

**INTRODUÇÃO À ANÁLISE
QUALITATIVA COMPARATIVA
E AOS CONJUNTOS
FUZZY (FSQCA)**

Admir Antonio
Betarelli Junior e
Sandro de Freitas
Ferreira

COLEÇÃO

Metodologias
de Pesquisa

**Introdução à análise qualitativa
comparativa e aos conjuntos
Fuzzy (fsQCA)**

Enap Escola Nacional de Administração Pública

Presidente

Francisco Gaetani

Diretora de Formação Profissional e Especialização

Iara Cristina da Silva Alves

Diretor de Educação Continuada

Paulo Marques

Diretor de Inovação e Gestão do Conhecimento

Guilherme Alberto Almeida de Almeida

Diretor de Pesquisa e Pós-Graduação Stricto Senso

Fernando de Barros Filgueiras

Diretora de Gestão Interna

Camile Sahb Mesquita

Editor: Flávio Cireno Fernandes (Enap). *Revisão:* Luiz Augusto Barros de Matos e Renata Fernandes Mourão. *Projeto gráfico e editoração eletrônica:* Vinicius Aragão Loureiro. *Revisão gráfica:* Ana Carla Gualberto Cardoso.

Introdução à análise qualitativa comparativa e aos conjuntos *Fuzzy* (fsQCA)

*Admir Antonio Betarelli Junior
Sandro de Freitas Ferreira*

Brasília
Enap
2018

Ficha Catalográfica por: Daiane da Silva Yung Valadares – CRB1/2802

B5622i Betarelli Junior, Admir Antonio.

Introdução à análise qualitativa comparativa e aos conjuntos Fuzzy (fsQCA) / Admir Antonio Betarelli Junior, Sandro de Freitas Ferreira. -- Brasília: Enap, 2018.
128 p. : il. –

Inclui bibliografia.

ISBN: 978-85-256-0088-2

1. Método de Pesquisa. 2. Teoria dos Conjuntos - Análise qualitativa comparativa. 3. Estatística. I. Título.

CDU 001.8

Catalogado na fonte pela Biblioteca Graciliano Ramos da Enap

As opiniões emitidas nesta publicação são de exclusiva e inteira responsabilidade do(s) autor(es), não exprimindo, necessariamente, o ponto de vista da Escola Nacional de Administração Pública (Enap). É permitida a reprodução deste texto e dos dados nele contidos, desde que citada a fonte. Reproduções para fins comerciais são proibidas.

Enap Fundação Escola Nacional de Administração Pública

SAIS – Área 2-A

70610-900 – Brasília, DF

Telefones: (61) 2020 3096 / 2020 3102 – Fax: (61) 2020 3178

Sítio: www.enap.gov.br

Tiragem: 500 exemplares

SUMÁRIO

Introdução.....	7
Afinal, o que a QCA trata?	15
Vantagens e diferenças para o uso da QCA	22
QCA como potencial análise de política	28
Um pouco sobre o estado da arte com QCA	32
Noções e relações de conjuntos.....	37
Noções de conjunto e operações lógicas	37
Relações entre conjuntos: suficiência e necessidade.....	45
Medidas de consistência e cobertura.....	52
Consistência e cobertura das condições suficientes	54
Consistência e cobertura das condições necessárias	66
Exemplo numérico para consistência e cobertura	72
Trade off da consistência e cobertura de condições suficientes e necessárias	75
Tabelas verdade	79
Tabela verdade em csQCA.....	81
Tabela verdade em fsQCA	83
Tabelas verdade com “ruídos”	90
Diversidade limitada e remanescentes lógicos	92
Linhas contraditórias da tabela verdade	101

Exemplo de aplicações.....	105
Aplicação ilustrativa	105
Estudo selecionado	111
QCA na estruturação de dados.....	113
QCA para extrair inferências.....	115
Aplicação da csQCA aos dados do GSC-REDD	115

INTRODUÇÃO

Em parceria com o programa de pós-graduação da Escola Nacional de Administração Pública (Enap), este livro oferece um guia didático e prático para a aplicação de uma ferramenta metodológica de natureza configuracional que diretamente se associa com a Teoria de Conjuntos. Ou melhor, ao longo deste livro trataremos da análise qualitativa comparativa, comumente conhecida pelo acrônimo QCA¹, do inglês *qualitative comparative analysis*. A QCA foi inicialmente introduzida pelo cientista social americano Charles Ragin (1987) e é uma abordagem analítica para dados multivariados. Essa abordagem configuracional constitui uma família de técnicas – *csQCA*, *mvQCA*, *fsQCA*² e *outras* –, sendo a utilização de cada variante dependente do tipo de dados e do problema de pesquisa em mãos. Há um contínuo aprimoramento da QCA e cada vez mais ela está sendo difundida internacionalmente entre pesquisadores de diversos campos empíricos (LONGEST; VAISEY, 2008; RAGIN, 2000; SCHNEIDER; WAGEMANN, 2012).

Embora tenha surgido e sido aceita no âmbito acadêmico norteamericano, a QCA se tornou mais proeminente entre os estudiosos europeus, que mais ativamente ensinam, desenvolvem e aplicam essa abordagem. Wagemann e Schneider (2010) sugerem que um dos motivos para a recepção diferenciada da QCA pode estar no fato de que o âmbito acadêmico europeu, sendo mais fragmentado e caracterizado por clivagens metodológicas menos fortes que a sua contraparte americana, permitiu mais espaço à QCA como uma terceira via metodológica entre os métodos

¹ Em português, AQC. Doravante utilizamos a sigla QCA.

² As técnicas de QCA distinguem-se entre si, dependendo dos tipos de conjuntos para os quais a lógica *booleana* é empregada. Para conjuntos em que os casos são diferenciados apenas qualitativamente (pertencimento ou não pertencimento), aplica-se a *crisp-set QCA* (*csQCA*) ou a técnica *multi-value QCA* (*mvQCA*); quando, além da diferença qualitativa, há a gradação do pertencimento dos casos nos subconjuntos qualitativos, emprega-se a *fuzzy-set QCA* (*fsQCA*).

qualitativos e quantitativos. Algo almejado por Ragin (1987)³, que referiu a QCA como “indo além das estratégias qualitativas e quantitativas”.

Queremos salientar aos leitores e pesquisadores interessados em ingressar no estudo da QCA, que essa é uma abordagem relativamente nova, apresenta terminologia própria e a forma de pensar ou interpretar os resultados alcançados é distinta, especialmente quando comparada à dos métodos estatísticos mais tradicionais, sejam eles de análise exploratória ou inferencial. Não obstante suas vantagens, como em qualquer ferramenta metodológica, existem recomendações e cuidados quanto à sua aplicação. Ao contrário da maioria das técnicas quantitativas, a QCA independe de um conjunto de pressupostos, tais como: causalidade permanente; uniformidade de efeitos causais; e unidade homogênea (BERG-SCHLOSSER *et al.*, 2008, p. 9). Contudo, de acordo com Hudson e Kühner (2013), ela não deve ser tratada como substituta das abordagens dominantes⁴. Ao se apresentarem algumas das vantagens da QCA em relação às outras, isso não significa que as abordagens dominantes sejam imperfeitas ou que a QCA seja superior, mesmo em um contexto específico (HUDSON; KÜHNER, 2013).

