

Optimizing Consistency and Coverage in Configurational Causal Modeling

Michael Baumgartner* and Mathias Ambühl†

Abstract

Consistency and coverage are two core parameters of model fit used by configurational comparative methods (CCMs) of causal inference. Among causal models that perform equally well in other respects (e.g. robustness or compliance with background theories), those with higher consistency and coverage are typically considered preferable. Finding the optimally obtainable consistency and coverage scores for data δ , so far, is a matter of repeatedly applying CCMs to δ while varying threshold settings. This paper introduces a procedure called *ConCovOpt* that calculates, prior to actual CCM analyses, the consistency and coverage scores that can optimally be obtained by models inferred from δ . Moreover, we show how models reaching optimal scores can be methodically built in case of crisp-set and multi-value data. ConCovOpt is a tool, not for blindly maximizing model fit, but for rendering transparent the space of viable models at optimal fit scores in order to facilitate informed model selection—which, as we demonstrate by various data examples, may have substantive modeling implications.

1 Introduction

Over the past three decades, different variants of *configurational comparative methods* (CCMs) have gradually been added to the toolkit for causal data analysis in many disciplines, ranging from social and political science to business administration, evaluation science, and on to public health and psychology. CCMs are designed to investigate causal structures featuring conjunctural causation and equifinality, which tend to prevent pairwise (linear) depen-

*University of Bergen, Norway, michael.baumgartner@uib.no

†ConsultAG, Switzerland, mathias.ambuehl@consultag.ch

dencies among analyzed variables and, hence, induce problems for many standard methodological frameworks. While other methods search for causal relations as characterized by counterfactual or probabilistic theories of causation (e.g. Lewis 1973; Suppes 1970), CCMs trace causation as defined in the tradition of Mackie's (1974) INUS-theory. CCMs do not quantify effect sizes but place a Boolean ordering on sets of causes by grouping their elements conjunctively, disjunctively, and sequentially. And unlike the models produced by many other methods, CCM models do not relate variables to one another but concrete values of variables (cf. Thiem et al. 2016).

The most well-known CCM is *Qualitative Comparative Analysis* (QCA) (Ragin 1987; Cronqvist and Berg-Schlusser 2009; Ragin 2008). *Coincidence Analysis* (CNA) is a more recent addition to the family of CCMs (Baumgartner 2009; Baumgartner and Ambühl 2020). There are various differences between QCA and CNA—in the underlying methodological principles, in the implemented algorithms, or in the search targets—, but also important commonalities. Both methods process configurational data featuring crisp-set, fuzzy-set, or multi-value variables (Thiem 2014), which are called *factors* in CCM jargon. They both exploit relations of sufficiency and necessity for causal inference and output models accounting for the values taken by endogenous factors in terms of redundancy-free Boolean functions of exogenous factor values. And they share two of their core parameters of model fit, which constitute the topic of this paper: *consistency* and *coverage* (Ragin 2006). Informally, consistency reflects the degree to which the behavior of an outcome obeys a corresponding sufficiency or necessity relationship or a whole model, whereas coverage reflects the degree to which a sufficiency or necessity relationship or a whole model accounts for the behavior of the corresponding outcome. What counts as acceptable scores on these parameters is defined in threshold settings determined by the analyst prior to the application of QCA or CNA.

Among causal models that perform equally well with respect to other criteria, for example, robustness and compliance with case knowledge or background theories, the ones with the higher aggregate of consistency and coverage are considered preferable. This raises the question of how to systematically find the models with optimal consistency and coverage. Currently, neither the procedural protocols of QCA nor of CNA have answers to that question on offer. Rather, optimizing consistency and coverage is a matter of repeatedly running QCA and CNA on the data while varying relevant thresholds and comparing the fit scores of resulting models. Such a trial-and-error approach is neither guaranteed to recover all consistency and coverage optima, of which there may be many, nor is it efficient, as it may require a

multitude of data re-analyses. Variations in the thresholds may induce substantive changes in the issued models as well as in their fit scores. These changes may not be proportional to the threshold variations. That is, higher thresholds are not guaranteed to produce models with higher aggregates of consistency and coverage. In consequence, a wide range of threshold settings may have to be searched in fine-grained steps.

This paper shows that it is possible to identify optimal consistency and coverage scores of CCM models inferable from data δ independently of actually applying CCMs to δ . We introduce an explicit procedure, called *ConCovOpt*, that calculates all consistency and coverage optima for δ , within certain computational limitations, prior to CCM analyses. *ConCovOpt* is complemented by a second procedure, called *DNFbuild*, that purposefully builds models reaching optimal scores for crisp-set and multi-value data. For these data types, models are hence guaranteed to exist at the consistency and coverage optima. *ConCovOpt* can also be applied to fuzzy-set data, in which case optimas amount to upper bounds that cannot possibly be outperformed by actual models, but there is no guarantee that models *de facto* exist at those bounds. The upper bounds, thus, constrain the interval of threshold settings within which optimal actual models must be searched.

ConCovOpt is a tool, not for blindly maximizing model fit, but for systematically exploring the space of viable models with optimal fit. Sometimes, optimally fitting models will turn out to be the best models overall, while sometimes optimizing consistency and coverage is only possible at the price of overfitting or of compromising on robustness or compliance with background theories or case knowledge. Choosing the best model(s) among all viable models, which may be numerous (Baumgartner and Thiem 2017), is a delicate task that requires balancing various criteria. Consistency and coverage are only two of those criteria. But, independently of whether the analyst wants to anchor that choice in the data only or additionally draw on external sources of information, making an informed choice presupposes that the whole model space is brought to the analyst's attention. The purpose of *ConCovOpt* is to contribute to that objective.

The article is organized as follows. Section 2 reviews some conceptual preliminaries. In section 3, *ConCovOpt* is presented using a simple crisp-set data example. Section 4 applies it to large- N data. *DNFbuild* is introduced in section 5 on the basis of multi-value data; and a fuzzy-set application is discussed in section 6. Finally, section 7 puts *ConCovOpt* into proper methodological perspective. We implemented *ConCovOpt* and *DNFbuild* as R functions, which are extensively used in the R script that is available in the supplementary

material and that allows for replicating all calculations of this paper.

2 Conceptual preliminaries

We begin by introducing some conceptual and notational preliminaries of our ensuing discussion. As indicated above, CCMs study Boolean dependence relations between factors taking on specific values. Factors represent categorical properties that partition sets of units of observation (cases) either into two sets, in case of binary properties, or into more than two (but finitely many) sets, in case of multi-value properties. Factors representing binary properties can be *crisp-set* (*cs*) or *fuzzy-set* (*fs*); the former typically take on 0 and 1 as possible values, whereas the latter can take on any (continuous) values from the unit interval. Factors representing multi-value properties are called *multi-value* (*mv*) factors; they can take on any of an open (but finite) number of non-negative integers as possible values. Values of a *cs* or *fs* factor X are often interpreted as membership scores in the set of cases exhibiting the property represented by X , while the values of an *mv* factor Y designate the particular way in which the property represented by Y is exemplified.

As the explicit ‘Factor=value’ notation yields convoluted syntactic expressions with increasing model complexity, we subsequently use—whenever possible—a shorthand notation that is conventional in Boolean algebra and CCM modeling: membership in a set is expressed by italicized upper case and non-membership by lower case Roman letters. Hence, in case of *cs* and *fs* factors, we normally write ‘ X ’ for $X=1$ and ‘ x ’ for $X=0$. In case of *mv* factors and within explicit definitions, value assignments are always written out, using the ‘Factor=value’ notation; that is, we write ‘ $Y=v$ ’ for factor Y taking the value v .

CCM models may feature all the standard Boolean operations: negation $\neg X$ (‘not X ’), conjunction $X*Y$ (‘ X and Y ’), disjunction $X+Y$ (‘ X or Y ’), implication $X \rightarrow Y$ (‘If X , then Y ’), and equivalence $X \leftrightarrow Y$ (‘ X if, and only if, Y ’). In case of *cs* and *mv* factors, these operations are given a rendering in classical logic (see e.g. Lemmon 1965 for a canonical introduction). In case of *fs* factors, Boolean operations are rendered in fuzzy logic: negation $\neg X$ amounts to $1 - X$, conjunction $X*Y$ to $\min(X, Y)$, disjunction $X+Y$ to $\max(X, Y)$, an implication $X \rightarrow Y$ is taken to express that the membership score in X is smaller or equal to Y ($X \leq Y$), and an equivalence $X \leftrightarrow Y$ that the membership scores in X and Y are equal ($X = Y$).

The implication operator is used to define the notions of *sufficiency* and *necessity*, which are the two dependence relations exploited by CCMs: X is sufficient for Y if, and only if (iff),

$X \rightarrow Y$ ('if X is given, then Y is given'), and X is necessary for Y iff $Y \rightarrow X$ ('if Y is given, then X is given'). CCM models have the form $\Phi \leftrightarrow Y$, where Y is an endogenous factor value and Φ stands for an expression $X_1 * \dots * X_i + \dots + X_m * \dots * X_n$ in *disjunctive normal form (DNF)* such that all factors in that DNF are different (and logically, conceptually, and metaphysically independent) from one another and from Y . All in all, thus, CCM models explain Y in terms of a necessary disjunction of sufficient conditions of Y .

Sufficiency and necessity relations amount to mere association patterns. As such, they carry no causal connotations whatsoever, and, hence, most of these relations do not reflect causation. Still, some of them do. Regularity theories of causation (Mackie 1974; Graßhoff and May 2001; Baumgartner and Falk 2019) are designed to filter out those sufficiency and necessity relations that do track causation. According to regularity theories, an expression of the form $\Phi \leftrightarrow Y$ tracks causation only if Φ is *redundancy-free*, meaning that no conjuncts or disjuncts can be removed from Φ without violating the truth of $\Phi \leftrightarrow Y$. QCA and CNA differ in regard to how rigorously Φ needs to be freed of redundancies before it is amenable to a causal interpretation. In QCA, complete redundancy elimination as implemented in so-called *parsimonious* models is not mandatory—partial redundancy elimination as in *intermediate* or *conservative* models may suffice as well. By contrast, CNA automatically eliminates all redundancies. These differences are bracketed in the following.

