Analysing necessity and sufficiency with Qualitative Comparative Analysis: how do results vary as case weights change?

Barry Cooper & Judith Glaesser

Abstract

Ragin's Qualitative Comparative Analysis (QCA) and related set theoretic methods are increasingly popular. This is a welcome development, since it encourages systematic configurational analyses of social phenomena. One downside of this growth in popularity is a tendency for more researchers to use the approach in a formulaic manner - something made possible, and more likely, by the availability of free software. We wish to see QCA employed, as Ragin intended, in a self-critical manner. For this to happen, researchers need to understand more of what is going on behind the results generated by the available software packages. One important aspect of set theoretic analyses of sufficiency and necessity is the effect that the distribution of cases in a dataset can have on results. We explore this issue in a number of ways. We begin by exploring how both deterministic and nondeterministic data-generating processes are reflected in the analyses of populations differing in only the weights of types of cases. We show how and why weights matter in causal analyses that focus on necessity and also, where models are not fully specified, sufficiency. We then draw on this discussion to show that a recent textbook discussion of hidden necessary conditions is weakened as a result of its neglect of weighting issues. Finally, having shown that case weights raise a number of difficulties for set theoretic analyses, we offer suggestions, drawing on two imagined population datasets concerning health outcomes, for mitigating their effect.

Introduction

Ragin's Qualitative Comparative Analysis (QCA) provides a way of undertaking case-based configurational analysis, focusing on the necessary and sufficient conditions for an outcome. QCA is increasingly used to undertake systematic set-theoretic analyses of small to medium datasets and, occasionally, to analyse large survey datasets. While regression approaches typically aim to estimate the net effects of variables on an outcome, while controlling for other variables, QCA typically focuses on the consequences of conjunctions of conditions (Ragin 1987, 2000, 2008). Whether QCA is a method for establishing causal or descriptive knowledge is currently a matter of debate. Baumgartner (2014), who has developed the related Boolean technique of Coincidence Analysis (CNA), argues that the Boolean analysis of regularities is the correct route to the establishment of causes as difference-makers. Others

see QCA as a more descriptive approach. We have argued elsewhere (Cooper and Glaesser 2012) that QCA is best used with some additional method such as process-tracing when causal understanding is sought. Gerring (2012, pages 356-358) makes similar points, discussing what Ragin and others have written on this topic.

An important question to ask of any method is how it functions when confronted with a thought experiment in which we postulate some assumed causal structure which should, ceteris paribus, generate a particular set of regular relationships between explanatory conditions and some outcome. Assuming a particular causal relation, and an associated data generating process, we would want the results delivered by any analytic method to reflect these, at least in those cases where we correctly enter all the conditions into our analysis. In practice, of course, social scientists often analyse data via less than perfectly specified models, and it is therefore also important to explore, as we will for the case of QCA, how a method copes, given a postulated assumed causal structure, when some explanatory conditions are omitted from the analytic model and/or some redundant ones are included.

We write then on the assumption that simulations using simple examples are often useful for assessing the extent to which methods can deliver findings that reflect validly some assumed underlying causal structure (Baumgartner, 2014). They can also serve a useful purpose in increasing scholars' understanding of the potential pitfalls associated with any analytic algorithm, especially those that arise when, as is typical, we have less than perfect knowledge of the factors causally relevant to an outcome.

Our particular concern in this paper is whether and how weighting, i.e., the distribution of types of cases with certain characteristics in the sample and/or the population under study, affects analytic results when QCA is employed. The importance of this issue has become clear to us while undertaking QCAs of various datasets (Cooper, 2005; Glaesser and Cooper, 2012) and when reflecting on arguments made in Schneider and Wagemann (2012) concerning the sources of "hidden necessary conditions" in Boolean analyses. We will discuss case weights in the context of crisp set QCA. We will not address fuzzy set QCA but the reader familiar with QCA will be able to see that parallel issues arise when fuzzy sets are employed. The arguments we make are of general relevance, and not specific to small, medium or large n QCA. There will be some specific situations, however, where the use, in minimising a truth table, of a frequency threshold would affect the ways in which weights affect an analysis.

To follow our arguments, any reader new to QCA will need a brief introduction to the indices used in crisp set QCA. The paper has the following structure:

- Section 1 introduces essential elements of crisp set QCA.
- Section 2 explores the problems that the relative proportions of cases can produce for QCA users with varying degrees of knowledge of the full set of causes of some outcome who research, in terms of quasi-sufficiency and quasi-necessity, a simple deterministic structure.
- Section 3 considers a non-deterministic structure.
- Section 4 discusses related problems in Schneider and Wagemann's treatment of necessary conditions.
- Section 5, in the light of the problems pointed to in the earlier sections, considers possible ways QCA users might mitigate their consequences.

1. QCA (crisp set): bare bones of measures

For strict *sufficiency* of a condition, or a conjunction of conditions, X for Y, we need, wherever X is present, to find Y also present. This requires the subset of cases with the conditions X to be a subset of the cases with the outcome Y (the left-hand side of Fig. 1). For strict *necessity* of a condition Z for Y, on the other hand, we need, given the outcome Y, always to find the condition Z present. This requires the set of cases with the outcome Y to be a subset of the set of cases with the condition Z (the left-hand side of Fig. 2). In practice, subset relations in the social world are frequently not as perfect as these. The more realistic right-hand sides of the two figures show such approximations to sufficiency (Fig. 1) and necessity (Fig. 2). Ragin (2008) uses, for crisp sets, a simple measure of the closeness of such relations to strict subsethood. Consider sufficiency. On the right-hand side of Fig. 1, he takes the proportion of cases in X that fall within the boundaries of Y as a measure of the consistency of these data with sufficiency. Such imperfect subset relations are usually described with the terms quasi-sufficient (and, for necessity, quasi-necessary). In the case of Fig. 1, the right-hand side would give us a consistency of the order of 0.8, usually taken as large enough to support a claim of quasi-sufficiency.

Fig. 1 Strict and quasi-sufficiency of X for Y	Fig. 2 Strict and quasi-necessity of Z for Y
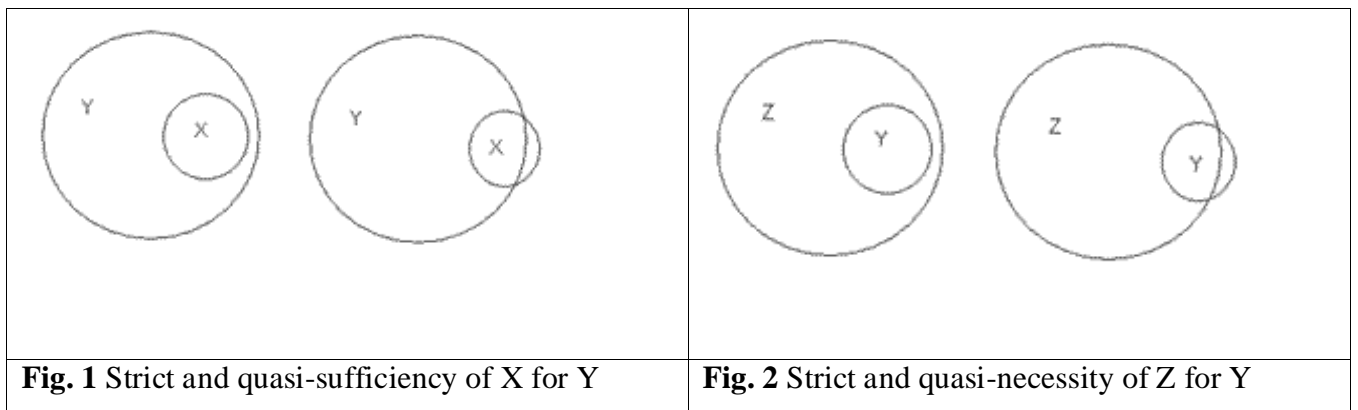
Alongside these measures of consistency, Ragin (2006, 2008) defines associated measures of *coverage*. These assess the empirical importance of a condition. In the case of sufficiency,

coverage assesses the extent to which a condition X "covers" the outcome set, i.e. how many cases of the outcome are "explained" by the condition. It is mathematically equivalent to the measure of the consistency with necessity of X for Y, i.e. how much of Y is also X. In Fig. 1 the coverage of X is low (and, correspondingly, X is not necessary or quasi-necessary for Y). For a necessary condition Z, coverage assesses the extent to which it is also sufficient for the outcome. Z's coverage is mathematically equivalent to the consistency with sufficiency of Z. In **Fig. 2**, the coverage of the necessary condition Z is low (and, correspondingly, Z is not sufficient or quasi-sufficient for Y).

