Advance Access publication December 24, 2015

Political Analysis (2016) 24:104–120
doi:10.1093/pan/mpv028

Enhancing Sensitivity Diagnostics for Qualitative Comparative Analysis: A Combinatorial Approach

Alrik Thiem
Department of Philosophy, University of Geneva,
Rue de Candolle 2|Bât. Landolt, 1211 Geneva, Switzerland
e-mail: alrik.thiem@unige.ch (corresponding author)

Reto Spöhel
Department of Engineering and Information Technology, Bern University of Applied Sciences,
Pestalozzistrasse 20, 3400 Burgdorf, Switzerland
e-mail: reto.spoehel@bfh.ch

Adrian Duşa
Department of Sociology, University of Bucharest,
Soseaua Panduri 90, 050663 Bucharest, Romania
e-mail: dusa.adrian@unibuc.ro

Edited by Jonathan Katz

Sensitivity diagnostics has recently been put high on the agenda of methodological research into Qualitative Comparative Analysis (QCA). Existing studies in this area rely on the technique of *exhaustive enumeration*, and the majority of works examine the reactivity of QCA either only to alterations in discretionary parameter values or only to data quality. In this article, we introduce the technique of *combinatorial computation* for evaluating the interaction effects between two problems afflicting data quality and two discretionary parameters on the stability of QCA reference solutions. In this connection, we challenge a hitherto unstated assumption intrinsic to exhaustive enumeration, show that combinatorial computation permits the formulation of general laws of sensitivity in QCA, and demonstrate that our technique is most efficient.

1 Introduction

The diffusion of Qualitative Comparative Analysis (QCA) throughout the social sciences continues to be accompanied by evaluations of the methodological properties of this relatively new procedure of causal inference.[1] Besides questions about issues concerning counterfactual assumptions (Schneider and Wagemann 2013; Baumgartner and Epple 2014; Baumgartner 2015), model ambiguities (Thiem and Duşa 2013a; Thiem 2014c; Baumgartner and Thiem 2015b), relations to regression and cluster analysis (Seawright 2005; Clark, Gilligan, and Golder 2006; Cooper and Glaesser 2011b; Vis 2012; Paine 2015; Thiem, Baumgartner, and Bol 2015), hypothesis testing (Braumoeller and Goertz 2000; Bol and Luppi 2013; Thiem 2014b; Braumoeller 2015), and causal complexity (Baumgartner 2009, 2013; Thiem 2015), the sensitivity of QCA has long been a topic of great interest to methodologists and applied users alike. For example, Goldthorpe (1997a, 7) had already conjectured almost two decades ago that "[i]f, on account of error in the

Authors' note: Supplementary materials for this article are available on the *Political Analysis* Web site (Thiem, Spöhel, and Duşa 2015). Previous versions of this article have been presented at the 1st and 2nd International QCA Expert Workshops, ETH Zurich, Switzerland. We thank Michael Baumgartner, Christian Rupietta, the participants at the aforementioned workshops, the editors of *Political Analysis*, and the three reviewers for their helpful comments.

[1]QCA has become an umbrella term by now for a family of configurational comparative methods. It currently subsumes four variants: crisp-set QCA, fuzzy-set QCA, multi-value QCA, and generalized-set QCA (Thiem 2014d).

original data, or in its treatment, even a single case happens to be placed on the 'wrong' side of a dichotomy, the analysis could well have a quite different outcome to that which would have been reached in the absence of such error."[2]

In response to these long-standing misgivings, several studies have recently begun to make attempts at evaluating the sensitivity of QCA in a more systematic manner (Skaaning 2011; Hug 2013; Bowers 2014; Lucas and Szatrowski 2014; Seawright 2014; Krogslund, Choi, and Poertner 2015). Research designs, methodological quality, and substantive conclusions have varied widely, but the *modus operandi* with respect to QCA's assessment has always been the same. After all solutions have been generated for a specific set or sets of artificial or empirical data under a series of controlled alterations in certain parameter values, the share of the reference solution among all identified solutions serves as a direct measure of how strongly the method reacts.

This article does not provide another sensitivity evaluation of QCA. Instead, its objective consists in scrutinizing the *modus operandi* of existing evaluations, and to develop an alternative approach. We call this new approach *combinatorial computation*.[3] The rationale behind the introduction of combinatorial computation is a set of three closely related observations with regard to existing studies: first, a questionable background assumption has been introduced but was never made explicit; second, a failure to identify mechanisms in the work flow of QCA that would have made the formulation of general laws of sensitivity possible has resulted from the exclusive focus on the mere syntactic structure of data-specific solutions; and third, the identification of final solutions has been unnecessarily resource-intensive.

By determining how solutions behave functionally in response to certain problems of empirical data analysis, we achieve three important goals with regard to the aforementioned observations: first, analyses of the effects of different assumptions respecting alterations in specific parameter values on the sensitivity of QCA are made possible; second, the formulation of general laws of sensitivity permits researchers to test hypotheses about solution stability in QCA; and third, the consumption of computational resources is drastically reduced to a tiny fraction of the conventional approach, whereby new opportunities for future methodological research arise.

The article is structured as follows. In Section 2, we provide an overview of the relevant literature and explain the research design in more detail. In Section 3, we motivate the development of a combinatorial approach to sensitivity diagnostics in relation to the work flow of QCA, and introduce all relevant theoretical concepts. Section 4 represents the main part, in which we derive the functional relations between QCA reference solutions and measurement error, the loss of data, respectively. We do so in relation to two different assumptions about the nature of these problems of empirical data analysis, and argue that one of them is considerably more plausible than the other. Section 5 concludes our study and identifies avenues for future research.

2 Research Design

Although evaluations of the sensitivity of QCA have become more systematical since Goldthorpe (1997a,b), researchers continue to carry out exploratory analyses, which have not been comparable across studies because they have drawn conclusions from specific data sets under disparate types of alterations in different parameters. Based on the categorization scheme provided by Thiem (2014a, 640), a few works can be classified as analyzing sensitivity to alterations in input parameters—aspects related to data quality (Hug 2013; Lucas and Szatrowski 2014) and the specification of the factor frame (Lucas and Szatrowski 2014; Krogslund, Choi, and Poertner 2015). The majority, however, deal with alterations in throughput parameters that are at the discretion of the researcher, including membership functions (Thiem 2014a), calibration thresholds (Seawright 2005; Skaaning 2011; Schneider and Wagemann 2012; Glaesser and Cooper 2014; Lucas and Szatrowski 2014; Krogslund, Choi, and Poertner 2015), inclusion cut-offs (Skaaning 2011; Schneider and Wagemann

[2]The performance of robustness tests with varying value assignments to contestable cases has long been common practice in applied QCA research (cf. Griffin et al. 1991, 130; Coverdill and Finlay 1995, 475). The macro-sociological study by Hicks, Misra, and Ng (1995, 341–2) is exemplary.

[3]We italicize concepts that are important to the content of this article at their first substantive appearance.

Table 1 Welfare state generosity among advanced industrial democratic countries

Block	*Country*	*Generosity of welfare state (W)*	*Strength of left parties (P)*	*Strength of unions (U)*	*Type of industrial system (C)*	*Sociocultural Homogeneity (S)*
1	Austria	1 (0.72)	1 (0.70)	1 (0.64)	1 (0.83)	1 (0.67)
	Denmark	1 (0.86)	1 (0.85)	1 (0.81)	1 (0.83)	1 (0.86)
	Finland	1 (0.76)	1 (0.56)	1 (0.86)	1 (0.83)	1 (0.72)
	Norway	1 (0.95)	1 (0.95)	1 (0.53)	1 (0.83)	1 (0.95)
	Sweden	1 (0.98)	1 (0.98)	1 (1.00)	1 (0.95)	1 (0.70)
2	Ireland	1 (0.67)	0 (0.11)	1 (0.63)	1 (0.67)	1 (0.84)
3	Belgium	1 (0.79)	1 (0.54)	1 (0.84)	1 (0.83)	0 (0.29)
4	Australia	0 (0.26)	0 (0.25)	0 (0.40)	0 (0.17)	0 (0.25)
	Canada	0 (0.26)	0 (0.00)	0 (0.06)	0 (0.05)	0 (0.10)
	France	0 (0.57)	0 (0.12)	0 (0.10)	0 (0.33)	0 (0.31)
	United States	0 (0.09)	0 (0.00)	0 (0.04)	0 (0.05)	0 (0.05)
5	Germany	0 (0.68)	0 (0.43)	0 (0.20)	1 (0.67)	0 (0.30)
	The Netherlands	0 (0.69)	0 (0.33)	0 (0.17)	1 (0.83)	0 (0.27)
	Switzerland	0 (0.53)	0 (0.34)	0 (0.13)	1 (0.67)	0 (0.10)
6	Japan	0 (0.52)	0 (0.00)	0 (0.04)	0 (0.33)	1 (0.95)
7	New Zealand	0 (0.56)	0 (0.40)	1 (0.54)	0 (0.17)	0 (0.15)

Notes: Original source (in brackets): Ragin (2000, 292, table 10.6); fuzzy sets dichotomized, Italy and United Kingdom dropped and coding changes on *W* for France, Japan, New Zealand, and Switzerland introduced by Grofman and Schneider (2009, 663); coding typos on *W* for Germany and the Netherlands by Hug (2013, 258).

