

# MODELING CONFIGURATIONAL EXPLANATIONS

Alessia Damonte

Dept of Social and Political Sciences  
Università degli Studi di Milano  
[alessia.damonte@unimi.it](mailto:alessia.damonte@unimi.it)

-- pre-print version of <https://doi.org/10.1017/ipo.2021.2> --

## Abstract

How can Qualitative Comparative Analysis contribute to causal knowledge? The article's answer builds on the shift from design to models that the Structural Causal Model framework has compelled in the probabilistic analysis of causation. From this viewpoint, models refine the claim that a 'treatment' has causal relevance as they specify the 'covariates' that make some units responsive. The article shows how QCA can establish configurational models of plausible 'covariates'. It explicates the rationale, operations, and criteria that confer explanatory import to configurational models, then illustrates how the basic structures of the SCM can widen the interpretability of configurational solutions and deepen the dialogue among techniques.

## Keywords

Explanation; *in*us causation; mediation; pruning; Qualitative Comparative Analysis; quasi-experimental designs; Structural Causal Model framework

## INTRODUCTION

The explanatory standing of Standard Qualitative Comparative Analysis (hereafter, QCA: Ragin, 2014, 2000, 2008; Rihoux and Ragin, 2009) seldom proves plain. Recently, Møller and Skaanig (2019: 81) maintained that QCA and set-theoretic methods ‘provide little in the way of distinguishing between logical and causal relationships’. Munck (2016: 777) clarified that ‘causation is not a logical relation but, rather, a relation between events or, more precisely, between changes in the properties of things’ and, as such, ‘must be understood ontologically’ – which he deemed beyond QCA. Although ‘causal theories should be built with due attention to the rules of logic, both qualitative and quantitative researchers are better off presenting their causal arguments in the general form  $\Delta X \rightarrow \Delta Y$ , meaning that a change in property  $X$  causes a change in property  $Y$ , and treating the covariation between  $X$  and  $Y$  as an essential means for checking whether their causal arguments are true’. In the social and political domains, convincing covariational models identify mediation effects (e.g. Seawright, 2019: 33) – which, once again, QCA would be unable to render.

Against this backdrop, the article addresses the question of whether Standard QCA as a technique is ill-equipped to contribute to causal knowledge. It considers the specificity of its analytic strategy in three sections. The first summarizes the state of the art in causal analysis, and brings attention to the move from design to models that the Structural Causal Model framework has developed to improve the counterfactual approach. In light of that move, the second section portrays QCA as a technique equipped for specifying models with credible explanatory import, with special diagnostics for omitted factors, irrelevant additions, and confirmation bias. The third section illustrates how the causal structures of the Structural Causal Model framework can illuminate the relationship between explanatory configurational findings. The conclusions highlight the special contribution that the explanatory usage of QCA can bring to the methodological agenda on causation in the social and political sciences.

## THE ISSUES WITH CAUSATION AS THE RESPONSE TO A STIMULUS

The Potential Outcome framework (hereafter, PO: Rubin, 1974, 1978; Rosenbaum and Rubin, 1983; Imbens, 2004; Morgan and Winship, 2015) provides the default choice for analyzing causal relationships. The framework narrows on a stimulus  $T$  with two mutually exclusive states, either realized (1) or not (0).  $T$  is recognized a causal effect when the claim stands that any  $i$ -th

unit under the realized stimulus  $T(1)$  returns a response  $Y_i(1)$  that would not have been yielded under the unrealized stimulus  $T(0)$ . The claim that the difference between the actual response  $Y_i(1)$  and its counterfactual  $Y_i(0)$  gauges the causal effect lies on a design that ensures statistically comparable samples of units except for the state of  $T$ . The design circumvents a reality in which each unit instead comes as a specific bundle of covariates  $u_i$  that blur comparability and make the calculation of the effect inaccurate.

Rubin's canon conventionally conceives of covariates as features that are exogenous to the stimulus–response relationship, yet bias the 'unit's propensity to receive'  $T(1)$  through some 'mechanism of self-selection' that the analysis holds as unknown. The experimental setting, the argument goes, rules out this bias by randomly exposing the units to the realized or the unrealized stimulus. Together, randomization and forced exposition defuse the self-selection mechanism and license two key tenets about the average net effect: first, the covariates do not confound it; second, the assignment mechanism bias is (strongly) ignorable. Beyond randomization, the canon maintains that unconfoundedness and strong ignorability can still be ensured by discounting the units' heterogeneity through their propensity score. Thus, estimations of the average treatment effect become valid on units that display a comparable propensity to self-select into the treatment.

Despite the efforts to make propensity scores as little 'model-dependent' as possible, the strategy adopted to deal with units' natural heterogeneity decides the credibility of the causal claim in observational studies and raises a modeling issue. In their structural rendering of the potential outcome, Winship and Morgan (1999: 668 ff) define the issue by decomposing the error term  $u_i$  of the individual response  $Y_i$  to the stimulus  $T_i$  into  $u_i = u_i^{T_0} + T_i(u_i^{T_1} - u_i^{T_0})$ . Their decomposition emphasizes that the set of covariates  $u_i^{T_1}$ , featuring the group under  $T_i = T(1)$ , is possibly different from that of the baseline group  $u_i^{T_0}$ . Moreover, they consider that the value of the assignment mechanism  $T_i$  ultimately depends on two sources of self-selection bias: the set  $Z_i$  of observed exogenous variables, and the set  $V_i$  of unobserved or missing variables. When the 'selection on the observables' occurs, the propensity score  $Pr(Z_i)$  can license the assumptions of unconfoundedness and ignorability of the assignment. Under the 'selection on the unobservables', instead, estimates call for further assumptions about the shape of the unobservables and, eventually, lose interpretability.

The key issue, in short, remains how to ensure that the relevant covariates are identified so that the unobservables only contain irrelevant 'noise'. To address it, the Structural Causal Model framework (hereafter, SCM: Pearl and Paz, 1987; Pearl and Verma, 1991; Pearl, 2009; Bareinboim and Pearl, 2016; Pearl and Mackenzie, 2018) invites researchers to shape the setting of counterfactuals

along the lines of Simon (1977).

While working on the definition of political power, Simon noted that conventional representations *via* single equations could not render the intuition of a relationship between two agents, *A* and *B*, such that *A*'s preferences *cause* *B*'s choices while the converse does not hold. To him, power relationships unveiled the intrinsic asymmetry of any causal relationship and called for a decomposition. 'To say that *X* is a cause of *Y* is to say that there is a certain order in which the equations must be solved – specifically, that we must first solve for *X* and insert its value in another equation which we then solve for *Y*. Correspondingly, to say that *P* causes *Q* is to say that we have a set of propositions (Boolean equations) such that we first determine the truth or falsity of *P* from some subset of these, and then use the truth value of *P* to determine the truth value of *Q*' (*ivi*: 50, notation adapted).

The SCM trades design for models that render structural assumptions of dependence. Any structural model consists of three sets – the exogenous variables  $\mathcal{U}$ , collecting the covariates; the endogenous variables  $\mathbf{V}: X, Y, Z, \dots$ , selected from  $\mathcal{U}$  as the causal model of the response; and the functions  $\mathbf{F}: \{f_{XZ}, f_{XY}, f_{YZ}, \dots\}$  that connect them. The model is then 'augmented' by a causal graph  $\mathcal{G}_M$  in which variables are 'nodes', functions are (missing) directed arrows or 'edges', and any consecutive edges build a 'path' between nodes. The outcome becomes the end-point of at least one path whose structure conveys the mechanism of data-generation – the 'mechanism' being the type of process to the outcome that the equation captures in some aspect of interest (Simon, 1977: 115). Then, the SCM framework identifies three ideal graphs to which any causal structures can be reduced – namely, the linear chain  $\mathcal{G}_l$ , the fork  $\mathcal{G}_f$ , and the collider  $\mathcal{G}_c$ .

