Comparative Political Studies
2016, Vol. 49(6) 742–774

DOI: 10.1177/0010414014565892
cps.sagepub.com

Still Lost in Translation! A Correction of Three Misunderstandings Between Configurational Comparativists and Regressional Analysts

Alrik Thiem[1], Michael Baumgartner[1], and Damien Bol[2]

Abstract

Even after a quarter-century of debate in political science and sociology, representatives of configurational comparative methods (CCMs) and those of regressional analytic methods (RAMs) continue talking at cross purposes. In this article, we clear up three fundamental misunderstandings that have been widespread within and between the two communities, namely that (a) CCMs and RAMs use the same logic of inference, (b) the same hypotheses can be associated with one or the other set of methods, and (c) multiplicative RAM interactions and CCM conjunctions constitute the same concept of causal complexity. In providing the first systematic correction of these persistent misapprehensions, we seek to clarify formal differences between CCMs and RAMs. Our objective is to contribute to a more informed debate than has been the case so far, which should eventually lead to progress in dialogue and more accurate appraisals of the possibilities and limits of each set of methods.

[1]University of Geneva, Switzerland
[2]University of Montreal, Canada

Corresponding Author:
Alrik Thiem, University of Geneva, Rue de Candolle 2 / Bât. Landolt, Geneva, 1211, Switzerland.
Email: alrik.thiem@unige.ch

Keywords
qualitative methods, quantitative methods, Boolean algebra, linear algebra, qualitative comparative analysis

Introduction

About a quarter-century ago, the publication of Charles Ragin's (1987) *The Comparative Method* sparked a debate in the literature on political and sociological research methods that has not lost one iota of its initial impetus to date. Quite the contrary, the two sequels *Fuzzy Set Social Science* (Ragin, 2000) and *Redesigning Social Inquiry* (Ragin, 2008) have brought the "Ragin Revolution" (Vaisey, 2009) to the current point of unprecedented commotion at which proponents and opponents are vying for the methodological high ground more fiercely than ever before.[1] The trigger of this revolution was the introduction of a novel method named Qualitative Comparative Analysis (QCA), which, after a slow yet solid start in the early 1990s, has passed the 100-articles-per-year mark in 2013—for the first time since its inaugural appearance in Ragin, Mayer, and Drass (1984).[2] According to its inventor, the primary motivation behind the development of QCA has been to "integrate the best features of the case-oriented approach with the best features of the variable-oriented approach" (Ragin, 1987, p. 84).

In this article, we do not endeavor to evaluate whether Ragin has succeeded on this front or not. Nor do we want to align ourselves in arguments over the vices and virtues of QCA. Instead, we pursue the following objective from the sidelines: to clear up three misunderstandings that have dominated the debate between representatives of configurational comparative methods (CCMs) such as QCA and those of regressional analytic methods (RAMs) ever since the publication of *The Comparative Method*.[3] These misunderstandings do not merely concern trivia at the methodological periphery but central issues. Contrary to expectations, however, the sources of these problems only partly reside in difficulties of communication *between* the two communities, but mainly in knowledge gaps and ambiguous definitions of concepts *within* these two communities. By filling these gaps and by clarifying concepts, we seek to clear the blockage in the debate, hopefully once and for all. We expect appraisals of the possibilities and limits of each set of methods to eventually also become more accurate in consequence. The three aspects to be addressed are listed in Table 1.

First, CCMs and RAMs build on disparate theories of mathematical structures whose syntax may often be equal but whose semantics remain incommensurable. While CCMs work under the axioms of a Boolean algebra, RAMs do so under those of a linear algebra.[4] This fundamental yet simple

Table 1. Three Aspects of the CCM–RAM Debate.

Aspect	CCMs	RAMs
Underlying algebraic system	Boolean	Linear
Associated hypothesis class	Implication	Covariation
Concept of causal complexity	Conjunction	Interaction

Note. CCM = configurational comparative method; RAM = regressional analytic method.

difference continues to be downplayed, misinterpreted, or even outright ignored (cf. Goertz & Mahoney, 2013a, pp. 280-81; 2013b, pp. 239-40). On the side of the proponents of RAMs, King, Keohane, and Verba (1994), for instance, have dismissed "Ragin's 'Boolean Algebra' approach" summarily as containing "no new features or theoretical requirements" (pp. 50, 87-91), but proponents of CCMs share some responsibility for the current state of affairs. Against the background of Schneider and Wagemann's (2012) otherwise laudable aim of "avoiding confusion and misinterpretation of set-theoretic methods" (p. 42), it is half unfortunate and half ironic to see the authors proclaim that "[t]he challenge in understanding set-theoretic methods is not so much in grasping the math" (pp. 16-17) when the small selection of Boolean math they introduce contains a considerable number of glaring errors.[5]

Second, particular classes of hypotheses demand the application of either CCMs or RAMs because the associations they posit are based on either a Boolean or a linear-algebraic framework. As Braumoeller and Goertz (2000, p. 847) note, however, most members of either camp remain oblivious of the inseparability of particular classes of hypotheses and the appropriate set of methods for building and testing them. Two quotes from articles that both have appeared in top methods journals are illustrative. While Katz, vom Hau, and Mahoney (2005) conclude that "regression methods and fuzzy-set methods cannot test the same hypotheses because the two approaches' contrasting understandings of causation lead them to formulate fundamentally different kinds of hypotheses" (p. 541), Clark, Gilligan, and Golder (2006) argue that "standard linear models that include interaction terms offer a better way" (p. 312) to test hypotheses about necessity and sufficiency. The opposition between these two statements could hardly be more direct.

Third, the issue of causal complexity has split the two communities (e.g., Brady, 2013; Clark et al., 2006; Vis, 2012). Some methodologists consider CCMs and RAMs to be closely related if not substitutable in analyzing causal complexity (Mahoney, 2008), others are of the opinion that the former have an edge over the latter (Wagemann & Schneider, 2010), and still others argue that CCMs are vastly inferior to RAMs (Clark et al., 2006). The order of these three aspects, from algebraic systems over hypotheses classes to the concept of causal complexity, is not random but reflects their level of generality. The mastery of analyzing causal complexity is impossible without an understanding of hypothesis classes, which is itself prevented by unfamiliarity with the formal algebraic rules that underlie them.[6] We thus contend that progress in the CCM–RAM debate is impossible without the joint dissolution of the confusion that surrounds these aspects individually.

The article is structured as follows. In the first section, we address the widely-held misconception that the same logic of inference applies to CCMs and RAMs by juxtaposing the formal differences in their respective languages. Surprisingly, these have so far never been laid out in the debate. Second, we systematize the different types of hypotheses each set of methods is associated with. It is usually recognized that CCMs can build and test hypotheses about relations of implication, and that RAMs are suitable for building and testing hypotheses about relations of covariation (e.g., Mahoney, 2007), but as yet an elaboration of these dissimilarities beyond the stage of mere recognition has not been presented. The differences between CCM conjunctions and RAM interactions are the topic of the third and final part, which integrates the core points of the preceding sections. In the conclusions, we recapitulate the argument, provide a short-term forecast of the direction in which the CCM–RAM debate will move over the coming years and issue some recommendations to influence its course.

Algebraic Systems

In their attempt to convince qualitative researchers of the universal applicability of the principles of RAMs, King et al. (1994) argue that "differences between the quantitative and qualitative traditions are only stylistic and are methodologically and substantively unimportant. All good research can be understood . . . to derive from the same underlying logic of inference" (p. 4). The authors insist on the existence of only one language for social-scientific inquiry and dismiss "Ragin's 'Boolean Algebra' approach" summarily as containing "no new features or theoretical requirements" (King et al., 1994, pp. 50, 87-91). We confute this assertion by showing that Boolean algebra provides a self-contained logic of inference for CCMs along with mathematical machinery that is neither reducible to nor reconcilable with the logic underlying RAMs. However, representatives of CCMs share responsibility for the current state of affairs by having misrepresented central elements of Boolean algebra. Our main contention in this section is the following:

> **Contention 1:** Different formal languages for drawing inferences exist. Boolean algebra establishes the language of CCMs, linear algebra that of RAMs. Notwithstanding syntactic commonalities, these languages are semantically incommensurable.[7]

The failure to appreciate the existence of an alternative language by proponents of RAMs has prevented progress in the debate on the most basic level. As Mahoney and Goertz (2006) point out, the dissimilarity even between simple binary operators such as the logical AND and the arithmetic TIMES has contributed to "substantial confusion across the two traditions" (p. 235). So as to remedy this unfortunate state of affairs, we lay out the principal differences between and commonalities of Boolean and linear algebras.[8] While commonalities are strictly limited to syntactic features of the formalisms in which they are expressed, fundamentally different semantic interpretations are called for. These differences are so profound that the two systems give rise to formal languages that defy all attempts at comprehensive translation. Contrary to King et al. (1994), we thus not only argue that the theoretical requirements of "Ragin's 'Boolean Algebra' approach" are at least as demanding as those of RAMs but also contradict Brady's (2013) more conciliatory verdict that "[l]anguage differences can be important, but they can be transcended through careful translation" (p. 253).