Schneider e Wagemann (2010) enumeraram as boas práticas e recomendações para a aplicação da QCA. Embora o leitor desconheça alguns termos próprios da abordagem nessas recomendações, preferimos destacá-las antecipadamente. De acordo com os autores, a QCA deve ser aplicada: (a) para propósitos originais de pesquisa; (b) em consonância com outras técnicas de análise de dados; (c) com justificativa explícita e detalhada da seleção dos casos; (d) com um número moderado de condições; (e) com condições e resultados estabelecidos por teoria ou pesquisa empírica; (f) com descrição do processo de calibragem dos escores de pertencimento aos conjuntos; (g) com a utilização apropriada

³ Esse livro foi atualizado por Ragin (2014).

⁴ Hudson e Kühner (2013) discutem algumas vantagens da QCA em relação à análise de *cluster*, análise de componentes principais (ACP), análise fatorial (AF) e método tradicional de regressão. Os autores pontuam que o tratamento de *outliers* (valores extremos ou discrepantes) é uma das principais vantagens do QCA sobre a análise de regressão.

das terminologias; (h) com análises em etapas separadas das condições necessárias e suficientes, de modo que ambas apresentem níveis apropriados; (i) com tabelas verdade minimizadas.

Outras recomendações serão explicitadas no decorrer do livro, especialmente quando discutirmos alguns exemplos e conforme avançarmos com os conceitos e instrumentos da QCA, seguindo Ragin (2008). Além disso, julgamos que o objetivo principal deste livro é somente introduzir o leitor no uso da QCA. Em razão disso, e uma vez que a QCA está ainda em desenvolvimento, é bem provável que o leitor terá que recorrer aos documentos suplementares ou aos mais recentes que a publicação deste livro. Todavia, qualquer pesquisador, interessado em uma nova abordagem analítica, necessita ser introduzido nos fundamentos principais da ferramenta. Esse é o pressuposto de trabalho subjacente ao conteúdo deste livro. Por outro lado, informamos também que essa abordagem configuracional já está bem documentada no exterior. Além dos principais livros escritos por Ragin (1987, 2000, 2008), junto com outras publicações, *e.g.*, Rihoux e Grimm (2006), Rihoux e Ragin (2009) e Schneider e Wagemann (2012), existe uma ampla gama de tópicos de pesquisas em diferentes disciplinas das Ciências Sociais aplicadas⁵ (veja www.compass.org para uma visão geral), tais como política, religião, democracia, tecnologia, crime, educação, regulação e estrutura de mercado, e economia regional e urbana. A QCA tornou-se, pois, uma plataforma para diversos estudos qualitativos e aplicados.

Notadamente, por emergir das Ciências Sociais, essa abordagem é capaz de tratar de casos ou fenômenos derivados de conceitos abstratos (SCHNEIDER; WAGEMANN, 2010). Entretanto, a sua expansão ou ramificação em outras áreas aplicadas de pesquisa parece ter sido impulsionada pelo seu aperfeiçoamento na variante de conjuntos *fuzzy* (*fuzzy set - fsQCA*). Ou melhor, a QCA passou a tratar os casos não somente em conjuntos dicotômicos, mas também pelo seu grau de pertencimento nos conjuntos,

⁵ Thiem (2015) fez um levantamento dos periódicos mais amigáveis para a publicação de estudos aplicados com o uso da QCA, que pode ser acessado em: <http://alrik-thiem.net/blog/ranking-most-qca-friendly-journals-some-observations/> .

permitindo considerar os aspectos qualitativos e quantitativos (nos aspectos de gradação) dos dados. É possível observar essa trajetória de inovação metodológica no trabalho de Rihoux *et al.* (2013), que mapearam 313 artigos publicados da Ciência Política e Sociologia entre o período de 1984 e 2011.

Apesar disso, muitos pesquisadores e formuladores de política desconhecem essa abordagem, especialmente para aqueles que buscam compreender fenômenos complexos⁶ ao substituir as abordagens tradicionais de estatística na análise e avaliação empírica de políticas. A utilização da QCA no Brasil é ainda incipiente, mas promissora, com alguns artigos escritos e teses desenvolvidas recentemente nas áreas de Ciência Política, Social e Econômica, *e.g.*, Gurgel (2011), Sandes-Freitas (2015), Melo (2015), Sandes-Freitas e Bizzarro-Neto (2015), Xavier (2016), Montenegro (2016), Ferreira (2017). A nossa percepção é que persiste uma grande lacuna quanto à utilização dessa ferramenta metodológica nos estudos das Ciências Sociais aplicadas no país, tendo em vista a sua relevância internacional, sua potencialidade em diversos ramos de pesquisa e os desenvolvimentos recentes com distribuições estatísticas e procedimentos mais formalizados (SANDES-FREITAS; BIZZARRO-NETO, 2015; SCHNEIDER; WAGEMANN, 2012; SMITHSON; VERKUILEN, 2006).

Uma vez que essa abordagem tem por base a Teoria de Conjuntos, uma forma analítica bem diferente da dos métodos dominantes (especialmente fundamentada na álgebra matricial e cálculo), alguns leitores com maior familiaridade em pesquisas qualitativas, sem perder de vista as contribuições quantitativas, devem ter mais facilidade no entendimento dos fundamentos das técnicas configuracionais, identificando as vantagens e potencialidades analíticas dessas ferramentas. Por ser incipiente e ainda bem desconhecida no Brasil, a publicação de trabalho com aplicações da QCA enfrenta certas restrições, embora alguma difusão das técnicas já esteja sendo veiculada mediante artigos científicos apresentados em congressos nacionais, *e.g.*, Associação

⁶ Conceituaremos em breve.

Nacional de Pós-Graduação e Pesquisa em Administração (Anpad), Encontro Nacional da Associação Brasileira de Estudos Regionais e Urbanos (Enaber), e Encontro Nacional de Economia Industrial e Inovação (Enei). A exceção dessa assertiva parece residir especialmente nas áreas das Ciências Sociais e Políticas. Há periódicos internacionais que recorrentemente publicam estudos aplicados com uso da QCA, conforme destacados por Thiem (2015).

