means that when all values for the outcome Y and the condition X are exactly equal (in an XY plot with fuzzy data all points would be on the diagonal line), then X is *both* a necessary and sufficient condition for Y.

Before moving to the calculation of consistency, it is useful to depict the set-theoretic relationships of necessity and sufficiency in visual terms. Recalling the illustrations from Chapter 3, the Euler diagrams in Figure 6.1 show these relationships for crisp sets, assuming data with binary values. X_1 is a perfectly consistent necessary condition, whereas as X_2 is a perfectly consistent sufficient condition for Y. The XY plots in Figure 6.2 show the equivalent set-theoretic relationships for fuzzy sets, with graded set-membership values. In these plots, the cases either all populate the area *below* or on the main diagonal (necessary condition X_1), or they are located on or *above* the diagonal line (sufficient condition X_2).

Figure 6.1 Necessary Condition and Sufficient Condition (Crisp Sets)

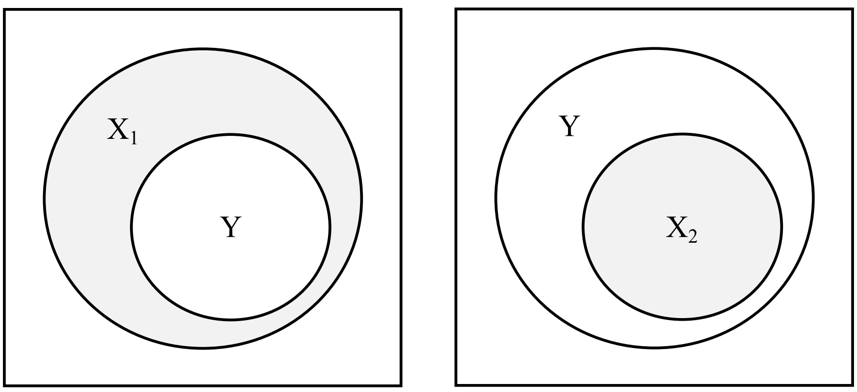
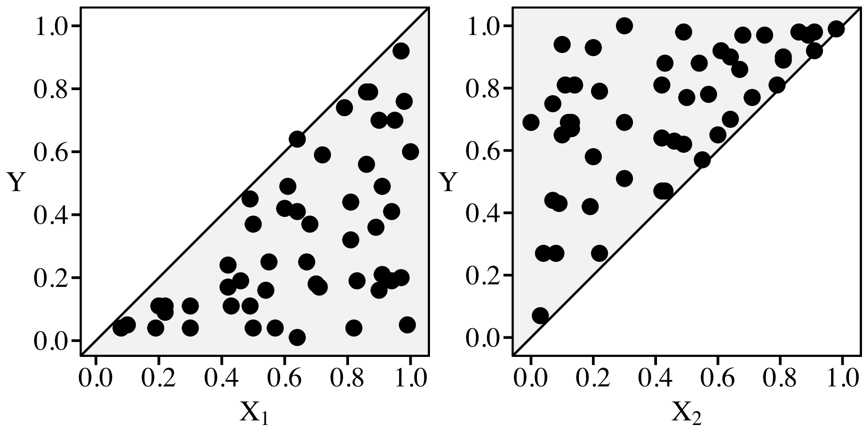


Figure 6.2 Necessary Condition and Sufficient Condition (Fuzzy Sets)



How can we express these relationships numerically? The consistency of necessary and sufficient conditions is calculated with the following two formulae, which differ only in the denominator (Ragin 2006a, 297; 2008):

$$\text{Consistency}_{\text{Necessity}}(Y_i \leq X_i) = \frac{\sum \min(X_i, Y_i)}{\sum Y_i}$$

$$\text{Consistency}_{\text{Sufficiency}}(X_i \leq Y_i) = \frac{\sum \min(X_i, Y_i)}{\sum X_i}$$

Aimed to detect formal subset relationships, each formula divides the sum of the minimum set membership scores for the condition and the outcome by the sum of the scores for the outcome (for *necessary* conditions), or by the sum of the scores for the condition (for *sufficient* conditions). For perfect subset relations, this calculation yields consistency scores of 1. For imperfect set relations, where one or more cases violate a statement of necessity or sufficiency, the resulting scores will be less than 1. The similarity in the formulae reflects the inverse relationship between necessity and sufficiency.

What is an appropriate level of consistency? At which point can we declare a condition or a combination necessary and/or sufficient for an outcome? Generally speaking, the closer consistency scores approximate 1, the more confident we can be about a set relation. For sufficient conditions, the analysis involves the minimization of the truth table (covered in Chapter 7). Here, the recommended *minimum* benchmark for the inclusion of truth table rows is 0.75 consistency (Ragin 2008, 46; Schneider and Wagemann 2012, 279). But it must be noted that this is but *one step* on the way to the QCA solution terms. Typically, the truth table analysis will include several rows with consistency levels that are higher than the threshold of 0.75. Thus the overall solution term will most often *exceed* the minimum consistency threshold. For necessary conditions, the analysis does not entail the truth table procedure and the theoretical focus typically rests on single conditions. Hence it is recommended to use a consistency threshold of 0.90 for necessary conditions (Schneider and Wagemann 2012, 278).

