Improving Research On Necessary Conditions: Formalized Case Selection for Process Tracing after QCA
Author(s): Ingo Rohlfing and Carsten Q. Schneider
Published by: Sage Publications, Inc. on behalf of the University of Utah
Stable URL: https://www.jstor.org/stable/23563606

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at https://about.jstor.org/terms
containing G that is sufficient for the outcome. It follows that in this example the existence of untenable assumptions is not an artifact of formal-logically untenable directional expectations. Also, strictly formally speaking, the fact of G being necessary for S does not indicate anything about the sufficiency of G or ~G for S, due to the asymmetric nature of logical statements. Of course, had we performed an analysis of necessity before that of sufficiency, we might eventually not have formulated the directional assumption ~G → S, which runs counter to the statement G ← S. Note, however, that incoherent easy counterfactuals can occur even if no directional expectations about the necessary condition (or its complement) are made, as demonstrated by Schneider and Wagemann (2012, chap. 9).

12. Assumptions can be contradictory without being simplifying.

13. Because from our previous analysis we know that L ← S and G ← S, we must specify the directional expectations ~L → ~S and ~G → ~S for the analysis of ~S. When analyzing S, one could formulate the directional expectations L → S and G → S (Ragin 2009b, 114), but, in our view, they do not have to be made. Note that, in the current example, the intermediate solution remains the same if we added the directional expectations L → S and G → S.

14. The consistency threshold (0.8) and the directional expectations (~D → S; ~U → S; ~I → S) are the same as above in the section on Contradictory Simplifying Assumptions.

15. The intermediate solution for S did rest on contradictory simplifying assumptions, though. These assumptions can be made as long as they are not also made in the analysis of outcome ~S or vice versa. We decide to bar contradictory simplifying assumptions from the analysis of ~S (below) and keep them for the analysis of outcome S.

16. The points we make are unaffected by the fact that Koenig-Archibugi uses fuzzy sets.

17. We do not report the parameters of fit as they are tangential to the argument we make here.

18. The above-described strategy of selecting entire logical remainder rows for counterfactual claims can be seen as an extreme form of formulating conjunctural directional expectations (CDEs).

19. More specifically, if CDE are a subset of directional expectations on single conditions (DE), then solutions obtained with CDE are a subset of (and more complex than) those derived with DE. We suggest, however, that CDE should not by default be a subset of DE, that is, they should not be restricted to being simplifying assumptions.

Improving Research On Necessary Conditions: Formalized Case Selection for Process Tracing after QCA

Ingo Rohlfing1 and Carsten Q. Schneider2

DOI: 10.1177/1065912912468269

Abstract

This paper aims at strengthening causal inference in necessary condition research. We demonstrate how process tracing based on purposefully selected cases can complement findings on cross-case patterns identified with Qualitative Comparative Analysis (QCA). Using an empirical example, we discuss the meaning of typical and deviant cases in analyses of necessity, develop formulas for identifying the most typical and most deviant cases, and detail the implications of so-called SUIN conditions for meaningful case selection. In addition, we clarify various viable variants of comparative process tracing and formulas for identifying the best-matching pairs of cases.

Introduction

Necessary conditions play a more prominent role in social science theory than is often acknowledged. Moreover, this type of condition has unique causal qualities: if it is absent, the outcome cannot occur but, conversely, when it is present, the outcome does not always occur (Harvey and Starr 1989: chap. 3). In recognition of their substantive relevance and vexing properties, the methodological

1University of Cologne, Germany
2Central European University, Hungary
processes operating within cases are largely neglected. This process tracing. We explain which forms of comparisons technique for set-relational research (see also Bol and Luppi in explain the challenges for case selection and process trac
valuable contributions to the improvement of Qualitative discussion. They are a manifestation of causal heteroge
cient but necessary for an outcome. omitted conditions.

whether a given distribution of cases that is not fully in with an fsQCA example briefly introduced in Section 2.
section concludes.

cessary condition research via QCA by developing prin- tute counterfactual reasoning, which is unavoidable in

Comparative Analysis (QCA) as the most formalized tech- neity in QCA (see Ragin 1987, chap. 2; 2000, 40). We

Mahoney, Kimball, and Koivu's (2009) notion of SUIN necessary condition (conditions with low coverage val

The meaning of typical and deviant cases in QCA-based research on necessary conditions is illustrated with an fsQCA example briefly introduced in Section 2. In Section 3, we discuss the most basic setting of single-case studies on the basis of a QCA solution that includes one necessary condition. This simple example allows us to highlight three important and general insights. First, we show what typical cases are in fsQCA of necessity and how to select the most typical case. Second, we elaborate on the role of inconsistency, that is, a less-than-perfect cross-case pattern of necessity, and relate it to deviant cases. This involves a discussion of how to select the most deviant case and how to employ knowledge about the potential source of deviance in process tracing. Third, we show that for trivial necessary condition (conditions with low coverage values), process tracing cannot shed light on how to enhance the condition's relevance by searching for omitted conditions.