The linear graph  $\mathcal{G}_l$  entails a model  $\mathcal{M}_l$  with four key relationships: of dependence of *Y* from *Z*; of dependence of *Z* from *X*; of 'likely' dependence of *Y* from *X* when the connecting functions grant transitivity; and, more important, of independence of *Y* on *X* conditional on *Z*.

$$\mathcal{G}_l := X \xrightarrow{f_{ZX}} Z \xrightarrow{f_{YZ}} Y$$

$$\mathcal{M}_l := \begin{cases} P(y|z) \neq P(y) \\ P(z|x) \neq P(z) \\ P(y|x) \neq P(y) \\ P(y|x, z) = P(y|z) \end{cases}$$

The relationships in  $\mathcal{M}_l$  convey that two nodes arranged in a linear chain are

independent conditional on the mediator in between that captures the whole of the relevant variation. Effective mediators, then, support the claim of a causal connection between  $X$  and  $Y$  by dissolving it.

The second prototypical shape is the fork. Two variables,  $X$  and  $Y$ , both descend from a third variable,  $Z$  – as in  $\mathcal{G}_f$ . In the corresponding model  $\mathcal{M}_f$ ,  $Z, Y$  are dependent;  $Z, X$  are dependent;  $Y, X$  are likely dependent and again become independent when conditioned on  $Z$ .

$$\mathcal{G}_f := X \xleftarrow{f_{XZ}} Z \xrightarrow{f_{YZ}} Y$$

$$\mathcal{M}_f := \begin{cases} P(y|z) \neq P(y) \\ P(x|z) \neq P(x) \\ P(y|x) \neq P(y) \\ P(y|x, z) = P(y|z) \end{cases}$$

The fork renders the confounder that makes the dependence of  $Y$  on  $X$  spurious and, again, explains it away. It shares all the features of the chain except the dependence between  $X$  and  $Z$ , which in  $\mathcal{G}_f$  runs in the opposite direction.

The last fundamental shape is the collider as in graph  $\mathcal{G}_c$ . Here,  $X$  and  $Y$  together determine the values of  $Z$ . Thus, the model  $\mathcal{M}_c$  renders that  $X, Z$  are dependent;  $Y, Z$  are dependent;  $X, Y$  are independent but display a dependence when conditioned on  $Z$ .

$$\mathcal{G}_c := X \xrightarrow{f_{ZX}} Z \xleftarrow{f_{ZY}} Y$$

$$\mathcal{M}_c := \begin{cases} P(z|x) \neq P(z) \\ P(z|y) \neq P(z) \\ P(x|y) = P(x) \\ P(x|y, z) \neq P(x|z) \end{cases}$$

The collider portrays the situation in which the output has two inputs. The shape makes knowing the value of the one input irrelevant when we know the value of the other and the output. Beyond that, a collider does not establish any backward causation – instead, it provides a further reason to handle the relationship of correlation and causation with caution.

In SCM, the three basic shapes with their sets of conditionalities offer criteria to accept or reject hypotheses about the relevant covariates beneath an

outcome. The acceptable hypothesis may not be univocal, though, as graphs with the same edges and '*v-structures*' arise the same testable implications, hence belong to the same equivalence class. Mostly, they offer 'inference rules for deducing which propositions are relevant to each other, given a certain state of knowledge', and conceptual tools to develop models where 'knowing *z* renders *x* irrelevant to *y*' (Pearl and Paz, 1987). Against this backdrop, the covariates are relevant that turn dependence into independence and *vice versa* (Kuroki and Pearl, 2014).

## MODELING COMPLEX CONDITIONALITY

Explanatory QCA, too, is interested in the bundle of relevant factors that, together, make the 'stimulus' effective and, hence, account for its relationship with an outcome. The viewpoint is slightly different, however.

The strategy follows Mackie (1965, 1980) in observing that, although we usually explain the burning of a house by a short-circuit, this explanation is 'gappy' unless we bring specific background features to the fore. The short-circuit initiated the fire *because*, for instance, it fired a spark on an oily rag, *and* there was enough oxygen in the room, *and* the sprinklers were broken. Thus, the causal response only takes place under the right conditions beyond the initiating factor. Moreover, the conditions and the 'initiating' factor stand on an equal footing. 'Cause' is just a conventional label for the anomaly in the field that attracts our attention: had the fire followed from a gas leak, we would have mentioned it instead of the sparkle. Besides, different particular conditions can unleash the same type of outcome in equivalent situations: a burning match *or* enough pressure would have done the same job as the spark. If we abstract the details away, eventually fire can be reduced to the consequence of 'heating', *and* 'combustible', *and* 'oxygen' being in the same place at the same time under 'no impediments'. The abridged formula pinpoints the types of elements that, together, sort the same type of effect irrespective of any other features of the context. Hence, such a formula is complete enough to travel and, once that the labels are properly assigned to actual things and events, specific enough to allow us to expect and explain the outcome at any timepoint and place.

The hallmark of this perspective lies in the relationship that the elements entertain with each other and their consequent. That any of them is in the right state is *insufficient* to the outcome: the causal power lies in their compound. Each of them, however, is *non-dispensable*: a component in the wrong state makes the compound fail. In turn, the compound is *sufficient* for the outcome, but its occurrence is an *unnecessary* event that only happens when all the components are properly arranged. In Mackie's terms, then,

each component is a partial, *inus* cause – an *insufficient* yet *necessary* part of an *unnecessary* yet *sufficient* compound.

Cartwright (2017) enhances the understanding with a functional account that makes further sense of the compound. She assumes the entities in the world ground capacities to do some job, such as, operate a change or preserve stability. Capacity, then, is a concept close to those of ‘potential’, ‘disposition’, ‘tendency’ in other accounts, and related to that of ‘causal power’. It entails a productivity that relationships of sheer association do not display: it is only at the intersection of the right capacities that something happens. The compound of these right capacities arises a ‘nomological machine’ – a ‘sufficiently stable arrangement’ (Cartwright, 2017) that makes a realization certain until all the relevant elements remain in the right state. Working nomological machines make sure that anything else is irrelevant before the same type of outcome across time and space. Neurons, engines, scientific labs are illustrations of these arrangements. Their relevant components can be understood as *inus* factors that, in the right team, perform basic tasks such as triggering ( $T = \{T_1, T_2, \dots\}$ ), enabling ( $E = \{E_1, E_2, \dots\}$ ), and shielding ( $S = \{S_1, S_2, \dots\}$ ) the special process to a state of an outcome.

Thus, a basic *inus* model can read  $T \cap E \cap S \rightarrow Y$  ( $\cap$  reading ‘and’,  $\rightarrow$  ‘is sufficient to’). With binary variables, the probabilistic illustration of this model would portray head (1) and tail (0) flips from three coins ( $T, E, S$ ), and an unmodeled endogenous data-generation process ensuring a bell ( $Y$ ) always rings (1) when all the three coins land on head, else leaving it silent (0). This model arises a finite sample space with eight possible realizations,  $\Omega_{TES} = \{TES, TES\bar{S}, T\bar{E}S, T\bar{E}\bar{S}, \bar{T}ES, \bar{T}E\bar{S}, \bar{T}\bar{E}S, \bar{T}\bar{E}\bar{S}\}$ . Of them, the first only,  $\omega_1 := TES$ , yields  $Y$  given the mechanism in place. QCA dubs the sample space ‘truth table’ and the realizations ‘primitive configurations’, but still understands them as possible states of the world and sets or partitions of the universe of reference  $\mathcal{U} = \{u_1, \dots, u_n\}$ .

The peculiarity of QCA lies in its eliminative nature. Its solutions arise from the dismissal of logically *irrelevant* conditions from sufficient primitives (Ragin 2014: 125 ff; cfr. Thiem, 2019). In itself, hence, pruning cannot determine causality. However, its protocol can be geared to probing the claim that a selection of conditions is a credible *inus* machine (Mackie, 1980: Appendix) when

- a) the same bundle of conditions accounts for both states of the outcome without contradictions;
- b) the solutions of each outcome are properly specified to their respective subpopulation.

**Requirement (a): establishing sufficiency for explanatory purposes**

As a technique, QCA relies on an algebra of set that preserves the equivalence to a first-order logic (Stone, 1936). Thus, the technique addresses Simon’s knowledge problem – the determination of the truth or falsity of  $P$  – as the problem of gauging a unit’s membership in the set of things that are or have  $P$  (e.g. Sartori, 1984; Ragin, 2008; Goertz and Mahoney, 2012; Ragin and Fiss, 2017).