Existem quatro programas computacionais específicos capacitados para operacionalizar diferentes variantes da QCA sobre dados multivariados, tais como: *fsQCA 3.0* (RAGIN; DRASS; DAVEY, 2017); *Tosmana 1.3.2* (CRONQVIST, 2017); *Stata*, com do-file *fuzzy* (LONGEST; VAISEY, 2008); e *R*, sendo QCA 1.0.3 (THIEM; DUSA, 2012a, 2012b) e QCA3 0.0.5 (HUANG, 2014)⁷. O programa *fsQCA 3.0* apresenta uma plataforma amigável para csQCA e fsQCA. Já o *Tosmana*, por exemplo, é utilizado geralmente para atender o processo de calibragem das âncoras qualitativas ou operacionalizar a csQCA e a mvQCA; a sua atualização permite também trabalhar com fsQCA. Por seu turno, o pacote em *R* operacionaliza csQCA e fsQCA, do mesmo modo que o comando do *Stata*. Ademais, à exceção do *Stata*, todos os demais programas são livres e gratuitos. Os manuais e informações sobre os comandos dos correspondentes programas computacionais estão devidamente publicados e podem também ser obtidos no sítio eletrônico: www.compass.org.

Apesar de existir certa diversidade de programas computacionais para QCA, que se assemelham com as técnicas orientadas às variáveis para a análise quantitativa de dados, como, por exemplo, análise de regressão em suas diversas formas, a QCA não pode ser reduzida apenas a uma técnica de análise de dados. Mais que priorizar o momento analítico em QCA, é importante interpretar qualitativamente os casos em estudo. Caso se limite somente ao momento analítico, o pesquisador se arrisca a entender superficialmente a natureza da QCA e deixar de empregar

⁷ Para uma lista completa de programas computacionais, consulte: <http://www.compass.org/software.htm>.

o seu potencial analítico. Um risco particular seria realizar uma análise puramente quantitativa ao avaliar a pesquisa baseada em QCA. Ao entender a análise sob um aspecto mais qualitativo, o pesquisador poderá perceber a necessidade de considerar diferentes padrões de avaliação (WAGEMANN; SCHNEIDER, 2010), não somente a presença do fenômeno da pesquisa, mas também a sua ausência.

Portanto, o apelo é tratar a QCA como uma abordagem de pesquisa, não somente como uma técnica analítica momentânea⁸ – *i.e.*, encontrar um padrão empírico nos dados (RAGIN, 2000; RIHOUX; RAGIN, 2009; WAGEMANN; SCHNEIDER, 2010). O entendimento da QCA como uma abordagem de pesquisa refere-se principalmente aos processos interativos de coleta de dados, como parte do processo de movimento “entre ideias e evidências” (RAGIN, 2000), à especificação do modelo, a uma visão holística dos casos, à seleção de casos e constante conceituação de condições e resultado, bem como a uma visão específica a respeito da natureza da causalidade⁹ (WAGEMANN; SCHNEIDER, 2010). Ou melhor, a abordagem de pesquisa corresponde aos processos antes e depois da análise dos dados e não somente é baseada em um algoritmo padrão de um programa apropriado. Geralmente o pesquisador despenderá maior parte do tempo e recursos se ele pretende gerar uma boa pesquisa baseada em QCA. A QCA como técnica de análise de dados é consideravelmente menos *time-consuming*. As aceções dos pesquisadores e o conhecimento que adquiriram antes e depois do momento analítico são importantes para tornar os resultados da análise mais robustos e plausíveis em QCA (SCHNEIDER; WAGEMANN, 2012). Todavia, em razão da característica didática deste livro, o nosso enfoque residirá essencialmente nos fundamentos técnicos da QCA.

⁸ A QCA como uma técnica analítica refere-se ao “momento analítico” (RAGIN, 2000), que denota quando os casos já foram selecionados e todas as condições bem como o resultado tenham sido calibrados.

⁹ A QCA permite que um pesquisador trate a complexidade causal, entendida como causalidade equifinal, conjuntural e assimétrica (WAGEMANN; SCHNEIDER, 2010).

Não estamos ignorando a QCA como abordagem de pesquisa, apenas a redirecionando para uma abordagem técnica e didática.

O uso adequado das terminologias e dos instrumentos próprios em QCA é importante porque permite ao pesquisador compreender melhor os resultados gerados e evitar confusão com as técnicas quantitativas (SCHNEIDER; WAGEMANN, 2010). Inicialmente apresentaremos uma visão geral da QCA, como pode ser aplicada para políticas públicas, suas vantagens e diferenças em relação aos métodos estatísticos mais tradicionais. Esse conteúdo é foco do segundo capítulo. Em seguida, o terceiro capítulo se dedica a discutir conceitualmente as noções e relações básicas de conjuntos. A QCA é baseada na Teoria de Conjuntos e utiliza as operações booleanas para estabelecer as relações lógicas entre os conjuntos. Não nos aprofundaremos na formalidade matemática envolvida nas relações de conjuntos, mas abordaremos os principais operadores lógicos (e, ou, não), a noção básica de subconjuntos e superconjuntos, pertencimento dos casos e a distinção entre os dois tipos de conjuntos. Ainda no capítulo 3, desenvolveremos a relação de necessidade e suficiência e introduziremos as medidas de consistência e cobertura. Mostraremos que a medida de consistência denota somente uma característica da relação de subconjunto e, por isso, introduziremos a medida de cobertura, que fornece uma expressão numérica para a importância empírica de uma determinada condição (ou uma combinação delas) para se relacionar ou produzir um resultado. Essas duas medidas também são muito úteis na análise das condições de necessidade e suficiência. Por seu turno, o capítulo 4 tem o propósito de discutir a tabela verdade, uma ferramenta central para pesquisa baseada em QCA, os remanescentes lógicos – que não têm instâncias empíricas –, bem como o processo de minimização e os diferentes termos de soluções oferecidos pela QCA. Conforme Legewie (2013), entender os conceitos e ferramentas é um pré-requisito para o uso da QCA, uma vez que os mesmos ajudam a compreender o que está acontecendo durante a análise e fornecem a base para a interpretação dos resultados. Por fim, no quinto capítulo realizaremos uma aplicação ilustrativa com a *fuzzy set – fsQCA*.

AFINAL, O QUE A QCA TRATA?

A QCA é fundamentada na Teoria de Conjuntos e em operações da álgebra booleana. Métodos baseados em Teoria de Conjuntos se enquadram nos chamados “métodos booleanos” (CARAMANI, 2009) ou “métodos lógicos” (MILL, 2002)¹⁰ e são frequentemente encontrados em disciplinas como a Filosofia e a Matemática, porém atuam marginalmente em áreas das Ciências Sociais aplicadas, comumente dominadas por métodos estatísticos, aplicações de Teoria de Probabilidades e álgebra matricial. Por essa razão, a Teoria de Conjuntos é menos familiar entre muitos pesquisadores.

Caracteristicamente, os métodos baseados na Teoria de Conjuntos permitem avaliar o pertencimento dos casos de uma estrutura de dados em determinados conjuntos, identificar as relações presentes entre certos conjuntos (que descrevem fenômenos teoricamente definidos), bem como interpretar essas relações em termos de propriedades de suficiência e de necessidade (SCHNEIDER; WAGEMANN, 2012). Por exemplo, se estamos tratando de uma estrutura de dados em que os casos são os países da América, então, podemos afirmar que o Brasil pertence ao conjunto de países do Mercosul (Mercado Comum do Sul), ao passo que os Estados Unidos, não.