Of course, these thresholds are *rules of thumb* and one should always use case knowledge and individual judgment when interpreting empirical data. Especially the set-theoretic requirement of 0.90 consistency for necessary conditions is fairly demanding. This means that it is not often reached in empirical settings, even when there are patterns in the data. For example, in a study of military coalition defection (Mello 2020), the condition "upcoming elections", which was deemed to be theoretically important, yielded a consistency of 0.83 and was thus formally speaking not necessary. However, a standard chi-square test of association showed that there was a statistically significant difference between groups that faced elections and those who did not, as 16 out of 18 cases showed the pattern (Mello 2020, 60).

Besides, one also has to take into account the *number of cases* involved. The smaller the number of cases, the higher the consistency thresholds should be set. For example, in a QCA study with only 12 cases, we would expect intimate knowledge of the selected cases and thus there should be very high or even perfect consistency. On the contrary, for a QCA study with 50 or even 80 cases, a lower level of consistency is anticipated and would certainly be acceptable (but consistency must still meet the minimum thresholds for set relations).

Calculating Consistency: An Example

Let us look at an example on how consistency is calculated. Suppose a group of office workers in a small company have all been tested positively for the coronavirus. After their infection with the COVID-19 illness, some of them developed serious symptoms, whereas others only showed mild symptoms, or none at all. How to explain this variation? Table 6.1 lists hypothetical data on the eight workers and their membership in the fuzzy sets *serious symptoms* (outcome), and two supposed explanatory conditions: *weak health* and *old age*. Weak health takes into account the severity of prior medical conditions, such as lung disease or diabetes, among other risk factors. Old age acknowledges that older adults are at a significantly higher risk of experiencing complications from the illness. Scores above 0.5 indicate that the person is *rather inside* the set, whereas a score of 1 refers to *full membership* in the respective set.

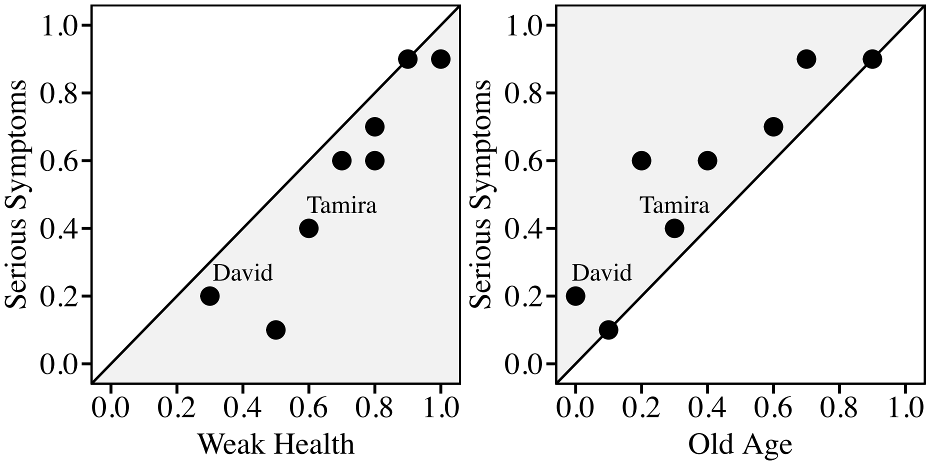
Table 6.1 Example: Coronavirus and Risk Factors

Person	Weak Health	Old Age	Serious Symptoms (Outcome)
Linda	1.0	0.7	0.9
George	0.9	0.9	0.9
Megan	0.8	0.6	0.7
Carlos	0.7	0.4	0.6
David	0.3	0.0	0.2
Sara	0.5	0.1	0.1
Tamira	0.6	0.3	0.4
Jamar	0.8	0.2	0.6

Before coming to the calculation of consistency, we can visualize the conditions and the outcome to see whether there are indications of set-theoretic relationships. Figure 6.3 shows separate XY plots for the conditions *weak health* and *old age* against the outcome *serious symptoms*. Two cases are labelled for illustrative purposes. In the left-hand plot, we can see that the condition weak health fits the pattern for a perfectly consistent *necessary condition*, because all of the cases are on or below the diagonal line. This indicates that all set membership values for weak health are equal to or larger than the respective set membership values for the outcome. In other words, the condition fully encloses the outcome (formally speaking, the condition is a *superset* of the outcome). When we look at the cases, we can see that the five people who are rather in the set *serious symptoms* also hold membership in *weak health* above 0.50. There is no person with serious symptoms without weak health. Yet, weak health is not sufficient for the outcome, as evidenced by the office worker Tamira, who has weak health but is rather outside the set serious symptoms.

Likewise, the right-hand plot shows that the condition old age fits the pattern for a perfectly consistent *sufficient condition*, since all cases are located on or above the diagonal line. This reflects a situation where all set membership values for the condition are smaller than or equal to the respective set membership values for the outcome. This means that the condition old age is a *subset* of the outcome, as the latter fully encloses the former. For the cases this means that whenever a person holds membership in *old age*, then they also show *serious symptoms.* Out of our group of workers, this applies to Linda, George, and Megan.

Figure 6.3 XY Plots: Risk Factors and Serious Symptoms



How to calculate consistency? To do that, we can apply the formulae introduced above to the data from our example. Table 6.2 shows the data plus the sums of the fuzzy-set membership values of our office workers for the two conditions and the outcome, as well as the minimum values across each condition and the outcome, as required for the calculation of consistency. For set-theoretic *necessity*, the outcome should be a perfect subset of the condition. For weak health (X_1) this is calculated by taking the sum of the minimum values across X_1 and Y and dividing this score by the sum of the values for the outcome. This means that we divide 4.4 by 4.4, which equals 1. This confirms that weak health is a perfect necessary condition for serious symptoms, as we already knew from the XY plot. For old age we divide 3.2 by 4.4, which equals 0.73. This is well below the threshold of 0.9 and hence old age cannot be considered a necessary condition for serious symptoms. This is reflected in the fact that the office workers Carlos and Jamar are both rather inside the set serious symptoms but rather outside the set old age (which we can see from Table 6.1).