Finally, it is important to note that it is usually configurations, or conjunctions of causes, that are the focus of analysis in QCA. These are represented by such expressions as (P*Q*~R) + (M*N) where the * indicates set intersection (logical AND), the ~ negates a set ("not R" in our example), and + indicates set union (logical OR). A conjunction such as P*Q*~R is allowed into the final solution if, say, at least 80% of the cases in the set P*Q*~R achieve the outcome under study. If it is also true that the configuration P*Q*R meets this 80% threshold for quasi-sufficiency, then the terms P*Q*~R and P*Q*R can be minimised to P*Q on the grounds that the absence or presence of R, *given the chosen threshold*, makes no relevant difference to the proportion of cases achieving the outcome.

2. A simple deterministic structure

We will explore the problems that may arise as a result of inadequate reflection on the analytic effects of case weights for those wishing to undertake causal analysis via QCA or similar techniques. We will initially work outwards from a postulated causal structure, showing how the distribution of cases in a dataset, given this structure, impacts on measures of consistency with sufficiency and necessity. It is important to hold in mind that we are assuming that this structure exists in the world but that the social scientist undertaking the analysis of its effects may have a far from perfect knowledge of it.

The first postulated structure links two factors, a condition T and an outcome O. Whenever T is present, and only when T is present, does O occur. We will also assume, however, that our imagined researchers may variously believe that T is the crucial cause of O, or that a factor S is the sole cause of the outcome, or that S and T together are important. For this reason, we set out four *empirical* relationships between the presence and absence of S, T and O that are compatible with the structure[1]. A plausible substantive interpretation of 1-4 might be that T

---

[1] We are assuming, to keep things simpler, that the causal relations between S, T and O are not a function of the distributions of the possible configurations of the two conditions S and T and their negations in the populations. In fact, they often will be: in cases, for example, where the configurations are real actors in conflict over positional goods. We are also bracketing out here any causal analysis of the origin of the pattern of weights in a truth table. Clearly there are causes behind the weight of cases in a population. Wars would affect, via

represents having high cognitive ability and S membership of the dominant ethnic group in a society where there are just two ethnic groups, while O is some career outcome.

$$S*T => O \hspace{3cm} \text{(Implication 1)}$$
$$S*\sim T => \sim O \hspace{3cm} \text{(Implication 2)}$$
$$\sim S*T => O \hspace{3cm} \text{(Implication 3)}$$
$$\sim S*\sim T => \sim O \hspace{3cm} \text{(Implication 4)}$$

Now, a researcher using QCA who correctly believes that T is the cause of O will find, on collecting data of adequate quality, that T is both sufficient and necessary for O, i.e. T <=> O. We now, however, explore the case of a researcher who believes that both T and S matter for the outcome. Such a researcher will collect data on S, T and O in order to construct a truth table to explore the evidence for implications 1-4. Such a truth table will have the form shown in Table 1 where A, B, C and D stand for the number of cases in the four configurations.

Table 1: Truth table representing implications 1-4.

Implication	S	T	Configuration	Number of cases	O: Consistency with sufficiency
1	1	1	ST	A	1
2	1	0	S~T	B	0
3	0	1	~ST	C	1
4	0	0	~S~T	D	0

Ignoring for the moment the numbers of cases in the rows, we could describe this table by ST + ~ST <=> O[2] (writing S*T as ST for convenience), or, after minimising the left-hand side, T <=> O. T is both sufficient and necessary for O, as we want. What happens if we begin to vary the number of cases per row? Table 2 shows two possible distributions of A-D.

Table 2: Two sets of values for A-D

Configuration	Outcome O	Number of cases	In dataset 1	In dataset 2
ST	Present	A	16	48
S~T	Absent	B	4	4
~ST	Present	C	17	5
~S~T	Absent	D	5	5

---

migration, the weights of ethnic groups, for example. Birth rates – and their causes – the weights of children in social classes. Random processes will affect weights in samples from populations. These matters are clearly also important, but not our concern here.

[2] This non-minimised relation also reminds us that, as Baumgartner (2008) notes, as long as there are at least two terms on the left-hand side, there is an asymmetry between sufficiency and necessity. It makes sense to think of causal influences running from left to right, but not from right to left, since knowing O does not tell us which of the two left-hand side terms are present.

Whatever the values of A-D[3], ST and ~ST will always be found to be sufficient for O. T will always be found to be necessary for O. The solution will remain T <=> O. But now consider a researcher who believes that only S matters and therefore collects no data on T. The truth table will reduce to just two rows, one for S (with A+B cases) and one for ~S (with C+D cases). The consistency with sufficiency of S is given by A / (A+B). For dataset 1 it is 0.8, and for dataset 2 it is 0.923. The consistency of S with necessity is given by A / (A+C). For dataset 1 it is 0.485; for dataset 2 0.906.

In the case of dataset 2 the result suggests that S is both quasi-sufficient and quasi-necessary for O. For dataset 1, on the other hand, there is good support for quasi-sufficiency but none for quasi-necessity. Remember the causal structure is fixed. Only the weights of cases have changed. Given that in much empirical research there is incomplete knowledge of the causal factors producing an outcome, this dependency of the solution on the weights of cases is clearly an important issue for users of QCA and related techniques. In this case, for dataset 2, the pattern of case weights has misled the analyst who believed S was the crucial cause.

Now, since we know, by our design of the thought experiment, the generative causal basis underlying 1-4, we know that S is not necessary, since implication 3, compatible with T <=> O, says ~ST => O. A problem exists concerning the assessment of the necessity of S for the researcher who believes S is the cause of O. Note that we can imagine population proportions which raise the consistency of S with necessity much higher. If we give C in dataset 2 the value of 1, then we have a consistency of 48 / (48+1), i.e. 0.980. It is worth looking at the range of possible values for the consistency with necessity of S in our example (where T <=> O). If we fix A at 100 and let C vary in relation to it (from 1 to 100 for illustration)[4] then it is easy to see that the consistency varies from a little over 0.99 to exactly 0.50. This is shown in Fig. 3.