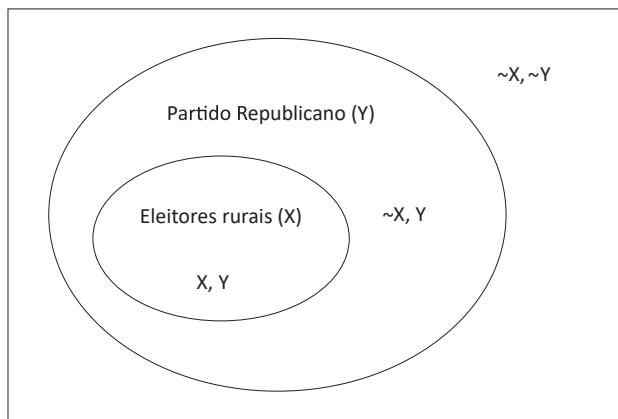
A partir do pertencimento dos casos, a QCA identifica como certos conjuntos, combinados ou não, estão presentes ou ausentes quando um fenômeno de interesse existe ou não, também tratado como um conjunto. Por meio de operações de álgebra booleana, como interseção, união e negação de conjuntos, estabelecem-se diretamente as relações entre os conjuntos. Se as relações são formadas, então podemos ter subconjuntos e superconjuntos. Em QCA a relação é conceituada como relação entre conjuntos e não deve ser confundida com o conceito de

¹⁰ A abordagem booleana pode ser encontrada nos trabalhos de John Stuart Mill, o qual se esforçou em sistematizar as investigações comparativas com o método de semelhanças e diferenças. A utilização da minimização booleana permite reduzir uma longa e complexa expressão para uma menor e mais parcimoniosa (por isso é chamada de minimização booleana) (RAGIN, 1987).

correlação, comum em métodos estatísticos mais tradicionais (RAGIN, 2000; SCHNEIDER; WAGEMANN, 2010).

A título de ilustração, conforme o exemplo de Schneider e Wagemann (2012), suponha que todos os cidadãos de pequenas cidades rurais nos Estados Unidos votaram no Partido Republicano. Essa relação denota uma relação de subconjunto, ou seja, o conjunto de todos os eleitores que vivem em pequenas cidades rurais (X) é um subconjunto de todos os eleitores republicanos (Y). Esse exemplo indica que todos os casos contidos em X , eleitores de pequenas cidades rurais, também exibem a propriedade Y , votam no Partido Republicano. A Figura 1 ilustra esses conjuntos (X, Y). Poderíamos entender que o conjunto X é suficiente para Y , porque X é subconjunto de Y . Todavia, nem todos os eleitores republicanos estão contidos em X (área $\sim X, Y$), ou seja, nada se pode dizer a respeito do comportamento dos eleitores que não residem em pequenas cidades rurais nos Estados Unidos. Eles podem ser eleitores do Partido Republicano (área $\sim X, Y$) ou não (área $\sim X, \sim Y$). A questão é, no entanto, que tais eleitores são irrelevantes quando se quer corroborar a alegação de que viver em uma cidade rural pequena é suficiente para votar no Partido Republicano. O fato de que existem outros tipos de eleitores no Partido Republicano indica simplesmente que existem outras condições suficientes para votar nos republicanos.

Figura 1 – Diagrama de venn para a relação de suficiência



Fonte: Adaptado de Schneider e Wagemann (2012).

Esse exemplo simples sugerido por Schneider e Wagemann (2012) reproduz a forma de interpretar as relações na Teoria de Conjuntos. Se existe X, então existe Y, mas a recíproca não necessariamente é verdadeira, uma vez que Y é superconjunto. Definimos o conjunto Y como o nosso fenômeno ou resultado de interesse, enquanto X, uma condição. A QCA extrapola a análise ao levar em conta um número maior de condições para Y. É provável que a existência (ausência) da combinação ou não de certas condições esteja relacionada com a existência (ausência) do nosso resultado de interesse. Como consequência, podemos ter diversas combinações lógicas relacionadas suficientemente com Y. Se acrescentarmos outra condição, Z (alta renda), no exemplo da Figura 1, em certas cidades rurais somente os moradores de alta renda poderão se relacionar com Y, enquanto em outras a condição Z não é suficiente para Y, ou seja, independentemente da renda, todos os moradores deste grupo de cidades rurais ainda votariam no Partido Republicano. Dessa maneira, pode haver combinações de vários conjuntos, em que condições únicas não exibem isoladamente relação com o resultado (como somente X para Y), mas somente em conjunto com outras condições (e.g., XZ para Y).

A QCA fornece um conjunto de combinações distintas de condições (**multicausalidade conjuntural**) que podem levar ao mesmo fenômeno (**equifinalidade**), evidenciando também combinações não relacionadas ao fenômeno (Y). Ou melhor, a **equifinalidade** é um cenário em que fatores alternativos podem produzir um único resultado, ao passo que a **multicausalidade conjuntural** refere-se às combinações de vários conjuntos ou condições únicas que podem exibir efeitos sobre o resultado de interesse da pesquisa. Juntas, a **multicausalidade conjuntural** e a **equifinalidade** conferem ao fenômeno uma **causalidade complexa**, justamente por sugerir diferentes caminhos teóricos ao fenômeno (RAGIN, 2000). Estritamente entenderemos que a **equifinalidade** e a **causação conjuntural** implicam a existência e a relevância causal das condições INUS e SUIN¹¹. Conforme

¹¹ A condição INUS significa “parte insuficiente, mas necessária de uma condição que, por sua vez, é desnecessária, mas suficiente para o resultado” (MACKIE, 1965). A condição SUIN significa “parte suficiente, mas desnecessária de um fator que é insuficiente, mas necessário para um resultado” (MAHONEY; KIMBALL; KOIVU, 2009).

Ragin (1987), o termo conjuntural denota que as alegações causais são consideradas como momentâneas ou não permanentes.

A noção de **multicausalidade conjuntural** é central em QCA, pois (a) muitas vezes uma combinação de condições, ao invés de uma única condição, leva a um resultado (fenômeno); (b) muitas vezes diferentes combinações de condições existem e levam a um resultado; e (c) os resultados são sempre contextuais, no sentido de que podem resultar da presença, mas também da ausência de uma condição particular e dependendo da sua combinação com as demais condições (RIHOUX; RAGIN, 2009). Não obstante, como decorrência da Teoria de Conjuntos, a causalidade complexa é também assimétrica, ou seja, a multicausalidade conjuntural que se relaciona com a ausência de um resultado não é a recíproca da multicausalidade conjuntural associada com a presença do resultado. Dito em outras palavras, a interpretação causal da assimetria é que a explicação para a não ocorrência do resultado não pode ser automaticamente derivada da explicação para a ocorrência do resultado. Por exemplo, uma autocracia não é simplesmente o oposto de uma democracia; riqueza não é simplesmente o oposto de pobreza. É preciso identificar separadamente as combinações de condições que se relacionam com a existência e com a ausência de um fenômeno (resultado) (SCHNEIDER; WAGEMANN, 2012).