2012; Krogslund, Choi, and Poertner 2015), and frequency cut-offs (Skaaning 2011; Lucas and Szatrowski 2014; Krogslund, Choi, and Poertner 2015).[4]

While much progress has been achieved in analyzing the effects of missing data and measurement error in other, more established methodological frameworks such as regression analysis (e.g., Wooldridge 2002, 70–81; Gelman and Hill 2007, 529–43), Hug (2013) presents the only serious attempt so far with respect to QCA. Following Hug, our input parameters include measurement error in the endogenous factor and the loss of data due to the list-wise deletion of cases. We also use the same data set for the purpose of illustration, and employ crisp-set QCA. Besides ensuring a cumulative generation of knowledge by continuity in basic design aspects, the advantage of crisp-set QCA is that it has been more established, unlike fuzzy-set QCA, many of whose properties are still controversially debated (e.g., Eliason and Stryker 2009; Cooper and Glaesser 2011a). Contrary to Hug, however, we allow the two problems in the quality of data to interact with fundamental parameters in QCA that are at the discretion of the analyst. We add the inclusion cut-off and the frequency cut-off as two important throughput parameters so as to allow for the possibility that the latter group of parameters amplifies or attenuates the effect of the former. Moreover, we cover both the conservative solution type as well as the parsimonious one. So as to set the stage for the remainder of this article, a brief recapitulation and discussion of Hug's (2013) study is in order.

Using the data set in Table 1, Hug first replicates the results from Grofman and Schneider (2009) by identifying the conditions under which countries exhibit a generous welfare state.[5] Among the factors hypothesized to be associated with certain generosity levels of a welfare state (*W*: 1 generous, 0 not generous) are the strength of left parties (*P*: 1 strong, 0 not strong) and the strength of unions (*U*: 1 strong, 0 not strong), the type of industrial system (*C*: 1 corporatist, 0

[4]An overview of some of these issues with a different categorization scheme is provided by Maggetti and Levi-Faur (2013). The concept of *inclusion* is synonymous with what is commonly known in QCA as *consistency*.

[5]This data set had originally been presented by Ragin (2000, 292). Afterward, it was deliberately modified a first time by Grofman and Schneider (2009, 663) (variables dichotomized; Italy and the United Kingdom dropped; outcome values of France, Japan, New Zealand, and Switzerland changed), and then accidentally a second time by Hug (2013, 258) (coding typos for Germany and the Netherlands). Also note that Hug says he dropped Australia and Italy (p. 258), but it was Italy and the United Kingdom.

Table 2 Solution details for outcome $W^{\{1\}}$

Outcome	*Path*	*Inc*	Cov_r	Cov_u	*Cases*
$W^{\{1\}}$	$P^{\{1\}}U^{\{1\}}C^{\{1\}}$	1.000	0.857	0.143	BE; AT, DK, FI, NO, SE
	$U^{\{1\}}C^{\{1\}}S^{\{1\}}$	1.000	0.857	0.143	IE; AT, DK, FI, NO, SE
S_1	m_1	1.000	1.000		

not corporatist), as well as the presence of sociocultural homogeneity (S: 1 yes, 0 no). The target population is comprised of the set of advanced industrial democratic countries.[6]

The conservative solution for these data, S_1, consists of a single model, m_1, which is presented in Table 2. It says that the conjunction of strong left parties, strong unions, and a corporatist industrial system and the conjunction of strong unions, a corporatist industrial system, and sociocultural homogeneity are alternative causal paths for the existence of a generous welfare state.[7] Moreover, these two paths are also necessary for the outcome. Whenever at least one of them changed or disappeared, or when another path joined the model, the solution was classified by Hug as being different from the reference solution.[8]

For analyzing the sensitivity of QCA, Hug implements the following procedure. In a first pair of analyses, the volume of the data is reduced by deleting cases; in the second pair of analyses, the data volume is held constant but the endogenous factor is corrupted. More precisely, in the first test series, each of the $\binom{16}{1} = 16$ cases in Table 1 is systematically deleted before the solution is generated; in the second series, each of the $\binom{16}{2} = 120$ possible pairs of cases is deleted; in the third test series, the value on the endogenous factor for each of the 16 cases is systematically corrupted from 1 to 0 and *vice versa*; and in the fourth series, the values on the endogenous factor for each of the 120 possible pairs of cases are systematically corrupted from 1 to 0 and *vice versa*.

Although no stochastic component is involved, Hug calls his approach "seemingly Monte Carlo simulation" (p. 257).[9] More appropriately, it is known as *exhaustive enumeration* or *brute-force method* because all unique possibilities for changing a given number of parameter values in the analysis of a set of outputs of interest are realized in a systematic and controlled way (cf. Nievergelt 2000). In the remainder of this article, we contrast the method of exhaustive enumeration, which has also been applied in various variations by Krogslund, Choi, and Poertner (2015), Lucas and Szatrowski (2014), and Skaaning (2011), with combinatorial computation to argue that the latter outperforms the former in all relevant respects.

3 Preparatory Explanations and Definitions

In this section, we first spell out the succession of phases in QCA to motivate the introduction of combinatorial computation. Subsequently, all concepts relevant to realizing this method mathematically are defined.

3.1 *Methods of Sensitivity Diagnostics and the Work Flow of QCA*

So as to understand why we argue that combinatorial computation outperforms exhaustive enumeration and permits the formulation of general laws of sensitivity, it is expedient to begin by

[6]We use curly-bracket notation instead of upper-/lowercase notation because it is unambiguous.
[7]We have used the package QCA 1.1-4 for the R environment to regenerate the solution (Thiem and Duşa 2013b,c; Duşa and Thiem 2014; R Development Core Team 2014).
[8]Note that two different QCA models m_1 and m_2 derived from two different data sets δ_1 and δ_2 are compatible if the causal claims entailed by m_1 and m_2 stand in a subset relation. This important fact has not been taken into consideration by any of the evaluations cited above. Instead, different solutions have been treated as if they were incompatible.
[9]Generally, in a Monte Carlo simulation, n randomly changed data sets would be created, QCA would be run on each of them, and the number of times x in which the reference solution was retained would be counted, with x/n representing the estimate for the unknown probability of retaining the reference solution. Depending on the desired confidence level, fewer or more analyses would have to be performed.

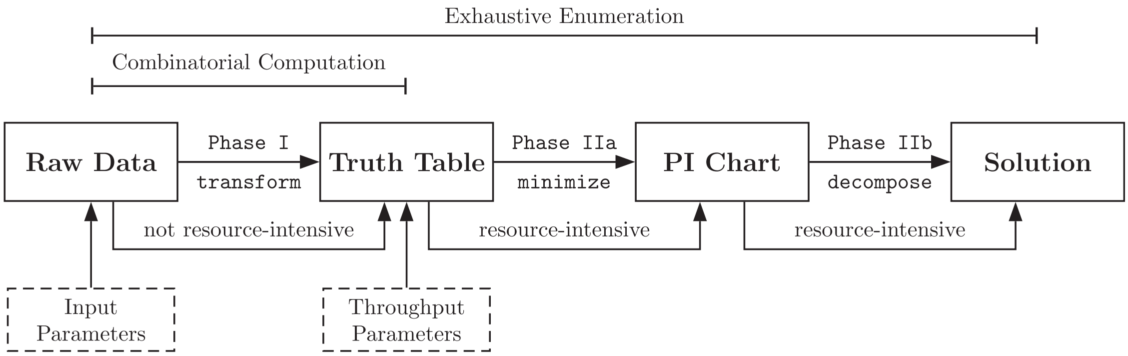


Fig. 1 Flow chart of QCA, approaches to calculating retention probabilities of QCA reference solutions, and resource intensity at each phase.

Table 3 Truth table for data in Table 1

Block	*P*	*U*	*C*	*S*	*OUT*	*Cases*
1	1	1	1	1	1	AT, DK, FI, NO, SE
2	0	1	1	1	1	IR
3	1	1	1	0	1	BE
4	0	0	0	0	0	AU, CA, FR, US
5	0	0	1	0	0	CH, DE, NL
6	0	0	0	1	0	JP
7	0	1	0	0	0	NZ
8–16	...	...	...	...	?	—

Note: OUT = output value.

situating either approach in the work flow of QCA, whose procedural protocol has three phases: the transformation of the raw data into a *truth table* in Phase I; the *minimization* of the function described by this truth table for constructing a *prime implicant chart* in Phase IIa; and the decomposition of this chart into a set of all equally well-fitting models, called *solution*, in Phase IIb. Phases IIa and IIb are usually treated as inextricable sub-phases of that stage in which all redundancies are eliminated to yield causally interpretable *models*. The flow chart of this protocol is visualized in Fig. 1.

For arriving at the truth table, one factor in the data must be declared the *endogenous factor*. All other factors remain as *exogenous factors*. For each unique combination of levels across all these exogenous factors—called a *min-term*—an *output value* is determined, based on the proposition that the min-term is sufficient for that level of the endogenous factor which has been set as the *outcome*. If a min-term has not been instantiated as often as required, it is classified as a *remainder* and assigned the output value "?"; if it has been instantiated as often as required but the proposition turns out to be false, it is classified as *negative* and receives the output value "0"; and if it has been instantiated as often as required and the proposition turns out to be true, it is classified as *positive* and assigned the output value "1". Each truth table thus consists of a matrix of min-terms to which a column of output values is appended. The truth table for the data set in Table 1 is presented in Table 3. The last row of this table subsumes the nine remainders.