---

[3] As long as they each are above zero and we have adequate evidence for the implication claims 1-4.
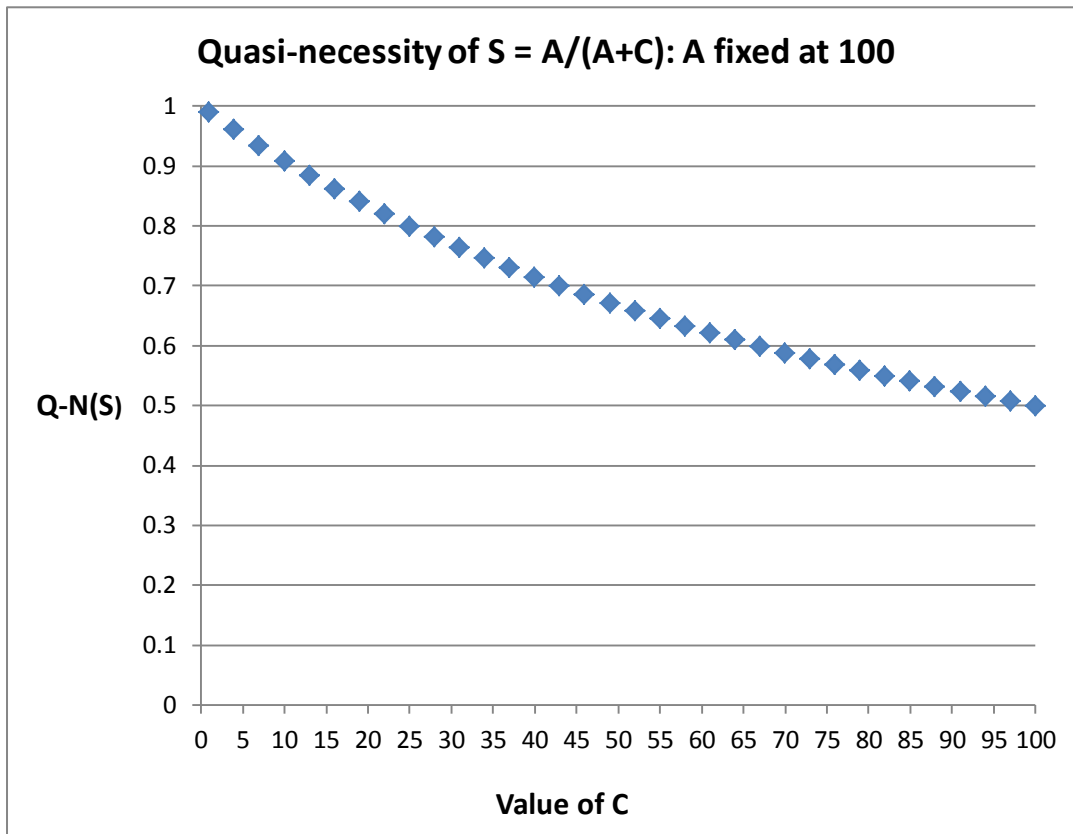[4] The crucial issue here is the ratio of A to C, not the absolute sizes of A or C.

Fig. 3 Quasi-necessity of S as the ratio of A to C varies.

It is often argued that the necessity (or sufficiency coverage) of a factor should only be assessed where the factor has first been shown to be sufficient (Ragin 2008). However, given less than perfect knowledge of the causal factors operating on an outcome, the distribution of cases, as we have seen, also affects consistency with sufficiency. Again, consider the researcher who believes S is the crucial cause. Returning to datasets 1 and 2, we can see that the consistency with sufficiency of S is given by A / (A+B). In dataset 1 it is 16 / (16+4), or 0.800. In dataset 2, it is 48 / (48+4), or 0.923. If we change the value of A to 2 in dataset 2, this consistency becomes 2 / (2+4), or just 0.333. If we fix A at 100 and, illustratively, let B vary from 1 to 100, then the consistency of S with sufficiency varies from 0.99 to 0.5, as did its consistency with necessity with varying C. Now, in this simple deterministic world, both B and C can be varied independently of A and of each other without affecting implications 1-4. We assume our generative causal structure remains intact under these changes. We can for example, having fixed A at 100, and leaving the irrelevant D at 5, reverse the values of B and C in two new datasets 3 and 4 (Table 3). When we do this we switch the consistencies of S with sufficiency and necessity, as Table 3 records:

Table 3: Two more datasets

Configuration	Percent with Outcome	Numbers of cases	Dataset 3	Dataset 4
ST	100	A	100	100
S~T	0	B	1	100
~ST	100	C	100	1
~S~T	0	D	5	5
		Consistency (necessity) of S	0.50	0.99
		Consistency (sufficiency) of S	0.99	0.50

We have shown that, given a fixed causal structure, and a researcher with incomplete knowledge of which causes matter, a wide range of results for consistencies are possible, simply as a result of changes in the weights of types of cases in the dataset. This suggests that, for any researcher thinking causally rather than descriptively, QCA analyses that ignore the effects of case weights on consistency formulas may be problematic. To this point we have considered a deterministic structure but allowed its analysis in terms of quasi-sufficiency and quasi-necessity. In the next section we will assume that the causal structure is not fully deterministic.

3. A non-deterministic causal scenario

We continue by considering a simple non-deterministic structure in which there are two causal factors, S and T, and an outcome O. The assumption is again that there is a particular structure in the world, but that the researcher has imperfect knowledge of its features. The postulated structure is such that, for the four possible configurations of S and T and their negations, all else being equal, the proportions gaining the outcome O are those shown in Table 4[5]. Here, differently from the earlier structure, we are assuming that both S and T are causes of O. We will also assume, however, that a researcher might suspect that S alone, or T alone, or both S and T may be causes of O.

Table 4: Consistencies with sufficiency for 4 configurations

S	T	Configuration	Proportion gaining the outcome (O)
1	1	ST	0.99
1	0	S~T	0.65
0	1	~ST	0.2
0	0	~S~T	0.1

[5] This fictitious structure, for our current purposes, is best seen as incorporating a random element (as in, e.g., ST causes the probability of gaining the outcome to be 0.99). It would also be possible to see it as having omitted some relevant causal factor but, given the ways in which imperfect model specification can also affect the assessment of sufficiency (see section 2 and what follows in this section), this possible account of it is best avoided in this paper.

Table 5: Three invented populations

S	T	A: numbers	B: numbers	C: numbers
1	1	100	200	200
1	0	100	20	1800
0	1	100	100	100
0	0	100	100	100

Table 6: Consistencies above 0.8 with sufficiency and necessity (coverage in brackets)

Set	Population A: consistency with sufficiency	Population A: consistency with necessity	Population B: consistency with sufficiency	Population B: consistency with necessity	Population C: consistency with sufficiency	Population C: consistency with necessity
~T						0.844 (0.621)
T				0.905 (0.727)		
~S						
S	0.820 (0.845)	0.845 (0.82)	0.959 (0.876)	0.876 (0.959)		0.979 (0.684)
~ST						
S~T						0.837 (0.65)
ST	0.990 (0.510)		0.990 (0.822)	0.822 (0.99)	0.990 (0.142)	
~S~T						

For the sake of our illustration of the ways weights may affect various researchers' findings we will consider three possible datasets, shown in Table 5[6]. Table 6 sets out all indices of consistency with sufficiency and necessity that pass a threshold of 0.8, adding the coverage figures in brackets. For completeness, given that any researcher may or may not know correctly what the true causes of O are, we have included in the table all of S, T, their negations and conjunctions of these.

First, we can see from Table 6 that, considering quasi-sufficiency, the results from all three populations tell us that ST is quasi-sufficient for O at a very high level (0.99). Any researcher correctly believing that S and T are both causes here would have found this result to be unaffected by the different weighting of cases in the three datasets. The researcher believing just S to be the cause of O would have found S to be sufficient, using a threshold of 0.8, in datasets A and B but not C. If s/he had happened to use a stricter threshold of 0.9, then S would have been found to be sufficient in just dataset B. Weighting matters, unless we assume a researcher with full knowledge of the causes operating.

Measures of necessity are also affected by the case weights, of course. For a researcher who enters both S and T in their QCA model, and finds the solution for sufficiency of ST => O, the consistency with necessity of ST will vary across the three datasets from 0.510 in dataset A, through 0.822 in B, to 0.142 in C. Separate assessments of the necessity of the two factors

---

[6] Again, we are assuming, to keep things simpler, that the causal relations between S, T and O are not a function of the distributions of the configurations in the populations. (See footnote 1.)

are also weights-dependent. T only passes the 0.8 threshold in dataset B. S passes it in all three datasets to differing degrees[7].