Boolean and linear algebras are mathematical objects that consist of a first set $\mathcal{I}$ comprising two distinguished identity elements typically denoted by "1" and "0", and a second set $\mathcal{O}$ consisting of two binary operations "+" and "$*$" as well as one unary operation "−". For any Boolean algebra, $\mathcal{O}$ is defined over $\mathcal{I}$ such that the following laws hold:

- commutativity of "+" and "$*$"

$$x + y = y + x \qquad x * y = y * x, \tag{BA$_1$}$$

- associativity of "+" and "*"

$$\left(x + y\right) + z = x + \left(y + z\right) \qquad \left(x * y\right) * z = x * \left(y * z\right), \tag{BA$_2$}$$

- distributivity of "+" over "*" and of "*" over "+"

$$x + \left(y * z\right) = \left(x + y\right) * \left(x + z\right) \qquad x * \left(y + z\right) = \left(x * y\right) + \left(x * z\right), \tag{BA$_3$}$$

- as well as these special laws

$$x + 1 = 1 \qquad x * 0 = 0, \tag{BA$_4$}$$

$$x + 0 = x \qquad x * 1 = x, \tag{BA$_5$}$$

$$x + x = x \qquad x * x = x, \tag{BA$_6$}$$

$$x + (-x) = 1 \qquad x * (-x) = 0, \tag{BA$_7$}$$

$$(-x) * (-y) = -(x + y) \qquad (-x) + (-y) = -(x * y). \tag{BA$_8$}$$

Although this set is unnecessarily large for providing a minimal definition, we explicitly list (BA_1) to (BA_8) to facilitate comparison.[9] In contradistinction, for any linear algebra, $\mathcal{O}$ is defined over $\mathcal{I}$ such that the following laws are respected:

- commutativity of "+" and "*"

$$x + y = y + x \qquad x * y = y * x, \tag{LA$_1$}$$

- associativity of "+" and "*"

$$(x + y) + z = x + (y + z) \qquad (x * y) * z = x * (y * z), \tag{LA$_2$}$$

- distributivity of "*" over "+"

$$x * (y + z) = (x * y) + (x * z), \tag{LA$_3$}$$

- as well as these special laws

$$x + 1 = x + 1 \qquad x * 0 = 0, \tag{LA$_4$}$$

$$x + 0 = x \qquad x * 1 = x, \tag{LA$_5$}$$

$$x + x = 2 * x \qquad x * x = x^2, \tag{LA$_6$}$$

$$x + (-x) = 0 \qquad x * (-x) = -x^2, \tag{LA$_7$}$$

$$(-x) * (-y) = x * y \qquad (-x) + (-y) = -(x + y). \tag{LA$_8$}$$

These purely syntactic definitions only partially determine the types of elements that $\mathcal{I}$ and $\mathcal{O}$ should contain, any of which that satisfy the respective constraints constitute a Boolean or a linear algebra. However, these definitions render a single interpretation shared by both algebraic systems impossible. Most fundamentally, consider the laws of distribution in (BA_3) and (LA_3), both of which are required for minimal definitions. The Boolean "+" and "$*$" are mutually distributive, whereas the linear-algebraic "$*$" distributes over "+" but not vice versa. Furthermore, not both of the Boolean null element laws in (BA_4) correspond to their linear-algebraic counterparts in (LA_4). While x is redundant in the operation $x + 1$ under a Boolean algebra, this is not true under a linear algebra. Analogously, one occurrence of x is redundant in the Boolean operations $x + x$ and $x * x$, whereas this does not hold for the linear case.[10] Finally, dissimilarities are most pronounced between the Boolean laws of complementarity in (BA_7) and the so-called De Morgan laws in (BA_8), and their linear counterparts in (LA_7) and (LA_8). The Boolean operation $x + (-x)$ yields the value 1, but the syntactically identical operation yields the value 0 in the linear case. Hence, it follows that "1" and "0" have reversed functional effects in Boolean and linear algebras in some contexts but equal ones in others. In consequence, there cannot possibly exist one determinate semantic interpretation that satisfies all minimal laws of either system simultaneously. Even translatable elements such as commutativity and associativity, or the Boolean identity laws in (BA_5) and (LA_5), do not deduct from the semantic incommensurability of the general frameworks in which Boolean and linear algebras are embedded. No conceivable interpretation of the former satisfying (BA_1) to (BA_8) could possibly also satisfy (LA_1) to (LA_8), and vice versa.

The most common formal languages that satisfy the axioms of a Boolean algebra are *set theory, propositional logic*, and *switching-circuit theory*.[11] In set theory, the identity elements are interpreted in terms of the *universal set* "**U**" ("1") and the *empty set* "$\varnothing$" ("0"), the operations in terms of *union* "$\cup$" ("+"), *intersection* "$\cap$" ("$*$"), and *complement* "$\overline{x}$" ("$-$"). In propositional logic, the identity elements designate the truth values *true* "*T*" and *false* "*F*", while the operations denote *disjunction* "$\vee$", *conjunction* "$\wedge$", and *negation* "$\neg$". In switching-circuit theory, the identity elements stand for the transmittance of *closed* and *open switches*, and the operations for *switches in parallel, switches in series*, and a *change in switch transmittance*. In contrast, the identity elements are interpreted in terms of the corresponding integers 1 and 0 in formal languages that satisfy the conditions of a linear algebra, while "+", "$*$", and "$-$" assume the meaning of the arithmetic operations *addition, multiplication*, and *subtraction*. These integers have nothing at all in common with universal or empty sets or truth values. Analogously, arithmetic operations radically differ in purpose and effect from set, logic, or switching operations.

RAMs usually follow the standard arithmetic interpretation of a linear algebra, but the interpretation of the Boolean algebra that underlies CCMs has unfortunately often meandered between a set-theoretic and a logical interpretation, sometimes for no other reasons than keyboard convenience (cf. Schneider & Wagemann, 2012, pp. 54-55). From a conceptual perspective, mixing interpretations of a Boolean algebra is unproblematic although this practice has partly been responsible for the current state of confusion. We apply a consistent rendering of the Boolean algebra employed by CCMs in terms of propositional logic in the remainder of this article.

The central corollary of the semantic differences induced by Boolean and linear algebras is that CCMs and RAMs model causes and effects as disparate entities.[12] The objects *x, y*, and *z* in (BA_1) to (BA_8) represent *conditions* and *outcomes* for CCMs, whereas *x, y*, and *z* in (LA_1) to (LA_8) are understood to be *regressors* and *regressands* by RAMs.[13] This difference is not merely one of terminology as Schneider and Wagemann (2012, p. 55) suggest. A condition or an outcome always refers to one concrete value that a variable takes on. In contrast, a regressor or a regressand always refers to the variable itself. For example, RAMs deal with countries' *degree* of social heterogeneity or their *degree* of electoral district magnitude, whereas CCMs process countries that show a *high degree* of social heterogeneity or a *large magnitude* of electoral districts. The distinction is subtle yet fundamental. Almost all contributions to the CCM–RAM debate of the last two decades have missed or misinterpreted this crucial difference. So as to keep these entities syntactically apart while remaining close to established notational conventions, we denote variables by italicized capital letters, for example, X, and conditions and outcomes by italicized capital letters to which a value indicator is appended in superscript, for example, $X^{\{\cdot\}}$. As Boolean algebra is limited to bivalent variables, we additionally introduce the following notational simplification for unary operations on simple terms: "$\neg X^{\{1\}}$" will be replaced by "$X^{\{0\}}$" and "$\neg X^{\{0\}}$" by "$X^{\{1\}}$".[14]

Further non-fundamental operations are definable based on fundamental operations. As a matter of fact, formal languages of Boolean and linear algebra typically feature a host thereof. Important non-fundamental operators in propositional logic are the left-to-right arrow "$\Rightarrow$", and the likewise common but non-standard right-to-left arrow "$\Leftarrow$".[15] They denote an *implication* and can be defined in two equivalent ways as given by definitions (DF_1) and (DF_2), (DF_3) and (DF_4), respectively. The corresponding notation with generic operators is provided in addition (in square brackets):

$$X^{\{1\}} \Rightarrow Y^{\{1\}} \overset{\text{def}}{=} \neg\left(X^{\{1\}} \wedge Y^{\{0\}}\right) \quad \left[x * (-y) = 0\right], \tag{DF$_1$}$$

$$\overset{\text{def}}{=} X^{\{0\}} \vee Y^{\{1\}} \quad \left[-x + y = 1\right], \tag{DF$_2$}$$

$$X^{\{1\}} \Leftarrow Y^{\{1\}} \overset{\text{def}}{=} \neg\left(X^{\{0\}} \wedge Y^{\{1\}}\right) \quad [-x * y = 0], \tag{DF$_3$}$$

$$\overset{\text{def}}{=} X^{\{1\}} \vee Y^{\{0\}} \quad [x + (-y) = 1]. \tag{DF$_4$}$$

In natural language, $X^{\{1\}} \Rightarrow Y^{\{1\}}$ reads as "If $X^{\{1\}}$ is the case, then $Y^{\{1\}}$ is the case as well". According to (DF_1) and (DF_2), this sentence is equivalent in meaning to "It is not the case that $X^{\{1\}}$ and $Y^{\{0\}}$ are given" as well as to "$X^{\{0\}}$ or $Y^{\{1\}}$ is given". However, the most common phrasing is "$X^{\{1\}}$ is sufficient for $Y^{\{1\}}$". In contrast, $X^{\{1\}} \Leftarrow Y^{\{1\}}$ reads as "Only if $X^{\{1\}}$ is the case, then $Y^{\{1\}}$ is the case as well". According to (DF_3) and (DF_4), equivalent phrasings are "It is not the case that $X^{\{0\}}$ and $Y^{\{1\}}$ are given" and "$X^{\{1\}}$ or $Y^{\{0\}}$ is given", but the most common wording is "$X^{\{1\}}$ is necessary for $Y^{\{1\}}$".

Unfortunately, methodologists and users of CCMs have infused the implication operator with qualities it simply does not possess. For example, Schneider and Wagemann (2012, pp. 51-53) introduce the three fundamental operations "$\vee$", "$\wedge$", and "$\neg$" as suitable for constructing complex sets, whereas the implication operator is said to be appropriate for analyzing (*causal*) relations between sets. Similarly, Rihoux and De Meur (2009) introduce the operator as expressing "the (*usually causal*) link between a set of conditions on the one hand and the outcome we are trying to "explain" on the other" (p. 35, emphasis added). Yet, it is obvious from (DF_1) and (DF_2) that an implication states nothing beyond a negated conjunction in which the consequent is negated, or a disjunction in which the antecedent is negated. Implications are not in any way more amenable to causal interpretation than any of the three fundamental operations.