A truth table such as Table 3 provides the input to Phase II when minimization proper commences. This process is usually associated with the Quine–McCluskey (QMC) algorithm (cf. Duşa and Thiem 2015, 94–97).[10] In Phase IIa of QMC, all redundant factors are eliminated by pairing two positive min-terms, and subsequently their descendants, or a positive min-term with a remainder and their descendants if the parsimonious solution type is used. When elimination is no longer

[10]Algorithms other than QMC are in use, but fs/QCA (Ragin and Davey 2014)—currently the most popular QCA software by far—implements QMC with only minor modifications.

possible, the surviving terms, called *prime implicants*, are set in relation to all positive min-terms in a prime implicant chart. This chart is then decomposed to yield all minimally necessary sets of prime implicants that cover the set of all positive min-terms.

For producing the conservative solution, QCA forces QMC to treat remainders as negative min-terms. The model set generated by this particular solution type therefore stands in a one-to-one relation with the set of positive min-terms. In other words, two truth tables based on the same data lead to the same conservative solution set of models if, and only if, they have the same set of positive min-terms.[11] More generally, alterations in input parameters which lead to changes in the data that do not affect the truth table will not affect the QCA solution, irrespective of the solution type. Likewise, alterations in throughput parameters that do not affect the truth table will not affect the QCA solution, irrespective of the solution type.

To arrive at the solution, exhaustive enumeration implements the full procedural protocol of QCA, including the resource-intensive Phases IIa and IIb. However, as a set of models obtained for a given set of data depends only on the truth table derived from those data, and in particular only on the set of positive min-terms for the case of the conservative solution type, phases IIa and IIb need not be executed if the goal is to find the probability to which such a reference solution will remain unchanged. This observation is decisive, and it informs a combinatorial approach to sensitivity diagnostics in QCA.

3.2 *Definitions*

While Section 3.1 sketched out the rationale behind an alternative to exhaustive enumeration, it is expedient at this juncture to formalize all relevant concepts in order to realize such an alternative mathematically. First, let δ denote a given data set, $\mathbf{T}_\delta$ the truth table resulting from δ, and $\mathscr{S} = \{m_1, m_2, \ldots, m_j\}$ the reference solution set of models derived from $\mathbf{T}_\delta$ after Phase IIb. In agreement with previous evaluation practice, we consider a solution $\mathscr{S}^*$ as being different to some reference solution $\mathscr{S}$ if the former is not identically equal with the latter. Furthermore, let $\theta_{\mathscr{S}}$ stand for the retention probability of $\mathscr{S}$. For example, δ could be the data presented in Table 1, $\mathbf{T}_\delta$ the truth table in Table 3, and $\mathscr{S}$ the solution S_1 in Table 2, with m_1 being its sole element.

As $\mathscr{S}$ stands in a one-to-one relation with the set of positive min-terms in $\mathbf{T}_\delta$ under the conservative solution type, the retention probability $\theta_{\mathscr{S}}$ is the probability that this set remains unchanged. In the following, when we say that a min-term is *included* in $\mathscr{S}$, we thus mean that it must be positive in $\mathbf{T}_\delta$; when a min-term *enters* $\mathscr{S}$, we mean that it changes from negative to positive; and when it *leaves* $\mathscr{S}$, it changes from positive to negative, or to a remainder.

We denote the number of cases (rows) in δ by n, the number of exogenous Boolean factors in δ by k, the number of cases of min-term ℓ in δ that exhibit the outcome of interest, $F_o^{\{1\}}$, by $c_\ell^\oplus$, and the number of cases of min-term ℓ in δ that exhibit the negation of the outcome of interest, $F_o^{\{0\}}$, by $c_\ell^\ominus$. $F_o^{\{1\}}$ and $F_o^{\{0\}}$ are the two levels of the endogenous Boolean factor F_o. Applied to the concrete data example in Table 1 again, $n = 16$ as there are 16 rows, $k = 4$ because P, U, C, and S are the exogenous Boolean factors, and $W^{\{1\}}$ represents the level of the endogenous factor W for which the analysis is to be carried out. Accordingly, $W^{\{0\}}$ is the negation of $W^{\{1\}}$ and *vice versa*. Finally, block 1 in Table 3 would be associated with $\ell = 1$, with $c_1^\oplus = 5$ and $c_1^\ominus = 0$ because this min-term shows five cases of $W^{\{1\}}$ but none of $W^{\{0\}}$.

The derivation of $\mathbf{T}_\delta$ from δ also depends on two important throughput parameters. Let α denote the inclusion cut-off that determines the ratio of cases of $F_o^{\{1\}}$ to those of $F_o^{\{0\}}$ below which some min-term ℓ is not positive anymore, and let β denote the frequency cut-off that determines the bound of observed cases of some min-term ℓ, irrespective of their level on F_o, below which ℓ is classified as a remainder.[12] In addition, let m denote the number of min-terms in $\mathbf{T}_\delta$ that are not remainders.[13]

[11] Note that not all computer programs for QCA identify the full model set of a solution, in consequence of which these sets may not be the same even though truth tables are (Thiem and Duşa 2013a; Thiem 2014c; Baumgartner and Thiem 2015b).

[12] $0.5 < \alpha \leq 1$, $\beta \in \mathbb{N}$, and $1 \leq \beta \leq \max_{1 \leq \ell \leq m}(c_\ell^\oplus + c_\ell^\ominus)$, $\mathbb{N}$ being the set of all non-negative integers.

[13] Thus, m depends on δ as well as β, and $m \leq 2^k$. If $m = 2^k$, $\mathbf{T}_\delta$ contains no remainders and is said to be *saturated*.

A min-term ℓ is thus included in $\mathscr{S}$ if, and only if, $c_\ell^{\oplus}+c_\ell^{\ominus}\geq\beta$ and $c_\ell^{\oplus}\geq\alpha(c_\ell^{\oplus}+c_\ell^{\ominus})$. For notational convenience, let A_0 in Equation (1), A_1 in Equation (2), and $\mathrm{A}_?$ in Equation (3) denote all pairs of integers that lead to negative min-terms, positive min-terms, and remainders, respectively:

$$\mathrm{A}_0 \stackrel{\text{def}}{=} \{(c^{\oplus},c^{\ominus})\in\mathbb{N}^2 \mid c^{\oplus}+c^{\ominus}\geq\beta \text{ and } c^{\oplus}<\alpha(c^{\oplus}+c^{\ominus})\}, \tag{1}$$

$$\mathrm{A}_1 \stackrel{\text{def}}{=} \{(c^{\oplus},c^{\ominus})\in\mathbb{N}^2 \mid c^{\oplus}+c^{\ominus}\geq\beta \text{ and } c^{\oplus}\geq\alpha(c^{\oplus}+c^{\ominus})\}, \tag{2}$$

$$\mathrm{A}_? \stackrel{\text{def}}{=} \{(c^{\oplus},c^{\ominus})\in\mathbb{N}^2 \mid c^{\oplus}+c^{\ominus}<\beta\}. \tag{3}$$

A min-term ℓ is then included in $\mathscr{S}$ if, and only if, $(c_\ell^{\oplus},c_\ell^{\ominus})\in\mathrm{A}_1$, and it leaves $\mathscr{S}$ as the result of corrupting i cases with outcome $F_o^{\{1\}}$ and j cases with outcome $F_o^{\{0\}}$ if, and only if, $(c_\ell^{\oplus}-i+j, c_\ell^{\ominus}-j+i)\notin\mathrm{A}_1$.

Generally, we call a change in the quality of δ as a result of measurement error through a corruption on F_o or the loss of data through the deletion of cases from δ a *perturbation*. If a method of sensitivity analysis generates perturbations independently of each other, we speak of an *independence-in-perturbation assumption* (IPA). Conversely, if perturbations are not generated independently of each other but are tied to a fixed number of cases *ex ante*, we speak of a *dependence-in-perturbation assumption* (DPA). DPA has been the implicit background assumption in all previous sensitivity evaluations that have used exhaustive enumeration. For example, in Hug (2013), exactly one/two case/s out of sixteen was/were corrupted, and one/two case/s out of sixteen was/were deleted.

4 Combinatorial Computation under DPA and IPA

In this section, we show how exhaustive enumeration could be replaced by combinatorial computation while maintaining DPA as the principal assumption about the dependencies between perturbations. However, we eventually criticize this assumption for lacking plausibility. We propose another variant of combinatorial computation which is based on IPA. While we develop this method in detail with respect to the conservative solution type of QCA, we show that combinatorial computation can be extended to parsimonious solutions, with some limitations for the case of data loss. Finally, we present the results of performance tests for all three methods—exhaustive enumeration, combinatorial computation under DPA, and combinatorial computation under IPA.

4.1 *Replacing Exhaustive Enumeration by Combinatorial Computation*

The method of exhaustive enumeration is based on the assumption of a fixed number of perturbations, what we have defined as DPA in Section 3. Given DPA, how could a combinatorial approach replace exhaustive enumeration and avoid having to pass through the resource-intensive Phases IIa and IIb of QCA's procedural protocol?

Before going into the formalities, let us begin with an example by way of developing a conception of the problem. Suppose we wanted to compute the number of times in which corrupting the value of exactly two cases on the endogenous factor in the data given in Table 1 would lead to a solution different from the reference solution S_1 when $\alpha=0.75$ and $\beta=1$. For these throughput parameters, the reference solution would change if, and only if, either two cases from block 1 were corrupted or at least one of Ireland, Belgium, Japan, and New Zealand was corrupted.