As combinações das condições podem ser específicas e distintas de um caso para o outro. Dessa maneira, é possível compará-las e avaliá-las qualitativamente, identificando padrões similares e distintos entre os casos para um determinado resultado da pesquisa. Em geral, a QCA é mais propícia à pesquisa que explora casos heterogêneos e pode ser adequada para representar as causas cada vez mais complexas e diversas de fenômenos sociais e políticos retratados como consequências (BYRNE, 2012; EBBINGHAUS, 2012). Portanto, a QCA é adequada para analisar causalidade complexa em termos de condições (e combinações de condições) necessárias e suficientes; para explorar os padrões complexos de relacionamento entre conjuntos que descrevem a contrapartida empírica de proposições teoricamente definidas; e dar suporte ao

pesquisador na interpretação de padrões, potencialmente causais, mostrados pelos casos em investigação. Consiste de análises orientadas para a compreensão de casos, percebidos de maneira holística como configurações de características analiticamente relevantes (WAGEMANN; SCHNEIDER, 2010). Ou melhor, a técnica favorece uma detalhada análise de dentro dos casos com as comparações formalizadas entre os casos de análise (LEGEWIE, 2013) – fatores que explicam o porquê para certo resultado podem estar ausentes ou serem diferentes entre os casos (SCHNEIDER; WAGEMANN, 2012).

A QCA aborda as hipóteses teóricas que preveem como múltiplos fatores irão operar em conjunto em níveis específicos para produzir um resultado. A natureza qualitativa permanece nessa abordagem, enquanto a forma comparativa entre os casos por meio dos múltiplos caminhos, que se relacionam com um fenômeno de pesquisa, permite ao pesquisador reavaliar as teorias e oferecer algum alcance e relevância acadêmica (RAGIN, 1987; SANDES-FREITAS; BIZZARRO-NETO, 2015). Por envolver uma forma qualitativa e comparativa – análise empírica –, a QCA representa, pois, uma abordagem híbrida entre pesquisas de cunho qualitativo e quantitativo (SCHNEIDER; WAGEMANN, 2012).

A causalidade na QCA não é empiricamente definida como uma soma ponderada de influências determinísticas sobre a variável de resultado, mas como, pela própria natureza da abordagem, uma relação entre conjuntos, validada parcimoniosamente em termos de suficiência e de necessidade¹², procurando semelhanças e diferenças entre os casos que compartilham o mesmo resultado. As relações causais são modeladas como relações de subconjuntos e superconjuntos (SCHNEIDER; WAGEMANN, 2012). Destarte, os resultados obtidos pela QCA não provam estritamente as relações causais. Em vez disso, os mesmos revelam padrões de relações entre os conjuntos, proporcionando assim um apoio para a existência de tais relações causais (SCHNEIDER; WAGEMANN, 2010).

¹² Condições suficientes e necessárias, identificadas pela QCA, tratam-se dos respectivos padrões de associação que fazem sentido teórico e empírico (LEGEWIE, 2013).

A QCA também não funciona como um processo de *push-button*, mas depende dos esforços dos usuários para refletir sobre se os padrões identificados poderiam descrever um vínculo de causalidade (LEGEWIE, 2013; RIHOUX, 2009). A interpretação como relação de causalidade ainda deve ser cuidadosa, pois, em QCA, por vezes, certas relações de conjuntos podem ser ontológicas, ou seja, dois eventos estão ligados porque um constitui o que o outro é, em vez de causá-lo (GOERTZ; MAHONEY, 2005); ou o vínculo pode exibir causalidade falsa, quando dois eventos estão associados simplesmente porque ambos são efeitos de uma terceira condição não observada (BRADY, 2008). Como bem destaca Schneider e Wagemann (2012, p. 12), “se houver boas razões para acreditar que o fenômeno de interesse é mais bem entendido em termos de relações entre conjuntos, então isso representa um forte argumento para o uso de QCA”. Segundo os autores, “não é o número de casos que indica a adequação ou não da aplicação da QCA; seu uso é apropriado mesmo quando há um grande número de casos ou quando os pesquisadores estão interessados em relações entre conjuntos em vez de correlações”.

O argumento empírico deve ser subordinado ao argumento teórico (SCHNEIDER; WAGEMANN, 2012). A título de ilustração, consultando-se as literaturas teórica e empírica sobre as causas dos crimes violentos (fatais, que incluem suicídio, homicídio, latrocínio; e não fatais, estupro, roubo) e não violentos (furto, falsificação), é possível identificar, como condições potenciais, fenômenos de natureza individual (nível de escolaridade, cor, histórico criminal), econômica (níveis de renda domiciliar *per capita*, proporção de pobres, desigualdade de renda, taxa de desemprego), social (*background* familiar, interações sociais), demográfica (urbanização, proporção da população masculina jovem) e de outras dimensões. Destarte, é possível encontrar as seguintes configurações: a presença simultânea de baixo desenvolvimento econômico e de baixa escolaridade como condições suficientes para alta taxa de homicídio; alternativamente, não excludente à anterior, poderia

estabelecer que alta urbanização é condição suficiente para alta taxa de homicídio (FERREIRA, 2017).

Há duas principais variantes de conjunto para operacionalizar em QCA: os conjuntos *crisp* (*crisp set* - *csQCA*) e conjuntos *fuzzy* (*fuzzy set* - *fsQCA*). O *csQCA* corresponde a um conjunto binário convencional com apenas duas categorias de informações (0 ou 1) (RAGIN, 2006). Intuitivamente a Teoria de Conjuntos nessa variante (*csQCA*) traz consigo uma noção dicotômica fundamental: pertencer ou não pertencer (RAGIN, 1987). Por outro lado, o *fsQCA*, *fuzzy set* QCA, fornece meios adequados de acomodar complementariedades complexas e relacionamentos não lineares entre as proposições (condições causais) (GANTER; HECKER, 2014). Ou melhor, em *fsQCA*, os casos possuem graus de pertencimento nos conjuntos, variando entre 0 e 1. Portanto, a grande vantagem do conjunto *fuzzy* reside na possibilidade de escalonar diferentes escores de pertencimento e, por conseguinte, fornecer pertencimentos parciais ou completos (RAGIN, 1987). Assim, um caso não necessariamente precisa ser um membro pleno/efetivo ou um não membro pleno de um conjunto, mas pode ser um membro parcial. Contudo, isto não implica que o uso de conjuntos dicotômicos (*crisp*) deveria ser completamente evitado. Se um conceito se apresentar como uma dicotomia pura, pode ser integrado sem problemas em uma aplicação da *fsQCA* (SCHNEIDER; WAGEMANN, 2012).