The implication operator forms the basis of the *equivalence* operator "$\Leftrightarrow$", which is alternatively defined by (DF_5) to (DF_8):

$$X^{\{1\}} \Leftrightarrow Y^{\{1\}} \overset{\text{def}}{=} \left(X^{\{1\}} \Rightarrow Y^{\{1\}}\right) \wedge \left(X^{\{1\}} \Leftarrow Y^{\{1\}}\right), \tag{DF$_5$}$$

$$\overset{\text{def}}{=} \neg\left(X^{\{1\}} \wedge Y^{\{0\}}\right) \wedge \neg\left(X^{\{0\}} \wedge Y^{\{1\}}\right), \tag{DF$_6$}$$

$$\overset{\text{def}}{=} \left(X^{\{0\}} \vee Y^{\{1\}}\right) \wedge \left(X^{\{1\}} \vee Y^{\{0\}}\right), \tag{DF$_7$}$$

$$\overset{\text{def}}{=} \left(X^{\{0\}} \wedge Y^{\{0\}}\right) \vee \left(X^{\{1\}} \wedge Y^{\{1\}}\right). \tag{DF$_8$}$$

It carries the meaning of "if, and only if", so that "$X^{\{1\}} \Leftrightarrow Y^{\{1\}}$" reads as "$X^{\{1\}}$ is the case if, and only if, $Y^{\{1\}}$ is the case as well". According to (DF_6) and (DF_8), equivalent phrasings are "It is neither the case that $X^{\{1\}}$ and $Y^{\{0\}}$ are given nor that $X^{\{0\}}$ and $Y^{\{1\}}$ are given" and "$X^{\{0\}}$ and $Y^{\{0\}}$ is given or $X^{\{1\}}$ and $Y^{\{1\}}$ is given", but the most common wording is "$X^{\{1\}}$ is sufficient and necessary for $Y^{\{1\}}$". In this connection, it is important to note that logical equivalence "$\Leftrightarrow$" is not the same as arithmetic equality "=". The latter is an operator that relates expressions referring to identical mathematical objects, as in "4 = 2 + 2", whereas the former is an operator that relates expressions with identical truth values. As "4" and "2 + 2" are neither true nor false but simply alternative names that refer to the number 4, the expression "$4 \Leftrightarrow 2 + 2$" is ill-formed. The same holds for "$X^{\{1\}} = Y^{\{1\}}$" because "$X^{\{1\}}$" and "$Y^{\{1\}}$" do not refer to mathematical objects but to conditions that can be true or false. We will come back to this crucial difference in the section on causal complexity.

Definitions (DF_1) and (DF_2) exemplify the unbridgeable semantic dissimilarities that exist between Boolean and linear algebras also for non-fundamental operations. The two generic expressions $x * (-y) = 0$ and $-x + y = 1$ are not only well-formed in Boolean but also in linear-algebraic syntax. Notwithstanding this commonality, their truth conditions are entirely different. In linear algebra, $x * (-y) = 0$ holds if, and only if, at least one of x or y is 0, whereas $-x + y = 1$ holds if, and only if, $y = x + 1$. As a result, it is indeterminate how implication and equivalence should be translated into linear algebra. Such a translation would have to give preference to either definition (DF_1) or (DF_2), but when read linear-algebraically, neither $x * (-y) = 0$ nor $-x + y = 1$ express the meaning of sufficiency or necessity. With regard to the first definition, it follows that 0 is sufficient for every element of $\mathcal{I}$ and that every element of $\mathcal{I}$ is sufficient for 0; analogously for necessity. As to the second definition, it follows that every element of $\mathcal{I}$ is sufficient for its successor ($x + 1$) and necessary for its predecessor ($y - 1$). However, claims of sufficiency or necessity neither concern the relationship between 0 and other elements of the algebra nor between the successors or predecessors of the algebra's elements. Instead, they concern the (in-)existence of certain constellations of attributes, which is expressed by the Boolean $\neg(X^{\{1\}} \wedge Y^{\{0\}})$ and $X^{\{0\}} \vee Y^{\{1\}}$.

The conclusion from our exposition above can only be that *Boolean algebra and linear algebra, despite occasional similarities in syntax, give rise to languages that are semantically incommensurable*. Unsurprisingly, none of the works which have propagated the unity of political methodology (e.g., Brady, 2013; Gerring, 2012; King et al., 1994; Mahoney, 2008) has been accompanied by an argument of how to reconcile the minimal set of axioms of each algebraic system under a generalized system.

Hypothesis Classes

Many political methodologists expressly consider (*causal*) *inference*—the formulation and testing of (causal) hypotheses with empirical data—to be the primary goal of social research (Gerring, 2012; Goertz & Mahoney, 2012; King et al., 1994).[16] Although the topic of hypothesis formulation is part and parcel of elementary training in research design, uncertainty about the types of hypotheses associated with each set of methods persists (Braumoeller & Goertz, 2000, p. 847). For instance, Katz et al. (2005) conclude that "regression methods and fuzzy-set methods cannot test the same hypotheses because the two approaches' contrasting understandings of causation lead them to formulate fundamentally different kinds of hypotheses" (p. 541), whereas Clark et al. (2006) argue that "standard linear models that include interaction terms offer a better way to test asymmetric hypotheses" (p. 312).[17] Before the difference between RAM interactions and CCM conjunctions will be discussed in the next section, we provide a basic taxonomy of the different types of hypotheses social scientists formulate in this part, based on the following claim:

> **Contention 2:** The type of a hypothesis determines the appropriate set of methods. An implication hypothesis links a condition with an outcome to form a proposition about the location of observations across a grid of spaces. A covariation hypothesis links a regressor with a regressand to form a proposition about the direction of the marginal rate of change over a plane around their joint arithmetic average. The former is based on a Boolean algebra and therefore associated with CCMs. The latter is based on a linear algebra and therefore associated with RAMs.[18]

The focus here is on simple conditions and regressors.[19] A three-level system structures the different types of hypotheses. On the first level, two hypothesis *classes* can be distinguished. Each of these classes comprises two *functions* on the second level, and to each of these functions, there exist two *arguments* on the third level that together form a pair of functional *substitutes*. This scheme is presented diagrammatically in Figure 1.

The two functions of implication hypotheses are *sufficiency* and *necessity*, whose two arguments are *absence* and *presence*. The two functions of covariation hypotheses are *positivity* and *negativity*, whose two arguments are

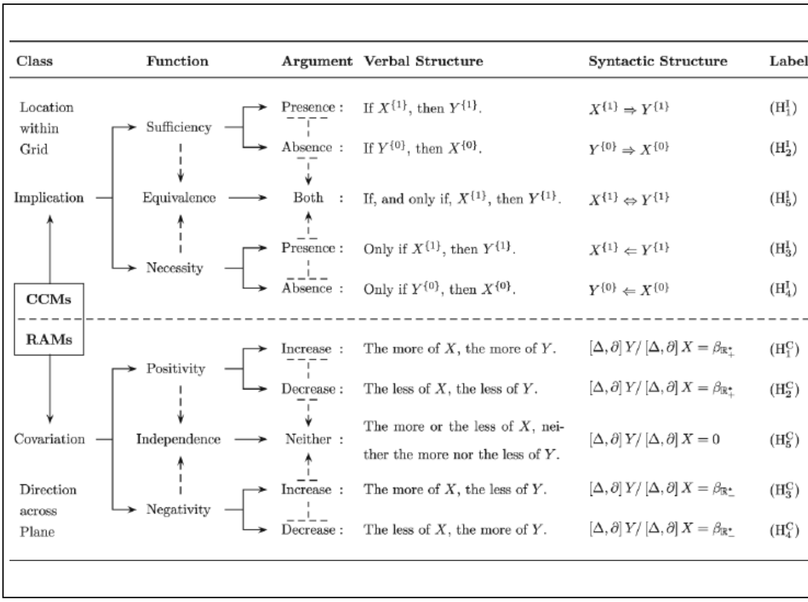

Figure 1. Hypothesis classes, functions, and arguments.
CCM = configurational comparative method; RAM = regressional analytic method.

increase and *decrease*. An implication hypothesis that features both sufficiency and necessity as functions to which absence and presence or presence and absence are supplied gives rise to an *equivalence* hypothesis.[20] A covariation hypothesis that features neither positivity nor negativity as functions, and thus neither increase nor decrease as arguments, generates an *independence* hypothesis.[21] In the remainder of this section, we elaborate on each combination of classes, functions, and arguments, first for implication and subsequently for covariation hypotheses.

A very common implication hypothesis involves the presence of condition $X^{\{1\}}$ and outcome $Y^{\{1\}}$ as arguments to the function of sufficiency. Strangely though, explicit sufficiency hypotheses seem difficult to find in the social sciences (cf. Goertz, 2003, p. 73). A few exceptions exist nonetheless. Gleditsch (1995), for instance, maintains that "nuclear deterrence may be interpreted as a *sufficient condition* for peace" (p. 543), and Landry, Davis, and Wang (2010) argue that electoral "competition defined as choice between candidates still is sufficient to engage voters" (p. 782).

Verbally, hypotheses of this type are thus usually put forward in one of the phrasings given by $(\mathrm{H}_1^{\mathrm{I}})$:

$$\text{If } X^{\{1\}}, \text{ then } Y^{\{1\}} \text{ / } X^{\{1\}} \text{ is sufficient for } Y^{\{1\}}. \qquad (\mathrm{H}_1^{\mathrm{I}})$$

Hypothesis type $(\mathrm{H}_1^{\mathrm{I}})$ can be syntactically rendered as $X^{\{1\}} \Rightarrow Y^{\{1\}}$ or, following (DF_1) and (DF_2), as $\neg(X^{\{1\}} \wedge Y^{\{0\}})$ and $X^{\{0\}} \vee Y^{\{1\}}$. Its functional yet less natural substitute is given by $(\mathrm{H}_2^{\mathrm{I}})$:

$$\text{If } Y^{\{0\}}, \text{ then } X^{\{0\}} \text{ / } Y^{\{0\}} \text{ is sufficient for } X^{\{0\}}. \qquad (\mathrm{H}_2^{\mathrm{I}})$$

This type is syntactically codified as $Y^{\{0\}} \Rightarrow X^{\{0\}}$ or again, following (DF_1) and (DF_2) prior to invoking the law of commutativity in (BA_1), as $\neg(X^{\{1\}} \wedge Y^{\{0\}})$ and $X^{\{0\}} \vee Y^{\{1\}}$. Note that although $(\mathrm{H}_1^{\mathrm{I}})$ proceeds from the presence of the condition, whereas $(\mathrm{H}_2^{\mathrm{I}})$ does so from the absence of the outcome, both hypotheses involve exactly the same proposition.