There are $\binom{5}{2}=10$ pairs of cases that correspond to the first possibility. To compute the number of pairs of cases corresponding to the second possibility, we can observe that there are fifteen pairs of cases that contain Ireland, fifteen pairs that contain Belgium, fifteen pairs that contain Japan, and fifteen pairs that contain New Zealand. This seems to suggest that, altogether, there are $4\cdot 15=60$ pairs of cases that correspond to the second possibility. However, all pairs that contain two of the four problematic cases mentioned above have been counted twice in this calculation. The exact number of pairs of cases is only $4 \cdot 15 - \binom{4}{2} = 54$. As both possibilities are mutually exclusive, the two corresponding numbers can be added to find that exactly $10 + 54 = 64$ pairs of cases would lead to a change in the reference solution if the values of two cases on the endogenous factor were corrupted.

We now present an abstraction of the above example in relation the conservative solution type. With respect to measurement error in the endogenous factor, DPA implies that F_o is corrupted for exactly D out of n cases in δ. A number of D such perturbations lead to a solution different from $\mathscr{S}$ if, and only if, they change the set of positive min-terms in $\mathbf{T}_\delta$. Informally, the probability $\theta_{\mathscr{S}}$ that $\mathscr{S}$ is retained is thus given by Equation (4):

$$\theta_{\mathscr{S}} = \frac{\text{number of sets of corruptions on } D \text{ out of } n \text{ cases that do not change the solution}}{\text{number of sets of corruptions on } D \text{ out of } n \text{ cases}}. \tag{4}$$

In the following, when saying that a particular set of corruptions on D cases *affects* a particular min-term ℓ, we mean that corrupting these D cases on F_o will cause ℓ to either enter $\mathscr{S}$ if it was not included before or leave $\mathscr{S}$ if it was previously included.

The denominator of Equation (4) is a simple binomial coefficient; more complex is the numerator. To count these sets, the *inclusion–exclusion principle* must be invoked (cf. Hohn 1966, 261–63). Let $q^D_{\mathscr{S}}$ denote the number of sets of corruptions on D cases that change $\mathscr{S}$, and q^D_ℓ the number of sets of corruptions on D cases that affect ℓ. More generally, let $q^D_{\ell_1,\ldots,\ell_t}$ denote the number of sets of corruptions on D cases that affect the t min-terms $\ell_1, \ell_2, \ldots, \ell_t$ simultaneously. Exploiting the inclusion–exclusion principle, $q^D_{\mathscr{S}}$ is then given by Equation (5):

$$q^D_{\mathscr{S}} = \sum_{\ell} q^D_\ell - \sum_{\{\ell_1,\ell_2\}} q^D_{\ell_1,\ell_2} + \sum_{\{\ell_1,\ell_2,\ell_3\}} q^D_{\ell_1,\ell_2,\ell_3} - \ldots + \ldots - \ldots \pm \sum_{\{\ell_1,\ldots,\ell_D\}} q^D_{\ell_1,\ldots,\ell_D}, \tag{5}$$

where each sum runs over all unordered subsets of all m non-remainder min-terms. For example, the second sum has exactly $\binom{m}{2}$ terms and the last sum has exactly $\binom{m}{D}$ terms. To get some intuition for why Equation (5) is true, note Equation (6):

$$q^D_{\mathscr{S}} \leq \sum_{\ell} q^D_\ell, \tag{6}$$

which holds with equality if there is no set of corruptions on D cases that affects two or more min-terms at once. If there is such a set, however, then it is counted twice on the right-hand side but only once on the left-hand side of Equation (6), causing its statement to hold with strict inequality. Subtracting the term starting with the second sum in Equation (5) corrects for this possible over-counting. Yet, if there is at least one set of corruptions on D cases that affects three or more min-terms, then this correction term introduces a similar effect in the opposite direction. The term starting with the third sum in Equation (5) corrects for this possible under-counting, and so forth.[14]

Consider now a fixed min-term ℓ, of which there are $c^{\oplus}_\ell$ cases of $F^{\{1\}}_o$ and $c^{\ominus}_\ell$ cases of $F^{\{0\}}_o$ in δ. If $(c^{\oplus}_\ell, c^{\ominus}_\ell) \in \mathrm{A}_1$, then ℓ is currently included in $\mathscr{S}$, and the number of sets of corruptions on D cases that affect it (corrupting the outcome of all cases in such a set causes ℓ to leave $\mathscr{S}$) is given by Equation (7):

$$q^{\triangleright}_\ell(c^{\oplus}_\ell; c^{\ominus}_\ell; D) = \sum_{\substack{(i,\,j):\\ i+j \leq D;\ (c^{\oplus}_\ell - i + j,\, c^{\ominus}_\ell - j + i) \notin \mathrm{A}_1}} \binom{c^{\oplus}_\ell}{i} \binom{c^{\ominus}_\ell}{j} \binom{n - c^{\oplus}_\ell - c^{\ominus}_\ell}{D - i - j}, \tag{7}$$

[14]No set of corruptions on D cases can affect more than D min-terms.

where i ranges from 0 to $c_\ell^\oplus$ and j ranges from 0 to $c_\ell^\ominus$.[15] Each summand is the number of ways to choose a set of D cases that contains exactly i from the $c_\ell^\oplus$ cases of ℓ with $F_o^{\{1\}}$, exactly j from the $c_\ell^\ominus$ cases of ℓ with $F_o^{\{0\}}$, and therefore exactly $D-i-j$ cases from the remaining $n-c_\ell^\oplus-c_\ell^\ominus$ cases in δ.

Similarly, if $(c_\ell^\oplus, c_\ell^\ominus)\in \mathrm{A}_0$, then ℓ is currently not included in $\mathscr{S}$, and the number of sets of corruptions on D cases that affect it (corrupting the outcome of all cases in such a set causes ℓ to enter $\mathscr{S}$) is given by Equation (8):

$$q_\ell^{\triangleleft}(c_\ell^\oplus; c_\ell^\ominus; D) = \sum_{\substack{(i,\,j):\\ i+j\le D;\ (c_\ell^\oplus - i + j,\, c_\ell^\ominus - j + i)\in \mathrm{A}_1}} \binom{c_\ell^\oplus}{i}\binom{c_\ell^\ominus}{j}\binom{n-c_\ell^\oplus-c_\ell^\ominus}{D-i-j}. \tag{8}$$

Combining these two cases, Equation (9) is obtained:

$$\sum_\ell q_\ell^D = \sum_{\substack{1\le \ell\le m:\\ (c_\ell^\oplus, c_\ell^\ominus)\in \mathrm{A}_1}} q_\ell^{\triangleright} + \sum_{\substack{1\le \ell\le m:\\ (c_\ell^\oplus, c_\ell^\ominus)\in \mathrm{A}_0}} q_\ell^{\triangleleft}. \tag{9}$$

By extension, the terms of the next large sum in Equation (5) are given by Equation (10):

$$q_{\ell_1,\ell_2}^D = \sum_{\substack{(i_1, j_1, i_2, j_2):\\ \text{see text}}} \binom{c_{\ell_1}^\oplus}{i_1}\binom{c_{\ell_1}^\ominus}{j_1}\binom{c_{\ell_2}^\oplus}{i_2}\binom{c_{\ell_2}^\ominus}{j_2}\binom{n-c_{\ell_1}^\oplus-c_{\ell_1}^\ominus-c_{\ell_2}^\oplus-c_{\ell_2}^\ominus}{D-i_1-j_1-i_2-j_2}, \tag{10}$$

where the inner sum runs over all choices for the indices (i_1, j_1, i_2, j_2) such that $i_1+j_1+i_2+j_2\le D$ and for $t=1,\ 2$ it holds that if $(c_{\ell_t}^\oplus, c_{\ell_t}^\ominus)\in \mathrm{A}_1$, then i_t and j_t must satisfy $(c_{\ell_t}^\oplus - i_t + j_t,\, c_{\ell_t}^\ominus - j_t + i_t)\notin \mathrm{A}_0$; and if $(c_{\ell_t}^\oplus, c_{\ell_t}^\ominus)\in \mathrm{A}_0$, then i_t and j_t must satisfy $(c_{\ell_t}^\oplus - i_t + j_t,\, c_{\ell_t}^\ominus - j_t + i_t)\in \mathrm{A}_1$.[16]

By the same token yet more generally, the individual summands of the q-th term in Equation (5) are given by Equation (11):

$$q_{\ell_1,\ldots,\ell_q}^D = \sum_{\substack{(i_1, j_1, \ldots, i_q, j_q):\\ \text{see text}}} \left(\prod_{t=1}^{q}\binom{c_{\ell_t}^\oplus}{i_t}\binom{c_{\ell_t}^\ominus}{j_t}\right)\cdot\binom{n-\sum_{t=1}^{q}(c_{\ell_t}^\oplus + c_{\ell_t}^\ominus)}{D-\sum_{t=1}^{q}(i_t+j_t)}, \tag{11}$$