Since CCM-processed data δ tend to feature various deficiencies (e.g. fragmentation, noise, etc.), expressions of type $\Phi \leftrightarrow Y$ that adhere to the strict standards of the equivalence operation (' \leftrightarrow ') often cannot be inferred from δ . To relax these standards, that is, to approximate strict sufficiency and necessity relations, Ragin (2006) introduced the consistency and coverage measures into the QCA protocol, which have subsequently also been imported into CNA (Baumgartner and Ambühl 2020). As the implication operator is defined differently in classical and in fuzzy logic, the two measures are defined differently for crisp-set and multi-value data, which both have a classical footing, and for fuzzy-set data. *Cs-consistency* (con^{cs}) and *cs-coverage* (cov^{cs}) of $X \rightarrow Y$ are defined as follows, where ' $|\dots|_\delta$ ' represents the cardinality of the set of cases instantiating the enclosed expression in the data δ :

$$con^{cs}(X \rightarrow Y) = \frac{|X*Y|_\delta}{|X|_\delta} \quad cov^{cs}(X \rightarrow Y) = \frac{|X*Y|_\delta}{|Y|_\delta} \quad (1)$$

Fs-consistency (con^{fs}) and *fs-coverage* (cov^{fs}) of $X \rightarrow Y$ are defined as follows, where n is

the number of cases in δ :

$$\text{con}^{fs}(X \rightarrow Y) = \frac{\sum_{i=1}^n \min(X_i, Y_i)}{\sum_{i=1}^n X_i} \quad \text{cov}^{fs}(X \rightarrow Y) = \frac{\sum_{i=1}^n \min(X_i, Y_i)}{\sum_{i=1}^n Y_i} \quad (2)$$

Whenever the values of X and Y are restricted to 1 and 0 in the crisp-set measures, con^{cs} and cov^{cs} coincide with con^{fs} and cov^{fs} , but for binary factors with values other than 1 and 0 and for multi-value factors that does not hold. Nonetheless, we will in the following not explicitly distinguish between the *cs* and *fs* measures because our discussion will make it sufficiently clear which of them is at issue.

What counts as acceptable scores on these measures is defined in threshold settings chosen by the analyst prior to the application of QCA or CNA. While QCA only accepts a consistency threshold, CNA requires both a consistency and a coverage threshold. Moreover, the implementation of these thresholds differs in important ways in the two methods. In QCA, a consistency threshold is imposed only on conjunctions of all exogenous factors (so-called *minterms*) in the course of the generation of *truth tables*, which are intermediate calculative devices for QCA. The final models issued may or may not meet the chosen threshold. In CNA, thresholds for both consistency and coverage are used as authoritative model building constraints. The thresholds define what counts as sufficient and necessary conditions, to the effect that models not meeting them cannot be built.

Despite these differences, in both QCA and CNA models with higher consistency and coverage are preferred over models with lower scores on these measures, provided they fare equally well in other respects (e.g. robustness). The following section introduces our procedure, ConCovOpt, calculating consistency and coverage optima.

3 The optimization procedure

The goal of ConCovOpt is to identify both optimal and maximal consistency and coverage scores—*con-cov optima* and *con-cov maxima*, for short—the distinction being that an optimum optimizes at least one of consistency and coverage, whereas a maximum optimizes their aggregate. The procedure is given configurational data δ and a set of outcomes \mathbf{O} in δ as input. By suitably synthesizing δ for the modeling of every $Y \in \mathbf{O}$, ConCovOpt first identifies output values for Boolean functions, so-called *rep-assignments*, which reproduce the behavior of Y as closely as possible, and, by calculating consistency and coverage scores

	P	O	A	C
E75	1	0	1	0
E87	1	1	0	1
P81	1	0	0	0
P90	1	0	0	0
N77	1	1	1	0
E80	0	1	0	0
E92	0	1	0	0
P75	0	0	1	1
P87	0	0	0	1
P92	0	0	0	1
N80	0	0	0	0

(a)

conf.	P	O	A	C	n	exo-groups	$\varphi(C=1)$	φ_1	φ_2
σ_1	1	0	1	0	1	$\{\sigma_1\}$	0	0	0
σ_2	1	1	0	1	1	$\{\sigma_2\}$	1	1	1
σ_3	1	0	0	0	2	$\{\sigma_3\}$	0	0	0
σ_4	1	1	1	0	1	$\{\sigma_4\}$	0	0	0
σ_5	0	1	0	0	2	$\{\sigma_5\}$	0	0	0
σ_6	0	0	1	1	1	$\{\sigma_6\}$	1	1	1
σ_7	0	0	0	1	2	$\{\sigma_7, \sigma_8\}$	0, 1	0	1
σ_8	0	0	0	0	1				

(b)

Table 1: Subtable (a) is the data matrix for the United States from Giugni and Yamasaki (2009, 476) with outcome C . Subtable (b) is a configuration table with labels in column ‘conf.’, case frequencies in ‘ n ’, imperfect pairs with gray shading, as well as added exo-groups, rep-list $\varphi(C=1)$, and rep-assignments φ_1 and φ_2 .

for these rep-assignments, it then infers all con-cov optima and maxima that CCM models of Y can possibly reach.

We introduce the procedure using the very simple cs data example in Table 1a drawn from Giugni and Yamasaki (2009, 476), who investigate the policy impact of different social movements between 1975 and 1995. The exogenous factors are high protest activity (P), public opinion favorable to the movement (O), and powerful institutional allies (A), with values 0 and 1 representing ‘no’ and ‘yes’ for all factors. The endogenous factor C takes the value 1 whenever a movement manages to significantly change a country’s policy, and 0 otherwise. The authors analyze the data for various western countries separately; Table 1a features the data for the United States.

We begin by searching for con-cov optima for outcome C (i.e. $C=1$) in Table 1a. The notion of a con-cov optimum shall be defined as follows:

Con-cov optimum. An ordered pair $\langle con, cov \rangle$ of consistency and coverage scores is a con-cov optimum for outcome $Y=\nu$ in data δ iff, prior to applying a CCM, it can be excluded that a model of $Y=\nu$ inferred from δ scores better on one element of the pair and at least as well on the other, whereas it cannot be excluded that a model of $Y=\nu$ reaches $\langle con, cov \rangle$.

If models of Y (i.e. $Y=1$) inferred from data δ can be modeled with perfect consistency and

coverage, $\langle 1, 1 \rangle$ is the only con-cov-optimum for Y in δ . Outcome C in Table 1a, however, does not have a con-cov optimum of $\langle 1, 1 \rangle$. The reason is that the cases P87, P92, N80 feature the same configuration of the exogenous factors—the configuration p^*o^*a —while C is given in P87 and P92 and c is given in N80. The configurations $p^*o^*a^*C$ and $p^*o^*a^*c$ constitute imperfect configurations, or what we will call an *imperfect pair*.¹

Imperfect pair. An imperfect pair for $Y=\nu$ in data δ is a pair of configurations $\{\sigma_i, \sigma_j\}$ in δ such that $Y=\nu$ is instantiated in one element of the pair and $Y \neq \nu$ in the other, while all other factors in δ take constant values in both σ_i and σ_j .

In cs (and mv) data δ , there is a tight logical connection between imperfect pairs for an outcome Y and $\langle 1, 1 \rangle$ being the con-cov optimum for Y in δ : Y has a con-cov optimum of $\langle 1, 1 \rangle$ in δ iff there does not exist an imperfect pair for Y in δ . The reason is that a model with perfect consistency and coverage expresses Y as a strict Boolean function of the other factors in δ , and such a function issues exactly one output for every input. If there does not exist an imperfect pair for Y , every input (i.e. every configuration of factors other than Y) can be straightforwardly mapped either onto Y or onto y . But if there exists an imperfect pair, there exists an input to which no determinate output can be assigned, meaning Y cannot be expressed as a strict Boolean function of the other factors in δ . On average, the more imperfect pairs Y has in δ the lower the con-cov optima for Y in δ .²

The existence of imperfect pairs indicates that there are varying causes of Y in the uncontrolled causal background. The variation of Y in an imperfect pair must have some cause or other; but that cause cannot be among the other factors in δ because they are constant in the pair. Since varying latent causes are a source of confounding, it is standardly recommended to try to resolve imperfect pairs prior to a CCM analysis; and there are various approaches on offer for how to do this (e.g. Rihoux and De Meur 2009). Of course, as suppressing the variation of latent causes, especially in observational studies, is very difficult, these approaches may be incapable of improving the data quality. For the purposes of this paper, we will hence assume that the quality of all our example data has been improved as far as possible, meaning that the remaining imperfect pairs cannot be resolved.

¹In the QCA literature, such configurations are often labelled *contradictory* (e.g. Rihoux and De Meur 2009, 46-49). We find this terminology misleading because a contradiction cannot possibly be realized. But, as Table 1a vividly demonstrates, it is very well possible for p^*o^*a to be realized in combination with both C and c .

²Not only the amount of imperfect pairs affects the con-cov optima but also the number of cases instantiating the elements of imperfect pairs.

A first step towards determining con-cov optima for an outcome Y is to identify imperfect pairs for Y in δ . To do this in a methodical manner, we re-organize the data such that the instantiated configurations are rendered more transparent by synthesizing all cases in δ instantiating the same configuration in a single row of what we call a *configuration table*. A configuration table CT of data δ merges multiple rows of δ in which all factors have identical values into one row, such that each row of CT corresponds to one determinate configuration of the factors in δ .³ The configurations in CT are labeled and the number of cases instantiating each configuration is stored in a frequency column. The first three (line-separated) columns of Table 1b amount to a configuration table of the data in Table 1a.

A configuration table then allows for splitting the configurations into groups in which all factors other than a scrutinized outcome Y , that is, all factors exogenous with respect to Y , take constant values. We shall speak of *exo-groups*, for short.

Exo-group. An exo-group of an outcome $Y=v$ in a configuration table CT is a group of configurations in CT with constant values in all factors in CT other than Y .