In Section 4, the notion of SUIN causes is added to the discussion. They are a manifestation of causal heterogeneity in QCA (see Ragin 1987, chap. 2; 2000, 40). We explain the challenges for case selection and process tracing that arise with SUIN causes and provide suggestions on how to master them in practice. In Section 5, the discussion is extended to case selection for comparative process tracing. We explain which forms of comparisons are (not) viable and propose guidelines and formulas for identifying the best-matching pair of cases for each form of comparison. One central insight is that, in principle, with comparative process tracing, researchers can substitute counterfactual reasoning, which is unavoidable in single-case studies, with empirical evidence. The final section concludes.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Consistency</th>
<th>Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>STOCK</td>
<td>0.89</td>
<td>0.72</td>
</tr>
<tr>
<td>UNI</td>
<td>0.81</td>
<td>0.67</td>
</tr>
<tr>
<td>MA</td>
<td>0.72</td>
<td>0.72</td>
</tr>
<tr>
<td>~OCCUP</td>
<td>0.71</td>
<td>0.75</td>
</tr>
<tr>
<td>BARGAIN</td>
<td>0.68</td>
<td>0.57</td>
</tr>
<tr>
<td>~EMP</td>
<td>0.64</td>
<td>0.74</td>
</tr>
<tr>
<td>OCCUP</td>
<td>0.58</td>
<td>0.66</td>
</tr>
<tr>
<td>EMP</td>
<td>0.58</td>
<td>0.61</td>
</tr>
<tr>
<td>~BARGAIN</td>
<td>0.50</td>
<td>0.78</td>
</tr>
<tr>
<td>~MA</td>
<td>0.50</td>
<td>0.60</td>
</tr>
<tr>
<td>~UNI</td>
<td>0.31</td>
<td>0.50</td>
</tr>
<tr>
<td>~STOCK</td>
<td>0.24</td>
<td>0.42</td>
</tr>
<tr>
<td>MA + STOCK</td>
<td>0.92</td>
<td>0.68</td>
</tr>
<tr>
<td>MA + UNI</td>
<td>0.91</td>
<td>0.66</td>
</tr>
</tbody>
</table>

Own calculation, based on data from Schneider et al. (2010).
The Empirical Example

Schneider, Schulze-Bentrop, and Paunescu (2010) aim at identifying necessary conditions for high export performance of the high-tech sector in nineteen OECD countries from 1990 to 2003 (N = 76). The outcome is membership in the set of countries with a high share of high-tech sector exports relative to overall exports (EXPORT). Schneider et al. focus on six conditions: high unemployment protection (EMP), high coverage of collective bargaining (BARGAIN), high share of university-trained citizens (UNI), high share of occupation-trained citizens (OCCUP), high share of stock market capitalized indigenous firms (STOCK), and high share of cross-border mergers and acquisitions (MA).

The analysis of necessity reveals that neither condition nor its complement passes the conventional consistency threshold of 0.9 (Ragin 2000b) for being necessary for EXPORT (see Table 1).

Going beyond the necessity of single conditions, Schneider et al. provide plausible arguments as to why condition MA could act as a functionally equivalent condition to conditions UNI and STOCK. All three conditions provide different ways of acquiring a broad knowledge base within a company (Schneider et al. 2010, 250-251), which we denote by K. Schneider et al. take K as a higher-order condition that is hypothesized to be necessary to being a successful high-tech country. The outcome EXPORT is thus the result of causally heterogeneous necessary conditions (Ragin 2000, 321-323). This means that members of the outcome EXPORT are also members of either MA, and/or UNI, and/or STOCK. Conceptually seen, it follows that MA, UNI, and STOCK are all SUIN conditions (Mahoney et al. 2009, 126). Each condition is an unnecessary but sufficient attribute of K, which, in turn, is necessary but not sufficient for outcome EXPORT. Indeed, the expressions MA + UNI (0.91) and MA + STOCK (0.92) both pass the conventional consistency threshold. Moreover, they are empirically nontrivial, as indicated by their coverage values of 0.68 and 0.66, respectively (Schneider et al. 2010, 255). In the following sections, we build on these results to illustrate systematic case selection for process tracing in set-theoretic multi-method research. Terminologically, we generally speak of a term or terms that are necessary for the outcome because a single condition can be necessary in the same way as multiple conditions joined by the logical OR operator.

Types of Cases and Single-Case Selection

Once necessary terms have been identified with the help of fsQCA, the standard XY plot is useful to discern the consistency of the statement of necessity. The distinction between consistent (below diagonal) and inconsistent (above diagonal) cases is valuable and makes use of the nuanced measurement of fuzzy-set memberships (Ragin 2000, chap. 6). It is, however, important to remember that in fsQCA, differences in kind trump differences in degree (Ragin 2000, 157-159). Our discussion of process tracing is therefore based on what we call an enhanced XY plot (Schneider & Rohlfing forthcoming). The enhanced XY plot blends the conventional XY plot with a 2 x 2 matrix as it is known from crisp-set QCA (cf. Braumoeller & Goertz 2000), thereby establishing qualitative differences via the 0.5 anchor in X and Y, respectively (Ragin 2000, 8). This allows for a straightforward visualization of the differences in degree and in kind between cases.