Hypotheses about the necessity of a condition for an outcome are considerably more common than those of $(\mathrm{H}_1^{\mathrm{I}})$.[22] For instance, Fortin (2012) concludes that "effective state capacity seems to be a necessary—but not sufficient—condition for democracy" (p. 904) and North and Weingast (1989) are convinced "that one necessary condition for the creation of modern economies dependent on specialization and division of labor . . . is the ability to engage in secure contracting across time and space" (p. 831). Hypotheses of this type are generally phrased in either of the two forms given by $(\mathrm{H}_3^{\mathrm{I}})$:

$$\text{Only if } X^{\{1\}}, \text{ then } Y^{\{1\}} \text{ / } X^{\{1\}} \text{ is necessary for } Y^{\{1\}}. \qquad (\mathrm{H}_3^{\mathrm{I}})$$

They are usually denoted by the Boolean-algebraic expression introduced in (DF_3) as $X^{\{1\}} \Leftarrow Y^{\{1\}}$, and equivalent in content with the type given by the two forms in $(\mathrm{H}_4^{\mathrm{I}})$:

$$\text{Only if } Y^{\{0\}}, \text{ then } X^{\{0\}} \text{ / } Y^{\{0\}} \text{ is necessary for } X^{\{0\}}. \qquad (\mathrm{H}_4^{\mathrm{I}})$$

The Boolean-algebraic expression $Y^{\{0\}} \Leftarrow X^{\{0\}}$ codifies the content of $(\mathrm{H}_4^{\mathrm{I}})$. Note again that, although $(\mathrm{H}_3^{\mathrm{I}})$ proceeds from the presence of the condition, whereas $(\mathrm{H}_4^{\mathrm{I}})$ does so from the absence of the outcome, both hypothesis types convey the same proposition. The conjunction of $(\mathrm{H}_1^{\mathrm{I}})$ and $(\mathrm{H}_3^{\mathrm{I}})$ or of $(\mathrm{H}_2^{\mathrm{I}})$ and $(\mathrm{H}_4^{\mathrm{I}})$ yields an equivalence hypothesis. Similar to bare sufficiency hypotheses, however, examples of this type seem extremely rare.

One exception that involves a conjunction of conditions can be found in the literature on altruistic behavior. For instance, Gross (1994) holds that "[b]ecause Jewish rescue in Western Europe was largely a collective undertaking, individual interests and motivations were a necessary but not a sufficient cause of successful action, requiring in addition suitable social and political conditions" (p. 489). Most naturally, this type is phrased in one of the two forms given by $(\mathrm{H}_5^{\mathrm{I}})$:

$$X^{\{1\}} \text{ if, and only if, } Y^{\{1\}} \;/\; X^{\{1\}} \text{ is necessary and sufficient for } Y^{\{1\}}. \qquad (\mathrm{H}_5^{\mathrm{I}})$$

The expression $X^{\{1\}} \Leftrightarrow Y^{\{1\}}$ introduced in (DF_5) is used to denote this type. Hypotheses $(\mathrm{H}_1^{\mathrm{I}})$ to $(\mathrm{H}_4^{\mathrm{I}})$, thus, constitute the set of fundamental implication hypotheses. Every compound such as equivalence can be constructed by connecting them. The four fundamental forms of implication hypotheses, and by extension also all compounds, are therefore semantically tied to a Boolean algebra.

In contradistinction to implication hypotheses, social scientists are considerably more practiced in using and interpreting covariation hypotheses.[23] A very common type passes the increase in the regressor X and the increase in the regressand Y as arguments to the function of positivity. These hypotheses are usually phrased as given by $(\mathrm{H}_1^{\mathrm{C}})$:[24]

$$\text{The more of } X, \text{ the more of } Y. \qquad (\mathrm{H}_1^{\mathrm{C}})$$

Syntactically, $(\mathrm{H}_1^{\mathrm{C}})$ can be translated as $\Delta Y / \Delta X = \beta \; {}^{*}_{+}$ for discrete and $\partial Y / \partial X = \beta \; {}^{*}_{+}$ for instantaneous changes, where β denotes the marginal effect of X on Y and ${}^{*}_{+}$ the set of real positive numbers. A functional substitute is the less naturally occurring type coupling positivity and decrease as given by $(\mathrm{H}_2^{\mathrm{C}})$:

$$\text{The less of } X, \text{ the less of } Y. \qquad (\mathrm{H}_2^{\mathrm{C}})$$

Note that the syntactic representation remains the same regardless of the arguments passed to the function of positivity. Hypotheses positing negativity are generally phrased most naturally in the form given by $(\mathrm{H}_3^{\mathrm{C}})$:

$$\text{The more of } X, \text{ the less of } Y. \qquad (\mathrm{H}_3^{\mathrm{C}})$$

Syntactically, $(\mathrm{H}_3^{\mathrm{C}})$ can be codified as $\Delta Y / \Delta X = \beta \; {}^{*}_{-}$ for discrete and $\partial Y / \partial X = \beta \; {}^{*}_{-}$ for instantaneous changes, where ${}^{*}_{-}$ designates the set of real negative numbers. Its functional substitute less naturally combines a decrease in X with an increase in Y as given by $(\mathrm{H}_4^{\mathrm{C}})$:

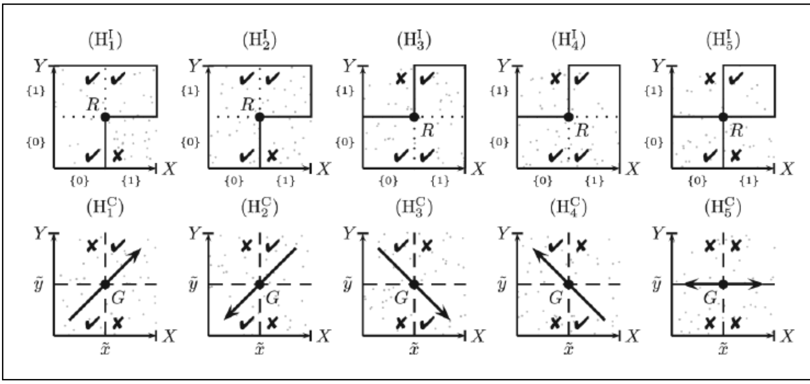

Figure 2. Visualization of basic implication and covariation hypotheses.

The less of X, the more of Y. $(\mathrm{H}^{\mathrm{C}}_4)$

The simultaneous negation of $(\mathrm{H}^{\mathrm{C}}_1)$ and $(\mathrm{H}^{\mathrm{C}}_4)$ or of $(\mathrm{H}^{\mathrm{C}}_2)$ and $(\mathrm{H}^{\mathrm{C}}_3)$ yields an independence relation. Although rarely formulated this way by researchers, such hypotheses accord to $(\mathrm{H}^{\mathrm{C}}_5)$:

The more or the less of X, neither the more nor the less of Y. $(\mathrm{H}^{\mathrm{C}}_5)$

This type can be expressed linear-algebraically as $\Delta Y/\Delta X = 0$ for discrete and $\partial Y/\partial X = 0$ for instantaneous changes. In point of fact, $(\mathrm{H}^{\mathrm{C}}_5)$ represents the most common type in applications of RAMs. Figure 2 visualizes all eight fundamental hypothesis types introduced above and the two compound forms. Squares indicate the spaces that the respective implication hypothesis generates, whereas arrows denote the direction of change around the joint arithmetic average at point $G(\tilde{x}, \tilde{y})$ —the center of gravity through which any regression line has to pass. Point R in the panel row above is not the equivalent of G for implication hypotheses. Both may coincide if the calibration of the condition and outcome is based on empirical measures of central tendency, but more often than not they will be different, and considerably so.[25] Crosses indicate areas where observations falsify the hypothesis in relation to the grid of space, the joint arithmetic average, respectively, whereas tick marks signify areas where they corroborate it.[26] For example, the plot for $(\mathrm{H}^{\mathrm{I}}_5)$ in the rightmost panel of the top row shows that, following (DF_6) and (DF_8), observations in the two spaces of the grid delineated by $X^{\{1\}} \wedge Y^{\{0\}}$ and $X^{\{0\}} \wedge Y^{\{1\}}$ falsify the hypothesis, whereas observations in the spaces delineated by $X^{\{0\}} \wedge Y^{\{0\}}$ and $X^{\{1\}} \wedge Y^{\{1\}}$ do not.

In summary, implication hypotheses always require variables to take on specific values that can be true or false for any given object. Covariation hypotheses only require variables to stand in some functional relationship. It is therefore uninformative to hypothesize that "the more of $X^{\{1\}}$, the more of $Y^{\{1\}}$", or, conversely, that "if X, then Y". A country having a *high* degree of social heterogeneity cannot simultaneously have a higher degree of social heterogeneity, just as the degree of social heterogeneity of a country cannot be necessary for the number of legislative parties.[27] A condition cannot increase or decrease (unlike regressors or regressands), and a regressor cannot be true or false (unlike conditions and outcomes). In consequence, *implication hypotheses are based on a Boolean algebra and therefore associated with CCMs, whereas covariation hypotheses are based on a linear algebra and therefore associated with RAMs*. Claiming that the latter are superior to the former for building or testing implication hypotheses (e.g., Clark et al., 2006; King et al., 1994) is thus tantamount to ignoring semantics.

Causal Complexity

The third area of misunderstandings to be addressed concerns the difference between Boolean and linear-algebraic products. Following terminological conventions, we refer to the former as *conjunctions* and to the latter as *interactions*. More specifically, we demonstrate that the main difference between these two constructs resides in the fact that implication hypotheses involving conjunctions give rise to Boolean expressions that delineate multi-dimensional grids of spaces, whereas covariation hypotheses involving interactions give rise to linear-algebraic expressions that delineate discrete or continuous multi-dimensional planes. Contrary to received wisdom, the degree of complexity of either construct is irrelevant to this difference. Both conjunctions and interactions can be of any order within the constraints set by the number of arguments to their higher-level functions. Before we reanalyze an influential study in this connection, a review of the applied and methodological bodies of literature reveals the current state of confusion.

Amenta and Poulsen (1996), for instance, argue that

> social spending outcomes are due to complex interactions .... Because of multicollinearity and losses of degrees of freedom . . . these interactions are sometimes ignored. Qualitative comparative analysis offers a solution . . . (p. 55-56)

Similarly, Davidsson and Emmenegger (2013) emphasize that the "small number of observations would not allow for the inclusion of multiple interaction terms. In contrast, fsQCA can deal with complex causality even if the number of cases is relatively small" (p. 349). Heikkila (2004), in contrast, draws on QCA to identify "the interaction terms among variables . . . , which can complement the predicted interaction effects from the logit model" (p. 109). And for Grandori and Furnari (2008), the "choice of the data analysis method was driven by the need to detect interaction effects . . . " (p. 473), but "three-way interactions currently represent a limit for regression analysis applications. . . . For these reasons, . . . we found Boolean comparative analysis . . . the most suitable method for our purposes".[28] In summary, applied work seems to regard conjunctions and interactions as substitutes, but as the latter often create problems of a technical and/or interpretative nature, CCMs are considered an attractive alternative to RAMs.