The imperfect pairs for Y (i.e. $Y=1$) in data δ can then be directly read off the list of Y 's exo-groups: exo-groups with more than one element such that Y is instantiated in one element and not instantiated in another element correspond to imperfect pairs. To illustrate with our example, the exo-groups of C are listed in the fourth (line-separated) column of Table 1b. While σ_1 is the only configuration in Table 1b featuring P^*o^*A , meaning that $\{\sigma_1\}$ is a singleton exo-group of C , there are two configurations featuring p^*o^*a , viz. σ_7 and σ_8 , which constitute an exo-group with two elements $\{\sigma_7, \sigma_8\}$. As C is instantiated in one element of that group and c in the other, $\{\sigma_7, \sigma_8\}$ amounts to an imperfect pair for C ; and as C has no other exo-groups with more than one element, it is C 's only imperfect pair.

In order for a CCM model, which, to recall, has the form $\Phi \leftrightarrow Y$, to have highest possible consistency and coverage, its redundancy-free DNF Φ must reproduce the instantiation behavior of the outcome Y as closely as possible. The notion of *reproducing the behavior* of an outcome as closely as possible will be of crucial importance for ConCovOpt. It must be understood somewhat differently for *cs* and *mv* data, on the one hand, and *fs* data, on

³Note that a configuration table is *not* a QCA truth table. A configuration table merely synthesizes cases to configurations and shows the actual value distribution of the outcome, whereas a truth table replaces the outcome distribution by a function column indicating for each configuration (minterm) whether it is sufficient for the outcome according to a chosen consistency threshold. As ConCovOpt does not aim to optimize the fit scores relative to some chosen consistency threshold but to find absolute con-cov optima it has to operate on the actual outcome distribution.

the other. In case of *cs* and *mv* data, we say that Φ reproduces the behavior of an outcome as closely as possible iff Φ returns the value 1 for every exo-group in which the outcome is constantly instantiated, 0 for every exo-group in which it is constantly non-instantiated, and either 0 or 1 for every exo-group with a varying instantiation of the outcome. Applied to our example, this means that a Φ —whichever concrete DNF this may be—reproduces the behavior of C as closely as possible iff Φ returns 1 for exo-groups $\{\sigma_2\}, \{\sigma_6\}$; 0 for $\{\sigma_1\}, \{\sigma_3\}, \{\sigma_4\}, \{\sigma_5\}$; and either 0 or 1 for $\{\sigma_7, \sigma_8\}$. Taken together, these value assignments yield what we will call the *rep-list* (reproduction list) $\varphi(C=1)$ for outcome C in Table 1b.

Rep-list. A rep-list $\varphi(Y=\nu)$ for an outcome $Y=\nu$ assigns all the values reproducing the behavior of $Y=\nu$ as closely as possible to every exo-group of $Y=\nu$.

Moreover, to an assignment that returns a value from a rep-list for every exo-group we will refer as a *rep-assignment* (reproduction assignment).

Rep-assignment. A rep-assignment φ for an outcome $Y=\nu$ assigns exactly one value from a rep-list $\varphi(Y=\nu)$ to every exo-group of $Y=\nu$.

Whatever concrete factor values CCMs may ultimately incorporate in DNFs accounting for outcome Y , it is clear—prior to applications of CCMs—that a DNF not returning a rep-assignment does not reach a con-cov optimum for Y . At the same time, as will be shown in section 4, some rep-assignments yield non-optimal consistency or coverage scores. That is, returning a rep-assignment for Y is necessary but not sufficient for a DNF to reach a con-cov optimum for Y . In order to identify those rep-assignments that actually yield con-cov optima, all possible rep-assignments must be built from $\varphi(Y=\nu)$ and the consistency and coverage scores they induce tested for optimality.

The number of rep-assignments that can be built from a rep-list $\varphi(Y=\nu)$ is equal to the number of combinatorially possible value distributions drawn from $\varphi(Y=\nu)$, that is, to $\prod_{i=1}^n |\varphi(Y=\nu)_i|$, where n is the number of exo-groups and $|\varphi(Y=\nu)_i|$ the cardinality of the set of possible values assigned to exo-group i . In our example, the complete set of rep-assignments is easily built, as there is only one exo-group with more than one value in the rep-list $\varphi(C=1)$. Hence, outcome C has a total of two rep-assignments, φ_1 and φ_2 , which are featured in the last two columns of Table 1b. φ_1 and φ_2 coincide except for the fact that they contain the values 0 and 1, respectively, for exo-group $\{\sigma_7, \sigma_8\}$. φ_1 induces perfect consistency but does not cover the instance of C in σ_7 , whereas φ_2 covers the instance of C

in σ_7 but violates perfect consistency in σ_8 .

It only remains to be determined which of all rep-assignments actually reach con-cov optima. To this end, consistency and coverage scores are calculated for all rep-assignments. In case of *cs* and *mv* data, this can be done by plugging the values of a rep-assignment φ_i and the corresponding instantiation behavior of outcome $Y=\nu$ into the definitions con^{cs} and cov^{cs} in Expression (1). In our example, this means that columns ' φ_1 ' and ' φ_2 ' of Table 1b yield the X -values of con^{cs} and cov^{cs} , column ' C ' the Y -values, and column ' n ' the case frequencies. We get the following consistency and coverage scores:

$$\varphi_1 : con = 1 ; cov = 0.5 \quad (3)$$

$$\varphi_2 : con = 0.8 ; cov = 1 \quad (4)$$

Due to the imperfect pair in exo-group $\{\sigma_7, \sigma_8\}$, it is impossible, in principle, for a CCM model of C inferred from the data in Table 1a to score better on consistency and coverage. φ_1 outperforms φ_2 in consistency and φ_2 outperforms φ_1 in coverage. As neither of the two scores better than the other on one measure and at least as well on the other, they are both con-cov optima. φ_1 optimizes consistency, φ_2 optimizes coverage. At the same time, φ_2 clearly outperforms φ_1 in the aggregate of consistency and coverage, which we take to be the product of consistency and coverage (i.e. the *con-cov product*). That is, φ_2 has the better overall model fit; it reaches a *con-cov maximum*:

Con-cov maximum: An ordered pair $\langle con, cov \rangle$ of consistency and coverage scores is a con-cov maximum for outcome $Y=\nu$ in data δ iff $\langle con, cov \rangle$ is a con-cov optimum for $Y=\nu$ in δ with the highest aggregate, i.e. product, of consistency and coverage.

Of course, the product of consistency and coverage is only one option among many to aggregate consistency and coverage. Assessing the overall model fit based on the con-cov product amounts to giving equal weights to consistency and coverage, which, while standard in CNA, may not be endorsed by all QCA methodologists (who tend to have a preference for consistency). We do not want to take a stance on that issue here, but, instead, invite a reader who wants to assess the overall model fit by giving unequal weights to consistency or coverage to view the simple con-cov product in the above definition as a placeholder for any preferred function aggregating consistency and coverage. That is, a con-cov maximum might alternatively be defined as a con-cov optimum with maximal score on $con^{0.75} \cdot cov^{0.25}$, or on

Procedure 1 (ConCovOpt) consistency and coverage optimization

Input: configurational data δ with a set of outcomes \mathbf{O} in δ .

Output: all con-cov optima and maxima for all outcomes in \mathbf{O} .

- (1) Aggregate δ in a configuration table CT.
 - (2) For every outcome $Y=\nu$ in \mathbf{O} , split the configurations in CT into exo-groups of $Y=\nu$.
 - (3) Build the rep-list $\varphi(Y=\nu)$.
 - (4) Build all rep-assignments φ_1 to φ_m from $\varphi(Y=\nu)$.
 - (5) Calculate the consistency and coverage scores of φ_1 to φ_m .
 - (6) Eliminate all rep-assignments with scores that do not reach a con-cov optimum.
 \Rightarrow The scores of the remaining rep-assignments correspond to all *con-cov optima* for $Y=\nu$ in δ , the optima with highest aggregate are the *con-cov maxima*.
-

$(0.25 \cdot con) + (0.75 \cdot cov)$, or on $min(con, cov)$, etc.⁴ While all of these alternative definitions identify φ_2 as con-cov maximum for outcome C, they may select different con-cov maxima in other examples. Still, to avoid unnecessary complications, we shall subsequently only work with the con-cov product as our aggregation function of choice.

In sum, without having applied CCMs to Guigni and Yamasaki's (2009) data, we have identified optimal and maximal consistency and coverage scores for them. Before we search for actual CCM models for our example, let us assemble the different procedural steps. To this end, one generalization is still needed. For simplicity, all data analyzed in this paper comprise a single outcome only, but, of course, configurational data may feature *multiple outcomes*. If that is the case, exo-groups, rep-lists, and rep-assignments must be formed and consistency and coverage scores calculated for each outcome separately. For generality, we thus let the input of ConCovOpt be data δ along with a set of outcomes \mathbf{O} in δ . If no prior knowledge is available as to which values of which factors in δ are possible outcomes, ConCovOpt can simply be run by setting \mathbf{O} equal to all values of all factors in δ . ConCovOpt is presented in Procedure 1.

Let us now apply CCMs to our example data in order to find actual models returning φ_1 and φ_2 . At a consistency threshold anywhere between 1 and 0.67, QCA produces the following parsimonious (QCA-PS) and conservative models (QCA-CS),⁵ the latter of which

⁴Accordingly, our R implementation of ConCovOpt in the supplementary material is accompanied by a function `selectMax()` that accepts any function aggregating consistency and coverage as criterion to select con-cov maxima.