Figure 1 shows that the enhanced XY plot consists of six zones (or areas or cells). In the following, we elaborate on the different analytic information contained in cases from different zones. For presentational purposes, we offer the enhanced XY plot for condition MA² in Figure 1.

Typical Cases

In QCA, typical cases are those that conform to the set-relational pattern of interest. For fsQCA, this means that typical cases are located in Cell 2 of the enhanced XY plot. They are below the secondary diagonal and thus consistent with the set-theoretic statement of necessity. Furthermore, from a qualitative perspective, they are more in than out of both the condition and the outcome. By using their fuzzy-set membership in X and Y, further differentiations can be made among the qualitatively identical cases in Zone 2.

We argue that the ideal typical case for discerning the mechanisms behind a pattern of necessity has full membership in the necessary term and the outcome. For applied research, it follows that the most typical case is the one closest to the upper-right corner of the XY plot. In this context, it is important to note that the most typical case is determined with regard to its proximity to the ideal typical case and not with regard to how cases are distributed within Cell 2. The reason for this selection rule is that all cases in the same cell are qualitatively identical and those in Cell 2 are therefore all typical cases. How much they come close to the ideal typical case is expressed by their distance to the upper-right corner.

In applied QCA, it might be difficult to determine the most typical case by visual inspection alone. We therefore propose a formula that draws on each case’s fuzzy-set memberships to distinguish more typical from less typical cases. The typicalness of a case i for necessity, denoted as \( N_{Ti} \), is calculated by

\[
N_{Ti} = \frac{X_i - Y_i}{Y_i}
\]

For all cases located in Zone 2, \( X_i \) and \( Y_i \) denote the set-membership score of a given case in the condition and the outcome. The numerator expresses a case’s proximity to the secondary diagonal. Cases with larger membership
in $Y$ are more valuable for process tracing because analyses of necessity are $Y$-oriented (Goertz & Starr 2003). This reasoning is captured by the denominator, which standardizes the distance to the secondary diagonal by the membership in $Y$. Hence, the most typical case is the one with the smallest value for all $N_i$. The application of formula $N$ to condition $MA$ identifies the United Kingdom in 1995 as the most typical case (formula score of 0.01 with $MA = 0.88$ and $EXPORT = 0.87$).

In making a claim of necessity, we effectively argue that in the absence of the condition, the outcome would also be absent. Hence, the focus of process tracing should be on within-case evidence that lends credence to this claim. In the analysis of single typical cases, this can only be a counterfactual because we only select cases with both $X$ and $Y$ present. Of course, this is not a simple cross-case counterfactual because we already know that the cross-case evidence supports the claim of necessity. Instead, we should collect process-tracing evidence lending support to the claim that the pattern of necessity is, indeed, causal. In process tracing on the condition $MA$, for instance, the supporting evidence could be interview statements from managers from successful companies stating that “We wouldn’t be successful if we were not for realizing mergers and acquisitions because this ensures a broad knowledge for our company.” Although we still lack process-tracing evidence on a case that is not a member of either $MA$ or $EXPORT$ (a case in Zone 4 of Figure 1), such statements would substantially increase confidence in the necessity of $MA$.

**Deviant Cases Consistency**

Any case located in Areas 1, 5, and 6 of the enhanced $XY$ plot is inconsistent with the statement that $X$ is necessary for $Y$. However, the only cases that should be chosen for single-case process tracing are those in Zone 6, which we label deviant cases consistency in kind. They are members of $Y$ and nonmembers of $X$, which is puzzling because we should not observe the outcome when the necessary condition is absent.

The ideal deviant case is located in the upper-left corner of Cell 6 and has full membership in $Y$ and full nonmembership in $X$. Thus, the most deviant case is the one closest to this corner. If the data at hand does not contain the ideal deviant case, one should calculate the degree of deviance for all cases $i$ in Zone 6, $N_{iY}$, with the following formula:

$$N_{iY} = 1 - \frac{Y_i - X_i}{Y_i}$$

The numerator now includes the difference between the membership in $Y$ and $X$. The greater the difference, the more distant is the case from the secondary diagonal. This
is exactly what makes a case both puzzling and interesting. Again, we standardize the difference with the membership in the outcome because the higher the membership in Y, the more appropriate this case is for process tracing. To achieve a uniform interpretation of scores across our formulas, we subtract the ratio from 1. Thus, the most deviant case is the one with the smallest value for \( N_Y \).

Applying the formula to all deviant cases in Zone 6, Japan 2003 (\( MA = 0.06 \); \( EXPORT = 0.94 \)) qualifies as the most deviant case for the condition \( MA \) with a score of 0.06.