A partir dessas duas principais variantes, segundo Schneider e Wagemann (2010), a QCA deve ser usada para: resumir os dados; checar se os dados são coerentes com as alegadas relações entre os conjuntos; testar hipóteses e teorias; dar uma rápida visão global sobre as suposições básicas da análise; desenvolver novos argumentos teóricos; e criar tipologias empíricas. Especialmente sobre a elaboração de tipologia, a abordagem oferece uma visualização de como os casos são classificados em várias dimensões, não somente em escala unidimensional de pertencimento. Se, por exemplo, estados de bem-estar variassem em duas dimensões – proteção do mercado de trabalho e pagamentos de transferências –, então haveria

quatro diferentes tipos ideais de estado de bem-estar: alta proteção do mercado de trabalho com altos pagamentos de transferência; alta proteção do mercado de trabalho com baixos pagamentos de transferência; baixa proteção do mercado de trabalho com altos pagamentos de transferência; e baixa proteção do mercado de trabalho com baixos pagamentos de transferência (SCHNEIDER; WAGEMANN, 2012).

Vantagens e diferenças para o uso da QCA

A QCA confere ao pesquisador a possibilidade de encontrar distintas combinações de condições que, por seu turno, geram diferentes caminhos teóricos para um determinado resultado. A diversidade e complexidade dos casos multidimensionais e as relações causais entre condições e resultados geram desafios para aqueles que almejam tirar lições para políticas públicas a partir de técnicas estatísticas de caráter exploratório, descritivo ou inferencial. Notadamente, essas técnicas tradicionais têm pontos fortes, foram testadas e testadas e, portanto, também se beneficiam de um forte grau de familiaridade entre os formuladores de políticas, profissionais e pesquisadores (HUDSON; KÜHNER, 2013).

Os modelos de regressões tradicionais limitam-se à formulação de formas funcionais específicas que denotam as influências de certas variáveis independentes sobre a dependente (fenômeno de interesse). Nessas abordagens multivariadas, os coeficientes estimados para cada variável explicativa, após os testes de significância estatística e análise descritiva dos dados, denotam os efeitos líquidos, aditivos e lineares sobre a variável-resposta. Os escores são geralmente obtidos pela soma ponderada entre os coeficientes estimados e os valores das variáveis aleatórias de cada unidade amostral. Portanto, caso o pesquisador queira obter um resultado para cada unidade amostral, o mesmo é decorrente de uma estimativa média da amostra como um todo. Por essa razão, as influências das variáveis explicativas apresentam a mesma configuração

para cada unidade amostral, justamente por utilizar os mesmos coeficientes estimados.

Haveria, portanto, apenas um caminho teórico estabelecido por uma regressão tradicional (unifinalidade e singular). Mesmo que se ampliassem caminhos alternativos na estrutura de regressão, ou seja, a separação e seleção de variáveis para explicar o mesmo resultado (variável dependente), as influências médias ainda permaneceriam e os efeitos interativos seriam limitados. Esse aspecto oculta as especificidades das unidades amostrais e somente captura as variações em nível do vetor aleatório. Por conseguinte, valores semelhantes de variáveis características podem levar à má interpretação de casos individuais quando considerados os escores médios a partir dos coeficientes estimados, limitando a tipologia entre os mesmos. Além disso, a causalidade em regressões tradicionais é simétrica, ou seja, reduções nos valores das variáveis explicativas representam reduções (aumentos) na variável dependente, caso a relação entre as mesmas seja positiva (negativa) (GURGEL, 2011; RAGIN, 2014; RIHOUX; RAGIN, 2009). A estrutura geral dos modelos estatísticos multivariados, que visam explorar relações complexas, torna-se problemática em um contexto de multicolinearidade. Na presença de dependência linear entre variáveis, a interpretação do efeito é questionada, devido à dificuldade de isolar os efeitos sobre a variabilidade da medida de resposta (MIETHE; DRASS, 1999).

De acordo com Hudson e Kühner (2013), se os formuladores de política estão interessados em conhecer especificamente caso a caso, então analisar as configurações específicas que vão além dos escores médios torna-se importante. Tal assertiva se estende para os agrupamentos por índices, cujas classificações não funcionam muito bem se comparadas com as tipologias empíricas com base no pertencimento dos casos em um conjunto *fuzzy* (EBBINGHAUS, 2005). “Todos os casos têm múltiplas dimensões, algumas das quais provavelmente serão **antitéticas** – boas classificações devem ser capazes

de refletir isso”. Essa limitada tipologia parece também se estender quando a pesquisa utiliza a análise de *cluster*. Essa técnica estatística agrupa as observações semelhantes conforme um algoritmo e métrica, sem restrição no tamanho do vetor aleatório. Por formar uma partição final com parcimonioso número de grupos, a técnica ignora diferenças específicas entre as observações. Ou melhor, por vezes, a análise de *cluster* não confere atenção ao significado conceitual ou político das variações nos dados, buscando somente padrões ou grupos similares. Na análise ideal do conjunto *fuzzy*, essas variações são fundamentais e, portanto, produzem classificações mais vinculadas aos conceitos, que são diferentes dos resultados de uma análise de *cluster* (HUDSON; KÜHNER, 2013).

Embora a análise de componentes principais (ACP) ou análise fatorial (AF) permitam elucidar melhor as especificidades das observações por suas tipologias, os escores computados também carregam o aspecto médio, como discutido sobre os modelos de regressão. São técnicas dependentes dos coeficientes de correlação e, portanto, é difícil de descobrir as variações significativas de caso para caso. O cálculo dos escores dos fatores é altamente dependente das características da distribuição das observações.

Em contrapartida, a QCA ainda é uma abordagem desconhecida ou pouco familiar para muitos estudiosos, formuladores de políticas e profissionais. Na QCA, o conceito de causalidade tem um sentido específico que não pode ser confundido com o conceito de causalidade na estrutura de regressão, pois recorre a múltiplas configurações possíveis (multicausalidade) potencialmente presentes (ou ausentes) para um mesmo fenômeno (equifinalidade). A causalidade na QCA não é empiricamente definida como uma soma ponderada de influências determinísticas sobre a variável de resultado, mas como, pela própria natureza da técnica, uma relação entre conjuntos, validada pela relação de suficiência e necessidade (RAGIN, 1987). QCA oferece uma maneira mais sistemática de analisar causalidade complexa e relações lógicas entre fatores causais e um resultado (SCHNEIDER; WAGEMANN, 2010).