In the methodological literature, Clark et al. (2006) argue that CCMs are dispensable because RAMs with interactions offer a superior means for testing (probabilistic) hypotheses about necessity and sufficiency relations, a view that appears to have recently convinced a number of scholars (Brady, 2013, pp. 258-263; Fiss, Sharapov, & Cronqvist, 2013, p. 192; Hug, 2013, p. 257). Mahoney and Goertz (2006) concede that "[t]his is not a completely unreasonable view . . . , for the logical AND is a first cousin of multiplication" (p. 235), to the effect that "as statistical comparativists start to use saturated interaction models in which all possible interactions are assessed and simplified in a top-down manner, we would essentially see an integration of QCA techniques and statistical methods" (Mahoney, 2008, p. 425). Griffin and Ragin (1994, p. 11) hold that QCA and logit regression are in fact alike, the only real difference being that the former is better at handling causal complexity. This counterargument to Clark et al. (2006) is supported by Vis (2012, p. 173) as well as Wagemann and Schneider (2010, p. 384), who consider interactions more problematic in interpretation and QCA as better suited for identifying causal complexity. Somewhere in between these standpoints, Grofman and Schneider (2009) argue that "once we have completed QCA we can use what we have learned to mimic its results with more traditional methods such as binary logistic regression . . . " (p. 669). To summarize, methodologists' opinions diverge greatly. Some consider CCMs and RAMs to be closely related if not substitutable in analyzing causal complexity, others believe CCMs to have an edge over RAMs, and still others not only take the opposite view but even argue that CCMs are vastly inferior to RAMs. In this section, we clarify the relation between conjunctions and interactions by integrating the key points from the two preceding sections. More specifically, we make the following assertion:

Contention 3: Conjunctions and interactions are incommensurable constructs that represent disparate concepts of causal complexity and therefore can neither be substituted for each other nor remedy the shortcomings of the counterpart. The former are Boolean-algebraic products of conditions that define complex grids of spaces, whereas the latter are linear-algebraic products of regressors that define complex planes.

As Clark et al. (2006; hereafter CGG) present the most unequivocal dismissal of CCMs, we develop our argument on the basis of their influential essay. The authors present a seemingly attractive argument: "to determine whether X_1 and/or X_2 is necessary, sufficient, or necessary and sufficient for Y" (p. 320), the general interaction model with two regressors X_1 and X_2 and a regressand Y given in Equation (1) suffices:[29]

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2 + \varepsilon. \quad (1)$$

Based on this model, CGG present eight combinations of coefficients, with each of which a specific "valid conclusion" concerning the type of implication between each regressor and the regressand is associated (CGG, Table 3, p. 322).[30] For example, the authors argue that the statistical insignificance of β_1 and β_2 in conjunction with the statistical significance of β_3 would warrant the conclusion that X_1 and X_2 are individually necessary but not sufficient for Y, whereas from the significance of β_1 and β_2 together with the insignificance of β_3, the conclusion could be drawn that X_1 and X_2 are individually sufficient but not necessary for Y.

Three problems arise with this logic. First, all of CGG's conclusions about necessity and sufficiency relations on the basis of regression coefficients run counter to the fundamentals of Boolean algebras and Boolean-algebraic hypothesis formulation. Implication hypotheses are posited that establish associations between regressors and regressands but, as illustrated in (DF_1) to (DF_8), CCMs model causes and effects in terms of variables taking on specific values. Implicational hypotheses about variables that are functionally related through arithmetic equality are ill-formed. Second, and in temporary disregard of the previous point, coefficient combinations are presented only for conclusions about single variables, but not for binary operations such as $X_1 \wedge X_2$ or $X_1 \vee X_2$, although these expressions are intrinsic to CCMs. Third and final, CGG note that they have omitted negative coefficients for ease of presentation but do not indicate how such coefficients should actually be interpreted with respect to necessity and sufficiency. Irrespective of how this would be done, we have shown above that the unary operation "−" does not travel in translation between linear and Boolean algebras.

In spite of these problems, the authors continue to reanalyze the argument made by Duverger (1954) about the structure of a country's party system.

They distill three implication hypotheses from Duverger's work, namely that multi-member districts ($MMD^{\{1\}}$) and high social heterogeneity ($SHG^{\{1\}}$) are individually necessary for a multi-party system ($MPS^{\{1\}}$), and that the conjunction of high social heterogeneity and multi-member districts is sufficient for a multi-party system (CGG, 2006, pp. 322-324).[31] Following (DF_3) and (H_3^I), the first two hypotheses can be expressed as (H_1): $MMD^{\{1\}} \Leftarrow MPS^{\{1\}}$ and (H_2): $SHG^{\{1\}} \Leftarrow MPS^{\{1\}}$. The third hypothesis constructs a causally complex argument by involving a conjunction, and can be expressed as (H_3): $SHG^{\{1\}} \wedge MMD^{\{1\}} \Rightarrow MPS^{\{1\}}$. It is surprising that CGG do not point out that (H_1) to (H_3) merge into a single equivalence hypothesis of type (H_5^I); if multi-member districts and high social heterogeneity are each hypothesized to be necessary for a multi-party system, then their conjunction must be necessary as well: $MMD^{\{1\}} \wedge SHG^{\{1\}} \Leftarrow MPS^{\{1\}}$. But if this conjunction is simultaneously hypothesized to be sufficient for a multi-party system according to (H_3), (H_1) to (H_3) integrate into $SHG^{\{1\}} \wedge MMD^{\{1\}} \Leftrightarrow MPS^{\{1\}}$. In other words, Duverger's work would have led to a single yet causally complex equivalence hypothesis. So as to stay close to CGG's original study and for ease of presentation, however, we continue to focus on each hypothesis separately.

To test (H_1) to (H_3), CGG estimate a model with dichotomized regressors for which "the connection between multiplicative interaction models and testing for necessary and/or sufficient conditions is clearest" (p. 325). This model is presented in Equation (2), where NP is the number of parties; MMD is a binary variable with integer 1 indicating multi-member districts and 0 single-member districts; and SHG is a binary variable with integer 1 indicating high social heterogeneity and 0 low social heterogeneity.[32] The dichotomization thresholds are derived both through data-based and theoretical criteria. A value of 1 is applied to differentiate single-member from multi-member districts, the sample median of 1.2775 to distinguish high from low social heterogeneity, and a lower bound of three parties is set to identify multi-party systems:

$$NP = \beta_0 + \beta_1 MMD + \beta_2 SHG + \beta_3 MMD * SHG + \varepsilon. \tag{2}$$

The authors interpret their results, which show β_1 and β_2 not to be significantly different from zero in contrast to β_3 (CGG, Table 4, p. 324), as corroborating their hypotheses. More specifically, they conclude that "an increase in the heterogeneity of a country is not expected to increase the size of the party system in countries with single-member districts" (p. 324). Conversely, an increase in district size is not expected to increase the size of the party system in countries whose social heterogeneity is low. These conclusions clearly follow hypothesis type (H_5^C) and the use of RAMs is

Table 2. Inclusion Score Tests For Outcome $MPS^{\{1\}}$ and Negated Outcome $MPS^{\{0\}}$.

			$MPS^{\{1\}}$				$MPS^{\{0\}}$		$MPS^{\{1\}}$		
C	$MMD^{\{\cdot\}}$	$SHG^{\{\cdot\}}$	*n*	*I* (∧, ⇒)	*n*	*I* (∨, ⇒)	*I* (∧, ⇒)	*I* (∨, ⇒)	*n*	*I* (∧, ⇐)	*I* (∨, ⇐)
1	0	0	9	0.333***	38	0.289***	0.667	0.711†††	20	0.150***	0.550**
2	0	1	11	0.000***	36	0.333***	1.000	0.667††	20	0.000***	0.600
3	1	0	18	0.444***	43	0.465***	0.556*	0.535***	20	0.400***	1.000
4	1	1	16	0.562*	45	0.378***	0.438***	0.622**†	20	0.450***	0.850
5	0	—	20	0.150***			0.850		20	0.150***	
6	1	—	34	0.500***			0.500***		20	0.850	
7	—	0	27	0.407***			0.593*		20	0.550**	
8	—	1	27	0.333***			0.667†		20	0.450***	

Note. MPS = multi-party system; *C* = condition; MMD = multi-member districts; SHG = social heterogeneity; *n* = number of cases; *I*(∧, ⇒) inclusion score for conjunction and sufficiency, and so on.
**p* < 0.10 for alternative hypothesis *Incl* < 0.75; ***p* < 0.05 for *Incl* < 0.75; ****p* < 0.01 for *Incl* < 0.75;
†*p* < 0.10 for *Incl* > 0.5; ††*p* < 0.05 for *Incl* > 0.5; †††*p* < 0.01 for *Incl* > 0.5.

appropriate. The conditioning on one value of the interacting regressor does not change this fact but only limits the domain over which the proposition about the (discrete) marginal rate of change between the regressor in question and the regressand is to assume validity.

At the same time, however, CGG see "strong evidence that both multi-member districts and social heterogeneity are necessary, but not sufficient, for more legislative parties" (p. 325). If "more legislative parties" is interpreted to mean "multi-party systems", as CGG explicitly do when they argue that "multi-member districts are necessary, but not sufficient, for multipartism" (p. 324), then these conclusions follow hypothesis type (H_3^I). However, if CGG had wanted to test (H_1) to (H_3), then, according to the definition of a Boolean implication, they should have tested for the absence of observations in specific locations within the grid of spaces defined by the thresholds used to dichotomize the regressors. For example, for (H_1) to receive probabilistic corroboration, significantly more cases must be observed that show $MPS^{\{1\}}$ in conjunction with $MMD^{\{1\}}$ than those showing $MPS^{\{1\}}$ in conjunction with $MMD^{\{0\}}$ according to (DF_3).[33] Similarly, (H_3) can be rewritten as $\neg(SHG^{\{1\}} \wedge MMD^{\{1\}} \wedge MPS^{\{0\}})$ according to (DF1), and can thus only be upheld if significantly more cases show $SHG^{\{1\}}$ and $MMD^{\{1\}}$ in conjunction with $MPS^{\{1\}}$ than those showing $SHG^{\{1\}}$ and $MMD^{\{1\}}$ in conjunction with $MPS^{\{0\}}$.[34]

In Table 2, we not only reassess (H_1) to (H_3) but also provide a comprehensive battery of tests for both $MPS^{\{1\}}$ as well as its Boolean negation $MPS^{\{0\}}$—a two-party system—using conventional CCM inclusion score tests.[35] Results are presented for conjunctive ($I(\wedge, \Rightarrow)$) and disjunctive sufficiency ($I(\vee, \Rightarrow)$) with respect to $\text{MPS}^{\{1\}}$ and $\text{MPS}^{\{0\}}$ as well as conjunctive ($I(\wedge, \Leftarrow)$) and disjunctive necessity ($I(\vee, \Leftarrow)$) with respect to $\text{MPS}^{\{1\}}$.[36] For computing inclusion scores, we apply the inclusion ratios for sufficiency and necessity given in Equations (3) and (4), respectively, where m_i is the membership of case i in the condition and the outcome (Smithson & Verkuilen, 2006, pp. 11, 65-68):

$$I\left(C \Rightarrow \text{MPS}^{\{\cdot\}}\right) = \frac{\sum_{i=1}^{n} \min\left(m_i\left(C\right), m_i\left(\text{MPS}^{\{\cdot\}}\right)\right)}{\sum_{i=1}^{n} m_i\left(C\right)}, \quad (3)$$

$$I\left(C \Leftarrow \text{MPS}^{\{\cdot\}}\right) = \frac{\sum_{i=1}^{n} \min\left(m_i\left(C\right), m_i\left(\text{MPS}^{\{\cdot\}}\right)\right)}{\sum_{i=1}^{n} m_i\left(\text{MPS}^{\{\cdot\}}\right)}. \quad (4)$$

With k variables j of p_j values, there exist $\prod_{j=1}^{k}\left(p_j + 1\right) - 1 = 3^2 - 1 = 8$ (complex) conditions C, each of which is observed n times. As I is a proportion, binomial tests can be performed to adjudicate between rival hypotheses. We use the QCA-package for the *R* environment to this end (Dușa & Thiem, 2014; Thiem & Dușa, 2012; 2013a; 2013b).