⁵The relevant replication code is available as an R script in the supplementary material.

is also the model published by Giugni and Yamasaki (2009, 479):

$$\text{QCA-PS: } p^*A + P^*O^*a \leftrightarrow C \quad \text{con} = 1 ; \text{cov} = 0.5 \quad (5)$$

$$\text{QCA-CS: } p^*o^*A + P^*O^*a \leftrightarrow C \quad \text{con} = 1 ; \text{cov} = 0.5 \quad (6)$$

In both (5) and (6), not only the solution consistency but also the consistencies of all sufficient conditions (i.e. disjuncts) are 1. Our previous calculations attest that the QCA models reach a con-cov optimum, as they return rep-assignment φ_1 . At the same time, we now see that (5) and (6) do not reach a con-cov maximum. Rep-assignment φ_2 shows that it is possible to significantly improve on the overall model fit. But at a conventional consistency threshold of 0.75, standard QCA does not find a better scoring model—for two main reasons. On the one hand, QCA builds models *from the top down* by first searching for complete minterms satisfying a chosen consistency threshold and then eliminating redundancies. The minterm p^*o^*a of the exo-group $\{\sigma_7, \sigma_8\}$, however, only reaches a consistency of 0.667 and is therefore not further considered by QCA (in QCA jargon: it is coded ‘0’)—despite the fact that, as we shall see below, a proper part of that minterm indeed scores 0.75 on consistency. On the other hand, standard QCA does not accept a coverage threshold and, hence, cannot be ‘asked’ to build models with specific target scores on coverage.

CNA, by contrast, accepts separate thresholds for consistency of sufficient conditions and of whole models as well as for coverage of whole models. It can thus be ‘asked’ to build models at any target scores. Moreover, it builds models *from the bottom up* by first testing single factor values for compliance with chosen thresholds and by then gradually adding further factor values until threshold compliance is established.⁶ Dusa (2018) has recently presented a promising new algorithm for QCA called *CCubes* that also builds models from the bottom up and accepts the same types of thresholds as CNA.⁷ At a threshold setting of $\langle 1, 0.5 \rangle$, CNA and CCubes return the same model as QCA-PS, but at $\langle 0.75, 1 \rangle$ they find the following model realizing φ_2 :

$$\text{CNA / CCubes: } p^*o + P^*O^*a \leftrightarrow C \quad \text{con} = 0.8 ; \text{cov} = 1 \quad (7)$$

⁶For more on the difference between a top-down and bottom-up search see Baumgartner and Ambühl (2020).

⁷Dusa (2018) describes CCubes as a minimization algorithm that is optimized for speed but produces the same models as traditional QCA algorithms (e.g. Quine-McCluskey optimization). In fact, however, CCubes accepts much more fine-grained tuning parameters and its bottom-up approach may generate models that differ from standard QCA models, as the examples in this paper show. CCubes is not equivalent to standard QCA algorithms,

This is not the place to select among the different model candidates we have now recovered for Guigni and Yamasaki’s (2009) data, nor to substantively interpret them. What matters for our purposes is that computing con-cov optima and maxima by means of ConCovOpt prior to actually conducting CCM analyses has (at least) three important payoffs. First, it allows us to determine how close actually obtained models come to optimal and maximal fit. Second, it renders transparent whether the obtained models exhaust the space of con-cov optima or whether further models should be searched at different thresholds. Third, without having to try out a whole range of threshold settings, CCMs can be run by directly constraining them towards optimal thresholds.

4 Large- N crisp-set data

The data in Table 1a are very simple and, although the resulting optimal models differ significantly in overall fit, they have a considerable overlap in causal ascriptions, thus inducing only marginally different causal conclusions. To show that optimizing consistency and coverage can also make a substantive difference in causal conclusions, we now turn to a more intricate, large- N data example. Britt et al. (2000) investigate the determinants leading to the parental decision to terminate a pregnancy after a prenatal diagnosis of trisomy 21. Four exogenous factors are examined: existing children (C ; 0 := ‘none’, 1 := ‘1 or more’), maternal age in years (M ; 0 := ‘37 and under’, 1 := ‘38 and above’), prior voluntary abortions (A ; 0 := ‘none’, 1 := ‘1 or more’), and gestational age in weeks (G ; 0 := ‘16 and under’, 1 := ‘17 and over’). The endogenous factor is termination (T ; 0 := ‘continue’, 1 := ‘terminate’). The cases are 142 pregnant women receiving a trisomy 21 diagnosis at Wayne State University Clinic from September 1989 through October 1998. The complete data can be consulted in our replication script, the configuration table resulting from that data is in Table 2.

As is frequently the case in large- N data, there are numerous imperfect pairs, highlighted with gray shading. Instead of first calculating con-cov optima and maxima and only afterwards looking at concrete models, we proceed in reverse order for this example. We begin by presenting the model offered by Britt et al. (2000, 412). They choose a consistency threshold of 0.875. At this threshold, QCA builds two parsimonious models (only the second of which is mentioned by the authors):⁸

⁸The data analyzed by Britt et al. (2000) do not feature limited diversity, which is why all QCA solution types coincide.

conf.	<i>C</i>	<i>M</i>	<i>A</i>	<i>G</i>	<i>T</i>	<i>n</i>	exo-groups	$\varphi(T=1)$	φ_{pub}	φ_{max}
σ_1	1	1	0	0	1	27				
σ_2	1	1	0	0	0	1	$\{\sigma_1, \sigma_2\}$	0, 1	1	1
σ_3	1	1	1	0	1	7				
σ_4	1	1	1	0	0	1	$\{\sigma_3, \sigma_4\}$	0, 1	1	1
σ_5	1	0	0	0	1	11	$\{\sigma_5\}$	1	1	1
σ_6	1	0	1	1	1	9	$\{\sigma_6\}$	1	1	1
σ_7	0	0	1	0	1	1	$\{\sigma_7\}$	1	1	1
σ_8	1	1	1	1	1	3				
σ_9	1	1	1	1	0	1	$\{\sigma_8, \sigma_9\}$	0, 1	0	1
σ_{10}	0	1	0	1	1	1	$\{\sigma_{10}\}$	1	1	1
σ_{11}	0	0	0	1	1	8	$\{\sigma_{11}\}$	1	1	1
σ_{12}	1	0	0	1	1	19				
σ_{13}	1	0	0	1	0	6	$\{\sigma_{12}, \sigma_{13}\}$	0, 1	0	1
σ_{14}	1	0	1	0	1	7	$\{\sigma_{14}\}$	1	1	1
σ_{15}	1	1	0	1	1	11				
σ_{16}	1	1	0	1	0	3	$\{\sigma_{15}, \sigma_{16}\}$	0, 1	0	1
σ_{17}	0	0	0	0	1	4				
σ_{18}	0	0	0	0	0	2	$\{\sigma_{17}, \sigma_{18}\}$	0, 1	0	1
σ_{19}	0	1	1	1	1	1	$\{\sigma_{19}\}$	1	1	1
σ_{20}	0	1	1	0	1	1				
σ_{21}	0	1	1	0	0	1	$\{\sigma_{20}, \sigma_{21}\}$	0, 1	0	1
σ_{22}	0	0	1	1	1	6	$\{\sigma_{22}\}$	1	1	1
σ_{23}	0	1	0	0	1	11	$\{\sigma_{23}\}$	1	1	1

Table 2: Configuration table for the data in Britt et al. (2000, 412) with (highlighted) imperfect pairs, exo-groups, rep-list $\varphi(T=1)$, and the rep-assignments φ_{pub} and φ_{max} .

$$\text{QCA: } c^*G + C^*g + m^*A + c^*M^*a \leftrightarrow T \quad con = 0.98 ; cov = 0.70 \quad (8)$$

$$c^*G + C^*g + m^*A + M^*a^*g \leftrightarrow T \quad con = 0.98 ; cov = 0.70 \quad (9)$$

A handful of QCA re-runs with only slightly varied consistency thresholds shows that the results are highly volatile, yielding many different models with different overall fit. Identifying a con-cov maximum for these data calls for a systematic approach. We thus apply ConCovOpt with $\mathbf{O} = \{T=1\}$. The result of step (1) is the configuration table in Table 2. Step (2) yields 9 singleton exo-groups entailing determinate values for the rep-list in step (3). In all 7 non-singleton exo-groups the outcome T varies, meaning that DNFs reproducing the behavior of T as closely as possible can return either 0 or 1 for those groups. In total, the resulting rep-list $\varphi(T=1)$ induces $2^7 = 128$ rep-assignments in step (4). In step (5), consis-

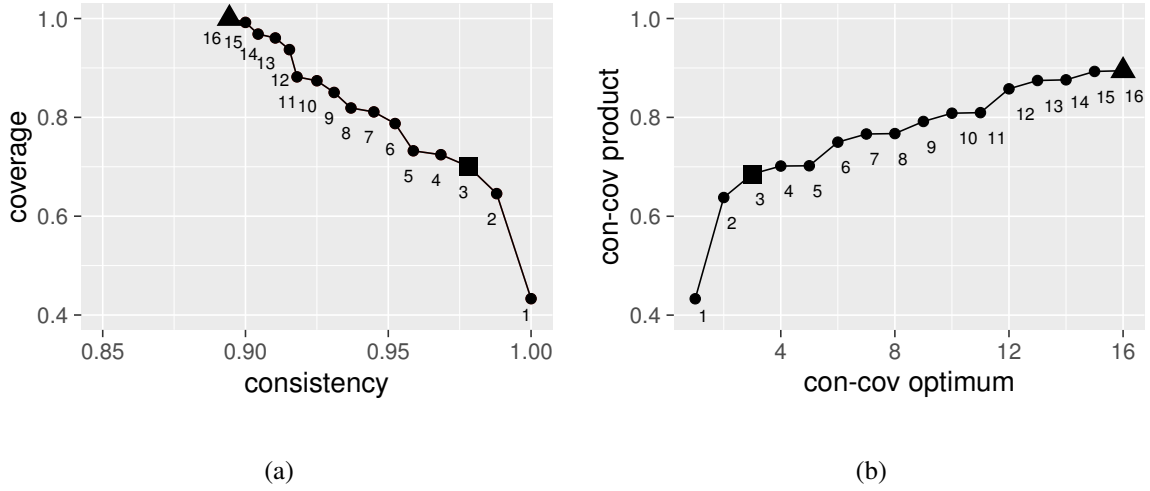


Figure 1: Plot (a) shows the 16 con-cov optima for outcome T in the data of Britt et al. (2000), plot (b) the con-cov products of each optimum. \blacktriangle is the con-cov maximum, \blacksquare the published model.

tency and coverage scores are calculated for all of them and those with non-optimal scores eliminated in step (6). 16 con-cov optima remain. They are plotted Figure in 1a.