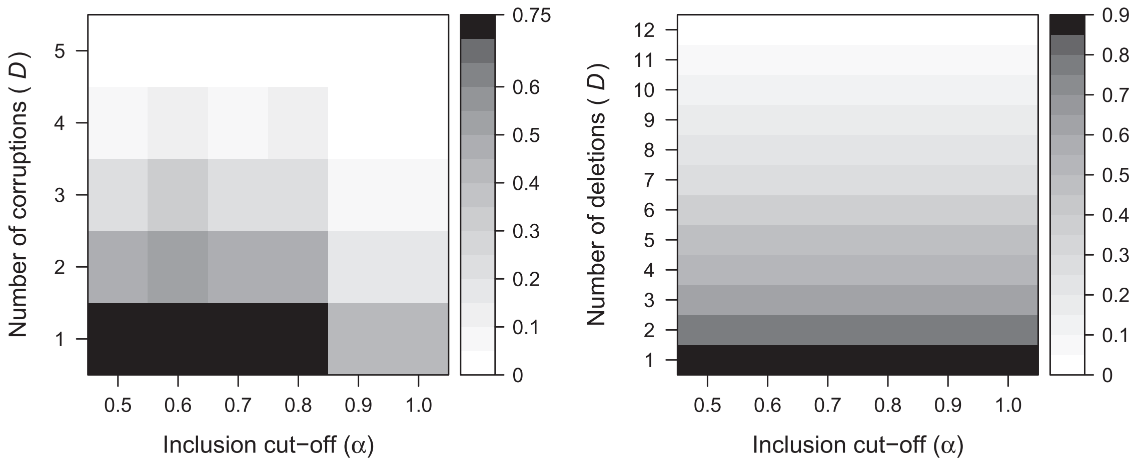
where the inner sum runs over all choices for the indices $(i_1, j_1, \ldots, i_q, j_q)$ such that $\sum_{t=1}^{q}(i_t+j_t)\le D$ and for each $t=1,\ldots,q$ the respective condition from the two conditions given above holds. Plugging Equation (11) into Equation (5) (for $q=1,\ldots,D$), and observing that the numerator of Equation (4) is $\binom{n}{D}$ minus the inclusion–exclusion term $q_{\mathscr{S}}^D$ given in Equation (5), we obtain a general method for computing the retention probability $\theta_{\mathscr{S}}$ without resorting to exhaustive enumeration.

The issue of data loss, where a fixed number D of cases are deleted from δ, can be treated in a similar fashion. To obtain $\theta_{\mathscr{S}}$, $(c_\ell^\oplus - i + j, c_\ell^\ominus - j + i)$ in Equations (7) and (8) only need to be replaced by $(c_\ell^\oplus - i, c_\ell^\ominus - j)$.[17] To summarize at this point: we have derived functional laws

[15]Here, i represents the number of cases of min-term ℓ whose value on the endogenous factor is corrupted from $F_o^{\{1\}}$ to $F_o^{\{0\}}$, and j represents the number of cases of min-term ℓ whose value on the endogenous factor is corrupted from $F_o^{\{0\}}$ to $F_o^{\{1\}}$. Consequently, $c_\ell^\oplus - i + j$ represents the number of cases of min-term ℓ still showing $F_o^{\{1\}}$ post-corruption. In the condition below the sum sign, instead of $(\ldots,\ldots)\notin \mathrm{A}_1$, we could also have written $(\ldots,\ldots)\in \mathrm{A}_0$, as a positive min-term that leaves $\mathscr{S}$ due to measurement error necessarily becomes a negative min-term, and not a remainder. As we shall see later, stating the formulas in the way we do here has the advantage of making it easier to adapt them to the case of data loss when positive min-terms can become remainders.

[16]This condition guarantees that ℓ_1 and ℓ_2 are indeed affected by corrupting both types of cases in each min-term as given by i_1, j_1, i_2, j_2.

[17]Before, $c_\ell^\oplus - i + j$ denoted the number of cases of min-term ℓ with positive outcome after corrupting the outcome of i positive and j negative cases. Now, $c_\ell^\oplus - i$ denotes the number of cases of min-term ℓ with positive outcome after deleting i positive and j negative cases.



(a) Measurement error **(b)** Loss of data

Fig. 2 Retention probabilities under DPA for data in Table 1.

between conservative QCA solutions and two problems affecting the quality of data in interaction with two fundamental throughput parameters. Contrary to Schneider and Wagemann (2012, 294), who argue that "it is difficult to formulate general laws of robustness in QCA," we have demonstrated that such laws can indeed be formulated; and they can be applied to any data set.

For instance, going back to our example data set in Table 1, Fig. 2 plots the retention probability of the QCA reference solution from Table 2 as a function of the inclusion cut-off α and the number of corruptions D in panel 2a and the number of deletions D in panel 2b, for $\beta = 1$.[18] Inclusion cut-offs range quasi-continuously from 0.5 to 1 as cut-offs below 0.5 imply that for some min-term more cases exist that do not show the outcome of interest than those that do. It is not reasonable to classify such min-terms as positive, as Krogslund, Choi, and Poertner (2015, 34) have done. So as to cover the full spectrum of retention probabilities down to under 5%, the number of corruptions ranges from 1 to 5, and the number of deletions from 1 to 12.

For all inclusion cut-offs below 0.81 and $D = 1$, the probability that S_1 is retained amounts to 75%.[19] As corruptions proliferate, the retention probability decreases significantly, but not uniformly so. At inclusion cut-offs between 0.61 and 0.66, it is smaller than at any other value below 0.8. At three corruptions, it does not even reach 20%. As higher inclusion cut-offs generally put more demands on the quality of the data for keeping the reference solution intact, probabilities above an inclusion cut-off of 0.8 are lower than anywhere else for each number of corruptions. The reason why they are invariant above this value lies in the simple fact that the most populated min-term, block 1 in Table 3, contains five cases. As soon as any single case out of these five is corrupted, the empirical inclusion score will drop to 0.8.

The pattern in Fig. 2b deviates from that in Fig. 2a. In particular, two differences are noticeable. First, retention probabilities do not vary with inclusion cut-offs but only with the number of deletions. The reason is that the original data in Table 1 are such that no block which corresponds to a positive min-term in Table 3 contains a case of $W^{\{0\}}$. In other words, no deletion affects any positive min-term's empirical inclusion score. And second, retention probabilities for data loss deteriorate at a much smaller rate than for measurement error. While five corruptions cause the retention probability to drop below 5% across the entire range of inclusion cut-offs, only at 12 deletions does data loss lead to similarly low figures.

[18]Note that the color range has been scaled for each figure individually. We have used lattice 0.20–29 to produce Figs. 2 and 3 in Section 4.2 (Sarkar 2008).
[19]For comparison, see table 4 in Hug (2013, 260).

While combinatorial computation affords clear advantages over exhaustive enumeration in terms of computational efficiency due to its avoidance of Phases IIa and IIb of QCA's procedural protocol, it still relies on DPA. But how plausible is DPA from an empirical point of view? After all, knowing that one specific case is perturbed usually neither increases nor decreases the conditional probability that one of the remaining cases is perturbed, yet the assumption of a fixed number of errors entails exactly such an effect. For $D=2$, perturbing one particular case out of n cases decreases the conditional probability that any other given case is perturbed from originally $2/n$ to $1/(n-1)$ if each case has the same *a priori* likelihood of experiencing a perturbation. However, why should this be the case unless one has definite knowledge about the presence of such dependencies? So as to avoid these strong implications, we propose a variant of combinatorial computation that substitutes IPA for DPA, in which each case is perturbed with some fixed probability *independently* of what happens to all other cases. As it turns out, IPA is not only more realistic than DPA, but combinatorial computation on the basis of the former instead of the latter is also more easily tractable from a mathematical point of view and more efficient still.

4.2 *Combinatorial Computation with IPA: Conservative Solutions*

For the case of measurement error, F_o is corrupted independently under IPA with probability p for each of the n cases in δ. Conceptually, this is similar to corrupting exactly a p-fraction of all n cases in δ if np is an integer, but it is, as argued above, more plausible from an empirical point of view. For a given min-term ℓ, the probability π_ℓ that exactly i of its $c_\ell^\oplus$ cases of $F_o^{\{1\}}$ and exactly j of its $c_\ell^\ominus$ cases of $F_o^{\{0\}}$ are corrupted is given by Equation (12)[20]:

$$\pi_\ell(i, c_\ell^\oplus; j, c_\ell^\ominus; p) = \binom{c_\ell^\oplus}{i} p^i (1-p)^{c_\ell^\oplus - i} \cdot \binom{c_\ell^\ominus}{j} p^j (1-p)^{c_\ell^\ominus - j}. \tag{12}$$

Consider now a fixed min-term ℓ, of which there are $c_\ell^\oplus$ cases of $F_o^{\{1\}}$ and $c_\ell^\ominus$ cases of $F_o^{\{0\}}$ in δ. If $(c_\ell^\oplus, c_\ell^\ominus) \in \mathrm{A}_1$, then ℓ is currently included in $\mathscr{S}$, and the probability $\pi_\ell^\triangleright$ that it leaves $\mathscr{S}$ in consequence of corruptions on F_o is given by Equation (13):

$$\pi_\ell^\triangleright(c_\ell^\oplus; c_\ell^\ominus; p) = \sum_{\substack{(i,\, j):\\ (c_\ell^\oplus - i + j,\ c_\ell^\ominus - j + i) \notin \mathrm{A}_1}} \pi_\ell, \tag{13}$$

where i ranges from 0 to $c_\ell^\oplus$ and j from 0 to $c_\ell^\ominus$.[21]