At a minimum, the hypothesis that a condition is sufficient for an outcome would never be upheld if I was not significantly greater than 0.5, in which case more evidence may exist for a sufficiency relation between the condition and the negation of the outcome. Moreover, inclusion scores of substantive significance should ideally exceed a value of 0.75 (Ragin, 2008, p. 46; Schneider & Wagemann, 2012, p. 129). In this case, there are 3 times as many observations that show the conjunction of the condition and the outcome as there are observations exhibiting the conjunction of the condition and the negation of the outcome.

It is incontestable that none of the four complex conditions C_1 to C_4 would be considered sufficient for $\text{MPS}^{\{1\}}$ by any standards of CCM research, which ultimately also means that (H_3) should be rejected. Contrary to what CGG argue, high social heterogeneity in conjunction with the presence of multi-member districts is not sufficient for the presence of a multi-party system. All that can be inferred with regard to (H_3) at conventional levels of statistical significance ($p < 0.1$) is that the ratio between the number of cases that show both conditions as well as the outcome and the number of cases that show both conditions but not the outcome amounts to less than 0.75, which is equally true for all inclusion scores that exceed 0.5 (only C_4) as well as those that do not (C_1, C_2, and C_3). The only notable score in this column is that of C_2. As it is 0, all cases that show the conjunction of the absence of multi-member districts and high social heterogeneity must be associated with two-party systems. The complementary inclusion score of 1 is not merely statistically indistinguishable from the threshold of 0.75 with 11 observations but significantly higher.

Similarly, (H_2) should be rejected because high social heterogeneity, at an inclusion score of 0.45, is far from passing as a necessary condition for a multi-party system. If anything, there is more evidence for the hypothesis that high social heterogeneity represents a necessary condition for a two-party system.[37] Only (H_1) receives corroboration, but not because the regression coefficients β_1 and β_2 for Model (1) were indistinguishable from zero while β_3 was statistically significant. Instead, it is corroborated because the inclusion test produces a score of 0.85, which is statistically indistinguishable from the threshold of 0.75. Because of the monotonicity of necessity relations with respect to the operation of disjunction, any condition that is added to multi-member districts will never detract from the inclusion score of the initial condition (Baumgartner, 2013b, p. 90). This behavior can be observed for the disjunction of multi-member districts and high social heterogeneity, for which the inclusion score remains at 0.85, and for the disjunction of multi-member districts and low social heterogeneity, for which it increases to unity.

Words do not readily convey the difference between the testing of causally complex hypotheses with CCMs and that with RAMs. Figure 3 thus provides a three-dimensional visualization of CGG's data, the corresponding regression plane defined by Model (2) together with its 95% confidence interval, and the implication space of hypothesis (H_3). Two different perspectives are provided for better orientation. Equation (2), which models the interaction between social heterogeneity and district magnitude, produces a regression plane of four points because each value of one variable can be combined with one of two values of the other variable. In contrast, the conjunction in (H_3) hypothesized to be sufficient for a multi-party system produces a grid of four spaces along the dichotomization thresholds of the three variables. The two most important spaces have been enclosed by gray-transparent rectangles.

Cases in the space created by $SHG^{\{1\}} \wedge MMD^{\{1\}} \wedge MPS^{\{1\}}$ corroborate (H_3), whereas those in the space given by $SHG^{\{1\}} \wedge MMD^{\{1\}} \wedge MPS^{\{0\}}$, and only those, act as falsifiers according to (DF_1). Countries with low social heterogeneity or single-member districts or both, regardless of the number of parties, do not contradict the claim that the combination of high social heterogeneity

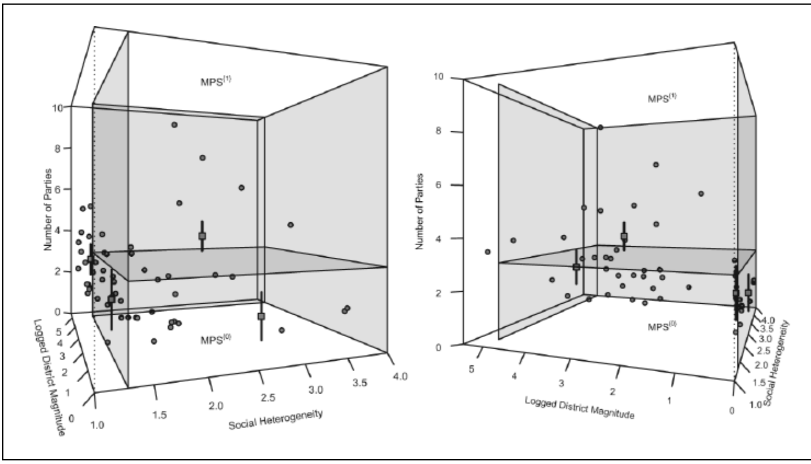

Figure 3. Predicted number of parties (squares) with 95% confidence intervals, and implication space of $SHG^{\{1\}} \wedge MMD^{\{1\}} \wedge MPS^{\{1\}}$ and $SHG^{\{1\}} \wedge MMD^{\{1\}} \wedge MPS^{\{0\}}$.
SHG = social heterogeneity; MMD = multi-member districts; MPS = multi-party system.

and the presence of multi-member districts is sufficient for multi-party systems, although, according to CGG's logic, any of these cases that show a multi-party system should undermine the hypothesis. Put again in graphical language, RAM researchers hunt for areas on regression planes that show discernible deviations from flatness (where one regressor has no effect on the regressand given a (range of) value(s) of the interacting regressor), whereas CCM researchers seek to collapse grids of implication spaces (where one condition has no verifiable effect on the outcome given the presence of the conjunction of other conditions).

In conclusion, irrespective of the setting of any of the 27 different constellations of regression coefficients in a RAM interaction model such as Equation (2), *conjunctions and interactions are incommensurable constructs modeling causal complexity that can neither be substituted for each other nor remedy the shortcomings of the counterpart*.[38] By extension, neither do saturated interaction models integrate conjunctions, as Mahoney (2008, p. 425) claims, nor can conjunctions be conceptually mimicked by interactions, as suggested by Grofman and Schneider (2009, p. 669).

Conclusion

Fundamental misunderstandings about algebraic systems, hypotheses classes, and the concept of causal complexity have been blocking progress in the debate between configurational comparativists and regressional analysts for more than a quarter-century. Contrary to expectations, the sources of these misunderstandings have not mainly resided in problems of communication between the two communities, but in knowledge gaps and ambiguous definitions of concepts within these two communities. The objective of this article has been to clear this blockage once and for all by showing how the differences arising under these aspects inform the distinct purposes of CCMs and RAMs.

It has first been demonstrated that CCMs are based on Boolean algebra, whereas RAMs work according to the laws of linear algebra, both of which give rise to semantically incommensurable languages despite occasional resemblances in syntax. If the debate is to progress at any rate, representatives of CCMs and those of RAMs cannot be spared from gaining more proficiency in the mathematical formalities of these languages. Casual dismissals of Boolean algebra by proponents of RAMs should not be tolerated any more, just as elementary misinterpretations such as those concerning the Boolean implication operator need to be finally overcome by proponents of CCMs.

It has then been argued that hypotheses formulated in social-scientific research for purposes of (causal) inference generally divide into implication and covariation hypotheses, the former of which posit implicational associations between conditions and outcomes, and the latter of which posit covariational associations between regressors and regressands. Just as "condition" is not merely a CCM term for "regressor", an "outcome" is something entirely different from a "regressand". When researchers design their projects and formulate hypotheses, they need to be aware of the consequences their decisions entail in this connection.

Finally, we have juxtaposed conjunctions and interactions so as to emphasize the disparate concepts behind these constructs, and to argue that CCMs and RAMs cannot substitute for each other in analyzing causal complexity. Graphically speaking, conjunctions delineate grids of spaces, whereas interactions produce discrete or continuous planes. Methodologists should thus stop arguing about the superiority of one set of methods in dealing with causal complexity, and instead begin to appreciate their distinct capabilities, leveraging respective strengths wherever apposite.

Linear algebra is essential to an understanding of RAMs, just as Boolean algebra is indispensable for comprehending the principles of CCMs. Most methods curricula for political scientists and sociologists at university departments around the world, however, assume students to be familiar with the fundamentals of only the former. Unless the horizon is broadened, we will see more incendiary works being published over the coming years that reinforce the misunderstandings addressed in this article. To prevent such methodological inertia or even regress, methods curricula should thus feature introductions to propositional logic as that branch of Boolean algebra closest to the social sciences. The tools of formal logic are not the exclusive domain of analytic philosophers, electrical engineers, or genetic biologists. If they want to employ or judge them, political scientists and sociologists must begin to eventually acquaint themselves more thoroughly with these tools as well.

Acknowledgments

Previous versions of this article have been presented at the Annual General Conference of the European Political Science Association, Edinburgh, 19-21 June 2014 and the Annual Meeting of the American Political Science Association, Washington, DC, 28-31 August 2014. We thank Barry Cooper, Gary Goertz, two anonymous reviewers of *CPS*, the editors of *CPS*, Benjamin Ansell and David Samuels, and the participants at the aforementioned conferences for thought-provoking comments and very helpful suggestions.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Alrik Thiem and Michael Baumgartner gratefully acknowledge financial support from the Swiss National Science Foundation, award number PP00P1_144736/1.