In light of these results, we can now say that the published model (9), which scores $\langle 0.98, 0.70 \rangle$, indeed reaches a con-cov optimum (φ_3 ; \blacksquare in Figure 1). However, with its con-cov product of 0.69 it is quite far away from a con-cov maximum. The con-cov products of all 16 optima are plotted in Figure 1b. The con-cov maximum for outcome T is $\varphi_{16} \langle 0.89, 1 \rangle$ (\blacktriangle). For transparency, we add the rep-assignment φ_{pub} (i.e. φ_3) realized by the published model and the assignment φ_{max} (i.e. φ_{16}) yielding the con-cov maximum to Table 2.

The value distribution of φ_{max} is striking: it assigns 1 to *every* exo-group, resulting in a (vacuous) tautology. In other words, a best fitting model for the data of Britt et al. (2000) entails that pregnancy is terminated in case of a trisomy 21 diagnosis whatever the values of the exogenous factors, that is, independently of the pregnant woman's existing children, prior abortions, age, and gestational age. To render this more concrete, let us now search for actual models at $\langle 0.89, 1 \rangle$. While standard QCA algorithms do not find a model reaching the con-cov maximum, CNA and CCubes have no problem finding such a model. At a threshold setting of $\langle 0.89, 1 \rangle$, the following tautologous model is returned:

$$\text{CNA / CCubes: } M + m \leftrightarrow T \quad \text{con} = 0.89 ; \text{ cov} = 1 \quad (10)$$

Of course, any other tautologous model, as $C + c \leftrightarrow T$ or $A + a \leftrightarrow T$, reaches the same fit scores, but (10) is the only model such that the two disjuncts M and m individually reach the consistency threshold of 0.89 as well—which is why it is the only model issued by CNA and CCubes. The overall fit of (10) is good by all CCM standards, and (10) also meets the consistency threshold used by Britt et al. (2000). It is the best fitting model for their data, and, for principled reasons, its fit scores cannot be outperformed. But of course, it is *not a causally interpretable* model. Causes are difference-makers of their effects, yet a tautology does not make a difference to anything.

This finding casts doubts on all causal conclusions Britt et al. (2000) have drawn from their data. 127 of all 141 women receiving a trisomy 21 diagnosis in their sample choose to terminate, making termination the canonical response to the diagnosis. No non-tautologous function of the exogenous factors can account for the outcome better than the tautologous model that entails termination no matter what. The data contain too little variation on the outcome T to conclude anything about its causes; in particular, there is no evidence that T is caused by any of the factors C , M , A , or G . A causal interpretation of the published model (9) is unwarranted by the data. This shows that a systematic search for con-cov maxima (prior to a CCM analysis) may have implications that go way beyond minor model adjustments or improvements. The optimization of consistency and coverage scores rendered possible by ConCovOpt may thoroughly change the conclusions drawn from a study.

5 Multi-value data

ConCovOpt is straightforwardly applicable to multi-value data. Although the factors in mv data can take more than two values, models for an mv outcome $Y=v$ have the same logical form as cs models: they account for $Y=v$ in terms of redundancy-free DNFs, which, irrespective of whether they feature cs or mv factors, are true or false, that is, only return 1 or 0. Hence, for an optimal mv DNF to reproduce the instantiation behavior of an outcome $Y=v$ as closely as possible the exact same conditions must be satisfied as in the cs case (see p. 10). It follows that rep-lists and rep-assignments can be built and evaluated for consistency and coverage in exactly the same way for mv data as for cs data.

To illustrate, we apply ConCovOpt to the mv data of Verweij and Gerrits (2015), who investigate the impact of different management strategies in response to unplanned events occurring during the implementation of a large infrastructure project in Maastricht. Their

data comprise 18 unplanned events (between 2009 and 2011) measuring the following exogenous factors: nature of the event (E ; 0 := ‘physical/remote’, 1 := ‘social/project/public’), nature of the management response (M ; 0 := ‘internal’, 1 := ‘external’), nature of the interaction between public and private managers (I ; 0 := ‘autonomous public’, 1 := ‘autonomous private’, 2 := ‘cooperation’). The endogenous factor is the satisfactoriness of the management response (S). To emphasize the data’s *mv* nature, we replace the original values of S (i.e. 0 and 1) by two (arbitrarily chosen) different ones; that is, we will say that S takes the value 3 if satisfactoriness is high and the value 2 if it is not high.

The input to ConCovOpt, hence, is Verweij and Gerrits’ complete data with $\mathbf{O} = \{S=3\}$ (see the replication script). In step (1), ConCovOpt synthesizes that data in the configuration table contained in Table 3. As there is only one outcome, step (2) forms one set of exo-groups, which are listed in column 4. There are two non-singleton exo-groups, $\{\sigma_3, \sigma_4\}$ and $\{\sigma_7, \sigma_8\}$, each inducing two possible values in the rep-list in step (3). Step (4) generates $2^2 = 4$ rep-assignments φ_1 to φ_4 . For transparency, we list them all in Table 3. In step (5), the value distributions of φ_1 to φ_4 and the instantiation behavior of outcome $S=3$ are plugged into con^{cs} and cov^{cs} . Step (6) checks the consistency and coverage scores for optimality, which check is positive for all scores. Overall, ConCovOpt identifies the following four con-cov optima for outcome $S=3$, the last of which, φ_4 , is a con-cov maximum:

$$\varphi_1 : con = 1 ; cov = 0.44 \quad (11)$$

$$\varphi_2 : con = 0.86 ; cov = 0.67 \quad (12)$$

$$\varphi_3 : con = 0.78 ; cov = 0.78 \quad (13)$$

$$\varphi_4 : con = 0.75 ; cov = 1 \quad (14)$$

Verweij and Gerrits (2015, 21) present a conservative solution, (15), that scores $\langle 1, 0.44 \rangle$ and, hence, realizes the optimum φ_1 . Independently of where in the interval $[1, 0.67]$ the consistency threshold is placed, QCA finds two parsimonious solutions, (16) and (17), which likewise realize φ_1 .

$$\text{QCA-CS: } M=0 * I=0 + E=0 * M=0 * I=2 + E=1 * M=1 * I=1 \leftrightarrow S=3 \quad (15)$$

$$\text{QCA-PS: } M=0 * I=0 + E=0 * I=2 + M=1 * I=1 \leftrightarrow S=3 \quad (16)$$

$$M=0 * I=0 + M=0 * I=2 + M=1 * I=1 \leftrightarrow S=3 \quad (17)$$

conf.	<i>E</i>	<i>M</i>	<i>I</i>	<i>S</i>	<i>n</i>	exo-groups	$\varphi(S=3)$	φ_1	φ_2	φ_3	φ_4
σ_1	0	0	2	3	1	$\{\sigma_1\}$	1	1	1	1	1
σ_2	1	0	1	2	2	$\{\sigma_2\}$	0	0	0	0	0
σ_3	1	1	0	2	2	$\{\sigma_3, \sigma_4\}$	0, 1	0	0	1	1
σ_4	1	1	0	3	3						
σ_5	1	0	0	3	1	$\{\sigma_5\}$	1	1	1	1	1
σ_6	1	1	1	3	1	$\{\sigma_6\}$	1	1	1	1	1
σ_7	1	1	2	2	1	$\{\sigma_7, \sigma_8\}$	0, 1	0	1	0	1
σ_8	1	1	2	3	2						
σ_9	0	0	1	2	4	$\{\sigma_9\}$	0	0	0	0	0
σ_{10}	0	0	0	3	1	$\{\sigma_{10}\}$	1	1	1	1	1

Table 3: Configuration table for the data in Verweij and Gerrits (2015), with re-coded factor *S*, exo-groups, rep-list, and all resulting rep-assignments.

That is, in the range of conventional threshold settings QCA only finds models realizing one of four con-cov optima. Instead of now applying CNA and CCubes, we next introduce a procedure for building CCM models realizing *any* con-cov optimum for *cs* and *mv* data.

A con-cov optimum is realized by a DNF that outputs either 1 or 0 for every exo-group. Any two DNFs that return the same output for all exo-groups have the same consistency and coverage. For convenience, let us call the set of exo-groups to which a rep-assignment φ_i assigns the value 1 the *positive group* of φ_i . One particularly interesting DNF returning φ_i then is what we label φ_i 's *canonical DNF*—in reference to canonical normal forms of logical expressions (Lemmon 1965, 198).

Canonical DNF. The canonical DNF returning a rep-assignment φ_i for outcome $Y=\nu$ is the disjunction of the configurations of all factors exogenous to *Y* in φ_i 's positive group.

Just as any logical expression is guaranteed to have a unique canonical normal form, so is every rep-assignment guaranteed to have exactly one canonical DNF, which moreover is easily built. For example, disjunctively concatenating the configurations of factors *E*, *M* and *I* in φ_4 's positive group in Table 3 yields the canonical DNF returning φ_4 ; it is a DNF that scores $\langle 0.75, 1 \rangle$ in accounting for $S=3$:

$$\begin{aligned}
& E=0 * M=0 * I=2 + E=1 * M=1 * I=0 + E=1 * M=0 * I=0 + \\
& E=1 * M=1 * I=1 + E=1 * M=1 * I=2 + E=0 * M=0 * I=0
\end{aligned} \tag{18}$$

In the same way, the canonical DNFs of the other con-cov optima for outcome $S=3$ can be

construed. For example, as the exo-group $\{\sigma_7, \sigma_8\}$ is not in φ_3 's positive group, removing the configuration of the exogenous factors in $\{\sigma_7, \sigma_8\}$, viz. $E=1 * M=1 * I=2$, from (18) is all it takes to build the canonical DNF returning φ_3 .