Essa causalidade complexa consiste de três dimensões: equifinalidade, causalidade conjuntural e causalidade assimétrica. Como mencionado, a equifinalidade refere-se à propriedade de que várias combinações de condições conduzem a um mesmo resultado; a causalidade conjuntural significa que as condições não necessariamente conduzem ao resultado de modo isolado uma da outra, mas que podem ser combinadas entre si para revelar padrões causais de um resultado; e a causalidade assimétrica trata do fato de que não somente a ocorrência do fenômeno requer análise separada, mas também sua ausência, porque a presença/ausência das condições podem produzir diferenças no resultado. Portanto, diferentes combinações podem levar a um determinado evento ou fenômeno, e os fatores podem ter efeitos diferentes, dependendo da situação em que os mesmos operam (FERREIRA, 2017; RAGIN, 1987; SANDES-FREITAS; BIZZARRO-NETO, 2015). Se os pesquisadores estão interessados no que produz um determinado fenômeno (causalidade), ou querem saber quais os caminhos diferentes de um dado fenômeno existente (tipologia), a QCA oferece a possibilidade única de combinar análise qualitativa detalhada em profundidade com comparações sistemáticas entre os casos, identificando padrões, bem como casos que se desviam desses padrões (LEGEWIE, 2013). Dito em outras palavras, a QCA confere ao pesquisador a possibilidade de encontrar configurações específicas por casos para a existência (ausência) de um fenômeno.

Em virtude disso, o tratamento de *outliers* (discrepantes) representa, pois, uma das principais vantagens em QCA, se comparada à análise de regressão tradicional (BERG-SCHLOSSER *et al.*, 2008). Como discutido, a análise de regressão múltipla estima os efeitos médios líquidos de variáveis independentes para explicar a mudança de uma variável dependente para uma amostra ou população de casos. Em outras palavras, o objetivo da análise de regressão é prover um modelo causal único que melhor se ajuste aos dados em todos os casos ao mesmo tempo, ignorando padrões específicos e distintos” (RAGIN, 2014). Nesse ponto, a QCA avança o pressuposto de **aditividade**, que postula que cada variável explicativa incluída em um modelo de regressão, *ceteris paribus*,

tem um efeito distinto e independente sobre a variável dependente. Em vez disso, a noção de causalidade conjuntural, inerente em QCA, aponta para o fato de que caminhos separados podem levar ao mesmo resultado simultaneamente.

Ademais, em QCA, os efeitos causais não são nem uniformes nem simétricos, como é o caso na análise de regressão – não uniforme no sentido de que a mesma condição em conjunto com outras condições pode levar a resultados diferentes; e não simétricos no sentido de que diferentes explicações devem ser encontradas pela ausência e presença em um resultado (BERG-SCHLOSSER *et al.*, 2008). Além da heterogeneidade dos casos, a QCA não descarta os casos discrepantes (*outliers*) como conceitualmente incomuns e de incômodos estatísticos. Não obstante, a regressão multivariada para análise de políticas públicas tem os seus méritos e não pode ser desconsiderada, especialmente pelos seus avanços ao incluir dados longitudinais, não somente transversais, ampliando os graus de liberdade, bem como pelas crescentes discussões metodológicas sobre estimativas não tendenciosas e ineficientes. Apesar disso, esses modelos parecem enfrentar dificuldades para o tratamento de fatores independentes, como diferenças culturais e/ou institucionais a nível macro, tanto quanto estrutura social em nível micro (HUDSON; KÜHNER, 2013). É cada vez mais frequente o uso de efeitos de interação multiplicativa em análise de regressão. Ainda assim, tendo em vista a dificuldade de interpretar dois, três ou até quatro modos de termos de interação, o que teria que ser considerado juntamente com todas as suas interações de componentes (HUDSON; KÜHNER, 2013; LONGEST; VAISEY, 2008), a QCA oferece uma técnica alternativa de análise de dados levando em conta os processos complexos de forma mais eficaz (RAGIN, 2008). Por exemplo, Vis (2012) resumiu algumas abordagens de regressão mais recentes que tentaram incluir as hipóteses de assimetria e equifinalidade dos resultados (BRAUMOELLER, 2003; CLARK; GILLIGAN; GOLDBER, 2006), porém com restrições. O Quadro 1 sintetiza as principais diferenças entre QCA e as técnicas estatísticas tradicionais.

Quadro 1 – Principais diferenças entre QCA e técnicas quantitativas

Técnicas Quantitativas Tradicionais	Análise Qualitativa Comparativa
Variáveis	Conjuntos
Variável dependente	Resultado
Variáveis independentes	Condições
Correlações	Relações entre conjuntos
Matriz de correlação	Tabela Verdade
Efeitos líquidos das variáveis	Caminhos causais
Relações de aditividade e lineares	Relações não aditivas
Causalidade múltipla ou singular	Causalidade conjuntural múltipla
Universalidade ou equifinalidade	Equifinalidade
Unifinalidade	Multifinalidade
Causalidade simétrica	Causalidade assimétrica
Análise dos efeitos das variáveis	Análise dos efeitos das configurações

Fonte: Adaptado de Rihoux e Ragin (2009); Ragin (2014).

A QCA também possui terminologias próprias, que se diferenciam das técnicas estatísticas. Como exemplo, tem-se que as variáveis explicativas são retratadas aqui como **condições**, a variável dependente é denominada **resultado**, as observações são definidas como **casos** e as equações são rotuladas como **soluções**. A combinação das condições causais ou conjuntos individuais de um caso é referida como uma configuração específica. Em uma **fórmula de solução**, o resultado e as condições causalmente relevantes são representados por letras que estão relacionadas com operadores booleanos. Esses conceitos serão mais bem postos nos próximos capítulos. As técnicas QCA e suas aplicações foram desenvolvidas com o propósito de tratar uma pequena quantidade de observações (*small-N*) e de fornecer uma abordagem macro comparativa (*macro-comparative*). A QCA permite um número mínimo de 10 casos. Tamanhos pequenos encontram-se entre 10 e 15 casos. Já uma situação intermediária situa-se entre 50 e 100 casos.

QCA como potencial análise de política

Os aspectos da QCA levantados na seção anterior fornecem algumas ideias de como essa abordagem pode auxiliar a avaliação de políticas. Rihoux *et al.* (2013), por exemplo, revisa alguns estudos aplicados com QCA que se relacionam com análise de política. Segundo Rihoux e Grimm (2006), nos últimos anos, um número crescente de cientistas políticos e sociais, gestores e analistas políticos tem optado por estudos envolvendo múltiplos casos (unidades de observação) como estratégia de pesquisa; isso se deve ao fato de que necessitam juntar acepções mais aprofundadas sobre os diferentes casos, capturando a complexidade presente nos casos, a fim de produzir potencialmente algum grau de generalização. Nas análises de política, muitas das unidades de análise – do ponto de vista de acadêmicos e políticos – são naturalmente limitadas em número, por exemplo, referem-se a regiões, países, setores econômicos, entre outros. Essas populações em número limitado (N-pequeno ou N-intermediário) são, em muitas instâncias, relevantes na perspectiva das políticas públicas, especialmente no contexto transnacional ou transregional (*e.g.*, dentro de macrounidades sociais).