How important are these effects of weights? The answer partly hinges on the position one takes on the purpose of QCA and related techniques. Baumgartner (2014) argues that the purpose of Boolean techniques is to gain causal knowledge via the analysis of regularities, and that this requires, ideally, a solution of the form $C_1 + C_2 + C_3 <=> O$, where the disjunction of conjuncts on the left-hand side is both sufficient and necessary for the outcome O. On this argument, if ST <=> O appears as the solution, with ST both sufficient *and* necessary for O, then this lends support to a causal claim. Using our threshold of 0.8, and when the researcher correctly believes both S and T matter, this solution would appear for population B, but not for A or C. The key point is perhaps that any causal conclusions based merely on an analysis of the regularities deriving from our postulated causal structure will vary with the weights characterising our dataset. If the researcher is assumed to know that S and T matter, the key problem in this particular scenario concerns the ways the assessment of necessity vary with case weights. Any researcher deciding on the necessity of a causal condition, or a combination of conditions, purely on the basis of the consistency with necessity taken from a particular population would seem to be effectively assuming that the distribution of cases in this particular population is, in some sense, privileged[8].

There is another view of this issue. Someone arguing that the sufficiency of ST is the crucial matter to uncover might argue that the varying assessments of nearness to necessity – in their guise of coverage – merely give us useful descriptive information about the empirical importance of this pathway in the three datasets. Ragin has argued this position, arguing that the best uses of the truth table analysis in QCA are "more descriptive", providing "a basis for causal interpretation at the individual case level"[9]. Here the analysis of sufficiency is prioritised and any subsequent assessment of the necessity of individual factors appearing in solutions should depend not only on the numerical index of consistency with necessity but should also entail "a good measure of corroborating substantive evidence and theoretical argumentation"[10].

---

[7] It is perhaps important to note that our arguments do not depend on the particular threshold (0.8) employed. Given appropriate case weights and/or postulated causal structures, the problems we discuss could also appear with stricter thresholds such as 0.9.

[8] We will not discuss samples from populations in this paper, but clearly much of what we write has implications for the use of QCA with samples, especially where some cases are more or less represented than others via sampling (see Glaesser and Cooper 2012). We should also note, concerning populations, that sometimes there will be only one population. Some implications of this for the analysis of necessity are considered in Section 5.

[9] Personal communication (email), 10/11/14.

[10] Personal communication (email), 10/11/14.

Considering in this section a non-deterministic causal structure, we have shown again that the case weights matter. In the case of *necessity*, even when we assumed, initially, that we knew that S and T are the important causal factors, and that we knew the causal consequences for the outcome of their various conjunctions, we found the assessment of necessity of some conditions and configurations varying across our three populations. We want now to show how these weighting problems considered so far are reflected in a recent discussion by Schneider and Wagemann of "hidden necessary conditions". We have previously pointed out some problems with their textbook discussion of "hidden necessary conditions" (Cooper et al. 2014). We will revisit these, but then develop a more general perspective on one of the issues they raise, drawing on our discussion earlier in this paper.

4. Schneider & Wagemann on "hidden necessary conditions" and inconsistent truth table rows

So far, in thought experiments, we have considered the problems weights might cause given two causal structures, one deterministic, one not. We will now discuss an example from a prominent textbook on set theoretic methods (Schneider and Wagemann 2012) where the generative causal structure is, as is often the case in empirical science, unknown to the analyst. The focus in this section is on assessing quasi-necessity.

Having previously discussed "hidden necessary conditions" due to incoherent counterfactuals in the context of limited diversity[11] Schneider and Wagemann (2012) begin their second section on "hidden necessary conditions" by arguing that such incoherent assumptions about remainders are not the only reason for the disappearance of necessary conditions. They note that necessary conditions can disappear even from a conservative solution term (their preferred term for what Ragin terms a complex solution). This **can** arise, they say, when "inconsistent truth table rows are included in the logical minimisation that contain the absence of the necessary condition" (225). By "inconsistent" rows, they refer to rows containing cases where some achieve the outcome and some do not. They illustrate this possibility with the hypothetical example reproduced here as Table 7.

---

[11] Cooper et al. (2014) also provide a critical discussion of Schneider and Wagemann's arguments and proposals re the treatment of necessary conditions given limited diversity.

Table 7: Schneider and Wagemann's (2012) Table 9.2

	Conditions			Outcome			
Row	A	B	C	Y	Cases with Y	Cases with ~Y	Consistency for Y
1	0	0	0	0	0	1	0
2	0	0	1	1	20	0	1
3	0	1	0	0	0	39	0
4	0	1	1	0	0	15	0
5	1	0	0	1	10	0	1
6	1	0	1	1	15	0	1
7	1	1	0	1	4	1	0.8
8	1	1	1	0	0	2	0

Schneider and Wagemann's argument proceeds as follows. Looking at row 7 of Table 7, it can be seen that, of these five cases that are AB~C, four have the outcome Y. This configuration is therefore quasi-sufficient for Y, with a consistency of 0.8. If we run a QCA on the truth table in Table 7 that allows row 7 to go forward (i.e. accepting a consistency of 0.8 as adequate for quasi-sufficiency) we obtain the following minimised solution:

$$A\sim C \; + \; \sim BC \; => Y \qquad\qquad \text{(Equation 1)}$$

The first term, A~C, has a consistency with sufficiency of 0.93 and the second, ~BC, of 1.0. As Schneider and Wagemann note, no single condition appears in both paths and one might be "tempted to conclude there is no necessary condition for Y" (226). They note, however, that separate tests for necessary conditions show ~B to be necessary for Y, with a very high consistency of 0.92[12]. They argue that, "based on the empirical evidence, we have good reasons to consider ~B to be a relevant necessary condition for Y". They argue (p. 226) that ~B disappears because it is "minimized from the sufficiency solution by matching row 5 of Table 9.2 (A~B~C) with the inconsistent row 7 (AB~C) into the sufficient path A~C. ... thus, the necessary condition disappears from the sufficiency solution because both the former and the latter are not fully consistent."

Our earlier discussion of the effects of the relative numbers of cases suggests this is far from the whole story. To see why, consider a small change in the truth table, one that removes the problem of inconsistency. Assume, contrary to what we see in Table 7, that all five cases in row 7 achieve the outcome Y. What then happens when we undertake a QCA that parallels that reported by Schneider and Wagemann? On the basis of changing just one case from not having to having the outcome, we obtain the same algebraic solution as in Equation 1 (though in this case both terms are strictly sufficient, i.e. have consistencies of one). And, once again, ~B is quasi-necessary for Y (with a consistency of 0.90 rather than 0.92). We have no inconsistent rows, but we still have the "problem". Clearly, ~B, when it is a quasi-necessary condition, can disappear from the solution even in the absence of the inconsistency in row 7.

[12] 45 of the 49 cases with the outcome Y have the condition ~B.

The problem of this hidden necessary condition is not then due to the inconsistency of row 7 per se. What then is its cause? Our earlier discussion suggests that the relative number of cases per row may be a crucial factor: analysing deterministic datasets 1 and 2, we found that a researcher focusing attention on the single condition S would find the condition S quasi-necessary for dataset 2, but not for dataset 1, simply as a result of the differing distributions of case numbers in the configurations. A glance at Schneider and Wagemann's truth table shows there are quite different numbers of cases in its various rows. It is therefore instructive to undertake another analysis. If we increase the number of cases in row seven to 40, but retain the proportion in this row achieving the outcome at 0.8, we obtain the same solution for sufficiency (with the first term having a consistency of 0.84 and the second of 1), but we now find that a test for the necessity of ~B returns a consistency of only 0.58. ~B is no longer a quasi-necessary condition. The "problem" of its being a hidden *necessary* condition disappears. The inconsistency of row 7 is not the fundamental problem. Even given this inconsistency, we can remove the problem of the disappearance of ~B by reweighting so that it is no longer a necessary condition in the first place. This is what we found earlier when we kept the proportion of cases within the configurations obtaining the outcome the same in populations A, B and C (see Section 3), but varied the case numbers. Our varying of the numbers had a marked effect on whether or not certain conditions or configurations of conditions passed the threshold for quasi-necessity.