Table 6.2 Calculating Consistency: Coronavirus Example

Person	Weak Health (X_1)	Old Age (X_2)	Serious Symptoms (Y)	Minimum (X_1, Y)	Minimum (X_2, Y)
Linda	1.0	0.7	0.9	0.9	0.7
George	0.9	0.9	0.9	0.9	0.9
Megan	0.8	0.6	0.7	0.7	0.6
Carlos	0.7	0.4	0.6	0.6	0.4
David	0.3	0.0	0.2	0.2	0.0
Sara	0.5	0.1	0.1	0.1	0.1
Tamira	0.6	0.3	0.4	0.4	0.3
Jamar	0.8	0.2	0.6	0.6	0.2
Sum	5.6	3.2	4.4	4.4	3.2

Given the data in Table 6.2, the calculation of consistency is as such:

$$\text{Consistency}_{\text{Necessity}}(X_1) = \frac{4.4}{4.4} = 1 \qquad \text{Consistency}_{\text{Necessity}}(X_2) = \frac{3.2}{4.4} = 0.73$$

For set-theoretic *sufficiency*, the relationship should be inverted, which means that the respective condition should be a subset of the outcome. We calculate the consistency of sufficient conditions by taking the sum of the minimum values across the condition and the outcome and dividing this score by the sum of the values for the condition. For *weak health* (X_1), we thus divide 4.4 by 5.6, which is 0.79. This means that the condition weak health can be considered a *weak* sufficient condition for the outcome serious symptoms, because the relationship is far from perfect, but above the minimum threshold of 0.75.[3] We can see this reflected in our data, where a close correspondence exists between the values for weak health and serious symptoms. Yet, values in the former frequently exceed those in the latter. In qualitative terms, Tamira is one person who is rather inside the set weak health but rather outside the set serious symptoms, which also speaks against weak health being a perfect sufficient condition. For the condition *old age* (X_2), we divide the sum of the minimum values across the condition and the outcome (3.2) by the sum of the values for the condition (3.2), which equals 1. Hence, for this hypothetical example, old age is a perfect sufficient condition for the outcome serious symptoms. There are three people in our group of workers who are rather inside this set (Linda, George, and Megan), and all three of them also show high fuzzy-set membership in the outcome serious symptoms. Based on the data in Table 6.2, we calculate this as such:

$$\text{Consistency}_{\text{Sufficiency}}(X_1) = \frac{4.4}{5.6} = 0.79 \qquad \text{Consistency}_{\text{Sufficiency}}(X_2) = \frac{3.2}{3.2} = 1$$

Set-Theoretic Coverage

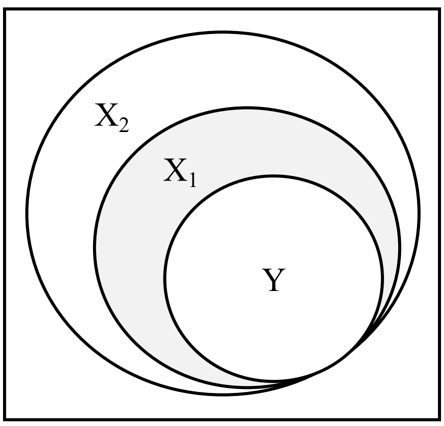
The measure of *coverage* is used to assess "the degree to which a cause or causal combination 'accounts for' instances of an outcome" (Ragin 2008, 44). Put differently, coverage is used to determine "the *relevancy* of a condition in empirical terms" (Mello 2017, 128). For sufficient conditions, coverage indicates "how much" of the empirical evidence is explained by a given condition or combination. For instance, we may identify a condition that, whenever present, leads to the outcome. Yet only a single case of our population shows this condition. Formally speaking, this may be a sufficient condition but its contribution to the explanation of the outcome is rather small.

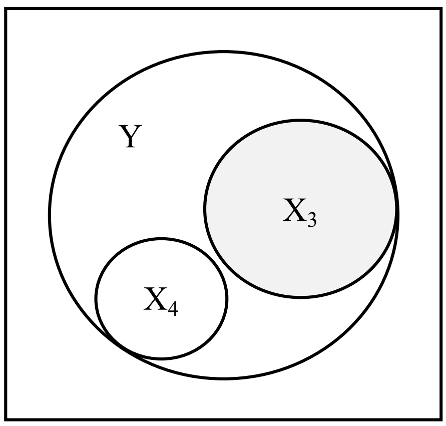
For necessary conditions, coverage helps to distinguish *relevant* from *trivial* necessary conditions. For example, a condition may formally fulfill the criteria for set-theoretic necessity (as when X is a perfect *superset* of the outcome Y), but this might still be a trivial finding because the condition is almost always present and thus it is difficult to establish a causal link between the condition and the outcome. For example, *oxygen* is nearly ubiquitous. Formally speaking, it may thus be a necessary condition for a range of phenomena, including the outbreak of civil war. However, stating "oxygen is a necessary condition for civil war" would have no analytical value because of the omnipresence of oxygen.