QCA affords two such gauges: crisp and fuzzy. The crisp gauge assigns binary truth values in line with the Boolean canon: the 0-membership in  $P$  means 1-membership in the negated set  $\bar{P}$ . The fuzzy gauge refines the assignment with that which, from an explanatory perspective, can be understood as an ambiguity penalty. Fuzzy scores incorporate such a classification error in the value assigned to each unit. They span from 0.00 to 1.00 and have their point of highest ambiguity at 0.50: thus, units scoring 0.50 are instances of neither  $P$  nor  $\bar{P}$ ; units with extreme values, instead, are sure instances of either the set or its negation. The difference in gauging slightly changes the operationalization of the three logical axioms on which QCA can build valid causal claim.

**Table 1.** Axioms, set renderings, and gauges in QCA

Axiom	Set rendering	Crisp gauge for unit $u_i$	Fuzzy gauge for unit $u_i$
<i>Negation</i>	$\bar{P} := \mathcal{U} \setminus P$	$\psi_{\bar{P}_i}^{cs} = 1 - \psi_{P_i}^{cs}$	$\psi_{\bar{P}_i}^{fs} = 1 - \psi_{P_i}^{fs}$
<i>Non-contradiction</i>	$\bar{P} \cap P := \emptyset$	$\psi_{P_i \cap \bar{P}_i}^{cs} = \min(\psi_{P_i}^{cs}, \psi_{\bar{P}_i}^{cs}) = 0$	$\psi_{P_i \cap \bar{P}_i}^{fs} = \min(\psi_{P_i}^{fs}, \psi_{\bar{P}_i}^{fs}) < 0.5$
<i>Excluded middle</i>	$\bar{P} \cup P := \mathcal{U}$	$\psi_{\bar{P}_i \cup P_i}^{cs} = \max(\psi_{P_i}^{cs}, \psi_{\bar{P}_i}^{cs}) = 1$	$\psi_{\bar{P}_i \cup P_i}^{fs} = \max(\psi_{P_i}^{fs}, \psi_{\bar{P}_i}^{fs}) > 0.5$

**Notes:**  $\psi^{cs}$  is for the crisp-set membership score,  $\psi^{fs}$  is for the fuzzy-set membership score.  
 $\psi_{P_i}$  refers to membership score that the  $i$ -th unit from a universe  $\mathcal{U}$  of size  $N$  takes in the set of instances sharing the condition  $P$  in a state.  $\psi_{\bar{P}_i}$  is the membership score of the same unit in the set of the condition in the negated state.  
 The overbar reads ‘not’. The alternative notation in QCA is the curl  $\sim$ , or the lowercase set name.  
 The backlash  $\setminus$  indicates the set difference.  
 $\cup$  reads ‘union’ and corresponds to the logical inclusive (weak) ‘or’. The alternative logical notation is the vee  $\vee$ . In QCA’s applications, the plus sign (+) is common.  
 $\cap$  reads ‘intersection’ and corresponds to the logical ‘and’. The alternative logical notation is the wedge  $\wedge$ . In QCA’s applications, the operator is a dot ( $\cdot$ ) and often omitted.  
 $\emptyset$  indicates the empty set; the corresponding logical notation is the empty curly braces  $\{\}$ . QCA conventionally renders it with a 0, although 0 is also assigned to the observed instance that is ‘fully out’ to a set.

As summarized in Table 1, the three rules establish, respectively, the *negation* of a state as its arithmetic complement in the universe of reference; the rule of *non-contradiction* as the impossibility for the same unit to be in a state and its negation at the same time; the rule of the *excluded middle* as the necessity of a unit to display either one or the other of the two possible states.



These gauges allow enforcing the conventional understanding of sets as ‘conceptually uniform’ partitions of the universe with respect to a condition’s state. The construction of sets as uniform partitions provides the ground of the set-theoretical gauge of sufficiency. The relationship is, quite conventionally, a conditional. The parameter that captures it in any QCA application is the ‘consistency of sufficiency’ (*S.cons* for short: Ragin, 2000, 2008; Duşa, 2018 names it ‘inclusion’). The parameter is defined as the ratio of the size of the intersection of the outcome and a primitive, and the size of the primitive itself:

$$S.cons_{\omega_j \subset Y} := \frac{|\omega_j \cap Y|}{|\omega_j|}.$$

It closely recalls Kolmogorov’s gauge of conditional probability as the ratio of success to trials of a certain kind (e.g. Hájek, 2011). Indeed, without any loss, the  $S.cons_{\omega_j \subset Y}$  can be rewritten as  $\pi(Y|\omega_j)$ , being  $\pi$  the size of the sets. In the special case of crisp QCA, the size of the set is the number of its instances and Kolmogorov’s conditional probability coincides with the consistency of sufficiency. In both cases, moreover, the parameter renders the long-honored regularity criterion that ‘if  $\omega_j$  causes  $Y$ , then any instance of  $\omega_j$  is an instance of  $Y$ ’. At the same time, it accounts for the challenge that contradictory evidence rises to the modal understanding of causal regularity as ‘if  $\omega_j$  causes  $Y$ , it cannot be the case that an instance of  $\omega_j$  is an instance of  $\bar{Y}$ .’

The regularity claim of the *S.cons* stands when the parameter takes either its highest value of 1.00 or its lower value of 0.00, indicating that the primitive draws a uniform partition of the positive or the negative outcome set. For explanatory purposes, violations of this subset relationship make the primitive ‘contradictory’. Contradictions weaken the claim that the team of conditions renders an *inus* machine, and suggest that the team is ill specified – possibly, due to some omitted components. From a logical perspective, a contradiction makes the *inus* hypothesis ‘false’ as its realizations cannot establish the subpopulation of  $Y$ -instances as a separate set from that of  $\bar{Y}$ -instances.

The diagnosis of the contradiction through the *S.cons* is straightforward with crisp scores (Rihoux and De Meur, 2009) but can prove harder with fuzzy scores. The violation can be downgraded to an acceptable inconsistency when it arises from units for which  $Y_i \geq \omega_{ji} + 0.1$  (Ragin, 2000); truly contradictory instances instead remain those crisp ‘consistency outliers’ for which  $\omega_{ji} > 0.5$  while  $Y_i < 0.5$  (Rohlfing and Schneider, 2013; Rubinson, 2013; Rohlfing, 2020). The misalignment of the crisp and the fuzzy consistencies depends on the fuzzy scores leaving arithmetic residuals in intersections: to witness,  $\psi_P^{fs} = 0.8$  yields  $\psi_{P_i \cup \bar{P}_i}^{fs} = 0.2$ , which indicates an empty intersection by the rules in Table 1, yet inflates the diagnostic of the *S.cons*. To overcome the problem, a corrected version of the parameter has been devised. The Proportional

Reduction of Inconsistency (*PRI*), calculated as

$$PRI_{\omega_j \subset Y} := \frac{|\omega_j \cap Y| - |\omega_j \cap Y \cap \bar{Y}|}{|\omega_j| - |\omega_j \cap Y \cap \bar{Y}|},$$

explicitly borrows the rationale of the Proportional Reduction of Error. The gauge of inconsistency is  $|\omega_j \cap Y \cap \bar{Y}|$ , which assigns higher penalties to fuzzy values of the outcome closer to 0.5. The *PRI*, hence, displays a steep fall in its values, often for primitives with *S.cons* equal to 0.85 or lower. The convention maintains that a configuration is usually sufficient to an outcome when its *S.cons* is 0.85 or higher and supported by a similar *PRI* – although model specifications may suggest otherwise (Schneider and Wagemann, 2012).