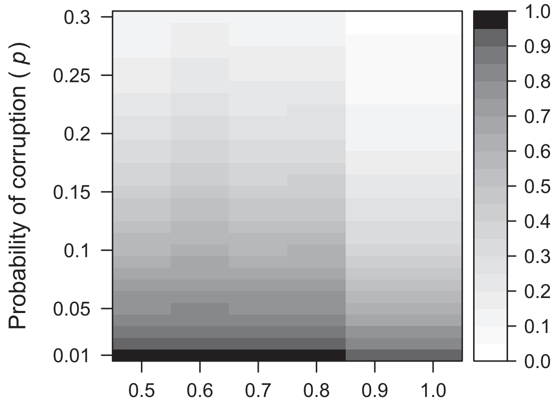
Similarly, if $(c_\ell^\oplus, c_\ell^\ominus) \in \mathrm{A}_0$, then ℓ is currently not included in $\mathscr{S}$, and the probability $\pi_\ell^\triangleleft$ that it enters $\mathscr{S}$ in consequence of corruptions on F_o is given by Equation (14):

$$\pi_\ell^\triangleleft(c_\ell^\oplus; c_\ell^\ominus; p) = \sum_{\substack{(i,\, j):\\ (c_\ell^\oplus - i + j,\ c_\ell^\ominus - j + i) \in \mathrm{A}_1}} \pi_\ell. \tag{14}$$

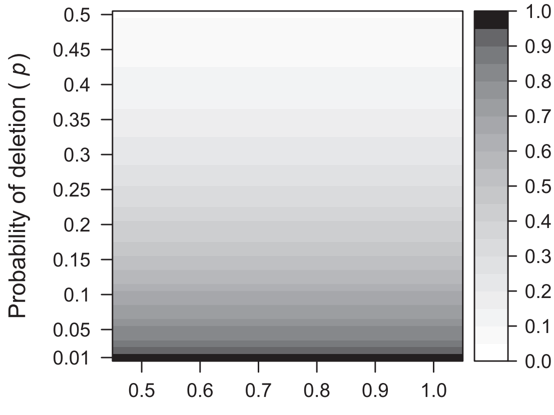
Because of the fact that the inclusion of each min-term ℓ into $\mathscr{S}$ in consequence of corruption, its exclusion from $\mathscr{S}$ respectively, occurs independently, the probability that $\mathscr{S}$ remains unchanged is

[20] As the two events are independent from each other, their joint probability is the product of their individual probabilities. The number of cases of $F_o^{\{1\}}$ that are corrupted is binomially distributed with parameters $c_\ell^\oplus$ and p, and the number of cases of $F_o^{\{0\}}$ that are corrupted is binomially distributed with parameters $c_\ell^\ominus$ and p.

[21] As before, i represents the number of cases of min-term ℓ whose value on F_o is changed from $F_o^{\{1\}}$ to $F_o^{\{0\}}$, and j represents the number of cases of min-term ℓ whose value on F_o is changed from $F_o^{\{0\}}$ to $F_o^{\{1\}}$. Because of the fact that the events in question are pairwise disjoint, the total probability is given by the sum of the individual probabilities of these events.



(a) Measurement error



(b) Loss of data

Fig. 3 Retention probabilities under IPA for data in Table 1.

Table 4 Polynomial terms of retention probability for the data set in Table 1

ℓ	$c^{\oplus/\ominus}$	$0.5 < \alpha \leq 0.6$	$0.6 < \alpha \leq 2/3$	$2/3 < \alpha \leq 0.75$	$0.75 < \alpha \leq 0.8$	$0.8 < \alpha \leq 1$
1	5	$(1-p)^5 + 5p(1-p)^4 + 10p^2(1-p)^3$	$(1-p)^5 + 5p(1-p)^4$	$\rightarrow$	$\rightarrow$	$(1-p)^5$
2	1	$1-p$	$\rightarrow$	$\rightarrow$	$\rightarrow$	$\rightarrow$
3	1	$1-p$	$\rightarrow$	$\rightarrow$	$\rightarrow$	$\rightarrow$
4	4	$1-p^4 - 4p^3(1-p)$	$\rightarrow$	$\rightarrow$	$1-p^4$	$\rightarrow$
5	3	$1-p^3 - 3p^2(1-p)$	$\rightarrow$	$1-p^3$	$\rightarrow$	$\rightarrow$
6	1	$1-p$	$\rightarrow$	$\rightarrow$	$\rightarrow$	$\rightarrow$
7	1	$1-p$	$\rightarrow$	$\rightarrow$	$\rightarrow$	$\rightarrow$

the product of the complementary probabilities of those given in Equations (13) and (14). The probability $\theta_{\mathscr{S}}$ that $\mathscr{S}$ is retained is therefore given by Equation (15):

$$\theta_{\mathscr{S}}(c^{\oplus}; c^{\ominus}; p) = \prod_{\substack{1 \leq \ell \leq m: \\ (c_\ell^{\oplus},\ c_\ell^{\ominus}) \in \mathrm{A}_1}} (1 - \pi_\ell^{\triangleright}) \cdot \prod_{\substack{1 \leq \ell \leq m: \\ (c_\ell^{\oplus},\ c_\ell^{\ominus}) \in \mathrm{A}_0}} (1 - \pi_\ell^{\triangleleft}), \tag{15}$$

where $c^{\oplus} = (c_1^{\oplus}, \ldots, c_m^{\oplus})$ and $c^{\ominus} = (c_1^{\ominus}, \ldots, c_m^{\ominus})$ denote the full vectors of respective case counts for positive min-terms and negative min-terms, respectively.

With regard to data loss, the only modification required consists in replacing the expression $(c_\ell^{\oplus} - i + j, c_\ell^{\ominus} - j + i)$ in Equations (13) and (14) by $(c_\ell^{\oplus} - i, c_\ell^{\ominus} - j)$, as shown before in Section 4.1 for DPA.[22]

Figure 3 plots the retention probability of the QCA reference solution in Table 2 under IPA, as a function of the inclusion cut-off α and the probability of corruption p in panel (a), and as a function of the inclusion cut-off α and the probability of deletion p in panel (b), again for $\beta = 1$. As before in Fig. 2, inclusion cut-offs range quasi-continuously from 0.5 to 1. The probability of corruption ranges from 1% to 30%; the probability of deletion from 1% to 50%. Similar to Fig. 2, the retention probability for the case of measurement error deteriorates at a higher rate than that

[22] See also footnote 17.

for data loss, but the retention probabilities for the latter are lower under IPA than under DPA, and higher under IPA than under DPA for the former.

To complement the information conveyed in Fig. 3a, Table 4 lists the concrete polynomials with respect to measurement error for the crucial ranges of α. The retention probability for a given range of α is obtained as the product of the terms across all rows. An arrow in a table cell means that the last polynomial to its left applies. For example, the retention probability for $0.75 < \alpha \leq 0.8$ and $p = 0.1$ is obtained as the product of seven factors, one for each min-term ℓ, the factor for $\ell = 1$ being $0.9^5 + 5 \cdot 0.1 \cdot 0.9^4 \approx 0.9185$ and the one for $\ell = 4$ being $1 - 0.1^4 = 0.9999$. Altogether, one obtains $\theta_{\mathscr{S}}(0.1) \approx 0.5836$.

As the values of the terms increase for min-terms with $c_\ell^\oplus > 0$ and decrease for min-terms with $c_\ell^\ominus > 0$ when going down the table, the retention probability as a whole neither increases nor decreases monotonously with α, regardless of the value of p. For example, the retention probability for $0.6 < \alpha \leq 2/3$ and $p = 0.3$ amounts to ≈ 0.09. On either side of this range of α, the probability is higher, by more than two percentage points up to $\alpha = 0.75$, and by more than five percentage points down to $\alpha > 0.5$. A very similar retention probability for the case of $0.6 < \alpha \leq 2/3$ and $p = 0.3$ is obtained, for example, when $\alpha > 0.8$ and $p \approx 0.23$. A higher inclusion cut-off along this range would thus require the probability of error to be seven percentage points lower for keeping the retention probability constant.

4.3 *Combinatorial Computation with IPA: Parsimonious Solutions*

Combinatorial computation under IPA is also possible for the parsimonious solution type, with some limitations for the case of data loss. With respect to measurement error, Equation (15) still applies for computing the retention probability of the parsimonious solution because the set of remainders never changes, and any corruption on F_o also affects the parsimonious solution. In other words, the retention probability of the parsimonious solution is exactly the same as the retention probability of the conservative solution for the same set of data δ and the same throughput parameter values.

For the case of data loss under the parsimonious solution type, it is only possible to derive the probability $\theta_{\mathbf{T}}$ that the truth table does not change, but not the probability $\theta_{\mathscr{S}}$ that the solution does not change, for the following reason: if a previously positive min-term turns into a remainder, this may leave the parsimonious solution unaffected because that particular min-term is reintroduced as a *simplifying assumption* by QMC at the beginning of Phase IIa.[23] Similarly, a previously negative min-term that turns into a remainder may not affect the parsimonious solution, either, because its structure may be such that QMC cannot turn it into a simplifying assumption.[24] Therefore, $\theta_{\mathbf{T}}$ only provides a lower bound for $\theta_{\mathscr{S}}$.