What are the potential reasons for deviance that are to be addressed in process tracing? It is useful to generally distinguish between model-related and model-unrelated reasons (Schneider & Rohlfing forthcoming). The latter consist of the misspecification of the population and the fallacious calibration of conditions (Ragin 1987, 2000). Model-related reasons, in turn, consist of the inclusion of too few (underspecification) or too many conditions (overspecification).

In research on necessity, deviant cases consistency can only be the result of underspecification, which, more specifically, means the omission of a SUIN condition. The rationale is that a deviant case consistency can be converted into a typical case by altering the solution such that the case becomes a member of a new solution. This, in turn, can only be achieved by adding a SUIN condition to the solution via the logical OR operator in which the deviant case holds a membership that exceeds its membership in Y. In conceptual terms and as previously mentioned, the expansion of a solution through logical OR requires that the omitted term be a functional equivalent for the original necessary term(s) (Ragin 2000, 241-244), that is, they must be SUIN conditions.

Applying this line of reasoning to our empirical example in which the original solution only comprises the condition \( MA \), suppose that process tracing in the most deviant case (Japan 2003) reveals that a high share of university-trained citizens (\( UNI \)) plays an important role in bringing about the outcome. Further imagine that a credible conceptual claim can be made that conditions \( MA \) and \( UNI \) are functional equivalents with regard to the more general construct “broad knowledge base within the company” \( K \) (see above).

Establishing the functional equivalence of \( MA \) and \( UNI \) is important, but this only satisfies the “SU” part of SUIN because \( MA \) and \( UNI \) are now taken as sufficient but unnecessary attributes of \( K \). The claim that \( K \) (the higher-order construct) is insufficient—the “I” in SUIN—has to be determined in a separate QCA that tests the sufficiency of \( K \). This yields an important insight: In the process of identifying a SUIN condition, researchers have to perform sufficiency tests on the cross-case level, which, in turn, means that SUIN conditions cannot be fully identified in deviant cases consistency alone.

Finally, to establish the necessity—the “N” in SUIN—of the hitherto omitted condition \( UNI \), both a cross-case and a within-case analysis are needed. On the cross-case level, one has to test for its necessity in those cases that are not members of \( MA \). We have cross-case evidence for the necessity of \( UNI + MA \) if \( UNI \) is always present when the outcome \( EXPORT \) is present and condition \( MA \) absent. The cross-case evidence needs to be supported with process-tracing evidence that lends credence to a counterfactual, which, in turn, substantiates the claim of necessity. If \( UNI \) is indeed necessary, the argument for typical cases above also applies here. The rationale is that adding \( UNI \) to \( MA \) via logical OR is meant to turn former deviant cases consistency into typical cases. Consequently, process-tracing evidence should permit the inference that the absence of \( UNI \) would lead to the absence of the outcome. If a researcher is able to make such an inference in the within-case analysis, the picture is completed, and it can be concluded that \( UNI \) is an omitted SUIN condition that has to be added to the solution.

A final but important note concerns the assumption that we make with regard to the similarity of deviant cases consistency. Pending evidence to the contrary, the most reasonable and parsimonious assumption is that all deviant cases consistency are deviant because of the omission of the same term. Consequently, if true, then adding this term to the solution will turn all deviant into typical cases. Our starting assumption might, of course, be wrong, but the procedure we propose here is sensitive to this possibility, for if there is indeed more than one omitted SUIN condition, some cases will remain deviant even after adding one omitted SUIN condition. Researchers are therefore advised to construct an enhanced \( XY \) plot after adding a new SUIN condition and to check whether there are still cases in Zone 6.

Noninsightful Types of Cases in Necessity Research

The enhanced \( XY \) plot divides the plot into six zones, but only typical cases (Zone 2) and deviant cases consistency (Zone 6) are relevant to single-case process tracing. In necessary condition research, we are interested in either the mechanisms linking the condition to the outcome (typical case), or in the reason that the outcome is present in the absence of a necessary condition (deviant case consistency). In light of this, process tracing in nonmembers of the outcome—that is, cases in Zones 3, 4, and 5—is futile as it cannot shed light on either research goal (Schneider & Rohlfing forthcoming) unless they are compared with cases that hold membership in Y (see below).

There are, however, important differences between cases from Zone 3 to 5 that are relevant with regard to comparative process tracing. Cases in Zone 3 and 5 lack any value for process tracing, that is, for single-case and comparative process tracing alike. While it is true that cases in Area 3 reduce the relevance of necessary conditions, they do not contradict the statement of necessity;
they simply do not represent any puzzle that can be addressed in necessary condition research. Cases in Cell 5, instead, are irrelevant in the QCA of necessity. Although they are not in line with the statement of necessity, they are not members of $X$ or $Y$.

In contrast, cases in Zone 4 can play a role in comparative process tracing. This is why we label them as individually irrelevant (IIR) cases. They are formally in accord with a pattern of necessity, but are substantively irrelevant for explaining how the presence of $Y$ requires the presence of $X$ because they are neither members of $X$ nor $Y$. Below, however, we explain that cases in Zone 4 are useful when compared with a typical case and with a deviant cases consistency.