### **Requirement (b): handling the overspecification of *inus* machines**

The additional relevant information in Standard QCA comes from the coverage of sufficiency, or *S.cov* for short, defined as

$$S.cov_{\omega_j \subset Y} := \frac{|\omega_j \cap Y|}{|Y|}.$$

The parameter indicates the empirical relevance of the realization  $\omega_j$  to the instances of the outcome *Y* (Ragin, 2006, 2008). It takes its highest value of 1.00 when the *j*-th realization is shared by all the instances of *Y* and tends toward 0.00 the more there are instances of the outcome set ( $\psi_{Y_i} > 0.50$ ) outside the configuration set ( $\psi_{\omega_{ji}} < 0.50$ ).

These uncovered instances or ‘coverage outliers’ may follow from omitted alternative paths to the outcome (e.g. Rohlfing and Schneider, 2013; Oana and Schneider, 2018). Usually, their presence is deemed of little or no threat to the standing of a consistent hypothesis: if we are only interested in the machine triggered by our model, any uncovered instances of *Y* can be a red herring. However, coverage outliers can have another and more concerning source. They are also diagnosed on the overspecification of the model that occurs when unrelated conditions are added. In explanatory QCA, the minimization algorithm offers a strategy to handle this threat while providing an argument in favor of the often-deplored practice of selecting on the dependent (e.g. King *et al.*, 1994: 130). Indeed, a renowned ‘paradox of confirmation’ illustrates the absurd conclusions that standard analyses can reach when blindly applied to instances selected on a ‘wrong’ independent (Salmon, 1989: 50).

The paradox portrays the case of the table salt that is believed to dissolve (*Y*) once put in hexed (*H*) water (*W*). The underlying *inus* model, then, reads  $HW \rightarrow Y$ . The paradox arises when natural diversity presents us with the primitives as in Table 2.

**Table 2.** Truth table from the model  $HW \rightarrow Y$

Id	H	W	Y
$\omega_1$	1	1	1
$\omega_2$	1	0	0
$\omega_3$	0	1	1
$\omega_4$	0	0	0

An analysis narrowing on  $\pi(Y|\omega_1)$  would consider the relationship sound, even by contrast with  $\pi(Y|\omega_4)$ . The proof of the irrelevance of **H** only emerges from the evidence that  $\pi(Y|\omega_1) = \pi(Y|\omega_3)$  and  $\pi(Y|\omega_2) = \pi(Y|\omega_4)$ , after the analytical distinction is made between hexed and non-hexed water. Under such improved model specification, the comparison of configurations to the same outcome pinpoints the irrelevant component as the one whose variation does not affect the state of the outcome. Its removal follows from a logical operation that the original pruning algorithm of QCA, the Quine-McCluskey, performs systematically. The algorithm compares primitives pairwise and drops the single different literal in otherwise identical trains. To witness: in Table 2, either  $\omega_1$  or  $\omega_3$  goes with  $Y$ . The two primitives differ by one literal only, that is, the state of **H**. We can rewrite the two as  $(H \cup \bar{H})WY$ . By the axiom of the Excluded Middle in Table 1,  $H \cup \bar{H} = U$ ; hence,  $H$  can be dismissed as a true ‘noisy’ background variation. The primitives to  $\bar{Y}$ , i.e.,  $\omega_2, \omega_4$ , suggest the same conclusion. Together, these ‘minimizations’ license the inference that the *in*us model is truer to the cases at hand when specified as  $W \rightarrow Y$ .

**Addendum to requirement (b): handling unobserved realizations**

The previous example has assumed a saturated truth table in which all the possible realizations were observed. Technically, observed realizations are those in which at least one unit has a crisp membership score of 1. In actual explanatory QCA, however, an *in*us model easily allows for a wider array of realizations than the units may afford from a certain universe. The problem is understood as ‘limited diversity’ and provides a further version of the curse of dimensionality. It is independent of the mere ratio of the number of cases and variables, and may not be properly addressed by adding cases: infinite instances of the same primitive in an analytic space of four still make three primitives unobserved and provide no analytic leverage.

The strategies to handle unobserved realizations are many, each addressing a possible source of the problem. Observed diversity may increase if we widen the space-time region of the analysis – the ‘scope condition’ for case selection (Marx and Dusa, 2011). The dimensionality of the analytic space can decrease if some gelling interactions are hardened into measures of coarser factors

(Berg-Schlosser and De Meur, 2009; Schneider, 2019). If inadvertently added, empirical constants may be dropped before the analysis of irrelevance as they double the analytic space but leave half of the primitives unobserved (Goertz, 2006).

The latter consideration has evolved into a whole step of the standard protocol. The ‘analysis of individual necessity’ calculates the same parameters as the analysis of sufficiency, but of single conditions and with a reverse meaning. Constants are degenerate necessary conditions, that is, limiting cases of supersets of an outcome-set. They arise when  $|Y| \leq |P|$  and entail a lower-triangular fit. The membership in the condition given that in the outcome renders the individual consistency of necessity

$$kN.cons_{Y \subset A} := \pi(A|Y) = \frac{|A \cap Y|}{|Y|}$$

whereas the membership in the outcome given that in the condition provides the individual coverage of necessity:

$$kN.cov_{Y \subset A} := \pi(Y|A) = \frac{|A \cap Y|}{|A|}$$

– where ‘individual’ refers to the  $k$ -th *inus* condition.

The Relevance of Necessity (*RoN*: Schneider and Wagemann, 2012) is a later addition to detect skewed *inus* components. It is calculated as the reciprocal of the (Boolean) complements in the individual coverage of necessity:

$$RoN_{Y \subset A} := \frac{1}{1 - \pi(Y|A)} = \frac{|\bar{A}|}{|\bar{A} \cup \bar{Y}|}.$$

Conventionally, the condition with *kN.cons* equal or greater than 0.9 but *RoN* lower than 0.5 could be dismissed as trivial. However, the crucial test and consistent with the *inus* rationale remains whether the model requires that special condition to prevent the rising of contradictions in the truth table (Damonte, 2018; Rohlfing, 2020).

However, the protocol of reference to handle limited diversity treats unobserved primitives as counterfactuals (Ragin, 2008; Schneider and Wagemann, 2012) and, ultimately, a problem of missing outcomes. The related counterfactual question asks whether any instance of these known yet unobserved configurations would have displayed  $\bar{Y}$  instead of  $Y$  (Ramsey, 1929). The standard protocol provides three answers – ‘conservative’, ‘parsimonious’, and ‘intermediate’ – and draws its conclusions under as many alternative assumptions:

- the conservative stipulates that no unobserved realization would have obtained; so, its minimizations only operate on observed diversity.
- the parsimonious maintains that any unobserved realization would have obtained that finds a perfect observed match but for a single *minimizand* (the component to be declared irrelevant).

- the intermediate corrects the parsimonious assumption with plausibility concerns. It requires the unobserved *minimizand* to be in the state that the *inus* theory considers ‘right’ to the outcome. The minimization under such ‘directional expectation’ provides the intermediate or ‘plausible’ solution.

Given all non-contradictory primitives, the conservative and the parsimonious solutions provide the tighter and looser boundaries of a ‘confidence solution space’ in which the intermediate solution usually offers the plausible estimate to the best of knowledge – with a *caveat*.

Concerns have been raised that directional expectations might introduce confirmation bias in solutions. Minimizations are geared toward disconfirming the relevance of a factor, not to establish it; hence, relying on directional expectations to drop a term rather uses the theory against itself. Instead, the blind application of the plausibility rules may arise a different version of the ‘paradox of confirmation’ in which the belief in a wrong theory prevents the dropping of irrelevant conditions from the solution. To witness, let us take the hexed salt example of Table 2, but now with an unobserved realization as is  $\omega_3^* := \bar{H}W$  in Table 3.

**Table 3.** Truth table from the model  $HW \rightarrow Y$  with unobserved  $\omega_3$

Id	H	W	Y
$\omega_1$	1	1	1
$\omega_2$	1	0	0
$\omega_3^*$	0	1	?
$\omega_4$	0	0	0

Let also assume that our directional expectations about  $H$  boil down to the belief that, teamed with the right *inus* factors  $\Phi$ , hexing becomes an *inus* component of the machine to  $Y$  when present. In short, we can expect  $H\Phi Y$ , but  $\bar{H}\Phi\bar{Y}$ .