Notes

1. Ragin's ideas have not only been discussed in influential textbooks (Brady & Collier, 2004; Gerring, 2001, 2012; Goertz & Mahoney, 2012; King et al., 1994) but have also been the center of several journals issues, including the special issue "Formal Methods of Qualitative Analysis" in *Sociological Methods & Research*, 1994, 23(1); the symposium on comparative macro-sociology in *Comparative Social Research*, 1997, 16(1); the symposium "QCA: Qualitative Comparative Analysis" in *Qualitative Methods*, 2004, 2(2); the "Symposium on Qualitative Comparative Analysis" in *Studies in Comparative International Development*, 2005, 40(1); the special issue "Fuzzy Sets and Social Research" in *Sociological Methods & Research*, 2005, 33(4); the exchange on case relevance and the special issue "Causal Complexity and Qualitative Methods" in *Political Analysis*, 2002, 10(2) and 2006, 14(3); the special issue on comparative research in *International Sociology*, 2006, 21(5); the symposia on Goertz and Mahoney (2012) in *Comparative Political Studies*, 2013, 46(2) and *Qualitative &*

Multi-Method Research, 2013, 11(1); the symposium "Qualitative Comparative Analysis at 25" in *Political Research Quarterly*, 2013, 66(1); the special issue "Innovative Methods for Policy Analysis: QCA and Fuzzy Sets" in *Policy & Society*, 2013, 32(4); the symposium "The Set-Theoretic Comparative Method" and the exchange on Hug (2013) and Thiem (2014a) in *Qualitative & Multi-Method Research*, 2014, 12(1/2); and the symposium "Qualitative Comparative Analysis" in *Sociological Methodology*, 2014, 44(1).

2. See http://www.compasss.org/bibdata.htm for a comprehensive bibliographical database of QCA-related publications (accessed September 20, 2014).
3. We define configurational comparative methods (CCMs) to include all case study methods that are based on Boolean-algebraic principles such as Event Structure Analysis (Griffin, 1993; Heise, 1989), all variants of QCA (Cronqvist & Berg-Schlosser, 2009; Ragin, 1987, 2000, 2008; Thiem, 2013, 2014b, 2014c; Vink & Vliet, 2009) and Coincidence Analysis (CNA; Baumgartner, 2009, 2013a), whereas regressional analytic methods (RAMs) comprise all basic parametric regression methods as introduced in standard textbooks (Fox, 2008; Wooldridge, 2008). CCMs thus form a subset of what is typically considered the set of "qualitative" methods and RAMs a subset of the set of "quantitative" methods. To avoid the impression that the field of social research methods is dominated by the CCM–RAM controversy, we explicitly acknowledge the existence of other debates between methodological communities such as Bayesian versus frequentist statisticians (Gill, 1999) or data versus algorithmic modelers (Breiman, 2001).
4. We use the indefinite article because "algebra" describes a discipline but "an algebra" is a concrete mathematical object.
5. Some of these errors are detailed later. Schneider and Wagemann use the term *set-theoretic methods* as a synonym for CCMs.
6. Additional misunderstandings exist, for instance, regarding the concept of equifinality (Goertz & Mahoney, 2012), but they are less fundamental and therefore not part of this article.
7. The syntax of a formal language $\mathcal{L}$ provides its symbols as well as the rules regulating the construction of complex well-formed expressions of $\mathcal{L}$ from atomic ones and the transformation of well-formed expressions into other well-formed expressions. In contrast, the semantics of $\mathcal{L}$ determine the interpretation of its well-formed expressions.
8. While crisp-set QCA and CNA use Boolean techniques, fuzzy-set QCA adds elements from fuzzy logic. Still, the minimization procedure that effectuates the actual analysis of the data follows Boolean-algebraic principles (the procedure described in Eliason and Stryker, 2009, represents an exception). For an axiomatic comparison of Boolean and fuzzy logic, see Buckley and Eslami (2002). Although we have also included mvQCA in the set of CCMs, we do not introduce the method's algebraic system as this requires a generalization of Boolean algebra that would go beyond the scope of this article. We refer interested readers to Dubrova (2002).

9. For a minimal definition, laws (BA_1), (BA_3), (BA_5), and (BA_7) suffice. All other laws are derivable (cf. Givant & Halmos, 2009, p. 10; McCluskey, 1965, p. 98).
10. This Boolean-algebraic law is called *idempotency.*
11. For a compact overview, see South (1974, pp. 12-13). For a more thorough introduction, see Hohn (1966). Unfortunately, representatives of CCMs have often spoken of "set theory", "formal logic", and "Boolean algebra" as if they were alternatives without explaining the relations between these systems (e.g., Marx, Rihoux, & Ragin, 2014, p. 115; Schneider & Wagemann, 2010, p. 404).
12. We apply the terms *cause* and *effect* rather loosely to mean *input* and *output* in applications of CCMs and RAMs for purposes of causal modeling. Needless to say, none of these methods can conclusively identify causality.
13. Other names for "regressor" and "regressand" are "independent variable" and "dependent variable", respectively. We avoid these terms for reasons of clarity.
14. The superscripts "$\{0\}$" and "$\{1\}$" must not be confused with the identity elements in $\mathcal{I}$.
15. The corresponding set-theoretical operations are denoted by the subset operator "$\subset$" ($\Rightarrow$) and the superset operator "$\supset$" ($\Leftarrow$).
16. A *hypothesis* is a proposition for answering a question or solving a problem that does not merely restate the question or problem, and that is free from ambiguities, internally consistent and precise enough to be testable (cf. Babbie, 2007, pp. 44, 47; Copi, Cohen, & Flage, 2007, pp. 347-353).
17. Clark, Gilligan, and Golder (CGG; 2006) use the term *asymmetric* to refer to (probabilistic) relations of necessity and sufficiency.
18. We deliberately use graphical language to facilitate comprehension.
19. We do not consider covariation hypotheses that specify the strength of the rate of change in addition to its direction because these are relatively rare. Nor do we explicitly address more complex functional forms such as higher-order polynomials of the same regressor or non-linear models. Similarly, we also do not consider implication hypotheses that specify the distribution of cases in addition to their location.
20. Recall from (DF_5) that an equivalence is a dual implication.
21. Instead of equivalence, we could also have used *independence* in implication hypothesis as the counterpart to independence in covariation hypotheses. However, such hypotheses are considerably less common in CCM research than hypotheses of equivalence.
22. For a rich overview of such hypotheses from political science, sociology, and economics, see Goertz (2003). Psychological studies in human concept learning indicate that this phenomenon might be due to the simple fact that equivalent sufficiency hypotheses of necessity hypotheses are of higher complexity and thus more difficult to process for humans (Feldman, 2000). For an exception, see Lichbach (1981).
23. Because of their commonality, and for reasons of space, we do not provide examples.
24. As all variables have to be numerically codified to be processable by RAMs, all types of hypotheses to follow apply to continuous and categorical variables alike.

In the case of categorical regressands, probabilities take the place of expected values.

25. Such calibration approaches are subsumed under the heading of *data-based transformations* (Verkuilen, 2005, pp. 481-483). The well-known Human Development Index is a typical example.
26. Data points are fictitious and have only been inserted for graphical enrichment.
27. In contrast, the hypothesis that "*some* degree of social heterogeneity is necessary for *some* number of legislative parties" is meaningful although its informational content is low. In this case, the determiner *some* denotes an exact yet unspecified value of each variable.
28. This selection of studies represents only a small part of the applied literature that prefers QCA to regression models with interaction terms because of the reasons mentioned. Very few authors such as Kalleberg and Vaisey (2005) seem to be aware that "[a] configuration and an interaction term are not conceptually the same thing nor are they mathematically equivalent" (p. 447).
29. A similar yet less detailed argument is provided by Brady (2013, pp. 258-263). In contrast to CGG, however, he proposes a RAM interaction model without constituent terms to test for necessary conditions, and a simple RAM model without interactions to test for sufficient conditions. Besides the fact that interaction models without constituent terms impose oft-unwarranted constraints on parameters, Brady's argument is beset by the same fundamental problems that we elaborate on below with regard to CGG's study.
30. Note the typos in the header of Table 3 in CGG (2006, p. 322); X should have been X_1 and Z should have been X_2.
31. For ease of presentation, we use upright instead of italicized capital letters in this section to denote variables.
32. For the variable "social heterogeneity" (SHG), we use the label "low" as synonymous with "not high".
33. Trivial cases of corroboration where only instances of two-party system ($\mathrm{MPS}^{\{0\}}$) exist are not considered.
34. Trivial cases of corroboration where only instances of low social heterogeneity ($\mathrm{SHG}^{\{0\}}$) or single-member districts ($\mathrm{MMD}^{\{0\}}$) exist are not considered.
35. Note that "inclusion" is equivalent to "consistency" (Ragin, 2006; Smithson, 2005). Usually, CCMs such as QCA and CNA proceed iteratively instead of performing all tests at once. As the primary concern here lies in testing implicational hypotheses, we leave the otherwise important topic of minimization in CCM research aside.
36. For reasons of space, we omit the test results for disjunctive necessity with respect to $\mathrm{MPS}^{\{0\}}$. Some tests are redundant given the results of other tests, but we include them for reasons of completeness and comparability.
37. This result is not displayed in Table 2. See the replication file for details.
38. Assuming that any one of the three coefficients in Model (2) can be equal to, smaller than or larger than zero, 27 combinations result.

Supplemental Material

Replication files for this article are available in its online appendix or from the corresponding author on request. The online replication materials are available at http://cps.sagepub.com/supplemental.

References

Amenta, E., & Poulsen, J. D. (1996). Social politics in context: The institutional politics theory and social spending at the end of the New Deal. *Social Forces*, *75*, 33-60.

Babbie, E. (2007). *The practice of social research* (11th ed.). Belmont, CA: Thomson/Wadsworth.

Baumgartner, M. (2009). Inferring causal complexity. *Sociological Methods & Research*, *38*, 71-101.

Baumgartner, M. (2013a). Detecting causal chains in small-n data. *Field Methods*, *25*, 3-24.

Baumgartner, M. (2013b). A regularity theoretic approach to actual causation. *Erkenntnis*, *78*, 85-109.

Brady, H. E. (2013). Do two research cultures imply two scientific paradigms? *Comparative Political Studies, 46*, 252-265.