Finally, we call cases in Zone 1 deviant cases consistency in degree. They are inconsistent with a statement of necessity, but, in contrast to cases in Zone 6, are members of $X$ and $Y$. Deviant cases consistency in kind are therefore more puzzling than deviant cases in degree, and the latter should only be considered for process tracing if the former are not available.

**Implications of SUIN**

**Conditions for Case Selection**

In the previous section, we discussed case selection and process tracing based on a QCA result consisting of just one single necessary condition. It might well be, however, that the original solution term already consists of two (or more) SUIN conditions. Indeed, in our empirical example, Schneider et al. (2010, 250-251) take the conditions $MA$ and $UNI$, and $MA$ and $STOCK$ as SUIN conditions and test the necessity of the expressions $MA + UNI$ and $MA + STOCK$, respectively. In the following, we focus on the expression $MA + UNI$.

The most important implication of SUIN conditions is that at least one typical case for each SUIN condition must be selected. If we want to fully understand the diverse ways in which the outcome requires the presence of a necessary term, it is evident that we must do process tracing for each SUIN condition separately. With regard to case selection, the specific requirement is that the selected typical cases are a member of only one SUIN condition. Cases that meet this requirement are called unique members. Typical cases that are members of more than one SUIN condition are joint members and the corresponding phenomenon is joint membership. Finally, cases that are not members of any SUIN condition are labeled nonmembers. Process tracing should be done in unique members; otherwise, causal inferences on necessity become unnecessarily protracted.

For illustration, consider the United Kingdom in 1995. It is a member of outcome $EXPORT$ and a unique member of $MA$. The counterfactual to be addressed in a single-case study is this: if condition $MA$ was absent, would the outcome be absent as well? For unique members of $MA$, the process-tracing evidence should lead us to answer "yes." If not, then $MA$ would be a spurious necessary condition. The point now is that if we select a joint member of $MA$ and $UNI$, for example, UK 1990, the answer to the counterfactual question (Would $EXPORT$ be absent if $MA$ was absent?) would be "no." This is because the presence of $UNI$ ensures the presence of $EXPORT$ for a case such as UK 1990. A single-case study on a joint member, therefore, would require a twofold counterfactual (Would $EXPORT$ be absent if $MA$ and $UNI$ were absent?), a situation that is best and easiest avoided by only choosing unique members for process tracing in typical cases.

**Case Selection for Comparative Process Tracing**

Single-case studies have their merits, but there are three arguments in favor of comparative process tracing. First, insights on causal mechanisms gathered in one typical case can be strengthened by studying another typical case (Goertz 2008) or, as we show below, an IIR case. Second, without the knowledge of the causal processes operative in a typical case, one can hardly discern the reason(s) why some cases are deviant. (Lieberman 2005). Third, comparative case studies relieve one from engaging in counterfactuals in order to assess the causal quality of terms.

In a QCA on necessity, the goal of process tracing is to investigate why the presence of the outcome requires the presence of the necessary condition. This focus yields the first insight for the choice of cases for comparative process tracing: Every comparison must involve at least one case that is a member of the outcome (Schneider & Rohlfing forthcoming). Based on this requirement, we identify three feasible forms of comparisons in an fSQCA of necessity. The first compares two typical cases and serves to probe the starting assumption that the same mechanism is operative in different typical cases. The second comparison includes a typical and an IIR case to substantiate the counterfactual ("Would $Y$ be absent, if $X$ were absent?") with an empirical analysis involving a consistent case that is neither a member of $X$ nor $Y$. The third variant centers on a deviant case for consistency and compares it with two additional cases, namely, a typical and an IIR case. We show that this double pairwise comparison is needed to determine omitted SUIN conditions in comparative process tracing. In the following, we expand on each form of comparison and provide formulas for identifying the best-matching pair of cases.

**Comparing Two Typical Cases**

The motivation for a comparison of two typical cases is to see whether, indeed, the same mechanism is operative in different cases in line with a pattern of necessity (see
Figure 2. Best-matching typical cases for necessary condition MA.

Goertz 2008). This is best done by providing empirical evidence that the same causal mechanism is in play in cases with different degrees of membership in X and Y. The two typical cases to be compared should span the broadest possible range of fuzzy-set memberships in X and Y. The ideal comparison thus brings together the case in the upper-right and lower-left corner of Zone 2 in our enhanced XY plot. Such a comparison matches the most typical case with the case that is barely more in than out of both X and Y (membership of 0.51 in Y and X)\(^1\).\(^3\) We acknowledge that the underlying measures for X and Y are likely to entail measurement error and that calibration decisions are uncertain (Schneider & Wagemann 2012, chap. 11; Skaaning 2011). It is precisely for this reason that we recommend the comparison of the most typical with what we call a just-so typical case. If we can show that the same mechanism is in play in both cases, it is plausible to conclude that the calibrations are correct and that the barely typical case rightly belongs to the group of typical cases.