Given Table 3,

- $\omega_1$  is the only realization to  $Y$ , and the conservative solution reads  $HW$ : the salt dissolved because it was hexed and in water.
- the parsimonious minimization matches  $\omega_1$  and  $\omega_3^*$  and yields  $W$  as the prime implicant: the salt dissolved because in water, regardless of hexing.
- the intermediate minimization considers that  $\omega_3^*$  is an implausible or ‘hard’ counterfactual instead: according to our directional expectations, it carries the *minimizand* in the wrong state. Thus,  $\omega_3^*$  is barred from the minimization with  $\omega_1$ . The resulting ‘plausible’ solution overlaps the conservative, and  $H$  is not dropped.

To some, the proven inability of the intermediate solutions to get always rid of known irrelevant components disqualifies it as valid, and leaves the parsimonious solutions as the only finding worth discussing (e.g. Thiem, 2019). To others, parsimonious solutions are too dependent on observed realizations, and their sufficiency far less ‘robust’ when tested on shrinking diversity for providing reliable results (e.g. Duşa, 2019). Besides, assumptions are required to make causal structures emerge (e.g. Fiss, 2011; Damonte, 2018; Schneider, 2019).

The debate assumes theory disregards the fact that data contain handy information to establish the plausibility of our *inus* theory in the cases. The values of the analysis of individual sufficiency provide indications on the tenability of our directional expectations in  $\mathcal{U}$ .

**Table 4.** Analysis of individual necessity of the conditions in Table 3

Conditions	Outcome: $Y$				Outcome: $\bar{Y}$			
	<i>kPRI</i>	<i>kN.cov</i> <i>kS.cons</i>	<i>kN.cons</i> <i>kS.cov</i>	<i>RoN</i>	<i>kPRI</i>	<i>kN.cov</i> <i>kS.cons</i>	<i>kN.cons</i> <i>kS.cov</i>	<i>RoN</i>
$W$	1.000	1.000	1.000	1.000	0.000	0.000	0.000	0.667
$\bar{W}$	0.000	0.000	0.000	0.333	1.000	1.000	1.000	1.000
$H$	0.500	0.500	1.000	0.500	0.500	0.500	0.500	0.500
$\bar{H}$	0.000	0.000	0.000	0.667	1.000	1.000	0.500	1.000

When calculated on the conditions of the hexed salt model as in Table 4, the *kPRI* and the *kS.cons* values show  $H$  being equally insufficient to  $Y$  and  $\bar{Y}$ , while the *RoN* values warn that expectations of necessity are untenable, too.

In short, the analysis of individual necessity may suggest whether directional expectations stand evidence and can be forced onto solutions.

### QCA AND LOCAL CAUSAL STRUCTURES

The previous section showed that explanatory QCA is equipped to identify *inus* compounds with established tools, although adapted to its special gauges; moreover, it can diagnose the challenges from both ill-selected conditions and untenable theoretical expectations. Once the sufficiency requirements are satisfied and the irrelevant conditions dismissed, the solutions provide a regularity answer to why the outcome occurred in some of the cases at hand and why it failed in others. The answer is, the complete realizations of the *inus* machine were at work in the positive instances of the outcome, and its incomplete or obstructed realizations prevented the generative process in the

negative instances. That the two answers make two halves of the same causal story ultimately depends on both being based on a single bundle of *inus* conditions (Verba, 1967; cfr. Schneider and Wagemann, 2012).

Still, the further question remains open whether QCA solutions can be granted a causal interpretation beyond the theory that drove the original definition of the *inus* machine. The shapes that the SCM assumes as causal – namely, the chain  $\mathcal{G}_l$ , the fork  $\mathcal{G}_f$ , and the collider  $\mathcal{G}_c$  – all capture conditionality as Kolmogorov's probability; hence, they can be applied to illuminate the relationship between intermediate and parsimonious solution, as in the example below.

### Drawing an *inus* model

The illustrative model accounts for the differences in the national perception of corruption – the outcome – with the differences in how effective the accountability constraints are perceived to be to the discretion of the policy-makers in the public sector – the explanatory factors.

The underlying theory connects corruption to Elinor Ostrom's second-order social trap in which perceptions play the role of triggers. The basic mechanism considers that high perceived corruption fuels distrust in the fair working of institutions; distrust, in turn, makes people convinced that resorting to corruption remains the safest way of accessing public services and benefits even when they hold the right to it. These 'tragedies,' in Ostrom's framework, can be fixed if communities restore fairness and delegate the task of detecting and sanctioning violations to an independent 'monitor'. The fixing can nevertheless fail, too, when the monitor is perceived as ineffective or complacent. Unsanctioned violations and forbearance trigger a 'second-order' social trap: the choice of free-riding control becomes individually rational, and corruption institutionalizes. In short, the mechanism suggests that the trigger fires when accountability designs are perceived as poor or incomplete; *vice versa*, a credible accountability design should provide the *inus* machine that preserves trust and keeps the perception of corruption low (e.g. Ostrom, 1998).

The next question asks which components constitute such an *inus* machine. The theories of public accountability distinguish between internal and external systems, and assign higher effectiveness to the latter. While internal managerial or somehow hierarchical lines of oversight may invite forbearance to avoid blame, external systems are expected to deter corrupt practices by making oversight public, and be more effective when the institutional design maintains the chances low that a single concern can capture the whole attention of the decision-maker. The mechanism also suggests that a special place in the *inus* machine should be granted to the perceived effectiveness of

the judicial system as the warrant against complacency and forbearance (e.g. Weingast, 1984; Mungiu-Pippidi, 2013; Damonte, 2017).

The operationalization borrows the raw explanatory conditions from the sub-indices of the Rule of Law Index maintained by the World Justice Project. These gauges are composite, but their contents are consistent with the underlying concept, and all point in the same direction (Lazarsfeld and Henry, 1968). From the dataset related to 2017, the following gauges are used:

- subindex 1.3. 'Government powers are effectively limited by independent auditing and review', for the condition  $\langle ATEC \rangle$ . The raw variable gauges the perception that comptrollers or auditors, as well as national human rights ombudsman agencies, have sufficient independence and the ability to check on and oversight of the government.
- subindex 1.5 'Government powers are subject to non-governmental checks', to calibrate the condition  $\langle ASOC \rangle$ . The raw variable gauges the perception that independent media, civil society organizations, political parties, and individuals are free to report and comment on government policies without fear of retaliation.
- subindex 3.1 'Publicized laws and government data' to calibrate the condition  $\langle APUB \rangle$ . The raw measure gauges the perception that basic laws and information on legal rights are publicly available, offered in everyday language, and accessible. It also captures the quality and accessibility of the information published by the government in print or online, and whether administrative regulations, drafts of legislation, and high court decisions are promptly accessible to the public.
- subindex 3.2 'Right to information' to calibrate the condition  $\langle RTA \rangle$ . The underlying raw measure gauges the perception that requests for relevant information from a government agency are timely granted, that responses are pertinent and complete, and that the cost of access is reasonable and free from bribes.
- subindex 7.6 'Civil justice is effectively enforced' as the condition  $\langle ENFOR \rangle$ . The raw variable gauges the perception of effectiveness and timeliness of the enforcing practices of civil justice decisions and judgments in practice.

The operationalization of the outcome  $\langle CLEAN \rangle$ , instead, relies on the Corruption Perception Index maintained by Transparency International. This again provides a suitable gauge of the perceived level of corruption of the administrative bodies, collected from surveys, and, since 2012, validated through a transparent methodology.

As the measures of the outcome and the *inus* factors come from surveys and discount the variations in the year of reference, no need for lagging the



effect is envisaged. The perception of the accountability of the administration and the perception of corruption in the public sector count as aggregate responses to the same state of the policymaking system.

The data of the Corruption Perception Index and the World Justice Project are all collected from a variety of world regions, although not from the same countries. When combined, the more comprehensive coverage is of the countries in the European Union, the European Free Trade Area, and the core Anglophone countries. Together, their administrative and institutional systems provide enough diversity to ensure that patterns will emerge. At the same time, they all are uninterrupted democratic systems, although at different degrees of maturity, which ensures the gauges of the conditions in the model can be given unambiguous interpretations.