Figure 6.4 illustrates these relationships. The left-hand Euler diagram shows an outcome Y and the conditions X_1 and X_2 that are perfect supersets of Y. Thus, formally speaking, both of these can be considered necessary conditions for the outcome. Yet, we see that X_2 occurs more often than X_1 (the box represents our universe of phenomena in this simple setting). Hence, in the absence of any other information, X_2 can be considered less relevant than X_1.

Now consider the right-hand Euler diagram. Here, we have an outcome Y and two perfect subsets: X_3 and X_4. This means that whenever a case holds membership in either of these conditions, it also holds membership in the outcome. Thus, both X_3 and X_4 are sufficient conditions for Y. But we also see that X_4 is smaller than X_3, which means that there are fewer cases with membership in it. If our aim was to provide an account of Y, then X_3 would be the more relevant condition, simply because it covers more instances of Y. However, in an applied setting, both might be part of a solution term that entails several pathways to the outcome. For each pathway, we can calculate its *raw coverage* and *unique coverage*. The former is the total coverage of this pathway, irrespective of any empirical overlap with other pathways. The latter is limited to the unique contribution of the individual pathway (including only those cases that are not also covered by other pathways). We will return to this point when we examine solution terms in Chapter 7.

Figure 6.4 Coverage in Necessary and Sufficient Conditions (Crisp Sets)





How are these set relations calculated? Mirroring the computation of consistency, the *coverage* of necessary and sufficient conditions is calculated with the following two formulae, which differ in their denominator (Ragin 2006a, 63; 2008):

$$\text{Coverage}_{\text{Necessity}}(Y_i \leq X_i) = \frac{\sum \min(X_i, Y_i)}{\sum X_i}$$

$$\text{Coverage}_{\text{Sufficiency}}(X_i \leq Y_i) = \frac{\sum \min(X_i, Y_i)}{\sum Y_i}$$

It is apparent from the formulae for consistency and coverage that these are inversely related. Yet, what must be noted is that consistency is the *primary* measure of fit and as such it should always be calculated and examined first. Coverage is only meaningful when a consistent set-theoretic relationship has been identified in the analysis. In that sense, one should always proceed stepwise, starting with consistency. As Ragin (2008, 55) says, it is "pointless" to examine and interpret coverage for a condition which is not a consistent subset or superset of the outcome. But once consistency has been established, then the calculation of coverage can help to assess the empirical relevance of the identified set-theoretic relationship.

Calculating Coverage: An Example

Again, let us illustrate the calculation of coverage using the data from the coronavirus example, as listed in Table 6.1. In the previous stage, we identified *weak health* (X_1) as a perfect necessary condition for the outcome *serious symptoms*. Let us now calculate its coverage. For this we take the sum of the minimum values (for each office worker) across the condition and the outcome and divide it by the sum of the values for the condition. This means we divide 4.4 by 5.6, which yields 0.79. This coverage value suggests a fairly close fit between the condition and the outcome. We will return to this example when calculating the relevance of necessity towards the end of the chapter.

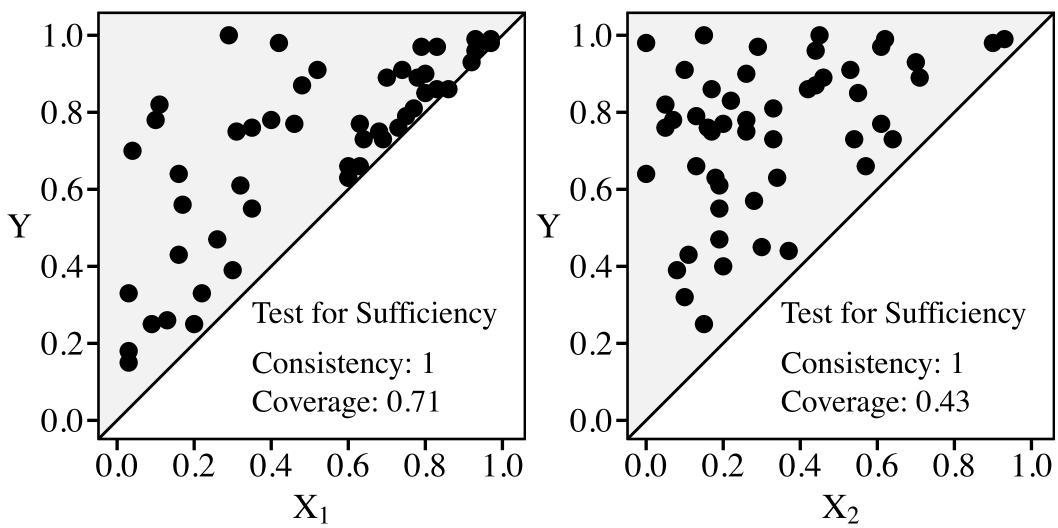
We also identified *old age* (X_2) as a sufficient condition for the outcome. To calculate its coverage, we divide the sum of the minimum values across the condition and the outcome (3.2) by the sum of the values for the outcome (4.4), which equals 0.73. This is a good coverage score. Yet, it also indicates that part of the outcome serious symptoms cannot solely be accounted for with the condition old age. For instance, the office workers Jamar and Carlos are both rather inside the outcome serious symptoms, but rather outside the set old age, suggesting that their weak health accounts for the outcome. Based on Table 6.1, we calculate coverage as follows:

$$\text{Coverage}_{\text{Necessity}}(X_1) = \frac{4.4}{5.6} = 0.79 \qquad \text{Coverage}_{\text{Sufficiency}}(X_2) = \frac{3.2}{4.4} = 0.73$$

The last step in the calculation means that the sufficient condition old age covers 0.73 of the *fuzzy-set membership values* of the outcome serious symptoms. This is an important distinction: on first observation, one may think that coverage reflects the *share* or *percentage* of the covered cases. But with fuzzy sets the calculation is based on cases that hold various degrees of set membership. Moreover, cases with less than 0.50 set membership are also included in the calculation. We can check this with a look at the data in Table 6.2. As we can see, there are five persons with membership above 0.50 in the outcome. Three of these are accounted for by the condition old age X_2. If it were just about the share of cases, the coverage of X_2 should be 0.60. However, as we calculated above, its actual coverage is 0.73.