Similar to the case of data loss under the conservative solution type (cf. Equation (13)), the probability that a fixed positive min-term ℓ turns into a negative min-term or a remainder is given by Equation (16):

$$\pi_\ell^{1 \triangleright \{0,?\}}(c_\ell^\oplus; c_\ell^\ominus; p) = \sum_{\substack{(i,\, j): \\ (c_\ell^\oplus - i,\ c_\ell^\ominus - j) \notin \mathrm{A}_1}} \pi_\ell; \tag{16}$$

but as negative min-terms may now also turn into remainders, the condition for each pair (i, j) in the counterpart to Equation (14) must be modified such that Equation (17) results:

$$\pi_\ell^{0 \triangleright \{1,?\}}(c_\ell^\oplus; c_\ell^\ominus; p) = \sum_{\substack{(i,\, j): \\ (c_\ell^\oplus - i,\ c_\ell^\ominus - j) \notin \mathrm{A}_0}} \pi_\ell. \tag{17}$$

[23] A simplifying assumption is a remainder that QMC temporarily converts to a positive min-term in Phase IIa in order to be able to use it for minimization. Simplifying assumptions need not be covered in the prime implicant chart in Phase IIb, unlike (proper) positive min-terms.

[24] A remainder that differs from a positive min-term on at least two positions cannot be used as a simplifying assumption.

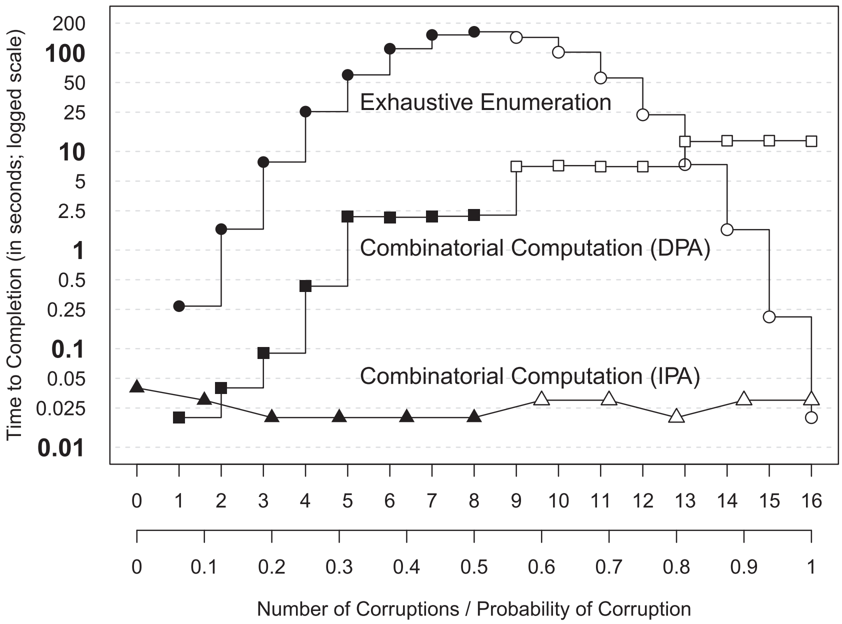


Fig. 4 Operation completion times of different methods of QCA sensitivity diagnostics for the analysis of measurement error.

Then, the probability $\theta_{\mathbf{T}}$ that $\mathbf{T}_{\delta}$ is retained is given by Equation (18):

$$\theta_{\mathbf{T}}(c^{\oplus}; c^{\ominus}; p) = \prod_{\substack{1 \leq \ell \leq m: \\ (c_{\ell}^{\oplus},\, c_{\ell}^{\ominus}) \in \mathrm{A}_1}} (1 - \pi_{\ell}^{1 \rhd \{0,?\}}) \cdot \prod_{\substack{1 \leq \ell \leq m: \\ (c_{\ell}^{\oplus},\, c_{\ell}^{\ominus}) \in \mathrm{A}_0}} (1 - \pi_{\ell}^{0 \rhd \{1,?\}}). \tag{18}$$

In summary, for measurement error on the endogenous factor, the retention probabilities for parsimonious solutions equal those for conservative solutions. With regard to data loss, however, it is only possible to derive a lower bound on the retention probability by computing the probability that the truth table will not change. Sometimes these probabilities will be equal; sometimes the probability that the solution is retained will be higher. At the current state of research, it cannot be determined which one is the case by purely combinatorial means.

4.4 *Performance Considerations*

Performance considerations with regard to analytical techniques are important for methodological research as these techniques determine which tests are feasible given time, and possibly other resource constraints. So as to address this issue, we have implemented a speed competition with respect to the data in Table 1. Figure 4 plots the operation completion times, in seconds, of each method—exhaustive enumeration (step function with circles), combinatorial computation under DPA (step function with squares), and combinatorial computation under IPA (continuous line with triangles)—for the case of measurement error. As the patterns for data loss are similar, we omit them.[25] For reasons of completeness, performance figures are shown over the full range of corruptions on the outcome factor, the probability of corruption, respectively, but statistics beyond eight corruptions, a probability of 0.5, respectively, are substantively irrelevant as even random value assignment would be preferable to measurement in these circumstances. It is also important to note that, for graphical reasons, the ordinate uses a logarithmic scale.

[25]Interested readers can produce these statistics as well as those for memory consumption using our replication code. Performance tests have been conducted on a regular end-user machine under Windows XP, with an Intel i5-3470S CPU 2.9 GHz processor.

Hug (2013) has conducted his data experiment with only up to two perturbations, which requires a rather small time investment of about 1.6 sec.[26] However, as exhaustive enumeration cycles through $\binom{n}{D}$ runs of the procedural protocol of QCA from Phase I to Phase IIb, the more serious the deficiencies in the data to be simulated, the longer the time and the larger the time increments needed to complete an operation. A sensitivity diagnosis at eight corruptions takes almost 3 min, 90 times as long as an analysis of the effects of two corruptions. The analysis of all cases, from one to eight corruptions, requires about 9 min.

In contrast to exhaustive enumeration, the resources consumed by a combinatorial approach under DPA scale approximately with $\binom{m}{D}$. The number of non-remainder min-terms m, however, is usually much lower than the number of cases n. The same relation applies to the constant factors involved. In combinatorial computation, for each set of at most D non-remainder min-terms, only a few arithmetic and logical operations are required, whereas exhaustive enumeration relies on a full run of QMC from Phase I to Phase IIb for each set of corruptions on D cases, irrespective of whether they change the solution or not. Particularly in situations of extreme model ambiguities, which may already occur for research designs involving as few as six to eight exogenous factors, this approach stretches operation completion times significantly.

At eight corruptions, combinatorial computation under DPA takes about 2.2 sec to complete its operation, whereas all scenarios up to four corruptions require less than half a second. In total, an analysis of one to eight corruptions is completed in under 10 sec. A method of combinatorial computation that retains the assumption of dependent perturbations thus outperforms exhaustive enumeration by a factor of 55.[27]

Once DPA is replaced by IPA, differentials in operation completion times between exhaustive enumeration and combinatorial computation become huge. Slight variations in the completion times of the latter method are substantively irrelevant and ascribable only to chance. Irrespective of the magnitude at which the probability of corruption is fixed, a combinatorial approach to analyzing the sensitivity of QCA reference solutions to measurement error for the data in Table 1 merely takes 0.02 sec under IPA. This method thus not only produces figures that are more realistic than those produced by combinatorial computation under DPA and exhaustive enumeration due to its more plausible assumption, but it also is about 50 times faster than combinatorial computation under DPA, and about 2600 times faster than exhaustive enumeration for a complete diagnosis of corruption scenarios at the point where data measurement is still preferable to random value assignment.

5 Discussion and Conclusions

The topic of sensitivity diagnostics has recently been put high on the agenda of methodological research into QCA. A significant number of studies have analyzed the reactivity of QCA to alterations in discretionary parameter values set by the researcher, but only a few have taken a closer look at the consequences of problems affecting the quality of data. However, almost all studies have relied on the method of exhaustive enumeration, whereby all unique possibilities for changing a given number of parameter values are systematically realized.

In this article, we have introduced a powerful alternative for evaluating the interaction effects between two problems afflicting data quality and two discretionary parameters. By employing a functional perspective on the stability of QCA reference solutions, we have developed the method of combinatorial computation for the analysis of measurement error and data loss. This method has not only proven computationally superior but, what is more, it also makes more plausible assumptions about the nature of these ubiquitous problems of empirical research.

Our study aspires to merely mark the beginning of a more systematic literature on sensitivity diagnostics in QCA. Future research should now extend its set-up to other QCA variants and

[26] We have integrated Hug's original functions in a single function for the purpose of performance testing, including improvements where appropriate. See the replication file for more details.

[27] Put differently, combinatorial computation under DPA consumes only about 1.8% of the resources required by exhaustive enumeration.

problems of empirical data analysis. In the short run, generalizations to multi-value QCA (Cronqvist and Berg-Schlosser 2009; Thiem 2013) and further research into the issue of data loss under the parsimonious solution type appear to be the two most promising avenues. Sensitivity evaluations of methods closely related to QCA, such as Coincidence Analysis, provide another possibility as corresponding software has now become available (Ambuehl et al. 2015; Baumgartner and Thiem 2015a). Last but not least, Monte Carlo simulations should be examined as a third alternative to exhaustive enumeration and combinatorial computation.

6 Funding

Alrik Thiem gratefully acknowledges financial support from the Swiss National Science Foundation, award number PP00P1_144736/1.

Conflict of interest statement. None declared.