After dropping the cases with missing values, the population suitable for the analysis includes 26 cases, whereas the specification of the model includes five explanatory conditions and reads

$$ATEC * ASOC * APUB * RTA * ENFOR \rightarrow CLEAN.$$

The raw values are reported in the online Appendix.

### Analysis and findings

Following theory and gauging, directional expectations are that each condition contributes to low perceived corruption (*CLEAN*) when present, and high perceived corruption (*clean*) when absent. The individual parameters of the calibrated measures, reported in Table 5, support all of them.

**Table 5.** *k*-parameters of fit

Conditions	Outcome: <i>CLEAN</i>			Outcome: <i>clean</i>		
	<i>kN.cons</i>	<i>RoN</i>	<i>kN.cov</i>	<i>kN.cons</i>	<i>RoN</i>	<i>kN.cov</i>
<i>atec</i>	0.106	0.685	0.170	0.760	0.956	0.916
<i>ATEC</i>	0.948	0.776	0.840	0.311	0.411	0.207
<i>asoc</i>	0.147	0.639	0.202	0.884	0.942	0.914
<i>ASOC</i>	0.937	0.893	0.915	0.228	0.460	0.167
<i>apub</i>	0.233	0.589	0.272	0.923	0.845	0.808
<i>APUB</i>	0.836	0.937	0.935	0.168	0.528	0.142
<i>rta</i>	0.361	0.571	0.377	0.932	0.756	0.732
<i>RTA</i>	0.744	0.950	0.936	0.207	0.599	0.196
<i>enfor</i>	0.203	0.610	0.252	0.924	0.894	0.861
<i>ENFOR</i>	0.888	0.934	0.940	0.197	0.503	0.157

The conditions' states are consistent with one outcome's state as expected

and symmetric in their set-relationships with the outcome and its negation. Moreover, their *kN.cons* is never trivial. Together, they yield the truth tables as in Table 6.

**Table 6.** Truth tables: observed realizations and consistency to the outcomes

$\omega$	<i>ATEC</i>	<i>ASOC</i>	<i>APUB</i>	<i>RTA</i>	<i>ENFOR</i>	n	<i>CLEAN S.cons</i>	<i>CLEAN PRI</i>	<i>clean S.cons</i>	<i>clean PRI</i>
26	1	1	0	0	1	2	1.000	1.000	0.348	0.000
30	1	1	1	0	1	2	1.000	1.000	0.291	0.000
24	1	0	1	1	1	1	1.000	1.000	0.706	0.000
32	1	1	1	1	1	10	0.999	0.999	0.135	0.000
25	1	1	0	0	0	1	0.662	0.204	0.858	0.665
17	1	0	0	0	0	1	0.333	0.009	0.994	0.991
01	0	0	0	0	0	9	0.137	0.002	0.996	0.996

Of 32 possible realizations, seven only are observed and neatly associated with either one outcome or the other with one exception ( $\omega_{25}$ ), which does not affect the analysis when run with crisp scores. Units concentrate in the two polar realizations: nine of 11 instances of the negative outcome are the best instances of  $\omega_1$  while ten out of 15 positive instances are best instances of  $\omega_{32}$ . Were the model already a well-specified *inus* machine, the concentration should be higher. Dispersion suggests alternative specifications and/or redundancies, which justifies minimizations.

The minimizations retrieve the solutions in Table 7.

**Table 7.** Implicants

Outcome	Solution	Id	Implicant	<i>S.cons</i>	<i>PRI</i>
<i>CLEAN</i>	Cons Int	1	<i>ENFOR * ATEC * ASOC * rta</i>	1.000	1.000
		2	<i>ENFOR * ATEC * APUB * RTA</i>	0.999	0.999
	Pars	3	<i>ENFOR</i>	0.940	0.932
<i>clean</i>	Cons Int	4	<i>enfor * apub * rta * ATEC</i>	0.923	0.886
		5	<i>enfor * apub * rta * asoc</i>	0.997	0.996
	Pars	6	<i>enfor</i>	0.861	0.832

**Notes:** Of the 15 instances of *CLEAN*,

- Prime Implicant 1 covers four, namely, AUT, BEL; AUS, CAN; PI 2 covers eleven: FRA; NZL, DEU, DNK, EST, FIN, GBR, NLD, NOR, SWE, USA.
- PI 3 covers all.

Of the 11 instances of *clean*,

- PI 4 covers two – namely, ITA; PRT.
- PI 5 covers ten: BGR, CZE, ESP, GRC, HRV, HUN, POL, ROU, SVN; ITA.
- PI 6 covers all.

The parsimonious solutions identify a single factor (the effectiveness of civil justice enforcement) whose variation captures the whole difference between the instances of the positive and the negative outcome. The intermediate solution instead overlaps the conservative; their prime implicants use all the conditions in the model, but differently specified to special subpopulations. They suggest that alternative *in*us machines are at work in groups of instances of the realized and unrealized outcome. The instances of the unrealized outcome contain one overdetermined case, in which the failure can be ascribed to one or the other of the two compounds.

Letting the substantive interpretation aside, the last open question asks which causal standing can be recognized to the information in the parsimonious and intermediate solutions.

### **Exploring the relationship between solution types**

Fiss (2011, Soda and Furnari 2012) dubs the conditions in the parsimonious term the ‘core’ element of the solution, while the conditions added under directional expectations are ‘peripheral’ contributors. The *S.cons* and *PRI* values reported in Table 6 prove that the core provides a worse explanation when alone than in conjunction with the peripheral terms. Hence, the peripherals are relevant to the outcome, although maybe not causally so. The in/equalities of the models  $\mathcal{M}_{(c)}$  that qualify the graphs  $\mathcal{G}_{(c)}$  in the SCM framework provide the suitable diagnostic device that clarifies the causal nature of their relationship.

Tables 8 and 9 report the conditionalities that identify the structures of a chain, a confounder, and a collider, computed for the dependencies between the core condition ( $z$ ), the peripheral conditions in each solution term ( $x_1, x_2$ ), and the outcome ( $y$ ) for each outcome state after turning the fuzzy scores into crisp. A fundamental causal structure is assigned to the solution term when all the identifying conditionalities are satisfied. The conditionalities are inevitably deterministic given the gauges and the single observation point. Nevertheless, in such a thin slice of the world, recognizable structures do emerge.

**Table 8.** Structures of the intermediate implicants to *CLEAN*

<b>Chain</b> $w \rightarrow z \rightarrow y$	$w = x_1$	$w = x_2$	<b>Confounder</b> $x_1 \leftarrow z \rightarrow x_2$	$x_1 \leftarrow z \rightarrow x_2$	<b>Collider</b> $x_1 \rightarrow z \leftarrow x_2$	$x_1 \rightarrow z \leftarrow x_2$
$P(y z)$	1.000	1.000	$P(x_2 z)$	0.267	$P(z x_1)$	1.000
$P(y)$	0.577	0.577	$P(x_2)$	0.192	$P(z)$	0.577
! =	TRUE	TRUE	! =	TRUE	! =	TRUE
$P(z w)$	1.000	0.800	$P(x_1 z)$	0.733	$P(z x_2)$	0.800
$P(z)$	0.577	0.577	$P(x_1)$	0.423	$P(z)$	0.577
! =	TRUE	TRUE	! =	TRUE	! =	TRUE
$P(y w)$	1.000	0.800	$P(x_2 x_1)$	0.000	$P(x_1 x_2)$	0.000
$P(y)$	0.577	0.577	$P(x_2)$	0.192	$P(x_1)$	0.423
! =	TRUE	TRUE	! =	TRUE	=	FALSE
$P(y w, z)$	1.000	1.000	$P(x_2 x_1, z)$	0.000	$P(x_1 x_2, z)$	0.000
$P(y z)$	1.000	1.000	$P(x_2 z)$	0.267	$P(x_1 z)$	0.733
=	TRUE	TRUE	=	FALSE	! =	TRUE
<b>TRUE</b>			<b>FALSE</b>		<b>FALSE</b>	

**Keys:** the nodes in the graphs are given the following values:  
 $y = CLEAN$ ;  $z = ENFOR$ ;  $x_1 = ATEC \cap APUB \cap RTA$ ;  $x_2 = ATEC \cap ASOC \cap rta$ .  
 $N = 26$  for each node.