While consistency is a straightforward metric, coverage can be more difficult to grasp. Essentially, coverage is about the fit between a condition and the outcome. The larger the gap between the values for the outcome and those for the respective condition, the lower the coverage score. This is illustrated in the two sufficient conditions shown in Figure 6.5. The left-hand condition X_1 has many cases that are clustered close to the main diagonal, indicating equal membership in the condition and the outcome. The condition X_1 has a *coverage* of 0.71 (and perfect consistency). For condition X_2 the relationship is different. Many cases are further removed from the main diagonal, which means that the values for the outcome far exceed those for the condition. Like the other condition, X_2 is also a perfectly consistent sufficient condition, but its *coverage* is only 0.43.

Figure 6.5 Sufficient Conditions with Varying Coverage (Fuzzy Sets)



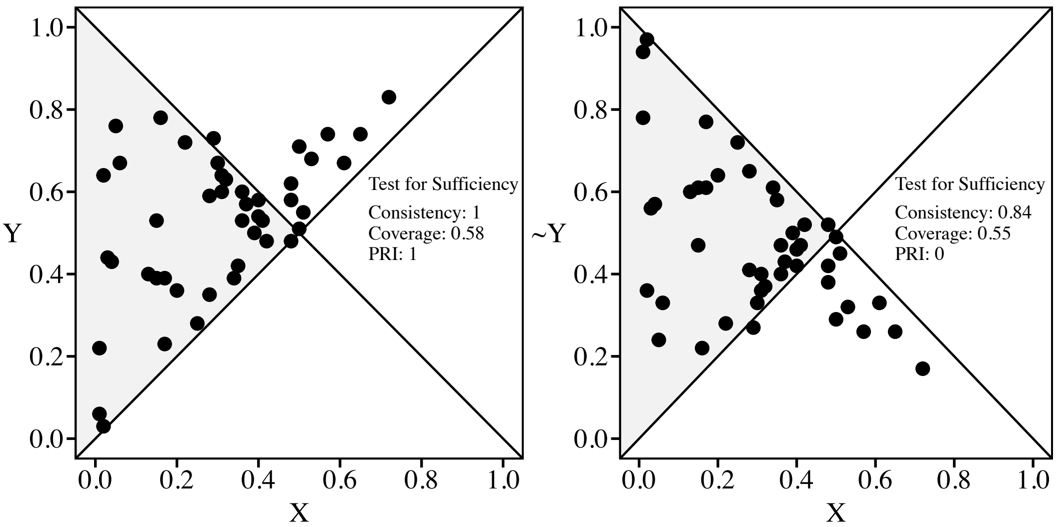
Proportional Reduction in Inconsistency

Proportional reduction in inconsistency (PRI) is a measure to identify *simultaneous subset relations* in the analysis of sufficient conditions. What does this mean? Simultaneous subset relations may happen with fuzzy-set data when a condition or combination of conditions is both a subset of the outcome and a subset of the non-outcome. This would be a *logical contradiction.* However, based solely on the measures of consistency and coverage it would be difficult to determine whether X should be treated as sufficient for the outcome or the non-outcome. To identify such situations and provide guidance for the correct interpretation of the set-theoretic relationship, the PRI measure was introduced into the fs/QCA software by Ragin (2006b) and first described in the textbook by Carsten Schneider and Claudius Wagemann (2012, 237-44).[4] The PRI measure has become a standard feature in most QCA software, including the "QCA" package for R developed by Adrian Duşa (2019, 134-36).

To illustrate simultaneous subset relations, let us look at an example with hypothetical data on 45 cases. Figure 6.6 shows XY plots for the condition X, the outcome Y, and the non-outcome ~Y. In the left plot, we can see that X is a sufficient condition for Y with a consistency of 1 and a coverage of 0.58. Yet, as shown in the right plot, X also appears to be a sufficient condition for the *non-outcome*, with a consistency of 0.84 and a coverage of 0.55.

How can this be? When we look at the empirical distribution of cases for the left plot, two observations stand out. First, apart from a few exceptions, most cases hold low scores in X with membership scores of less than 0.5, located on the left side of the cross-over. Second, most cases are placed inside the left triangle that is shaded in gray. The additional diagonal line divides the plot into four triangles. All cases inside the left triangle fulfill the criteria for being subsets of the outcome *and* the non-outcome, formally defined as X ≤ Y and X ≤ ~Y (Ragin 2011). This explains why consistency is high in both settings and thus formally passes the test for sufficiency. However, examining the data in these plots should raise doubts about treating X as a sufficient condition for ~Y because there are no cases with values above 0.5 in X that also show values above 0.5 in the non-outcome. This contrasts with the left plot, where we at least have a handful of cases that are rather inside both the condition and the outcome (> 0.5).

Figure 6.6 Simultaneous Subset Relations



Before coming to the calculation of PRI as a solution to the above problem, let us look at a simple example that involves just a few cases. Table 6.3 shows five hypothetical countries and their fuzzy-set membership values in the condition A, the outcome Y, and the non-outcome ~Y. On the right side, the table provides three metrics needed to calculate consistency, coverage, and PRI.