Brady, H. E., & Collier, D. (Eds.). (2004). *Rethinking social inquiry: Diverse tools, shared standards*. Lanham, MD: Rowman & Littlefield.

Braumoeller, B. F., & Goertz, G. (2000). The methodology of necessary conditions. *American Journal of Political Science*, *44*, 844-858.

Breiman, L. (2001). Statistical modeling: The two cultures. *Statistical Science*, *16*, 199-231.

Buckley, J. J., & Eslami, E. (2002). *An introduction to fuzzy logic and fuzzy sets*. Heidelberg, Germany: Physica.

Clark, W. R., Gilligan, M. J., & Golder, M. (2006). A simple multivariate test for asymmetric hypotheses. *Political Analysis*, *14*, 311-331.

Copi, I. M., Cohen, C., & Flage, D. E. (2007). *Essentials of logic* (2nd ed.). Upper Saddle River, NJ: Pearson.

Cronqvist, L., & Berg-Schlosser, D. (2009). Multi-Value QCA (mvQCA). In B. Rihoux & C. C. Ragin (Eds.), *Configurational comparative methods: Qualitative Comparative Analysis (QCA) and related techniques* (pp. 69-86). London, England: Sage.

Davidsson, J. B., & Emmenegger, P. (2013). Defending the organisation, not the members: Unions and the reform of job security legislation in Western Europe. *European Journal of Political Research*, *52*, 339-363.

Dubrova, E. (2002). Multiple-valued logic synthesis and optimization. In S. Hassoun & T. Sasao (Eds.), *Logic synthesis and verification* (pp. 89-114). New York, NY: Springer Science + Business Media.

Duşa, A., & Thiem, A. (2014). *QCA: A package for Qualitative Comparative Analysis, R* package version 1.1-4. Retrieved from: http://cran.r-project.org/package=QCA

Duverger, M. (1954). *Political parties*. New York, NY: John Wiley.
Eliason, S. R., & Stryker, R. (2009). Goodness-of-fit tests and descriptive measures in fuzzy-set analysis. *Sociological Methods & Research*, *38*, 102-146.
Feldman, J. (2000). Minimization of Boolean complexity in human concept learning. *Nature*, *407*, 630-633.
Fiss, P. C., Sharapov, D., & Cronqvist, L. (2013). Opposites attract? Opportunities and challenges for integrating large-n QCA and econometric analysis. *Political Research Quarterly*, *66*, 191-198.
Fortin, J. (2012). Is there a necessary condition for democracy? The role of state capacity in postcommunist countries. *Comparative Political Studies*, *45*, 903-930.
Fox, J. (2008). *Applied regression analysis and generalized linear models* (2nd ed.). London, England: Sage.
Gerring, J. (2001). *Social science methodology: A critical framework*. Cambridge, UK: Cambridge University Press.
Gerring, J. (2012). *Social science methodology: A unified framework* (2nd ed.). Cambridge, UK: Cambridge University Press.
Gill, J. (1999). The insignificance of null hypothesis significance testing. *Political Research Quarterly*, *52*, 647-674.
Givant, S., & Halmos, P. (2009). *Introduction to Boolean algebras*. New York, NY: Springer.
Gleditsch, N. P. (1995). Democracy and the future of European peace. *European Journal of International Relations*, *1*, 539-571.
Goertz, G. (2003). The substantive importance of necessary condition hypotheses. In G. Goertz & H. Starr (Eds.), *Necessary conditions: Theory, methodology, and applications* (pp. 65-94). Lanham, MD: Rowman & Littlefield.
Goertz, G., & Mahoney, J. (2012). *A tale of two cultures: Qualitative and quantitative research in the social sciences*. Princeton, NJ: Princeton University Press.
Goertz, G., & Mahoney, J. (2013a). For methodological pluralism: A reply to Brady and Elman. *Comparative Political Studies*, *46*, 278-285.
Goertz, G., & Mahoney, J. (2013b). Methodological Rorschach tests: Contrasting interpretations in qualitative and quantitative research. *Comparative Political Studies*, *46*, 236-251.
Grandori, A., & Furnari, S. (2008). A chemistry of organization: Combinatory analysis and design. *Organization Studies*, *29*, 459-485.
Griffin, L. (1993). Narrative, event-structure analysis, and causal interpretation in historical sociology. *American Journal of Sociology*, *98*, 1094-1133.
Griffin, L., & Ragin, C. (1994). Some observations on formal methods of qualitative analysis. *Sociological Methods & Research*, *23*, 4-21.
Grofman, B., & Schneider, C. Q. (2009). An introduction to crisp set QCA, with a comparison to binary logistic regression. *Political Research Quarterly*, *62*, 662-672.
Gross, M. L. (1994). Jewish rescue in Holland and France during the Second World War: Moral cognition and collective action. *Social Forces*, *73*, 463-496.

Heikkila, T. (2004). Institutional boundaries and common-pool resource management: A comparative analysis of water management programs in California. *Journal of Policy Analysis and Management*, *23*, 97-117.

Heise, D. R. (1989). Modeling event structures. *Journal of Mathematical Sociology*, *14*, 139-169.

Hohn, F. E. (1966). *Applied Boolean algebra: An elementary introduction* (2nd ed.). New York, NY: Macmillan.

Hug, S. (2013). Qualitative Comparative Analysis: How inductive use and measurement error lead to problematic inference. *Political Analysis*, *21*, 252-265.

Kalleberg, A. L., & Vaisey, S. (2005). Pathways to a good job: Perceived work quality among the machinists in North America. *British Journal of Industrial Relations*, *43*, 431-454.

Katz, A., vom Hau, M., & Mahoney, J. (2005). Explaining the great reversal in Spanish America. *Sociological Methods & Research*, *33*, 539-573.

King, G., Keohane, R. O., & Verba, S. (1994). *Designing social inquiry: Scientific inference in qualitative research*. Princeton, NJ: Princeton University Press.

Landry, P. F., Davis, D., & Wang, S. (2010). Elections in rural China: Competition without parties. *Comparative Political Studies*, *43*, 763-790.

Lichbach, M. I. (1981). Regime change: A test of structuralist and functionalist explanations. *Comparative Political Studies*, *14*, 49-73.

Mahoney, J. (2007). Qualitative methodology and comparative politics. *Comparative Political Studies*, *40*, 122-144.

Mahoney, J. (2008). Toward a unified theory of causality. *Comparative Political Studies*, *41*, 412-436.

Mahoney, J., & Goertz, G. (2006). A tale of two cultures: Contrasting quantitative and qualitative research. *Political Analysis*, *14*, 227-249.

Marx, A., Rihoux, B., & Ragin, C. (2014). The origins, development, and application of Qualitative Comparative Analysis: The first 25 years. *European Political Science Review*, *6*, 115-142.

McCluskey, E. J. (1965). *Introduction to the theory of switching circuits*. Princeton, NJ: Princeton University Press.

North, D. C., & Weingast, B. R. (1989). Constitutions and commitment: The evolution of institutional governing public choice in seventeenth-century England. *Journal of Economic History*, *49*, 803-832.

Ragin, C. C. (1987). *The comparative method: Moving beyond qualitative and quantitative strategies*. Berkeley: University of California Press.

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

Ragin, C. C., Mayer, S. E., & Drass, K. A. (1984). Assessing discrimination: A Boolean approach. *American Sociological Review*, *49*, 221-234.

Rihoux, B., & De Meur, G. (2009). Crisp-Set Qualitative Comparative Analysis (csQCA). In B. Rihoux & C. C. Ragin (Eds.), *Configurational comparative methods: Qualitative Comparative Analysis (QCA) and related techniques* (pp. 33-68). London, England: Sage.

Schneider, C. Q., & Wagemann, C. (2010). Standards of good practice in Qualitative Comparative Analysis (QCA) and fuzzy-sets. *Comparative Sociology*, *9*, 397-418.

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

Smithson, M. (2005). Fuzzy set inclusion: Linking fuzzy set methods with mainstream techniques. *Sociological Methods & Research*, *33*, 431-461.

Smithson, M., & Verkuilen, J. (2006). *Fuzzy set theory: Applications in the social sciences*. London, England: Sage.

South, G. F. (1974). *Boolean algebra and its uses*. New York, NY: Van Nostrand Reinhold.

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

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

Thiem, A. (2014b). Parameters of fit and intermediate solutions in multi-value Qualitative Comparative Analysis. *Quality & Quantity*, *49,* 657-674.

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

Thiem, A., & Duşa, A. (2012). Introducing the QCA package: A market analysis and software review. *Qualitative & Multi-Method Research*, *10*, 45-49.

Thiem, A., & Duşa, A. (2013a). *Qualitative Comparative Analysis with R: A users guide*. New York, NY: Springer.

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

Vaisey, S. (2009). QCA 3.0: The "Ragin Revolution" continues. *Contemporary Sociology*, *38*, 308-312.

Verkuilen, J. (2005). Assigning membership in a fuzzy set analysis. *Sociological Methods & Research*, *33*, 462-496.

Vink, M. P., & Vliet, O. (2009). Not quite crisp, not yet fuzzy? Assessing the potentials and pitfalls of multi-value QCA. *Field Methods*, *21*, 265-289.

Vis, B. (2012). The comparative advantages of fsQCA and regression analysis for moderately large-n analyses. *Sociological Methods & Research*, *41*, 168-198.

Wagemann, C., & Schneider, C. Q. (2010). Qualitative Comparative Analysis (QCA) and fuzzy-sets: Agenda for a research approach and a data analysis technique. *Comparative Sociology*, *9*, 376-396.

Wooldridge, J. M. (2008). *Econometric analysis of cross section and panel data* (2nd ed.). Cambridge, MA: MIT Press.

Author Biographies

Alrik Thiem is a post-doctoral researcher at the Department of Philosophy of the University of Geneva. His work addresses topics in the field of empirical social research methods, primarily configurational comparative ones such as Coincidence Analysis, Event Structure Analysis, and Qualitative Comparative Analysis, on which he has published widely.

Michael Baumgartner is a Swiss National Science Foundation professor at the Department of Philosophy of the University of Geneva. His research focuses on questions in the philosophy of science and logic. He has published numerous works on aspects of causation and causal reasoning, regularity theories, interventionism, determinism and logical formalization.

Damien Bol is a post-doctoral researcher at the Canadian Research Chair in Electoral Studies of the University of Montreal, where he coordinates the international project "Making Electoral Democracy Work". His research focuses on comparative political institutions, political behavior and political methodology.