**Table 9.** Structures of the intermediate implicants to *clean*

<b>Chain</b> $w \rightarrow z \rightarrow y$	$w = x_1$	$w = x_2$	<b>Confounder</b> $x_1 \leftarrow z \rightarrow x_2$	$x_1 \leftarrow z \rightarrow x_2$	<b>Collider</b> $x_1 \rightarrow z \leftarrow x_2$	$x_1 \rightarrow z \leftarrow x_2$
$P(y z)$	1.000	1.000	$P(x_2 z)$	0.909	$P(z x_1)$	0.182
$P(y)$	0.423	0.423	$P(x_2)$	0.423	$P(z)$	0.423
! =	TRUE	TRUE	! =	TRUE	! =	TRUE
$P(z w)$	0.182	0.909	$P(x_1 z)$	0.182	$P(z x_2)$	0.909
$P(z)$	0.423	0.423	$P(x_1)$	0.423	$P(z)$	0.423
! =	TRUE	TRUE	! =	TRUE	! =	TRUE
$P(y w)$	0.182	0.909	$P(x_2 x_1)$	0.091	$P(x_1 x_2)$	0.091
$P(y)$	0.423	0.423	$P(x_2)$	0.423	$P(x_1)$	0.423
! =	TRUE	TRUE	! =	TRUE	=	FALSE
$P(y w, z)$	1.000	1.000	$P(x_2 x_1, z)$	0.500	$P(x_1 x_2, z)$	0.100
$P(y z)$	1.000	1.000	$P(x_2 z)$	0.909	$P(x_1 z)$	0.182
=	TRUE	TRUE	=	FALSE	! =	TRUE
<b>TRUE</b>			<b>FALSE</b>		<b>FALSE</b>	

**Keys:** the nodes in the graphs are given the following values:  
 $y = clean$ ;  $z = enfor$ ;  $x_1 = ATEC \cap apub \cap rta$ ;  $x_2 = apub \cap asoc \cap rta$ .  
 $N = 26$  for each node.

In the example, the core and peripheral components in the implicants to both states of the outcome satisfy the conditionalities of the ‘chain’ structure. From this viewpoint, the core provides the mediating node between the peripheral conditions and the outcome. The conditionalities, moreover, support the claim that the core term provides neither the ‘confounding’ background common factor nor the ‘collider’ in any subpopulations – regardless of whether the peripheral terms display full set-independence [ $P(x_2|x_1) = 0.00$  and  $P(x_1|x_2) = 0.00$  in Table 8] or a slight dependence [ $P(x_2|x_1) = 0.091$  and  $P(x_1|x_2) = 0.091$  in Table 9].

These findings suggest that each of the solutions identified by the plausible minimizations renders the settings of a mechanism, and the core elements provide the ‘mediator’. Moreover, the shape suggests that the peripheral conditions do not offer alternative starting points, but equivalent backgrounds.

## CONCLUDING REMARKS

The article offers arguments and evidence that important reasons for discontent with QCA may apply to the inductive usage of the technique, yet are unjustified when addressed to its explanatory, theory-driven application (cfr. Schneider and Wagemann 2012, Thomann and Maggetti 2020).

When carefully implemented, explanatory QCA inevitably displays some commonalities with the probabilistic family. Both identify causality with the capacity to affect a key state of special units and consider causation as an asymmetric phenomenon. The quasi-experimental scholarship recognizes the issue as the difference in ‘propensities’ or ‘covariates’ presiding over the self-selection mechanisms to receive the stimulus. Explanatory QCA models the covariates and the stimulus as the team of *inus* conditions entailing the capacity to arise or maintain a state of the outcome. Thus, explanatory QCA offers a set-theoretic answer to the question asking which combination of conditions ensures the units’ response to a key factor and which ones make the units unresponsive instead. The PO may work it out as the unwelcome heterogeneity that biases the estimation of the effect. Just the opposite, explanatory QCA joins the SCM in considering settings as the special background that accounts for the firing of a trigger. Along this line, the solutions from explanatory QCA can provide the credible specification of the bundle of features that make units responsive or ‘inert.’

Moreover, the article shows that QCA solutions from an *inus* model can be tested for SCM causal structures, and meaningfully so. In the example, the parsimonious solution term – the ‘difference-making’ components of the model

– is proven to take the position of a mediator to the outcome. In turn, the peripheral conditions provide the ‘covariates’ that complete the *inus* machine and account for its effectiveness.

These considerations speak to the QCA scholarship interested in the debate on the standing of the intermediate solution. The findings suggest that the parsimonious solution remains a key component but seldom makes the whole of an explanation. In terms of the example in section 3, a civil justice perceived as effective supports the perception of low public sector corruption as it backs the belief that other accountability devices and holders are trustworthy, first. The sole perceived effectiveness of civil justice makes a ‘gappy’ explanation, as it does not clarify the ground on which such effectiveness stands.

These considerations possibly speak to the wider causal scholarship, too. They suggest the equivalence of SCM mediators and core QCA conditions, and of PO-SCM ‘covariates’ and peripheral conditions. The equivalences establish the relevance of configurational results to probabilistic models, as they offer logical devices to diagnose the misspecification of a graph and refine its composition. Moreover, the equivalence suggests the possibility of cumulating and improving causal knowledge by nesting and triangulating techniques. Hopefully, these considerations will contribute to widen the dialogue across causal strategies.

## REFERENCES

- Bareinboim E and Pearl J (2016) Causal inference and the data-fusion problem. *Proceedings of the National Academy of Sciences* 113, 7345–7352. <https://doi.org/10.1073/pnas.1510507113>
- Berg-Schlosser D and De Meur G (2009) Comparative research design: case and variable selection. In Rihoux B and Ragin CC (eds), *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques*. London, UK: Sage, pp. 19–32. <https://dx.doi.org/10.4135/9781452226569.n2>
- Cartwright N (2017) Causal powers. Why Humeans can’t even be instrumentalist. In Jacobs JD, ed. *Causal Powers*. Oxford, UK: Oxford University Press, pp. 9–23. <https://dro.dur.ac.uk/25421/>
- Damonte A (2017) Deploying administrative accountability to hinder systemic corruption: what do we know, and what can we expect? *Rivista Italiana di Politiche Pubbliche* 3, 417–441. <https://dx.doi.org/10.1483/88195>
- Damonte A (2018) Gauging the import and essentiality of single conditions in standard configurational solutions. *Sociological Methods & Research*. <https://dx.doi.org/10.1177/0049124118794678>
- Duşa A (2018) *QCA with R: A Comprehensive Resource*. Cham, CH: Springer. <https://doi.org/10.1007/978-3-319-75668-4>
- Duşa A (2019) Critical tension: sufficiency and parsimony in QCA. *Sociological Methods & Research*. Online First. <https://doi.org/10.1177/0049124119882456>
- Fiss PC (2011) Building better causal theories: a fuzzy-set approach to typologies in