Table 6.3 Simultaneous Subset Relations

Case	A	Y	~Y	min(A, Y)	min(A, ~Y)	min(A, Y, ~Y)
France	0.2	0.3	0.7	0.2	0.2	0.2
Greece	0.5	0.6	0.4	0.5	0.4	0.4
Italy	0.4	0.5	0.5	0.4	0.4	0.4
Portugal	0.6	0.7	0.3	0.6	0.3	0.3
Spain	0.3	0.4	0.6	0.3	0.3	0.3
Sum	2	2.5	2.5	2	1.6	1.6

Let us now calculate the consistency for A as a *sufficient condition* for the *outcome* Y and the *non-outcome* ~Y. We use the standard consistency formula introduced earlier in this chapter.

$$\text{Consistency}_{\text{Sufficiency}}(A \leq Y) = \frac{2}{2} = 1 \qquad \text{Consistency}_{\text{Sufficiency}}(A \leq {\sim}Y) = \frac{1.6}{2} = 0.8$$

Taking the sum values from the bottom line of Table 6.3, we divide the sum of the minimum values across A and Y (2) by the sum of the values for A (2), which equals 1. This means that A is a perfect sufficient condition for Y. What about the *non-outcome*? Here, we divide the sum of the minimum values across the condition and the non-outcome (1.6) by the sum of the values for A (2), which yields 0.8. Because the consistency value for A and ~Y is above the 0.75 threshold, we might also treat it as a sufficient condition for the non-outcome, especially if this score referred to a truth table row. Yet, doing so would mean that the condition equally leads to the outcome and its negation, which would be a logical contradiction. How to resolve this paradox? The measures of fit consistency and coverage do not help us in this scenario, but PRI detects the problem. The formula for the proportional reduction in inconsistency is (Ragin 2011; Schneider and Wagemann 2012, 242):

$$\text{PRI} = \frac{\sum \min(X_i, Y_i) - \sum \min(X_i, Y_i, {\sim}Y_i)}{\sum X_i - \sum \min(X_i, Y_i, {\sim}Y_i)}$$

We apply this formula to assess the relationship between A and Y as well as A and ~Y:

$$\text{PRI}_{(A,\,Y)} = \frac{(2\ -\ 1.6)}{(2\ - 1.6)} = \frac{0.4}{0.4} = 1 \qquad \text{PRI}_{(A,\,{\sim}Y)} = \frac{(1.6\ -\ 1.6)}{(2\ - 1.6)} = \frac{0}{0.4} = 0$$

The PRI value for A as a sufficient condition for Y is 1, whereas the PRI for A and ~Y is 0. While consistency and coverage are relatively similar for both outcome and non-outcome (see the summary in Table 6.4), the PRI scores give a clear indication that we should treat A only as a valid sufficient condition for Y and not as a sufficient condition for ~Y.

As a general rule, with statements of sufficiency one should always observe whether there is a substantial difference between PRI and consistency. When that is the case, one should examine the non-outcome for simultaneous subset relations. The problem of simultaneous subset relations *may* occur when there are too many cases with low values in a condition, especially if most of these are located in the left triangle depicted in Figure 6.6. Hence, as a precaution, it is always useful to create *histograms* of individual conditions and to plot these against the outcome before the actual analysis, to examine whether any of these may later cause analytical problems. Apart from theoretical considerations, this is another reason why one should strive to have a *reasonably even distribution* of membership values for each condition.

Table 6.4 Simultaneous Subset Relations and PRI

	A → Y	A → ~Y
Consistency	1	0.80
Coverage	0.80	0.64
PRI	1	0

Relevance of Necessity

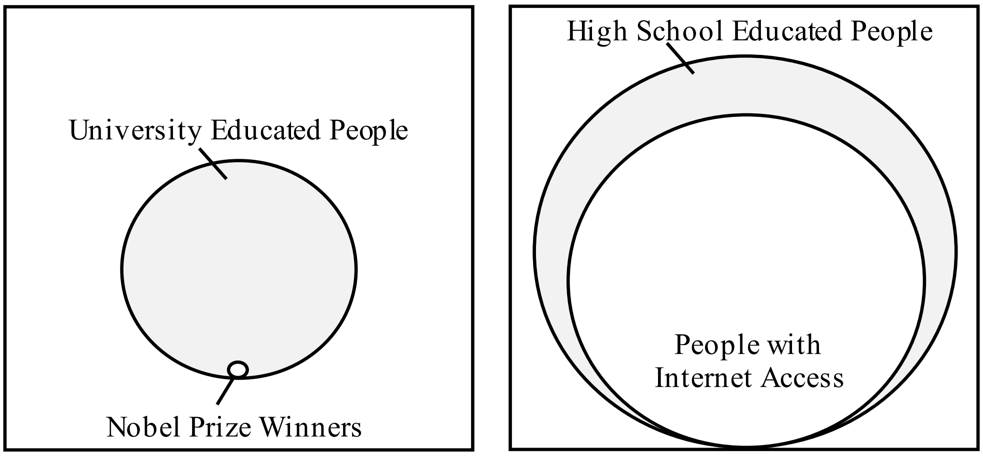
As introduced above, the standard coverage measure helps to distinguish trivial from relevant necessary conditions. The example used earlier was *oxygen* being a trivial necessary condition for the outbreak of *civil war*. In this example, the large difference between the "size" of the condition (or rather the frequency with which it appears) and the outcome will yield low coverage, indicating a trivial necessary condition. Put differently, in such a situation the data entails many instances where the condition is present, but very few or almost no cases where the outcome is also present. Formally, this may still indicate perfect set-theoretic consistency, but coverage would be low. This can happen especially when the condition is a widespread phenomenon and the outcome of interest is a rare event.