Of course, as canonical DNFs comprise all configurations of all exogenous factors in an optimum's positive group, they tend not to be redundancy-free, and consequently, not to be causally interpretable. For instance, if we remove $E=0$ from the first and the last disjuncts in (18), we are still left with a DNF with the same output for all exo-groups in Table 3, that is, a DNF returning φ_4 . The same holds if we continue to eliminate $E=1$ from all disjuncts. By contrast, if $I=2$ is eliminated from the first disjunct of (18), we are left with a DNF that scores $\langle 0.56, 1 \rangle$ in accounting for $S=3$ and, thus, no longer returns φ_4 . The reason is that a DNF featuring $E=0 * M=0$ as a separate disjunct (instead of $E=0 * M=0 * I=2$) outputs 1 for exo-group $\{\sigma_9\}$, whereas φ_4 assigns 0 to that group. In sum, while some factor values can be removed from (18) such that the remaining DNF still returns φ_4 , others are indispensable for returning φ_4 .

These considerations suggest that (18) can be turned into a *redundancy-free DNF* returning φ_4 by systematically removing factor values and checking whether the remainder still outputs the same as (18) for all exo-groups. All factor values for which this check is positive are redundant; all factor values for which the check is negative are indispensable (non-redundant). If all redundancies are removed from (18), we are left with this redundancy-free DNF: $I=2 + I=0 + M=1$. Accounting for $S=3$ on its basis yields the following model, which reaches the con-cov maximum for Verweij and Gerrits' data:

$$I=2 + I=0 + M=1 \leftrightarrow S=3 \quad con = 0.75 ; \quad cov = 1 \quad (19)$$

(19) has a peculiarity. Although the model as a whole reaches a consistency of 0.75, two of its component disjuncts do not. $M=1$ and $I=0$ are sufficient for $S=3$ with consistencies of 0.67 and 0.71 only. It follows that in order to find (19) with CNA and CCubes, the consistency threshold must be lowered to 0.66. Whether models some of whose components only reach a consistency of 0.67 should be considered acceptable is a question that requires further discussion, which, however, goes beyond the purposes of this paper. If (19) is considered unacceptable, our previous analysis has shown that there are other con-cov optima on offer for Verweij and Gerrits' data that score better than the published model. For example, by systematically eliminating redundancies from the canonical DNF returning φ_2 we find the

redundancy-free DNF in (20), all of whose disjuncts have consistencies of 0.75 or higher, such that (20) is returned by both CNA and CCubes at that consistency threshold:

$$I=2 + M=0 * I=0 + M=1 * I=1 \leftrightarrow S=3 \quad con = 0.86 ; \quad cov = 0.67 \quad (20)$$

(20) comes close to the two parsimonious QCA solutions (16) and (17). The only difference is that $I=2$ is conjunctively combined with $E=0$ in (16) and with $M=0$ in (17), creating a model ambiguity, whereas $I=2$ is a stand-alone disjunct in (20), which is non-ambiguous. Hence, replacing QCA's parsimonious solutions by redundancy-free DNFs returning φ_2 or φ_4 —systematically built via their canonical DNFs—not only increases the con-cov product from 0.44 to 0.58 and 0.75, respectively, but also resolves a model ambiguity—two clear advantages of (20) and (19) over (16) and (17).⁹

We end this section by assembling the steps building DNFs returning rep-assignments in procedural form, which we label *DNFbuild*, for short.

Procedure 2 (DNFbuild) build redundancy-free DNFs returning rep-assignments

Input: *cs* or *mv* configuration table CT and rep-assignment φ_i for outcome $Y=\nu$ in CT

Output: redundancy-free DNF(s) returning φ_i

- (1) Build the canonical DNF_{cano} returning φ_i by disjunctively concatenating the configurations of the factors exogenous to $Y=\nu$ in φ_i 's positive group.
 - (2) Eliminate all factor values from DNF_{cano} for which it holds that the result of the elimination, viz. DNF_{elim}, still returns φ_i (i.e. produces the same output as DNF_{cano} for all exo-groups in CT).
- ⇒ If no further factor values can be eliminated from DNF_{elim}, it is a redundancy-free DNF returning φ_i .
-

In case of *cs* and *mv* configuration tables, all φ_i 's have a canonical DNF, which is efficiently built by step (1) of DNFbuild. Step (2), by contrast, involves some intricacies. The reason is that different orders in which factor values are eliminated from DNF_{cano} may result in different redundancy-free DNFs. Generating all of these DNFs can be done more or less efficiently, but the running time of all algorithms that solve this problem grows exponentially with the number of factors and the number of values these factors can take. In the repli-

⁹At the same time it is worth noting that (20) and (19) do not contradict (16) and (17); rather, the former are mere submodels of the latter (Baumgartner and Ambühl 2020).

cation script, we use an algorithm called *ereduce* to generate all redundancy-free forms of DNF_{cano} .¹⁰ In the end, what matters for our current purposes is not the concrete implementation of step (2) but the fact that *DNFbuild* is guaranteed to find a redundancy-free DNF Φ_{rf} of every con-cov optimum $\langle h, k \rangle$. The equivalence $\Phi_{rf} \leftrightarrow Y=\nu$ then amounts to a CCM model accounting for $Y=\nu$ with consistency h and coverage k . That is, in case of *cs* and *mv* data, there exists a CCM model for every con-cov optimum.

6 Fuzzy-set data

This section, first, applies *ConCovOpt* to *fs* data, and second, shows that there is no guarantee that actual CCM models exist for every con-cov optimum in *fs* data and that calculating con-cov optima for *fs* data is more computationally demanding than for *cs* and *mv* data. As background for this discussion, we choose the study by Basurto (2013) who analyzes the autonomy among local institutions for biodiversity conservation in Costa Rica. The study aims to identify causes of, on the one hand, the emergence of autonomy between 1986 and 1998 and, on the other, the endurance of that autonomy between 1998 and 2006. Basurto investigates three groups of potentially causally relevant factors: local, national and international ones. In what follows, we focus on the local influence factors of high local communal involvement through direct employment (*E*), high local direct spending (*S*), and co-management with local or regional stakeholders (*C*), and we concentrate on the outcome of endurance of high local autonomy (*A*), with 0 representing ‘no’ and 1 ‘yes’ for all factors (Basurto 2013, 577). The data cover 16 Costa Rican biodiversity conservation programs; the factors are calibrated on a membership scale with increments of 0.2.

As input to *ConCovOpt* we, thus, use Basurto’s data (see the replication script) with $\mathbf{O} = \{A=1\}$. In step (1), *ConCovOpt* builds the configuration table in Table 4. In this example, each case corresponds to exactly one configuration—which is not uncommon for *fs* data. Since there, again, is only one outcome in \mathbf{O} , step (2) then builds one set of exo-groups, which are listed in column 4 of Table 4.

The main peculiarity of *fs* data is that configurations and outcomes are not merely instantiated or not instantiated in each case but instantiated with different set membership scores (i.e. to different degrees) in many different cases. One consequence is that there are often not

¹⁰In a nutshell, *ereduce* searches for minimal hitting sets in DNF_{cano} that prevent DNF_{cano} from being false in the data from which it is inferred.

conf.	E	S	C	A	n	exo-groups	$\varphi(A=1)$	φ_{max}	φ_1
σ_1	1.0	1.0	1.0	1.0	1	$\{\sigma_2\}$	1.0	1.0	1.0
σ_2	1.0	0.6	1.0	0.4	1				
σ_3	1.0	0.6	1.0	1.0	1	$\{\sigma_2, \sigma_3, \sigma_4\}$	0.4, 0.6, 1.0	0.6	0.6
σ_4	1.0	0.6	1.0	0.6	1				
σ_5	1.0	0.8	1.0	0.8	1	$\{\sigma_5\}$	0.8	0.8	0.8
σ_6	1.0	0.4	1.0	0.4	1	$\{\sigma_6\}$	0.4	0.4	0.4
σ_7	0.4	1.0	1.0	1.0	1	$\{\sigma_7\}$	1.0	1.0	1.0
σ_8	1.0	1.0	0.0	1.0	1	$\{\sigma_8\}$	1.0	1.0	1.0
σ_9	0.4	0.4	0.2	0.4	1	$\{\sigma_9, \sigma_{10}\}$	0.2, 0.4	0.4	0.4
σ_{10}	0.4	0.4	0.2	0.2	1				
σ_{11}	0.4	0.4	0.6	0.4	1	$\{\sigma_{11}\}$	0.4	0.4	0.4
σ_{12}	0.4	0.4	0.4	0.2	1	$\{\sigma_{12}, \sigma_{13}\}$	0.4, 0.6	0.6	0.4
σ_{13}	0.4	0.4	0.4	0.6	1				
σ_{14}	0.2	0.4	0.2	0.4	1	$\{\sigma_{14}\}$	0.4	0.4	0.4
σ_{15}	0.4	0.4	0.0	0.6	1	$\{\sigma_{15}, \sigma_{16}\}$	0.4, 0.6	0.6	0.4
σ_{16}	0.4	0.4	0.0	0.4	1				

Table 4: Configuration table for the data in Basurto (2013) with exo-groups, rep-list, and rep-assignments φ_{max} and φ_1 .

only imperfect *pairs* but imperfect *n-tuples*, where n is the number of different membership scores an outcome has in an exo-group. An example is exo-group $\{\sigma_2, \sigma_3, \sigma_4\}$ in Table 4. Since A has three different membership scores in that group, *viz.* $A=0.4$, $A=1.0$, and $A=0.6$, it corresponds to an imperfect triple. That the other 3 non-singleton exo-groups in Table 4 have only two elements is mere happenstance; *fs* exo-groups can have as many members as there are values the outcome can take. Still, just as CCM models for other data types, models for *fs* data have highest consistency and coverage if they reproduce the behavior of the outcome as closely as possible. That is, a DNF Φ of an *fs* model $\Phi \leftrightarrow Y$ scores the higher on consistency and coverage, the closer the value of Φ comes to the value of Y for every exo-group.