The fundamental problem is not due to inconsistencies per se (though these will certainly modify the way it appears) but to the weight of cases in the rows of truth tables, coupled with the manner in which the arithmetic of proportions works. Indeed we can make the problem much worse by running the analysis with just one case in row 7 – *a case which has the outcome* – and therefore with no inconsistency characterising this row. Doing this, we obtain the solution for sufficiency in Equation 1 (with both terms having consistencies of 1) while the consistency with necessity of ~B rises to an almost perfect 0.98, higher than the 0.92 that provided the basis for Schneider and Wagemann's arguments.

Clearly, Schneider and Wagemann, in choosing their example, have brought an important issue to our attention. In our view, however, in concentrating on just the inconsistency, they have not treated it in a general enough fashion. We also need to consider weights. To explore the consequences of weighting for their example, we will now consider how several consistency and coverage measures vary as we change the weighting of the rows in their table. We will present two graphs that illustrate the general patterns we find as we vary the number of cases in row 7 while holding the other case numbers constant. The first graph (Fig. 4) maintains Schneider and Wagemann's consistency of 0.8 for row 7. In the second (Fig. 5)

we move away from their focus on inconsistency and raise the consistency of row 7 to a perfect 1.0.

Fig. 4 shows how the consistency with necessity of ~B falls away as we increase the number of cases in row 7. Schneider and Wagemann choose to have 5 cases in this row - a small number in relation to the rows including ~B that achieved the outcome (respectively 20, 10 and 15). It is important to note that, given the particular distribution of cases in the other rows, the consistency with necessity of ~B falls away much more quickly than does that of the consistency of the solution for sufficiency itself. Once we have 25 cases in row 7 we find that the consistency with necessity of ~B is below 0.7 and the original problem of ~B being a (hidden) *necessary* condition begins to disappear.
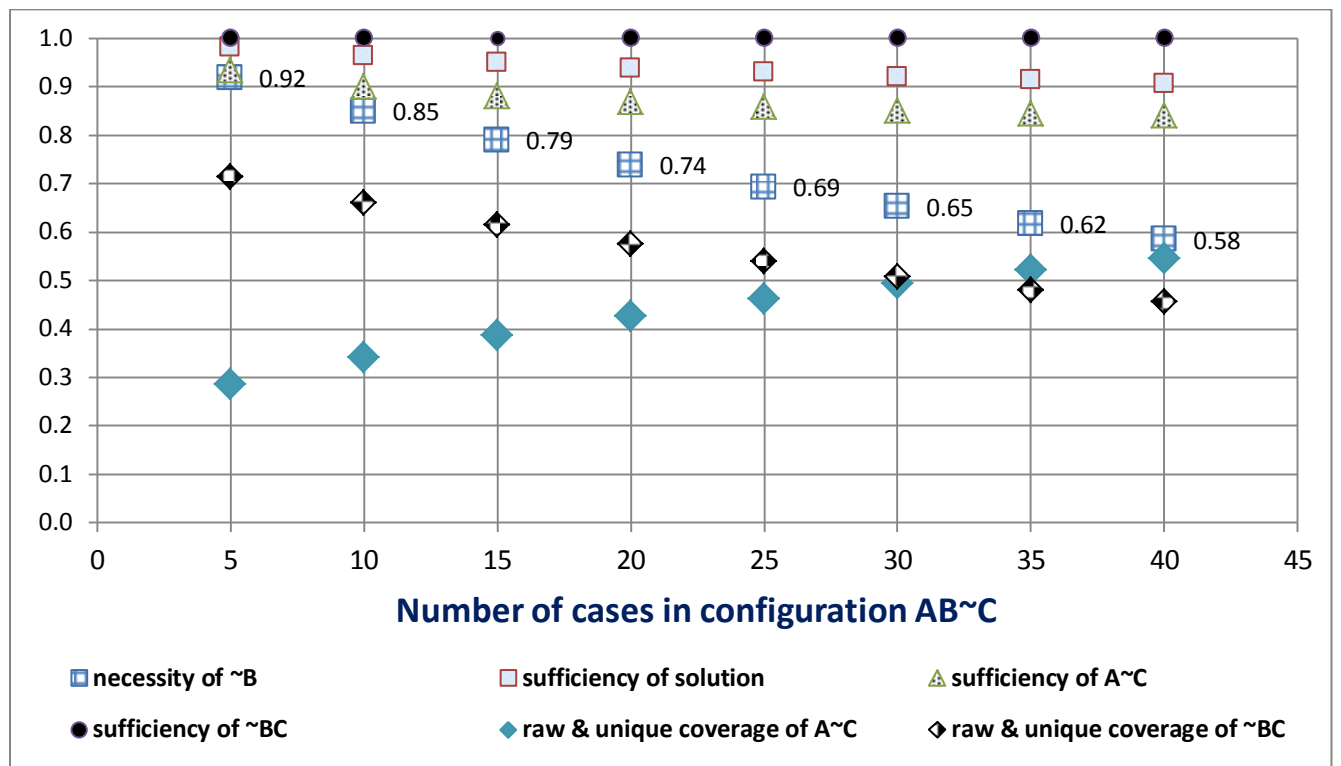

Fig. 4 Solution is A~C+~BC: Effect of changing the number of cases in AB~C (in which 80% achieve the outcome)

Fig. 5, where the AB~C row now has 100% consistency with sufficiency rather than 80%, shows a perhaps more striking pattern than Fig. 4. Here, looking at the situation where there are just 5 cases in row 7 (i.e. reproducing Schneider and Wagemann's case numbers), we have a consistency with necessity of ~B of 0.90 - very close to that of 0.92 in Schneider and Wagemann's original "inconsistent row" example. If we drop the cases in row 7 to 4 then, even without any inconsistency, we are back to their result: a consistency with necessity of ~B of 0.92 while ~B still does not appear in all the terms in the solution for sufficiency. The overall solution, A~C+~BC, maintains a perfect consistency of 1.0 here as the number of

cases in row 7 is varied. There is a sense in which the inconsistent case introduced by Schneider and Wagemann to make their argument concerning a hidden necessary condition is actually irrelevant. If we delete just this case from the dataset, the row AB~C (now with a consistency of exactly 1) goes forward into the solution A~C+~BC, and the consistency with necessity of ~B remains at 0.92[13]. Perhaps the most important general point to emphasise about Fig. 5 is that, given perfect sufficiency for all of the rows of Table 7, the necessity index for ~B changes as we vary the weight of cases in AB~C. The coverage indices for the two minimised terms in the solution also vary (but always add to 1). It makes sense to see coverage here as indicating the relative empirical importance of the two pathways A~C and ~BC as the weights change.