The formula that we propose for identifying the best-matching pair of typical cases is,

\[
N_{Ti} - N_{Tj} = \frac{1 - |X_{Ti} - X_{Tj}| + |Y_{Ti} - Y_{Tj}|}{Y_{Ti} + Y_{Tj}}
\]

The absolute difference between the membership of typical cases i and j in X and Y is captured by the expressions \(|X_i - X_j|\) and \(|Y_i - Y_j|\). The larger each difference, the more distant are the cases. We subtract the sum of the differences from 1 because the smallest score for the formula \(N_{Ti} - N_{Tj}\) should identify the best suitable pair of cases. In the final step, this score is divided by the sum of the cases’ membership in Y to take into account that greater membership in the outcome is more appropriate.\(^1\)^\(^4\)

Applied to the SUI condition MA, the formula produces two pairs of typical cases as the best matches. The United Kingdom 1999 and Switzerland 1999 display identical scores (\(MA = 1; EXPORT = 0.98\)) and are located in the upper-right corner and form the best possible pair together with Switzerland 1995 (\(MA = 0.99; EXPORT = 0.51\); see Figure 2).

Comparing a Typical and an IIR Case

Another variant of comparison contrasts a typical case with an IIR case. As explained, process tracing in a single typical case should help answer the counterfactual: Would Y be absent, if X was absent? The comparison of a typical case and an IIR case allows for an empirical investigation of this question rather than reliance on counterfactual reasoning. The researcher employs comparative process tracing to gain insights into whether and why the qualitative difference in the set membership in X is the reason behind the qualitative difference in the membership in Y.
One prerequisite for this type of comparison is that the two cases hold qualitatively different set memberships in $X$ and $Y$. In addition, they should belong to two truth table rows that are identical except for the values of the necessary condition $X$ and the outcome $Y$. If the comparison follows these rules, it is possible to determine via process tracing whether the absence of $X$ accounts for the absence of $Y$ and the presence of $X$ (in conjunction with other conditions from the truth table row) accounts for the presence of $Y$. If we were to select two cases that differ on more conditions than the necessary condition $X$, we might be unable to clearly attribute the absence of $Y$ to the absence of $X$.

Among all pairs of typical cases $i$ and IIR cases $j$ that meet these requirements, the ideal pair displays the maximum difference in set membership in both $X$ and $Y$. This means that the typical case should be a full member of $X$ and $Y$, whereas the IIR case should be a full nonmember of $X$ and $Y$. For applied research, it follows that the best-matching pair maximizes the difference on $X$ and $Y$. This effectively means that the comparison of a typical case and an IIR case follows the same logic as a comparison of two typical cases. The only difference is that one case now belongs to Zone 4 instead of Zone 2. Consequently, formalized case selection relies on the same formula:

$$N_{Ti} - N_{IIRj} = \left(1 - \left(X_{Ti} - X_{IIRj}\right)\right) + \left(1 - \left(Y_{Ti} - Y_{IIRj}\right)\right)$$

$$Y_{Ti} + Y_{IIRj}$$

Applying this reasoning to our empirical example, we find only two rows populated by typical and IIR cases that are similar except for their membership in the necessary condition $MA$ and the outcome $EXPORT$. Three typical cases (UK 1990, 1995, and 1999) are a member of row $MA^*EMP^*BARGAIN*UNI^*OCCUP*STOCK$ and two IIR cases (Canada 1990 and 1995) of row $\neg MA^*\neg EMP^*\neg BARGAIN*\neg UNI^*\neg OCCUP*STOCK$. Applying the formula to any possible pair of typical and IIR cases produces a score of 0.48 for the pair UK 1999 ($MA = 1$; $EXPORT = 0.98$) and Canada 1990 ($MA = 0.31$; $EXPORT = 0.28$). Figure 3 demonstrates that the formula identifies the best pair of cases because the selected cases are close to the secondary diagonal and, at the same time, achieve the maximum observable distance on $MA$ and $EXPORT$.

Comparing a Deviant Case Consistency with a Typical and an IIR Case

A third type of feasible comparison contrasts a deviant case consistency with a typical case and an IIR case. Departing from the previous two comparisons—and small-$N$ comparisons more generally (Tarrow 2010)—this is the only form of comparison that involves three types of cases. The rationale for this double pairwise comparison lies in the reason for the deviance of a deviant case consistency: the omission of a SUIN condition.
Membership in high-value merger & acquisitions (MA)

Each pairwise comparison addresses a different element of the SUIN term. The comparison of a typical case with a deviant case consistency addresses the “SU” in SUIN, that is, it serves the conceptual purpose of establishing functional equivalence between a solution term and an omitted SUIN condition. The “N” in SUIN is subject to a comparison of a deviant case consistency with an IIR case and serves a causal purpose by delivering process-tracing evidence for the necessity of the omitted term.

Our empirical example illustrates the manner in which knowledge of the mechanism underlying the original term MA facilitates exploratory process tracing in a deviant case consistency. Instead of relying on a counterfactual, comparative process tracing in a typical case and deviant case consistency can provide empirical evidence for the claim that MA contributes to a knowledge base (K). This, in turn, makes it easier to discern an omitted term that also contributes to a broad knowledge base, such as a high share of university-trained citizens (UNI).