- organization research. *The Academy of Management Journal* 5, 393–420. <https://doi.org/10.5465/amj.2011.60263120>
- Goertz G (2006) Assessing the trivialness, relevance, and relative importance of necessary or sufficient conditions in social science. *Studies in Comparative International Development* 41, 88–109. <https://doi.org/10.1007/BF02686312>
- Goertz G and Mahoney J (2012) Concepts and measurement: ontology and epistemology. *Social Science Information* 51, 205–16. <https://doi.org/10.1177/0539018412437108>
- Hájek A (2011) Conditional probability. In Bandyopadhyay PS and Forster MR (eds). *Philosophy of Statistics*. Amsterdam, NL: North-Holland, pp. 99–135. <https://doi.org/10.1016/B978-0-444-51862-0.50003-4>
- Imbens GW (2004) Nonparametric estimation of average treatment effects under exogeneity. *Review of Economics and Statistics* 86, 4–29. <https://doi.org/10.1162/003465304323023651>
- King G, Keohane RO and Verba S (1994) *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton, NJ: Princeton University Press. <https://doi.org/10.1515/9781400821211>
- Kuroki M and Pearl J (2014) Measurement bias and effect restoration in causal inference. *Biometrika* 101, 423–37. <https://doi.org/10.1093/biomet/ast066>
- Lazarsfeld PF and Henry NW (1968) *Latent Structure Analysis*. New York, NY: Houghton Mifflin Co.
- Mackie JL (1965) Causes and conditions. *American Philosophical Quarterly*, 2, 245–264. <https://www.jstor.org/stable/20009173>
- Mackie JL (1980) *The Cement of the Universe*. London, UK: Oxford University Press. <https://doi.org/10.1093/0198246420.001.0001>
- Marx A and Dusa A (2011) Crisp-set qualitative comparative analysis (csQCA), contradictions and consistency benchmarks for model specification. *Methodological Innovations Online*, 6, 103–148. <https://doi.org/10.4256/mio.2010.0037>
- Møller J and Skaaning SE (2019) Set-theoretic methods in democratization research: an evaluation of their uses and contributions. *Democratization*. 26, 78–96. <https://doi.org/10.1080/13510347.2018.1449208>
- Morgan SL and Winship C (2015) *Counterfactuals and Causal Inference*. New York, NY: Cambridge University Press.
- Munck G (2016) Assessing set-theoretic comparative methods: a tool for qualitative comparativists? *Comparative Political Studies* 49, 775–780. <https://doi.org/10.1177/0010414015626453>
- Mungiu-Pippidi A (2013) The Good, the Bad and the Ugly: Controlling Corruption in the European Union, <http://anticorrrp.eu/wp-content/uploads/2016/04/D3.4.1-The-Good-The-Bad-The-Ugly.pdf>
- Oana IE and Schneider CQ (2018) Setmethods: an add-on R package for advanced QCA. *The R Journal* 10, 507–33. <https://journal.r-project.org/archive/2018/RJ-2018-031/index.html>
- Ostrom E (1998) A behavioral approach to the rational choice theory of collective action. *American Political Science Review* 92, 1–22. <https://doi.org/10.2307/2585925>
- Pearl J (2009) *Causality: Models, Reasoning, and Inference*. New York, NY: Cambridge University Press.
- Pearl J and Mackenzie D (2018) *The Book of Why: The New Science of Cause and Effect*. New York, NY: Basic Books.
- Pearl J and Paz A (1987) Graphoids: a graph-based logic for reasoning about relevance

- relations. In Duboulay B, Hogg D and Steels L (eds), *Advances in Artificial Intelligence-II*. Amsterdam, NL: North-Holland Publishing Co, pp. 357–363. [http://ftp.cs.ucla.edu/pub/stat\\_ser/r53-L.pdf](http://ftp.cs.ucla.edu/pub/stat_ser/r53-L.pdf)
- Pearl J and Verma T (1991) A theory of inferred causation. In Allena J, Fikes R and Sandewall E (eds), *Principles of Knowledge Representation and Reasoning: Proceedings of the Second International Conference*. San Mateo, CA: Morgan Kaufmann, pp. 441–452. [https://doi.org/10.1016/S0049-237X\(06\)80074-1](https://doi.org/10.1016/S0049-237X(06)80074-1)
- Ragin CC (2000) *Fuzzy-Set Social Science*. Chicago, IL: University of Chicago Press. <https://press.uchicago.edu/ucp/books/book/chicago/F/bo3635786.html>
- Ragin CC (2006) Set relations in social research: evaluating their consistency and coverage. *Political Analysis*, 14, 291–310. <https://doi.org/10.1093/pan/mpj019>
- Ragin CC (2008) *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago, IL: University of Chicago Press. <https://press.uchicago.edu/ucp/books/book/chicago/R/bo5973952.html>
- Ragin CC (2014) *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*, 2nd Edn. Berkeley, CA: University of California Press. <https://www.ucpress.edu/book/9780520280038/the-comparative-method>
- Ragin CC and Fiss PC (2017) *Intersectional Inequality: Race, Class, Test Scores, and Poverty*. Chicago, IL: University of Chicago Press. <https://press.uchicago.edu/ucp/books/book/chicago/I/bo24957423.html>
- Ramsey FP (1929) General propositions and causality. In Mellor DH (ed), (1990) *F.P. Ramsey: Philosophical Papers*. Cambridge, UK: Cambridge University Press, pp. 145–63.
- Rihoux B and De Meur G (2009) Crisp-set qualitative comparative analysis (csQCA). In Rihoux B and Ragin CC (eds), *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques*. London, UK: Sage, pp. 33–68. <https://dx.doi.org/10.4135/978145226569.n3>
- Rohlfing I (2020) The choice between crisp and fuzzy sets in QCA and the ambiguous consequences for finding consistent set relations. *Field Methods* 32, 75–88. <https://doi.org/10.1177/1525822X19896258>
- Rohlfing I and Schneider CQ (2013) Improving research on necessary conditions: formalized case selection for process tracing after QCA. *Political Research Quarterly* 66, 220–235. <https://www.jstor.org/stable/23563606>
- Rosenbaum PR and Rubin DB (1983) The central role of the propensity score in observational studies for causal effects. *Biometrika* 70, 41–55. <https://doi.org/10.1093/biomet/70.1.41>
- Rubin DB (1974) Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66, 688–701. <https://doi.org/10.1037/h0037350>
- Rubin DB (1978) Bayesian Inference for causal effects: the role of randomization. *The Annals of Statistics*, 6, 34–58. <https://www.jstor.org/stable/2958688>
- Rubinson C (2013) Contradictions in fsQCA. *Quality & Quantity* 47, 2847–2867. <https://doi.org/10.1007/s11135-012-9694-3>
- Salmon WC (1989) *Four Decades of Scientific Explanation*. Pittsburgh, PA: University of Pittsburgh Press. <https://conservancy.umn.edu/handle/11299/185701>
- Sartori G (1984) Guidelines for concept analysis. In Id (ed), *Social Science Concepts: A Systematic Analysis*. Beverly Hills, CA: Sage, pp. 15–85.
- Schneider CQ (2019) Two-step QCA revisited: the necessity of context conditions. *Quality*



- and Quantity* 53, 1109–1126. <https://doi.org/10.1007/s11135-018-0805-7>
- Schneider CQ and Wagemann C (2012) *Set-Theoretic Methods for the Social Sciences*. Cambridge, UK: Cambridge University Press.
- Seawright J (2019) Statistical analysis of democratization: a constructive critique. *Democratization*, 26, 21–39. <https://doi.org/10.1080/13510347.2018.1511540>
- Simon H (1977) *Models of Discovery*. Dordrecht, NL: Reidel Publishing Company. <https://doi.org/10.1007/978-94-010-9521-1>
- Soda G and Furnari S (2012) Exploring the topology of the plausible: fs/QCA counterfactual analysis and the plausible fit of unobserved organizational configurations. *Strategic Organization* 10, 285–296. <https://doi.org/10.1177/1476127012452826>
- Stone MH (1936) The theory of representations of Boolean algebras. *Transactions of the American Mathematical Society*, 40, 37–111. <https://doi.org/10.2307/1989664>
- Thiem A (2019) Beyond the facts: limited empirical diversity and causal inference in QCA. *Sociological Methods & Research*. <https://doi.org/10.1177/0049124119882463> .
- Thomann E and Maggetti M (2020) Designing research with qualitative comparative analysis (QCA): approaches, challenges, and tools. *Sociological Methods & Research* 49, 356–86. <https://doi.org/10.1177/0049124117729700>
- Verba S (1967) Some dilemmas in comparative research. *World Politics* 20, 111–127. <https://doi.org/10.2307/2009730>
- Weingast BR (1984) The congressional bureaucratic system: a principal-agent perspective. *Public Choice* 44, 147–191. <https://doi.org/10.1007/BF00124821>
- Winship C and Morgan SL (1999) The estimation of causal effects from observational data. *Annual Review of Sociology* 25, 659–706. <https://doi.org/10.1146/annurev.soc.25.1.659>