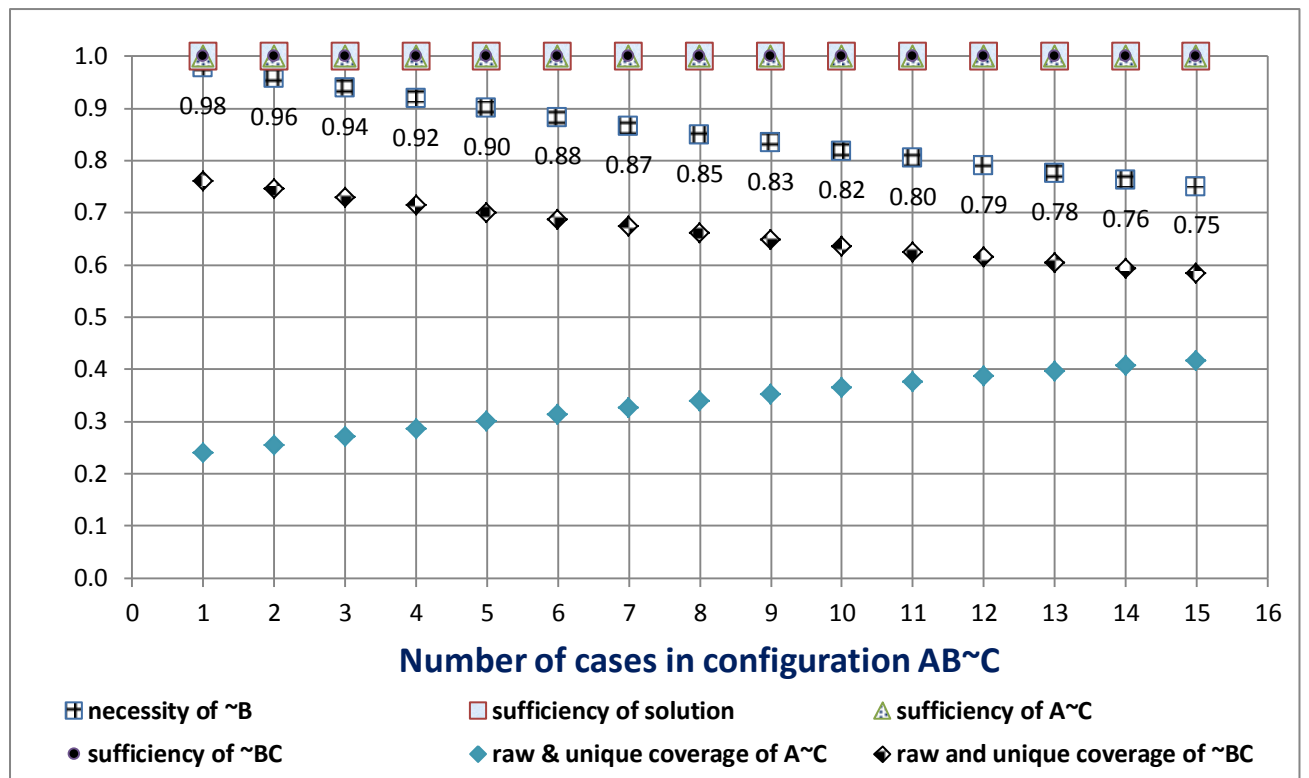**Fig. 5** Solution is A~C+~BC: Effect of changing the number of cases in AB~C (with 100% achieving the outcome)

We have shown by first considering hypothetical examples in sections 2 and 3 and then discussing Schneider and Wagemann's (2012) example that analytic work using QCA must pay more attention to the weights of cases in the rows of truth tables than has been, to now, usual. This is particularly important in the case of causally oriented analyses of necessity but, in addition, when we cannot be sure we have a full and correctly specified model of the causal structure we are exploring, the relative weights may also, as we have shown earlier,

---

[13] Note that the coverage index for the necessity of ~B (i.e. the extent to which ~B is sufficient for the outcome) remains at 0.978 (=45/46) for every point on both graphs. This is because the cases in row 7 are irrelevant to its calculation.

adversely affect analyses of sufficiency. Schneider and Wagemann (2012), addressing situations where the analysis of necessary conditions can go awry, have argued that the inconsistency in a row of a truth table is a crucial factor. We have argued, by generalising their example, that the crucial factor is actually not this inconsistency but the case weights characterising a truth table. Their problem of a hidden quasi-necessary condition both can arise in the absence of any such inconsistency and also fail to appear in the face of one.

Schneider and Wagemann, arguing that this problem is due to the inconsistent row, propose the "imperfect remedy" of increasing thresholds for consistency. They point out that had a stricter threshold for sufficiency of 1.0 been used to analyse Table 7, then the inconsistent row AB~C would not have been allowed into the minimisation process and the "necessary condition ~B would not thus have been logically minimised away". However, as they note, were the same threshold used for testing necessity, ~B would no longer be a necessary condition. We have shown that this problem is a more difficult one than they suggest, having its roots not so much in inconsistent rows as in the distribution of cases across rows (even where there are no inconsistencies at all). For this reason, even in truth tables with no inconsistent rows, the problem needs to be at the forefront of a QCA-user's mind. We therefore need more discussion of how to deal with this problem.

5. What to do about this problem?

We have considered the difficulties that can arise if the issue of case weights is not taken seriously. In the rest of the paper we will discuss what we might do to alleviate the problematic effects of varying weights across populations. We will focus, given the space available, on necessity, and consider it in the simplest context of having full knowledge of the causal structure[14]. We start by considering the usefulness here of measures of empirical importance such as coverage and relevance. We then develop an alternative approach that we first suggested in Cooper et al. (2014).

5.1 Measures of trivialness and relevance: a way forward?

Can the work that has been done on measures of coverage, triviality and relevance (Goertz 2006, Ragin 2006, Schneider and Wagemann 2012) offer us a way to address the issues raised earlier? To explore this question, we will employ a thought experiment involving two populations sharing the same causal structure. Consider Table 8. This, we will assume for the moment, is the full data from a small population on the relation between nourishment in pregnancy and the health of babies. We will also assume that there is an underlying causal

---

[14] In the case of sufficiency, our earlier discussion has shown the importance of having correct knowledge of causes.

relation between maternal nourishment and a baby's health. In this population the majority of mothers (69.94%) are adequately nourished (our X). Most babies (71.68%) are healthy (our Y).

Table 8: Society D

	Unhealthy baby (~Y)	Healthy baby (Y)	Totals
Adequately nourished mother (X)	25	3000	3025
Not adequately nourished mother (~X)	1200	100	1300
Totals	1225	3100	4325

Having an adequately nourished mother is close to being perfectly sufficient for the subsequent health of the baby. In addition, when the mother is not adequately nourished, over 90% of babies are unhealthy. Maternal nourishment clearly matters. It is also clear that having an adequately nourished mother is a quasi-necessary condition for being a healthy baby (consistency of 0.968, coverage of 0.992). Adequate nourishment is both quasi-necessary and quasi-sufficient for a healthy baby. Now we consider how some of the measures of the trivialness and/or relevance of a necessary condition behave with these data.

Goertz (2006) argues that a necessary condition becomes more relevant as it becomes more sufficient. Adequate nourishment passes this test, and, correspondingly, Ragin's (2006) coverage measure is high. Goertz also argues that the number of cases in the ~X,~Y cell allows us to assess the trivialness of a necessary condition. Zero cases in this cell indicate trivialness. We have 1200 cases here, far from zero. This is certainly large enough for us to be able to check, for example, that ~X tends to imply ~Y which we want to see when X tends to being necessary for Y.

Schneider and Wagemann (2012: 233-237) argue that X and Y being nearly constants is another basis for a necessary condition being trivial. They offer a formula for assessing trivialness and relevance (on a single scale) that addresses this problem, claiming (page 237) that this parameter: "… can be deemed a valid assessment of the relevance of a necessary condition. Low values indicate trivialness and high values relevance."[15] Applying the "Schneider-Wagemann formula[16] for relevance", where low scores on a 0-1 scale indicate trivialness, and high scores relevance, then we obtain for the necessary condition "adequate

---

[15] The treatment of trivialness / relevance as a single scale (where relevance indicates non-trivialness) differentiates their approach from that of Goertz. For him, relevance parallels Ragin's coverage, i.e. the degree to which a necessary condition is also sufficient. See footnote 12 on page 237 of their book.

[16] This is SUM(1-X_i) / SUM(1-MIN(X_i, Y_i)). The key point here is that cases in the X,Y cell of a 2x2 table contribute nothing to the summation. Schneider and Wagemann (2012) argue that their formula picks up a source of trivialness missed by Ragin's coverage measure.

nourishment" a score of 0.981. Adequate nourishment is, on their measure, a relevant necessary condition. All seems well so far, with the various parameters lining up.