Due to the instantiation by degrees in *fs* data, however, the notion of reproducing the behavior of an outcome as closely as possible cannot be spelled out exactly as for *cs* and *mv* data (cf. p. 10). While a rep-assignment for the latter data types must reproduce the behavior of the outcome by only assigning one of two values, 0 or 1, a rep-assignment φ for *fs* data may assign many more values, *viz.* the minimum of the conjuncts' scores in a conjunction and the maximum of the disjuncts' scores in a disjunction. The values φ can assign to a particular exo-group are constrained by the membership scores of the exogenous factors in

that group. More specifically, as φ is the output of a disjunction of conjunctions of factor values or their negations, it can only assign the membership scores of the positive or of the negated factors in that group, but no other values. To illustrate, take again the exo-group $\{\sigma_2, \sigma_3, \sigma_4\}$, in which $E=1.0$, $S=0.6$, and $C=1.0$. The conjunction $E*S*C$ issues 0.6 for that group, *viz.* $\min(1.0, 0.6, 1.0)$; correspondingly, $e*c$ and $E*C$ return 0.0 and 1.0, respectively, or $s*C$ outputs 0.4. But any value other than 0.0 (negation of E and C), 0.4 (negation of S), 1.0 (value of E and C) or 0.6 (value of S) cannot be assigned to that exo-group by φ , meaning that those are φ 's *possible values* for that group. Hence, a rep-assignment φ *reproduces the behavior* of an outcome as closely as possible if, and only if, it holds for every exo-group x that φ assigns one of its possible values to x that comes as close as possible to the outcome's membership scores in x . In the case of exo-group $\{\sigma_2, \sigma_3, \sigma_4\}$ that means that φ reproduces the behavior of A as closely as possible iff it assigns one of 0.4, 0.6, or 1.0 to that exo-group (but not 0.0); or for exo-group $\{\sigma_{12}, \sigma_{13}\}$ φ must assign either 0.4, which is equally close to $A=0.2$ and $A=0.6$ in that group, or 0.6, which is closest to $A=0.6$. Sometimes the closest possible value of φ exactly matches an outcome, sometimes it is far away from it. Sometimes only one of the possible values is closest to the outcome, sometimes multiple values are closest.

Based on this notion of reproducing the behavior of an outcome, step (3) builds the rep-list $\varphi(A=1)$ in Table 4 for Basurto's data. From this list, 24 rep-assignments for outcome A are built in step (4). Just as in *cs* and *mv* data, the values of the rep-assignments and of the outcome are then plugged into the definitions for consistency and coverage in step (5)—this time, of course, using the fuzzy-set definitions con^{fs} and cov^{fs} in Expression (2). After eliminating all rep-assignments inducing non-optimal scores, 7 con-cov optima remain, which are plotted in Figure 2a with their con-cov products in Figure 2b.

Basurto (2013) chooses a consistency threshold of 0.79 and produces an intermediate solution, (21), which coincides with QCA's conservative solution. At that consistency threshold, QCA issues the two parsimonious solutions (22) and (23).

$$\text{QCA-IS/CS: } C + E*S \leftrightarrow A \quad con = 0.79 ; \quad cov = 0.94 \quad (21)$$

$$\text{QCA-PS: } C + E \leftrightarrow A \quad con = 0.79 ; \quad cov = 0.94 \quad (22)$$

$$C + S \leftrightarrow A \quad con = 0.79 ; \quad cov = 0.96 \quad (23)$$

Contrary to our previous examples, these QCA models fall significantly short of a con-

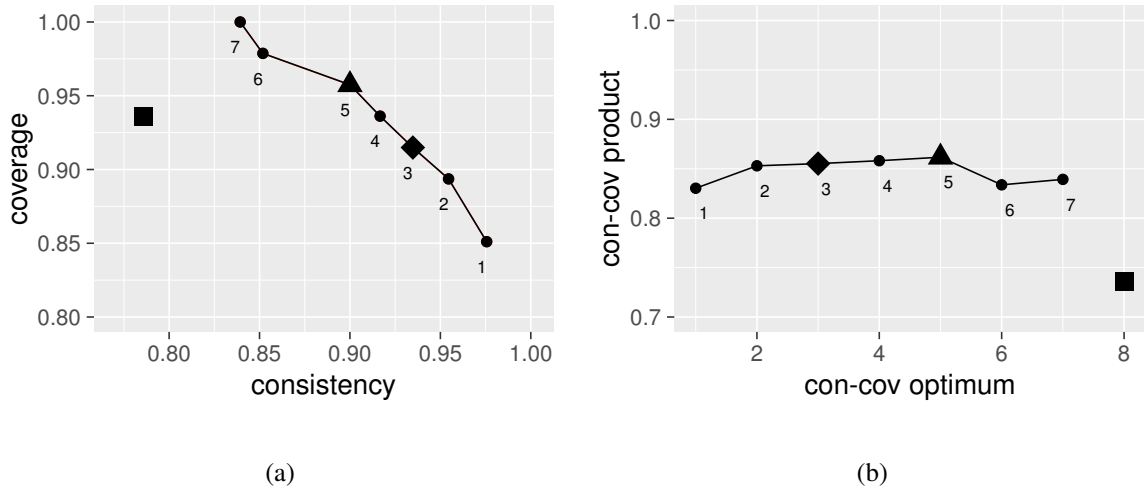


Figure 2: Plot (a) shows the 7 con-cov optima for outcome A in the data of Basurto (2013) and the con-cov score of the published model (■), plot (b) the con-cov products of each optimum and of the published model. ▲ is the con-cov maximum and ◆ the best optimum realizable by a CCM model.

cov optimum, let alone a con-cov maximum. Moreover, the models are not robust under variations of the consistency threshold. But thanks to ConCovOpt, we now have concrete threshold settings at which to search for models with optimal fit. The con-cov maximum of $\langle 0.9, 0.957 \rangle$ (▲ in Figure 2) is reached by the rep-assignment φ_{max} in Table 4. It turns out, however, that neither QCA nor CNA nor CCubes find a model at the threshold setting $\langle 0.9, 0.957 \rangle$, not even—as is possible in CNA and CCubes—if the consistency of individual disjuncts is allowed to fall short of 0.9. There simply does not exist a CCM model for outcome A at the con-cov maximum. This is not some idiosyncrasy of Basurto’s data. Con-cov optima for fs data frequently do not have actual CCM models realizing them.

In a cs and mv configuration table, the configuration Γ_i of exogenous factors in an exo-group $\{\sigma_i\}$ of an outcome Y is guaranteed not to be instantiated in any other exo-groups of Y . Hence, if Y is given in $\{\sigma_i\}$, Γ_i —properly freed of redundancies—can safely be included as a disjunct in a model of Y without affecting the model’s consistency in any exo-groups other than $\{\sigma_i\}$. It is therefore possible to modularly build CCM models for every rep-assignment φ_i along the lines of DNFbuild. The DNFbuild approach, however, does not work for fs data, because if Y is an fs outcome, the configuration Γ_i in exo-group $\{\sigma_i\}$ may be instantiated with some non-zero membership in many other exo-groups. In consequence, including Γ_i as a disjunct in a CCM model of Y is likely to affect the model’s consistency in exo-groups

other than $\{\sigma_i\}$. It may hence happen that a DNF only returns an optimal value for exo-group $\{\sigma_i\}$ provided it contains factor values that, at the same time, return a non-optimal value for another exo-group $\{\sigma_j\}$, meaning that no DNF returns optimal values for both $\{\sigma_i\}$ and $\{\sigma_j\}$.

To make this (abstract problem) concrete, compare the exo-groups $\{\sigma_9, \sigma_{10}\}$ and $\{\sigma_{15}, \sigma_{16}\}$ in Table 4. The rep-assignment φ_{max} assigns 0.4 to $\{\sigma_9, \sigma_{10}\}$ and 0.6 to $\{\sigma_{15}, \sigma_{16}\}$. Given the membership scores $E=0.4, S=0.4$ and $C=0.2$ in $\{\sigma_9, \sigma_{10}\}$, a DNF only returns 0.4 for $\{\sigma_9, \sigma_{10}\}$ if it includes E or S or $E*S$ (but not C) as a disjunct. But E, S and $E*S$ also return 0.4 for group $\{\sigma_{15}, \sigma_{16}\}$, while φ_{max} assigns 0.6 to that group. Given the membership scores $E=0.4, S=0.4$ and $C=0.0$ in $\{\sigma_{15}, \sigma_{16}\}$, a DNF would have to include e or s or $e*s$ in order to return 0.6 for $\{\sigma_{15}, \sigma_{16}\}$. But if those factor values are included, 0.6 is issued for $\{\sigma_9, \sigma_{10}\}$, which again is not the value assigned by φ_{max} . In sum, no DNF outputs 0.4 for $\{\sigma_9, \sigma_{10}\}$ and 0.6 for $\{\sigma_{15}, \sigma_{16}\}$, meaning that no actual CCM model will realize φ_{max} .

This problem is avoided if we do not require 0.6 to be issued for $\{\sigma_{15}, \sigma_{16}\}$ but 0.4, which indeed happens to be assigned by another rep-assignment, φ_1 (cf. Table 4), yielding the con-cov optimum $\langle 0.93, 0.91 \rangle$. Moreover, it turns out that there exists an actual DNF also reproducing all other values of φ_1 . At a threshold setting of $\langle 0.93, 0.91 \rangle$, both CNA and CCubes find the following model:¹¹

$$\text{CNA / CCubes: } S \leftrightarrow A \quad \text{con} = 0.93 ; \text{cov} = 0.91 \quad (24)$$

Apart from there being no guarantee that fs con-cov optima are realizable by actual models, the fs case also differs from ConCovOpt analyses of cs and mv data in its computational complexity. The most computationally costly step of ConCovOpt is step (5), which calculates fit scores of rep-assignments. The more rep-assignments are entailed by a configuration table, the more time-consuming that calculation. In case of cs and mv data, the amount of rep-assignments is only a function of the number of non-singleton exo-groups, which, in turn, is a function of the number of imperfect pairs in the data.¹² A cs or mv DNF optimally reproduces the behavior of an outcome $Y=\nu$ in every singleton exo-group by returning 1 if Y takes the value ν in that exo-group, or 0 otherwise. In consequence, no rep-assignment needs more than one value assignment to optimally capture the outcome's behavior in singleton exo-groups. By contrast, it might be that multiple output values of an fs DNF reproduce the

¹¹At a consistency threshold of 0.93, (24) is not found by standard QCA, but if that threshold is lowered to 0.9, (24) is also returned by QCA.