To take another example, consider the relationship between obtaining a university education and winning a Nobel prize. Albeit a few exceptions, a university education may be formally necessary for winning a Nobel prize, but this is a trivial finding as the set of people who won such a prize is tiny when placed in relation to those who completed a university education. This is illustrated in the left-hand Euler diagram in Figure 6.7. In fact, the depiction of Nobel prize winners should be even smaller than shown here. The illustration underlines that there is little to be gained from knowing that the condition university education is formally necessary for becoming a Nobel prize winner because there is such a large gap between the sizes of the respective sets. Hence this relationship points at a *trivial* necessary condition.

Apart from this common scenario, there can be circumstances where potentially trivial necessary conditions in the empirical data would *not* be detected with the existing formula for the calculation of coverage. Therefore, several alternative measures of necessary conditions' relevance and trivialness have been suggested (Braumoeller and Goertz 2000; Goertz 2006; Schneider and Wagemann 2012).

The source for such a form of trivialness is what can be termed a *constant* condition, as when a condition shows little variation across the observed cases. To stay with the education example, suppose we are interested in explaining the internet penetration rate, looking at the outcome *people with internet access*. In the US, this would be a large set of about 86% of the population. Now one of our conditions may be the set of *high school educated people*, which would be an even larger set that comprises about 90% of the US population. If we placed these sets in relation (and assume for illustration's sake that there is a perfect overlap between them), then it would turn out that being in the set of high school educated people is a necessary condition for having internet access. This is shown in the right-hand Euler diagram in Figure 6.7. The set-theoretic analysis would yield perfect consistency and high coverage scores. What is wrong with this? The problem in this scenario is that the condition comes close to being a constant. Hence, we could explain almost *any* outcome on the basis of high school education, even when there is no relationship. If our outcome is similarly omnipresent, then the standard coverage metric would not alert us to this issue in our data.

Figure 6.7 Two Types of Trivial Necessary Conditions



To identify this second type of trivial necessary conditions, Schneider and Wagemann (2012) put forth the "relevance of necessity" (RoN) measure, based on an earlier metric suggested by Gary Goertz (2006). This has become a standard indicator in testing for the relevance of necessary conditions (Oana et al. 2021). By conception, the RoN measure can yield values between 0 and 1, where lower scores indicate *trivialness*, and higher scores denote *relevance*. The more a condition X resembles a constant, the closer to 0 the RoN metric will be. The formula for the calculation of RoN is (Schneider and Wagemann 2012, 236):

$$\text{Relevance of Necessity} = \frac{\sum(1 - X_i)}{\sum(1 - \min(X_i, Y_i))}$$

How does this look like in practice? To illustrate the use of the RoN measure, let us take a simple example involving the condition X, the outcome Y, and hypothetical data on just four cases. Table 6.5 shows that, formally speaking, X is an almost perfect *superset* of Y because the values for X are almost always larger than or equal to the values for Y. On this basis, the condition X might be considered a necessary condition for Y.

The right-hand side of Table 6.5 further shows the results for the calculation of the *consistency* and *coverage* for necessary conditions, and the *relevance* of necessity. As expected, we can see that at 0.95 the set-theoretic consistency is very high, satisfying the formal threshold for necessary conditions (equal to or above 0.90). The coverage is lower, but at 0.59 it would not immediately prompt concern. However, we can see that the relevance of necessity indicator is closer to 0 than to 1, suggesting a potentially *trivial* necessary condition.

Table 6.5 Relevance of Necessity

	X	Y	Test for Necessity: X ⟵ Y	
Case 1	1.0	0.9	Consistency	0.95
Case 2	0.9	0.4	Coverage	0.59
Case 3	1.0	0.3	Relevance of Necessity	0.38
Case 4	0.3	0.4		

Why is that? This being a hypothetical example, we have no substantive knowledge of the underlying data. But what we can see is that X shows little variation, with three out of four cases at values equal to or close to 1. With data patterns like this, the consistency measure would always satisfy the criterion for a necessary condition. However, the RoN measure suggests that we should be *cautious* before treating it as a relevant necessary condition for the outcome. Ultimately, dealing with data patterns like this is a matter of interpretation. There can be situations where a condition is almost a constant, but still the condition may have substantive importance and relevance as a necessary condition. However, such an interpretation would have to be justified explicitly. As a rule of thumb, any *potential* necessary condition that meets the consistency benchmark of 0.9 should be checked for its coverage and relevance. If these metrics fall below 0.5, then this suggests that we may be dealing with a *trivial necessary condition*. In order to make an informed judgment on this, we should always examine the empirical distribution of our cases and their set-theoretic membership scores. A good way to do this are histograms and XY plots of the raw and calibrated data (on this, see also Chapter 10).

Notes

[1] Ragin (2008, 45).

[2] On the methodology of necessary conditions see Braumoeller and Goertz (2000), Goertz (2003), and Goertz and Starr (2003). More recent contributions include Thiem (2016), Vis and Dul (2018), and Dul et al. (2020).

[3] As noted, sufficiency is typically analyzed through the truth table procedure (see Chapter 7).

[4] On the PRI measure, see also Dușa (2019, 134-36).