References

Ambuehl, Mathias, Michael Baumgartner, Ruedi Epple, Alexis Kauffmann, and Alrik Thiem. 2015. *cna: A package for Coincidence Analysis, R package version 1.0-3*. http://cran.r-project.org/package=cna.

Baumgartner, Michael. 2009. Inferring causal complexity. *Sociological Methods & Research* 38(1):71–101.

———. 2013. Detecting causal chains in small-*n* data. *Field Methods* 25(1):3–24.

———. 2015. Parsimony and causality. *Quality & Quantity* 49(2):839–56.

Baumgartner, Michael, and Alrik Thiem. 2015a. Identifying complex causal dependencies in configurational data with Coincidence Analysis. *The R Journal* 7(1):176–84.

———. 2015b. Model ambiguities in configurational comparative research. *Sociological Methods & Research*. Advance online publication. DOI: 10.1177/0049124115610351.

Baumgartner, Michael, and Ruedi Epple. 2014. A Coincidence Analysis of a causal chain: The Swiss minaret vote. *Sociological Methods & Research* 43(2):280–312.

Bol, Damien, and Francesca Luppi. 2013. Confronting theories based on necessary relations: Making the best of QCA possibilities. *Political Research Quarterly* 66(1):205–10.

Bowers, Jake. 2014. Comment: Method games—A proposal for assessing and learning about methods. *Sociological Methodology* 44(1):112–7.

Braumoeller, Bear F. 2015. Guarding against false positives in Qualitative Comparative Analysis. *Political Analysis*. Advance online publication. DOI: 10.1093/pan/mpv017.

Braumoeller, Bear F., and Gary Goertz. 2000. The methodology of necessary conditions. *American Journal of Political Science* 44(4):844–58.

Clark, William Roberts, Michael J. Gilligan, and Matt Golder. 2006. A simple multivariate test for asymmetric hypotheses. *Political Analysis* 14(3):311–31.

Cooper, Barry, and Judith Glaesser. 2011a. Paradoxes and pitfalls in using fuzzy set QCA: Illustrations from a critical review of a study of educational inequality. *Sociological Research Online* 16(3):1–8.

———. 2011b. Using case-based approaches to analyse large datasets: A comparison of Ragin's fsQCA and fuzzy cluster analysis. *International Journal of Social Research Methodology* 14(1):31–48.

Coverdill, James E., and William Finlay. 1995. Understanding mills via Mill-type methods: An application of Qualitative Comparative Analysis to a study of labor management in Southern textile manufacturing. *Qualitative Sociology* 18(4):457–78.

Cronqvist, Lasse, and Dirk Berg-Schlosser. 2009. Multi-value QCA (mvQCA). In *Configurational comparative methods: Qualitative Comparative Analysis (QCA) and related techniques*, eds. Benoît Rihoux and Charles C. Ragin, 69–86. London: Sage Publications.

Duşa, Adrian, and Alrik Thiem. 2014. *QCA: A package for Qualitative Comparative Analysis, R package version 1.1-4*. http://cran.r-project.org/package=QCA.

———. 2015. Enhancing the minimization of Boolean and multi-value output functions with *e*QMC. *Journal of Mathematical Sociology* 39(2):92–108.

Eliason, Scott R., and Robin Stryker. 2009. Goodness-of-fit tests and descriptive measures in fuzzy-set analysis. *Sociological Methods & Research* 38(1):102–46.

Gelman, Andrew, and Jennifer Hill. 2007. *Data analysis using regression and multilevel/hierarchical models*. Cambridge, UK: Cambridge University Press.

Glaesser, Judith, and Barry Cooper. 2014. Exploring the consequences of a recalibration of causal conditions when assessing sufficiency with fuzzy set QCA. *International Journal of Social Research Methodology* 17(4):387–401.

Goldthorpe, John H. 1997a. Current issues in comparative macrosociology: A debate on methodological issues. *Comparative Social Research* 16:1–26.

——— 1997b. A response to the commentaries. *Comparative Social Research* 16:121–32.

Griffin, Larry J., Christopher Botsko, Ana-Maria Wahl, and Larry W. Isaac. 1991. Theoretical generality, case particularity: Qualitative Comparative Analysis of trade-union growth and decline. *International Journal of Comparative Sociology* 32(1–2):110–36.

Grofman, Bernard, and Carsten Q. Schneider. 2009. An introduction to crisp set QCA, with a comparison to binary logistic regression. *Political Research Quarterly* 62(4):662–72.

Hicks, Alexander, Joya Misra, and Tang Nah Ng. 1995. The programmatic emergence of the social security state. *American Sociological Review* 60(3):329–49.

Hohn, Franz E. 1966. *Applied Boolean algebra: An elementary introduction*. New York: Macmillan.

Hug, Simon. 2013. Qualitative Comparative Analysis: How inductive use and measurement error lead to problematic inference. *Political Analysis* 21(2):252–65.

Krogslund, Chris, Donghyun Danny Choi, and Mathias Poertner. 2015. Fuzzy sets on shaky ground: Parameter sensitivity and confirmation bias in fsQCA. *Political Analysis* 23(1):21–41.

Lucas, Samuel R., and Alisa Szatrowski. 2014. Qualitative Comparative Analysis in critical perspective. *Sociological Methodology* 44(1):1–79.

Maggetti, Martino, and David Levi-Faur. 2013. Dealing with errors in QCA. *Political Research Quarterly* 66(1):198–204.

Nievergelt, Jürg. 2000. Exhaustive search, combinatorial optimization and enumeration: Exploring the potential of raw computing power. In *SOFSEM 2000: Theory and practice of informatics*, eds. Václav Hlaváč, Keith G. Jeffery, and Jiří Wiedermann, 18–35. Berlin: Springer.

Paine, Jack. 2015. Set-theoretic comparative methods: Less distinctive than claimed. *Comparative Political Studies*. Advance online publication. DOI: 10.1177/0010414014564851.

R Development Core Team. 2014. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Ragin, Charles C. 2000. *Fuzzy-set social science*. Chicago: University of Chicago Press.

Ragin, Charles C., and Sean Davey. 2014. *fs/QCA: Fuzzy-set/Qualitative Comparative Analysis, version 2.5 [computer program]*. Irvine: Department of Sociology, University of California.

Sarkar, Deepayan. 2008. *Lattice: Multivariate data visualization with R*. New York: Springer.

Schneider, Carsten Q., and Claudius Wagemann. 2012. *Set-theoretic methods for the social sciences: A guide to Qualitative Comparative Analysis (QCA)*. Cambridge, UK: Cambridge University Press.

———. 2013. Doing justice to logical remainders in QCA: Moving beyond the standard analysis. *Political Research Quarterly* 66(1):211–20.

Seawright, Jason. 2005. Qualitative Comparative Analysis vis-à-vis regression. *Studies in Comparative International Development* 40(1):3–26.

———. 2014. Comment: Limited diversity and the unreliability of QCA. *Sociological Methodology* 44(1):118–21.

Skaaning, Svend-Erik. 2011. Assessing the robustness of crisp-set and fuzzy-set QCA results. *Sociological Methods & Research* 40(2):391–408.

Thiem, Alrik. 2013. Clearly crisp, and not fuzzy: A reassessment of the (putative) pitfalls of multi-value QCA. *Field Methods* 25(2):197–207.

———. 2014a. Membership function sensitivity of descriptive statistics in fuzzy-set relations. *International Journal of Social Research Methodology* 17(6):625–42.

———. 2014b. Mill's methods, induction and case sensitivity in Qualitative Comparative Analysis: A comment on Hug (2013). *Qualitative & Multi-Method Research* 12(2):19–24.

———. 2014c. Navigating the complexities of Qualitative Comparative Analysis: Case numbers, necessity relations, and model ambiguities. *Evaluation Review* 38(6):487–513.

———. 2014d. Unifying configurational comparative methods: Generalized-set Qualitative Comparative Analysis. *Sociological Methods & Research* 43(2):313–37.

———. 2015. Using Qualitative Comparative Analysis for identifying causal chains in configurational data: A methodological commentary on Baumgartner and Epple (2014). *Sociological Methods & Research* 44(4):723–36.

Thiem, Alrik, and Adrian Duşa. 2013a. Boolean minimization in social science research: A review of current software for Qualitative Comparative Analysis (QCA). *Social Science Computer Review* 31(4):505–21.

———. 2013b. QCA: A package for Qualitative Comparative Analysis. *The R Journal* 5(1):87–97.

———. 2013c. *Qualitative Comparative Analysis with R: A user's guide*. New York: Springer.

Thiem, Alrik, Michael Baumgartner, and Damien Bol. 2015. Still lost in translation! A correction of three misunderstandings between configurational comparativists and regressional analysts. *Comparative Political Studies*. Advance online publication. DOI: 10.1177/0010414014565892.

Thiem, Alrik, Reto Spöhel, and Adrian Duşa. 2015. *Replication Package for: Enhancing sensitivity diagnostics for Qualitative Comparative Analysis: A combinatorial approach*. http://dx.doi.org/10.7910/DVN/QE27H9, Harvard Dataverse, V1.

Vis, Barbara. 2012. The comparative advantages of fsQCA and regression analysis for moderately large-*n* analyses. *Sociological Methods & Research* 41(1):168–98.

Wooldridge, Jeffrey M. 2002. *Econometric analysis of cross section and panel data*. Cambridge, MA: MIT Press.