The ideal comparison matches the most typical and the most deviant case consistency located in the upper-right and upper-left corner of the XY plot, respectively. This pair of cases shares two important features. First, each case is a full member of Y, thus facilitating exploratory process tracing with the aim of identifying the respective SUIN conditions in the typical and deviant case. Second, the typical case is a full member of the original necessary condition, whereas the deviant case consistency is a full nonmember of the necessary condition. This criterion increases our confidence that the original solution term is necessary for Y in the typical case and that it cannot be necessary for Y in the deviant case consistency. Taking both requirements together, we propose the following formula for identifying the best pair involving a typical case T and deviant case D:

\[
N_T - N_D = \left(1 - \left(\frac{X_T - X_D}{Y_T - Y_D}\right)\right) + \left(2 - \left(Y_T + Y_D\right)\right)
\]

The expression \(X_T - X_D\) captures the requirement that the difference in X should be maximal. The difference is subtracted from 1 to ensure that the minimum score for \(N_T - N_D\) identifies the best-matching pair of typical and deviant cases. The sum of the membership in Y, \(Y_T + Y_D\), addresses the criterion that higher membership in Y of the typical case and deviant case consistency is preferable to lower membership. The sum is subtracted from 2 because lower scores should denote more adequate comparisons. Finally, the differences of the cases with respect to X and Y are again set in relation to the sum of memberships in Y. For our empirical example, two typical cases—Switzerland 1999 and UK 1999 both with \(MA = 1\) and \(EXPORT = 0.98\)—can be compared with Japan 1999 (\(MA = 0.08\), \(EXPORT = 0.96\)), producing a formula score of 0.07 (see Figure 4).
### Table 2. Summary of Post-QCA Case Studies in Studies of Necessity.

<table>
<thead>
<tr>
<th>Goal</th>
<th>Strategy</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>Building or testing hypothesis on causal mechanism</td>
<td>Single typical case $i$</td>
<td>$X_i - Y_i$</td>
</tr>
<tr>
<td>Comparison of two typical cases ($T_i, T_j$)</td>
<td>$\left</td>
<td>X_i - X_j \right</td>
</tr>
<tr>
<td>Comparison of typical case ($T$) and IIR case</td>
<td>$\left( 1 - \left( X_i - X_{IIR} \right) \right) + \left( 1 - \left( Y_i - Y_{IIR} \right) \right)$</td>
<td></td>
</tr>
<tr>
<td>Improving theory and QCA model</td>
<td>Single deviant case $i$</td>
<td>$\frac{Y_i - X_i}{Y_i}$</td>
</tr>
<tr>
<td>Comparison of typical case ($T$) and deviant case ($D$)</td>
<td>$\left( 1 - \left( X_i - X_{IIR} \right) \right) + \left( 2 - \left( Y_i + Y_D \right) \right)$</td>
<td></td>
</tr>
<tr>
<td>Comparison of deviant case ($D$) and IIR case</td>
<td>$\left( 1 - \left( X_D - X_{IIR} \right) \right) + \left( 1 - \left( Y_D - Y_{IIR} \right) \right)$</td>
<td></td>
</tr>
</tbody>
</table>

IIR = individually irrelevant

Following what we have argued about single-case analyses of deviant cases consistency, the comparison of this case with a typical case only covers the “SU” part of SUIN. The insufficiency of the omitted term $UNF$—the “I” in SUIN—is left for a separate QCA, much as demonstrating at a cross-case level that the omitted term qualifies as necessary when looking at all cases that are members of $Y$ and nonmembers of the original solution term $MA$. Still, a comparison of a deviant case consistency with an IIR case is of great value for delivering process-tracing evidence that the omitted term is necessary. In the discussion of a single-case analysis of deviant cases, we argued that this involves counterfactual reasoning on the basis of the insights gathered via process tracing in the deviant case. By comparing a deviant case consistency with an IIR case from Zone 3, this counterfactual can be replaced with empirical evidence.

The ideal comparison of a deviant case and the IIR case must meet four requirements. First, the deviant case is a full member of $Y$ whereas the IIR case is a full nonmember. Second, the deviant case is a full member of the omitted term whereas the IIR case is a full nonmember. Third, the two cases are members of two truth table rows that are identical except for the membership in the formerly omitted term and $Y$. Note that this comparison is identical to the comparison of a typical case with an IIR case. The comparison of a deviant case with an IIR case allows one to determine whether adding the omitted term is warranted before doing so. Consequently, the formula for this comparative design draws on the same formula, where $X$ now stands for the membership in the omitted term discerned via exploratory process tracing:

$$N_{IIR} - N_{IIR} = \frac{1}{\left( 1 - \left( Y_{IIR} - Y_{IIR} \right) \right) + \left( 1 - \left( Y_{IIR} - Y_{IIR} \right) \right)}$$

A look at our data shows that there are no deviant cases of consistency in kind and IIR cases that meet the requirements for this comparison. If the data at hand do not contain the appropriate cases for this comparison, one again has to rely on a counterfactual substantiated with process-tracing evidence gathered in the deviant case consistency.