Mello, Patrick A. (2021) *Qualitative Comparative Analysis: An Introduction to Research Design and Application,* Washington, DC: Georgetown University Press, Chapter 6.

References

Bol, Damien, and Francesca Luppi. 2013. "Confronting Theories Based on Necessary Relations: Making the best of QCA Possibilities." *Political Research Quarterly* 66 (1): 205-10.

Braumoeller, Bear F., and Gary Goertz. 2000. "The Methodology of Necessary Conditions." *American Journal of Political Science* 44 (4): 844–58.

Damonte, Alessia. 2018. "Gauging the Import and Essentiality of Single Conditions in Standard Configurational Solutions." *Sociological Methods & Research* online first, September 11 (doi: 10.1177/0049124118794678).

Dul, Jan, Erwin Van der Laan, and Roelof Kuik. 2020. "A Statistical Significance Test for Necessary Condition Analysis." *Organizational Research Methods* 32 (2): 385-95.

Duşa, Adrian. 2019. *QCA with R. A Comprehensive Resource*. Cham: Springer.

Eliason, Scott R., and Robin Stryker. 2009. "Goodness-of-Fit Tests and Descriptive Measures in Fuzzy-Set Analysis." *Sociological Methods & Research* 38 (1): 102–46.

Goertz, Gary. 2003. "The Substantive Importance of Necessary Condition Hypotheses." In *Necessary Conditions*, edited by Goertz, Gary and Harvey Starr. Lanham: Rowman & Littlefield, 65–94.

———. 2006. "Assessing the Trivialness, Relevance, and Relative Importance of Necessary or Sufficient Conditions in Social Science." *Studies in Comparative and International Development* 41 (2): 88–109.

Goertz, Gary, and Harvey Starr, ed. 2003. *Necessary Conditions: Theory, Methodology, and Applications.* Lanham: Rowman & Littlefield.

Haesebrouck, Tim. 2015. "Pitfalls in QCA's Consistency Measure." *Journal of Comparative Politics* (2): 65-80.

Mello, Patrick A. 2017. "Qualitative Comparative Analysis and the Study of Non-State Actors." In *Researching Non-State Actors in International Security: Theory & Practice*, edited by Kruck, Andreas and Andrea Schneiker. Oxon: Routledge, 123-42.

———. 2020. "Paths towards Coalition Defection: Democracies and Withdrawal from the Iraq War." *European Journal of International Security* 5 (1): 45-76.

Oana, Ioana-Elena, Carsten Q. Schneider, and Eva Thomann. 2021. *Qualitative Comparative Analysis Using R: A Beginner's Guide*. New York: Cambridge University Press.

Ragin, Charles C. 2000. *Fuzzy-Set Social Science*. Chicago, IL: University of Chicago Press.

———. 2006a. "Set Relations in Social Research: Evaluating Their Consistency and Coverage." *Political Analysis* 14 (3): 291–310.

———. 2006b. *User's Guide to Fuzzy-Set / Qualitative Comparative Analysis*. Irvine, CA: University of California.

———. 2008. *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago, IL: University of Chicago Press.

———. 2011. "PRI Consistency ("PRE" in fsQCA's Truth Table Spreadsheet)." Syracuse, NY: Institute for Qualitative and Multi-Method Research.

Ragin, Charles C., Susan E. Mayer, and Kriss A. Drass. 1984. "Assessing Discrimination: A Boolean Approach." *American Sociological Review* 49 (2): 221-34.

Rihoux, Benoît, and Gisèle De Meur. 2009. "Crisp-Set Qualitative Comparative Analysis (csQCA)." In *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques*, edited by Rihoux, Benoît and Charles C. Ragin. Thousand Oaks: Sage, 33-68.

Rihoux, Benoît, and Charles C. Ragin, ed. 2009. *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques.* Thousand Oaks, CA: Sage.

Schneider, Carsten Q., and Claudius Wagemann. 2010. "Standards of Good Practice in Qualitative Comparative Analysis (QCA) and Fuzzy-Sets." *Comparative Sociology* 9 (3): 397–418.

———. 2012. *Set-Theoretic Methods for the Social Sciences: A Guide to Qualitative Comparative Analysis*. New York, NY: Cambridge University Press.

Stoklasa, Jan, Pasi Luukka, and Tomáš Talášek. 2017. "Set-Theoretic Methodology Using Fuzzy Sets in Rule Extraction and Validation - Consistency and Coverage Revisited." *Information Sciences* 412-413: 154-73.

Thiem, Alrik. 2016. "Standards of Good Practice and the Methodology of Necessary Conditions in Qualitative Comparative Analysis." *Political Analysis* 24 (4): 478-84.

Veri, Francesco. 2018. "Coverage in Fuzzy Set Qualitative Comparative Analysis (fsQCA): A New Fuzzy Proposition for Describing Empirical Relevance." *Comparative Sociology* 17 (2): 133-58.

———. 2019. "Aggregation Bias and Ambivalent Cases: A New Parameter of Consistency to Understand the Significance of Set-theoretic Sufficiency in fsQCA." *Comparative Sociology* 18 (2): 229-55.

Mello, Patrick A. (2021) *Qualitative Comparative Analysis: An Introduction to Research Design and Application,* Washington, DC: Georgetown University Press, Chapter 6.

Vis, Barbara, and Jan Dul. 2018. "Analyzing Relationships of Necessity Not Just in Kind But Also in Degree: Complementing fsQCA With NCA." *Sociological Methods & Research* 47 (4): 872-99.