But now consider our focus - weighting. Consider Table 9, taken from a second society, where just the weight of cases in the ~X row happens to be different. The consistency with quasi-sufficiency of X and ~X for Y are unchanged. Consider the rows. Having an adequately nourished mother is again close to being perfectly sufficient for the subsequent health of the baby. Again, when the mother is not adequately nourished, over 90% of babies are unhealthy. Having an adequately nourished mother is a necessary condition for being a healthy baby (consistency of almost 1, coverage of 0.992). Again, adequate nourishment is quasi-necessary and quasi-sufficient for a healthy baby.

Table 9: Society E

	Unhealthy baby (~Y)	Healthy baby (Y)	Totals
Adequately nourished mother (X)	25	3000	3025
Not adequately nourished mother (~X)	12	1	13
Totals	37	3001	3038

We don't now have many cases in the ~X~Y cell, but enough, we think, to check that ~X tends to imply ~Y[17]. By this test, and coverage, the necessary condition X remains non-trivial and relevant. However, in Schneider and Wagemann's terms, both X and Y are nearly constants – the cases in this cell dominate - and, if we turn to their trivialness/relevance formula, then we get for the quasi-necessary "adequate nourishment" a low score of 0.342. Adequate nourishment, on their measure, is a trivial necessary condition. We can see that their formula is sensitive to the weights of the X and ~X rows, when nothing else, and certainly not the causal consequences of X and ~X, is changed. Table 10 summarises several measures for the two societies.

Table 10: Various indices for societies D and E.

Indices:	Society D	Society E
Consistency with necessity of X	0.968	0.9997
Coverage (re necessity) of X	0.992	0.992
S-W relevance measure for X	0.981	0.342
Consistency with sufficiency of X	0.992	0.992
Consistency with sufficiency of ~X	0.077	0.077

---

[17] Since, comparing Table 8 and Table 9, the 12 in the ~X,~Y cell in Table 9 forms a smaller proportion of the total cases, we might *perhaps* argue that the problem of a potentially trivial necessary condition X is greater for this second population – in the sense that this cell is nearer the zero that would indicate trivialness for Goertz (2006).

There is a small difference in the necessity of X for Y, but on Ragin's measure of importance (coverage, or the sufficiency of X for Y), the necessary condition is equally important in the two populations. The crucial difference is in the result from Schneider and Wagemann's formula. What should we make of this? Can it help us, for example, advise a policy maker? Two possible views follow.

1. A counterfactual thought experiment can be undertaken in which we shift the cases from the ~X row into the X row and then, assuming all else remains equal[18], allocate the health outcomes to these new cases of X that the original cases of X have. The point of this is to simulate what would happen if we were able, ceteris paribus, to remove maternal malnourishment from the population. Doing this for Table 8, we produce a reduction in the number of unhealthy babies from 1225 to 36 (a 97.08% reduction). For Table 9, the corresponding result is a 32.14% reduction[19], indicating a sense in which maternal nutrition is a less important factor in society E than D. This experiment suggests that the S-W formula can help us.

2. There is, however, an alternative view. The results of the thought experiment required us to have two populations to compare. It is easy to imagine a situation in which we only have an analysis of one population. If this had happened to be the population in Table 9 – and this seems just as likely a scenario as that in Table 8 – then the S-W formula would have reported that X is a trivial, not relevant necessary condition. In spite of this, in this population, all else being equal, changing the nutritional status of the mothers in the ~X row would produce a substantial reduction in the numbers of unhealthy babies. Twelve[20] of the twelve unhealthy babies in the row ~X would become healthy. The overall number of unhealthy babies would see a 32.14% reduction. In these senses, it isn't a trivial condition.

Both populations are possible scenarios. In each the generative causal structure is the same (i.e. X and ~X generate the same proportions of the two possible outcomes). Only the relative weights of X and ~X cases differ. The reason, of course, that in our first thought experiment we saw a greater reduction in unhealthy babies was simply because, while the generative causal structure remained the same, proportionally more of these appeared in the ~X row[21].

What conclusion might we draw from this section? One possible lesson might be that it is often worth introducing some policy-oriented thought experiments into the discussion of such

---

[18] This may not be the case in practice.
[19] This result is not sensitive to rescaling the tables so that they have the same total population.
[20] More precisely, 11.89.
[21] The coverage of ~X for ~Y is also much higher for society D than for society E: 0.980 versus 0.324. In Ragin's terms ~X has more empirical importance for ill health in *society* D than E, even though the (probabilistic) causal consequences of ~X *for any individual* remain the same.

matters as trivialness and relevance. Much discussion of these matters rightly focuses on the relative sizes of the sets X and Y, and related matters, and attempts to develop generally applicable formulae. But we know from other summarising measures that it is important to bear in mind the details of any particular case. The arithmetic average, for example, is a useful measure, *as long as* the distribution being summarised isn't bimodal, doesn't have many outliers, etc. In the same way as the latter features can cause problems for the interpretation of an average, we think that the weights of the rows of truth tables can cause problems when interpreting the meaning and implications of measures used in QCA. We must be careful when employing formulaic approaches.

5.2 Another approach

With that advice in mind, we will indicate one way of thinking about the problem that can keep things clear in one's mind when undertaking the analysis of truth tables. To illustrate this, we will return to Table 7. The suggested approach involves regarding each row as evidence for the sufficiency or otherwise of the configuration it represents, while ignoring at this stage finer details concerning the degree of consistency with sufficiency (or necessity)[22]. Assume, for the sake of argument, that the evidence for each configuration in Table 7 is considered good enough to treat it as warrant for the eight claims in Table 11.

Table 11: Eight statements concerning sufficiency (idealised version of Table 7)

Configurations with the outcome Y	Configurations with the outcome ~Y
AB~C => Y	ABC => ~Y
~A~BC => Y	~A~B~C => ~Y
A~BC => Y	~AB~C => ~Y
A~B~C => Y	~ABC => ~Y

Looking at the truth table in this way, we see that ~B is not strictly a necessary condition for Y, since AB~C=>Y. If we were to run a QCA on the *temporary* assumption that there is just one case per row, we would obtain a consistency with necessity for ~B for Y of 0.75, reflecting the fact it appears in three of the four rows to the left of Table 11. It can readily be seen, however, that, the higher the number of cases for the three lower rows to the left of Table 11 (~A~BC, A~BC, A~B~C) in relation to those for the row AB~C, the higher will be the reported consistency of necessity of ~B for Y, and the more serious the problem of this quasi-necessary condition being hidden in the solution for sufficiency will appear to be. If

---

[22] Ragin's (2008) truth table algorithm behaves similarly, in the analysis of sufficiency, in the way it minimises a truth table.

we allocate three cases each to the three lower rows on the left-hand side, leaving the row AB~C with just one, then consistency with necessity for ~B for Y is 0.90. If, however, we allocate 10 cases to AB~C, while leaving the lower three rows with just one case each, then the consistency with necessity for ~B for Y falls to 0.23.

We can apply this same approach to the simple dataset from Society D and E. In both cases we can argue, let's assume, that we have good enough evidence for the two (quasi-sufficiency) claims:

1. Adequately nourished mother => Healthy baby
2. NOT (Adequately nourished mother) => NOT (healthy baby)

The first of these is equivalent to the left-hand side of Table 11; the second to the right-hand side. If we were to pretend for a moment that these are both sufficient rather than (strong) quasi-sufficient claims, we could deduce from (2) the necessity of an "adequately nourished mother" for a healthy baby. This would enable us to say, having temporarily ignored the differing weights of X and ~X, that, *for both societies D and E*, our X is sufficient and necessary for Y (or, more strictly, quasi- in both cases). From this perspective, the weights of the cases in the rows of D and E make no difference to this conclusion. We know, however, that when these sufficiency relations are expressed in a particular simulated social setting characterised by some particular number of cases in the X and ~X rows, the quasi-necessity of X is, in some sense, of a different sort in D to that in E, in so far as getting rid of malnourishment would remove a larger percentage of unhealthy babies in D than in E.