### Conclusion

Necessary conditions play an important role in the social sciences. QCA is one of the most formalized techniques for their empirical analysis. Because QCA is a case-based method that works best with the in-depth knowledge of cases (Rihoux and Lobe 2009), one should do process tracing after a QCA to search for causal mechanisms that underlie set relations and to detect omitted terms. We have provided a comprehensive discussion of how QCA results can be used for intelligible and formalized case selection. The variants of process tracing after a QCA of necessity and corresponding formulas are summarized in Table 2. Our empirical illustration relied on fuzzy-set QCA as the more general variant of QCA. While the basic logic of case selection extends to csQCA (Schneider & Rohlfing forthcoming), our formulas only apply in fsQCA.

On a broader scale, our paper contributes to the advancement of multimethod research (MMR) by extending the perspective beyond regression analysis (Lieberman 2005; Rohlfing 2008) to set-theoretic...
methods. The topic of set-theoretic MMR involves more issues than we have discussed here. There are important differences between process tracing after a QCA of necessity, vs. a QCA of sufficiency. For example, sufficient solutions usually involve conjunctions and, in contrast to necessity, cases that reduce coverage are targets for process tracing (Schneider & Rohlfing forthcoming). Further exploration of these differences is a worthwhile endeavor to strengthen causal inferences in research on necessity and sufficiency alike.

Notes
2. Rihoux and Lobe's (2009) terms "upstream" and "downstream"—introduced in their discussion of the Qualitative Comparative Analysis (QCA) research process and the role of case knowledge therein—only partially overlap with ours.
3. We say "main purpose" because, even in post-QCA process tracing, one might find hitherto overlooked evidence for changes in the population, concepts, measurement, and calibration.
4. We use this example for illustrative purposes only.
5. As mentioned, MA alone is not necessary. We nevertheless start with a single condition to illustrate process tracing in the most basic setting of just one condition to identify potentially omitted terms.
6. Using simulated data, an online appendix to our paper demonstrates that our formulas produce plausible results (http://prq.sagepub.com/supplemental/).
7. If the formula nevertheless yields the same score for multiple cases, one should choose the one with higher membership in Y.
8. This holds true because the logical OR operator requires assigning cases the maximum set membership across all conditions combined by logical OR (Ragin 2000, 175).
9. Alternatively, they contribute to the trivialness of the necessary condition (Goertz 2006; Ragin 2006; Schneider and Wagemann 2012).
10. Pre-QCA studies of cases in Zone 3 might help to identify measurement and/or calibration error in either X (membership too easy) or Y (membership too difficult) and increase the relevance of X. These are not model-related modifications as we define them, though.
11. Goertz (2008, 11) calls this "choosing cases diversely."
12. Typical cases that are joint members should only be chosen if no unique members are available.
13. This selection strategy is the set-theoretic equivalent to the diverse case strategy (Seawright and Gerring 2008; Rohlfing 2012, chap. 3).
14. If two or more pairs of cases obtain the same score, researchers should choose the one with the largest difference in Y.
15. We know that the entire truth table row to which the typical case belongs is sufficient for the outcome. Every condition constituting this row is an INUS condition, and taking away any of these conditions might lead to the absence of the outcome (assuming that we analyze unique members and that the condition is not logically redundant). Consequently, case selection for comparative process tracing in the typical case and the IIR case must ensure that all INUS conditions are present in these cases.
16. Note that for process tracing on the mechanism behind the pattern of necessity, it only matters that cases are members of similar truth table rows, not how strong their membership in these rows are.
17. If two or more pairs of cases obtain the same score, researchers should choose the one with the largest difference in Y.
18. See the truth table in the online appendix (http://prq.sagepub.com/supplemental/).
19. If two or more pairs of cases obtain the same score, researchers should choose the one with the largest sum of membership scores in Y.
20. In the ideal scenario, the deviant cases consistency turn into typical cases if we were to add the omitted term to the solution.
21. If two or more pairs of cases obtain the same score, researchers should choose the one with the largest difference in Y.

Acknowledgements
The editors wish to thank Priscilla Álamos-Concha for her help in the fine-tuning of the manuscript.

Declaration of Conflict Interests
The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Charles Ragin acknowledges support from the National Science Foundation: NSF SciSIP and STS Programs, award #1057559.

References


Hicks, Alexander M. 1994. “Qualitative Comparative Analysis and Analytical Induction: The Case of the Emergence of the


Schneider, Carsten Q., and Clausdius Wagemann. 2007. QCA und fsQCA. *Ein einführendes Lehrbuch für Anwender und jene, die es werden wollen* [QCA and fsQCA. An introductory textbook for users and those who want to become so]. Opladen: Barbara-Budrich-Verlag.