¹²The amount of rep-assignments is 2^n , where n is the number of non-singleton exo-groups (see p. 15).

outcome's value in a single exo-group $\{\sigma_i\}$ equally optimally. This happens if the outcome's value is not itself among the possible values of the exogenous factors in $\{\sigma_i\}$ and multiple of those possible values are equally close to the outcome's value. In that case, multiple rep-assignments result from the singleton exo-group $\{\sigma_i\}$. It follows that the number of rep-assignments does not only grow with the number of imperfect configurations but also with the number of cases. This, in combination with the fact that rep-assignments may assign more than two values to exo-groups, yields that exhaustively searching for con-cov optima is much more computationally demanding for *fs* data.

Our R implementation of ConCovOpt can calculate the fit scores of about 10 million rep-assignments in reasonable time. For *cs* and *mv* data that means that ConCovOpt can process data of any sample size with up to 23 imperfect pairs. In case of *fs* data, however, the computational limit tends to be reached at intermediate-*N* sample sizes of 70 to 80 cases with 10 to 15 imperfect *n*-tuples (see the replication script for a corresponding benchmark test). To calculate con-cov optima also for large-*N* *fs* data, heuristics are called for. In our R implementation, we use an approximation method that induces ConCovOpt to calculate fit scores only for rep-assignments with values closest to the outcome's median. This is an efficient approach for finding many, but possibly not all, con-cov optima. Based on this approximation method, large-*N* *fs* data of up to 2000 cases become processable.

Overall, even though heuristics are needed to calculate con-cov optima for large-*N* data and there may be no models realizing certain optima, processing *fs* data by ConCovOpt prior to actual CCM analyses has a considerable payoff. It identifies a set of consistency and coverage settings at which concrete models can be searched in a goal-oriented manner. Without applying ConCovOpt to Basurto's data, we would have been in the dark as to where to search for optimal models and would have had to proceed in an inefficient trial-and-error manner. With ConCovOpt, we straightforwardly found model (24), which has significantly better fit than Basurto's published model (21). Moreover, that fit improvement substantively alters the causal conclusions to be drawn from Basurto's (2013) study. A causal interpretation of (24) suggests that the endurance of autonomy only depends on local direct spending. Contrary to Basurto's findings, there is no evidence in his data that the other exogenous factors might be difference-makers of autonomy endurance as well.

7 Discussion

We end this paper by putting ConCovOpt into proper methodological perspective. Most importantly, ConCovOpt is not intended as a tool for a “simple hunt for high values of consistency and coverage” (Schneider and Wagemann 2012, 148). As emphasized in the introduction, there are other criteria of model selection. That is, high consistency and coverage are one asset of a model among others, and models with higher consistency and coverage should not be unequivocally preferred if they are outperformed by rival models in other respects.

What is more, the problem of model overfitting is still underinvestigated in configurational causal modeling, and there is ample evidence that CCMs have a strong tendency to overfit if they are induced to do so by overly high consistency and coverage thresholds (see e.g. Braumoeller 2015; Arel-Bundock 2019). One (heuristic) indication that overfitting might be taking place is that the complexity of resulting models increases disproportionately to their increase in model fit. This phenomenon regularly occurs if CCMs are ‘forced’ to build models reaching con-cov maxima—an example is provided in the replication script. What would hence be needed is a tool analogous to, say, the Akaike Information Criterion (AIC) in statistical modeling that strikes a balance between model fit and simplicity. A general preference of models reaching con-cov maxima is blind and hazardous.

Still, we have discussed various examples in this paper for which ConCovOpt has helped to significantly increase the model fit without an increase in model complexity, thus steering clear of overfitting dangers. Moreover, systematically scanning the model space at optimal consistency and coverage scores has led to the resolution of model ambiguities in case of Verweij and Gerrits’ (2015) as well as Basurto’s (2013) data; and it has even called into question the causal interpretability of a whole data set, *viz.* in case of the study by Britt et al. (2000). All of this shows that rendering con-cov optima and maxima transparent may importantly affect the causal conclusions drawn from configurational data.

Hence, ConCovOpt is intended as a tool for systematically exploring the space of CCM models at optimal consistency and coverage scores. In recent years, the awareness in the CCM literature has grown that the space of viable models for analyzed data may be larger than anticipated (cf. e.g. Baumgartner and Thiem 2017). When the data quality is very high, meaning when there are no imperfect pairs (or n -tuples), uncovering the whole model space is a matter of applying the relevant CCM once, at one determinate threshold setting. But in the presence of imperfect pairs, multiple CCM-runs at various threshold settings are needed.

Slight changes in thresholds may greatly change the models and there may be no systematicity in how threshold changes affect model changes. ConCovOpt efficiently uncovers con-cov optima and maxima prior to the application of CCMs. This information makes it possible to apply CCMs by directly constraining them towards optimal thresholds, without having to go through a whole array of threshold settings. Also, it can be determined how close actually obtained models come to con-cov optima and whether the obtained models exhaust the space of con-cov optima or whether further models should be searched at different settings.

All of this is of importance not only to analysts primarily interested in data-driven causal inference but also to analysts (of which there are many) viewing CCMs as tools for inferences primarily rooted in available case knowledge. When it comes to model selection, all types of CCM analysts have to take model fit into account. Even the analyst who wants to choose models based on case knowledge does not draw causal inferences from case knowledge alone. Rather, she analyzes data by means of a CCM in order to be presented with a set of models to choose from. The ultimate purpose of ConCovOpt and DNFbuild is to contribute to the completeness of that set.

In sum, models reaching con-cov optima will sometimes turn out to be the best models overall, sometimes not, and sometimes they will be of methodological interest even without being causally interpreted. But in one way or another, transparency on the model space at consistency and coverage optima is univocally valuable for configurational causal modeling.

References

- Arel-Bundock, V. (2019). The double bind of Qualitative Comparative Analysis. *Sociological Methods & Research*, 1–20. doi: 10.1177/0049124119882460.
- Basurto, X. (2013). Linking multi-level governance to local common-pool resource theory using fuzzy-set Qualitative Comparative Analysis: Insights from twenty years of biodiversity conservation in Costa Rica. *Global Environmental Change* 23, 573–587.
- Baumgartner, M. (2009). Inferring causal complexity. *Sociological Methods & Research* 38, 71–101.
- Baumgartner, M. and M. Ambühl (2020). Causal modeling with multi-value and fuzzy-set Coincidence Analysis. *Political Science Research and Methods* 8, 526–542.
- Baumgartner, M. and C. Falk (2019). Boolean difference-making: a modern regular-

- ity theory of causation. *The British Journal for the Philosophy of Science*. doi: 10.1093/bjps/axz047.
- Baumgartner, M. and A. Thiem (2017). Model ambiguities in configurational comparative research. *Sociological Methods & Research* 46(4), 954–987.
- Braumoeller, B. F. (2015). Guarding against false positives in Qualitative Comparative Analysis. *Political Analysis* 23(4), 471–487.
- Britt, D. W., S. T. Risinger, V. Miller, M. K. Mans, E. L. Krivchenia, and M. I. Evans (2000). Determinants of parental decisions after the prenatal diagnosis of down syndrome: Bringing in context. *American Journal of Medical Genetics* 93(5), 410–416.
- Cronqvist, L. and D. Berg-Schlosser (2009). Multi-value QCA (mvQCA). In B. Rihoux and C. C. Ragin (Eds.), *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques*, pp. 69–86. London: Sage Publications.
- Dusa, A. (2018). Consistency Cubes: A fast, efficient method for exact Boolean minimization. *The R Journal* 10(2), 357–370.
- Giugni, M. and S. Yamasaki (2009). The policy impact of social movements: A replication through Qualitative Comparative Analysis. *Mobilization: An International Quarterly* 14(4), 467–484.
- Graßhoff, G. and M. May (2001). Causal regularities. In W. Spohn, M. Ledwig, and M. Esfeld (Eds.), *Current Issues in Causation*, pp. 85–114. Paderborn: Mentis.
- Lemmon, E. J. (1965). *Beginning Logic*. London: Chapman & Hall.
- Lewis, D. (1973). Causation. *Journal of Philosophy* 70, 556–567.
- Mackie, J. L. (1974). *The Cement of the Universe. A Study of Causation*. Oxford: Clarendon Press.
- Ragin, C. C. (1987). *The Comparative Method*. Berkeley: University of California Press.
- Ragin, C. C. (2006). Set relations in social research: Evaluating their consistency and coverage. *Political Analysis* 14, 291–310.
- Ragin, C. C. (2008). *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago: University of Chicago Press.
- Rihoux, B. and G. De Meur (2009). Crisp-set Qualitative Comparative Analysis (csQCA). In B. Rihoux and C. C. Ragin (Eds.), *Configurational Comparative Methods: Qual-*

itative Comparative Analysis (QCA) and Related Techniques, pp. 33–68. Thousand Oaks: Sage.

Schneider, C. Q. and C. Wagemann (2012). *Set-Theoretic Methods: A User's Guide for Qualitative Comparative Analysis (QCA) and Fuzzy-Sets in the Social Sciences*. Cambridge: Cambridge University Press.

Suppes, P. (1970). *A Probabilistic Theory of Causality*. Amsterdam: North Holland.

Thiem, A. (2014). Unifying configurational comparative methods: Generalized-set Qualitative Comparative Analysis. *Sociological Methods & Research* 43(2), 313–337.

Thiem, A., M. Baumgartner, and D. Bol (2016). Still lost in translation! A correction of three misunderstandings between configurational comparativists and regressional analysts. *Comparative Political Studies* 49, 742–774.

Verweij, S. and L. M. Gerrits (2015). How satisfaction is achieved in the implementation phase of large transportation infrastructure projects: A Qualitative Comparative Analysis into the A2 tunnel project. *Public Works Management & Policy* 20(1), 5–28.