5.3 A dilemma and a suggested response

Where does this leave us? We have shown, by analysing several simulated datasets, that the weights of cases in the rows of a truth table will affect QCA results concerning the consistency with necessity and, in some situations, sufficiency of a factor.

Should we worry about the fact that assessments of necessity, for example, vary as the weights of cases in rows of truth tables vary? We suspect that the position a scholar will take here will reflect their philosophical position. Someone of an empiricist frame of mind might want to argue that the assessment of necessity should accurately reflect just the observed patterns of data in whatever society is being studied. If the observed world, with our assumed causal structure in Section 2 (resulting in implications 1-4), looks like our dataset 2, and if we only enter S into our model, then S is quasi-sufficient and quasi-necessary for O, but if the world looks like our dataset 1 then it is only quasi-sufficient.

On the other hand, a realist (e.g. Bhaskar 1975, 1979, Pawson 1989) might argue that if we have adequate evidence from the study of the four types of case in datasets 1 and 2 in section 2 such that we can complete the 1 case per row table shown here as Table 12, and, better still, also have an understanding of plausible causal mechanisms that suggests T is causally relevant and S not, then, irrespective of the weights of cases[23] we can say that in both worlds observed via these datasets, T is sufficient and necessary for O, and S is not necessary for O. For the realist, we think, this would provide a single right answer.

Table 12: Four statements concerning sufficiency (from section 2)

Configurations with the outcome O	Configurations with the outcome ~O
ST => O	S~T => ~O
~ST => O	~S~T => ~O

As we saw in the previous section, we could also, in a similar way, argue for a single correct answer in the analysis of the societies D and E. However, this single answer doesn't capture the important "empirical" difference we saw between these societies when we carried out the thought experiment of removing malnourishment in mothers. We suggest the following resolution of this dilemma.

When analysing necessary conditions within a strictly causal perspective, it is informative to perform our one-case-per-row exercise. There seems to be a sense in which the answer this gives will be the right one, especially for a realist. If, in *both* societies D and E, there are causal processes that lead to (i) (nearly) all babies of malnourished mothers being unhealthy and (ii) (nearly) all babies of adequately nourished mothers being healthy, then it would seem right to privilege these two sufficiency claims (which parallel and actually arise from generative causal processes) - PLUS the implication of (i) that adequate maternal nourishment is also necessary for a baby's health - over any formulaic results such as those deriving from the S-W relevance formula (which showed adequate nourishment to be a relevant necessary condition in one of these societies but not in the other).

On the other hand, the S-W formula did capture an important difference of relevance to policy makers in the two societies. Similarly, Ragin's coverage measure for necessity also captures something crucial for policy makers about the empirical importance of nourishment in both societies (that it is not only necessary but also sufficient for a healthy outcome). We therefore also suggest a second stage in which the nature of necessary conditions that have passed the test of the one-case-per-row exercise are explored by means of the various approaches developed by Ragin, Goertz and Schneider and Wagemann but also by means of

---

[23] Beyond there being enough to warrant the causal claims.

various thought experiments of the sort we have used here[24]. In addition, wherever possible, reference to plausible causal mechanisms is also likely to be helpful when judging what the necessity of a condition actually comprises.

Ideally, of course, any scholar employing a Boolean approach will aim to establish a disjunction of configurations jointly sufficient and necessary for an outcome (Baumgartner, 2014). To achieve this s/he will need, as well as eliminating redundant factors, to avoid omitting key explanatory factors from analytic models. In the latter ideal scenario, the issues we have raised become less important, though the assessment of the necessity of individual conditions will still vary with weights. Knowledge of plausible generative mechanisms, whether derived from theory or from within-case studies, will be useful in approaching this goal. However, most scholars will not achieve this degree of perfect explanatory knowledge and therefore need, we believe, to take account of the issues we have discussed.

We have argued before (Cooper and Glaesser, 2011, Cooper et al. 2014) that there is more complexity in QCA than sometimes first meets the eye. Case weights add another layer of complexity to those we have previously discussed. Our view, confirmed by the exercise we have undertaken here, is that QCA users should always try to see beneath and around any formulaic approaches in the set theoretic field. We hope our discussion here, and any further work it encourages, will support such an approach.

Acknowledgements

References
Baumgartner, M.: Regularity theories reassessed. Philosophia 36, 327–354 (2008)
Baumgartner, M.: Detecting causal chains in small-n data. Field Methods 25, 3-24 (2013)
Baumgartner, M.: Parsimony and causality. Quality & Quantity (2014). doi: 10.1007/s11135-014-0026-7.
Baumgartner, M., Epple, R.: A coincidence analysis of a causal chain: The Swiss minaret vote. Sociological Methods & Research 43, 280-312 (2014)
Bhaskar, R.: A realist theory of science. Harvester, Brighton (1975)
Bhaskar, R.: The possibility of naturalism. Harvester, Brighton (1979)
Cooper, B.: Applying Ragin's crisp and fuzzy set QCA to large datasets: Social class and educational achievement in the National Child Development Study. Sociological Research Online 10, 2 (2005). http://www.socresonline.org.uk/10/2/cooper.html.
Cooper, B., Glaesser, J.: Paradoxes and pitfalls in using fuzzy set QCA: Illustrations from a critical review of a study of educational inequality. Sociological Research Online 16, 3 (2011). http://www.socresonline.org.uk/16/3/8.html

---

[24] Baumgartner, in his very important work on coincidence analysis (CNA) – a configurational method that uses an alternative minimisation algorithm to that employed in QCA – tends in his more fundamental papers effectively to assume one case per row (e.g. Baumgartner 2013, 2014). However, in his application, with Epple, of CNA to the case of the Swiss minarets referendum (Baumgartner and Epple 2014) this approach is combined with one that takes account of case numbers when calculating coverage, etc.

Cooper, B., Glaesser, J.: Qualitative work and the testing and development of theory: Lessons from a study combining cross-case and within-case analysis via Ragin's QCA. Forum: Qualitative Social Research / Qualitative Sozialforschung. 13, 2, Art. 4 (2012). http://www.qualitative-research.net/index.php/fqs/article/view/1776

Cooper, B., Glaesser, J., Thomson, S.: Schneider and Wagemann's proposed enhanced standard analysis for Ragin's Qualitative Comparative Analysis: Some unresolved problems and some suggestions for addressing them. COMPASSS WP Series 2014-77 (2014). http://www.compasss.org/wpseries/CooperGlaesserThomson2014.pdf

Duşa, A., Thiem, A.: QCA: Qualitative Comparative Analysis. R package version 1.0-5 (2013)

Gerring, J.: Social Science Methodology: A unified framework. Cambridge University Press, Cambridge, 2nd Edition (2012)

Glaesser, J., Cooper, B.: Gender, parental education, and ability: their interacting roles in predicting GCSE success. Cambridge Journal of Education 42, 4, 463–480 (2012)

Goertz, G.: Assessing the trivialness, relevance, and relative importance of necessary or sufficient conditions in social science. Studies in Comparative International Development 41, 88-109 (2006)

Pawson, R.: A measure for measures: A manifesto for empirical sociology. Routledge & Kegan Paul, London (1989)

Ragin, C.C.: The Comparative Method. Moving beyond Qualitative and Quantitative Strategies. University of California Press, Berkeley (1987)

Ragin, C.C.: Fuzzy-Set Social Science. University of Chicago Press, Chicago (2000)

Ragin, C.C.: Set relations in social research: Evaluating their consistency and coverage. Political Analysis14, 3, 291-310 (2006)

Ragin, C.C. Redesigning social inquiry: Fuzzy sets and beyond. University of Chicago Press, Chicago (2008)

Schneider, C.Q., Wagemann, C.: Set-theoretic methods for the social sciences. A guide to Qualitative Comparative Analysis. Cambridge University Press, Cambridge (2012)

Thiem, A., Duşa, A.: QCA: A package for Qualitative Comparative Analysis. The R Journal 5, 1, 87-97 (2013)