Por vezes, os estudos de caso costumam ser refutados pelo fato de serem mal selecionados, de utilizarem dados tendenciosos e/ou tratados como não científicos, ou como se pudessem provar praticamente qualquer coisa com eles. Todavia, é possível mostrar que, mediante o uso de métodos como a QCA, todas as etapas importantes da pesquisa de estudo de casos¹³ tornam-se mais transparentes e abertas à discussão e a críticas. Notemos que a transparência metodológica é especialmente relevante para formuladores e gestores de políticas que costumam avaliar o material desses estudos de caso.

A fim de contribuir para o desenvolvimento de uma agenda para análises de políticas aplicadas a partir do uso de uma abordagem de pesquisa como a QCA na sua variante de conjunto *fuzzy*, podemos listar algumas vantagens dessa abordagem (HUDSON; KÜHNER, 2013; RIHOUX;

¹³ Seleção de casos, planejamento do estudo de casos, seleção e operacionalização das variáveis, uso de dados e fontes, comparação de estudos de casos, generalização dos resultados empíricos encontrados, entre outros.

RAGIN, 2009). O Quadro 2 apresenta algumas das vantagens práticas dos novos métodos comparativos para a análise de políticas, conforme Rihoux e Grimm (2006). A abordagem de QCA, ao contrário das ferramentas estatísticas e econométricas tradicionais, possibilita a identificação de mais de um único caminho para um determinado resultado de política, pois mais de uma combinação de condições pode conduzir a um determinado resultado político. Esse aspecto é particularmente útil na prática política, até porque a experiência mostra que a eficácia das políticas geralmente depende do alcance nacional/regional, bem como das características específicas do contexto, como as diferentes tradições culturais, políticas e administrativas, que demandam planos de implementação específicos.

Quadro 2 – Vantagens práticas dos novos métodos comparativos

Vantagens	Descrição
Análise comparativa com número pequeno de casos	Permite comparar sistematicamente os programas de política em pesquisas com pequeno número de casos, em termos transnacional, transregional e trans-setorial (que costumam ser os domínios/jurisdições das políticas), intra e entre entidades políticas mais amplas ou grupos de países (e.g., UE, Mercosul, OECD), mas também intrapaíses (e.g., entre províncias de um país etc) ou intrarregiões (ex. entre municipalidades, regiões econômicas (e.g., Brasil, a RIDE)).
Pesquisa quase-experimental	Possibilita o envolvimento em uma abordagem de pesquisa quase-experimental; porque, por exemplo, permite que o analista/formulador/avaliador de políticas examine em quais condições (ou quais combinações de condições) uma política específica é ou não efetiva.
Transparência	É transparente, no sentido de que o analista de políticas pode modificar a operacionalização das variáveis para testes adicionais, incluir e/ou agregar "variáveis", etc.. Portanto, oferece uma análise pluralista e participativa.
Teste de modelos	Permite testar modelos causais alternativos (intervenção política) levando a uma produção favorável (ou desfavorável) de políticas e a resultados políticos favoráveis (ou desfavoráveis), distinguindo-se produtos de resultados.

Fonte: Adaptado de Rihoux e Ragin (2009); Ragin (2014).

Segundo Varone, Rihoux e Marx (2006), a QCA pode contribuir para enfrentar alguns desafios específicos da etapa da avaliação de políticas. Os autores identificaram quatro desafios: (a) a relação entre intervenções políticas e resultados da política e a identificação dos mecanismos causais que ligam as intervenções aos resultados; (b) a identificação do efeito líquido da intervenção política e a eliminação da complexidade dos fatores/condições causais; (c) a geração de evidência contrafactual (*i.e.*, tratar da questão condicional “Se...”); e (d) a evidenciação triangulada (procedimento frequentemente utilizado nas Ciências Sociais em que dois (ou mais) métodos são empregados em um estudo para verificar os resultados de uma mesma pesquisa). Além disso, a QCA oferece algumas respostas específicas a esses desafios, justamente por possibilitar uma comparação tripla: uma análise entre casos, uma análise intracasos e uma comparação entre realidade empírica e tipos-ideais teóricos. As abordagens de QCA podem ser utilizadas como métodos para estudar e analisar processos políticos. É possível utilizá-las, por exemplo, para descobrir, mediante comparações sistemáticas, sob que condições determinadas unidades macrossociais (nacionais, subnacionais e supranacionais) poderiam implementar com sucesso um determinado conjunto de compromissos políticos, e desenvolver generalizações e recomendações políticas para essas mesmas unidades e para outras que compartilham seu contexto.

A QCA permite ainda traduzir a complexidade, identificada em estudos de casos, em fórmulas comparáveis e estabelecer inferências a respeito das condições causais. Esse processo pode ser efetivamente aplicado a um processo político, assim como pode requisitar o envolvimento de especialistas e coordenadores, principalmente na etapa de estruturação e organização da informação (seleção de fatores/condições). O último capítulo deste livro ilustrará um exemplo no âmbito de políticas públicas do estudo de Korhonen-Kurki *et al.* (2014).

Hudson e Kühner (2013) sintetizam quais os aspectos contributivos que o uso da QCA na sua variante de conjuntos *fuzzy* pode possuir em termos de uma agenda para análise de políticas aplicada. O Quadro 3 reporta algumas das principais vantagens da QCA, qualificadas pelos

autores como complemento-chave para as técnicas de análise de política mais estabelecidas. O ponto central da pesquisa reside na classificação de casos por tipologias, tanto quanto na identificação de ligações causais entre combinações de condições e resultado. Esse ponto envolve temas relevantes para a análise de políticas públicas aplicadas, nomeadamente política social, política de saúde, economia política, ciência política, sociologia política e estudos de desenvolvimento. Por exemplo, Lee (2013) buscou uma configuração para níveis excepcionais de empregos na Coreia e no Japão. Com foco no salário mínimo, na proteção do emprego e na proteção do rendimento do desemprego, a QCA encontrou duas configurações de políticas públicas separadas nos dois países: um baixo salário mínimo legal combinado com a fraca proteção ao emprego para trabalhadores temporários foi suficiente para o emprego não padronizado no caso do Japão; na Coreia, o emprego não padronizado é explicado pelo baixo salário mínimo legal combinado com uma forte proteção do emprego para trabalhadores permanentes.

Quadro 3 – QCA e análise de política aplicada

QCA permite ...	Implicação para pesquisa de políticas aplicadas
... descrição dos casos	A combinação sistemática de informações qualitativas e quantitativas permite um melhor conhecimento dos casos, mesmo em especificações intermediárias e grandes N.
	Força os pesquisadores a considerarem cuidadosamente a seleção de casos, a justificativa para a calibração de conjunto fuzzy e a inclusão de dimensões / condições.
	Permite que os pesquisadores se envolvam em análises formalizadas e replicáveis de várias fontes de informação.
... identificação orientada para a diversidade de grupos dos casos	A visualização de casos como configurações força os pesquisadores a identificar dimensões-chave para descrever / explorar casos.
	A combinação de dimensões usando princípios booleanos de negação e interseção alivia os efeitos de compensação.
	Um melhor tratamento das dimensões antitéticas leva a grupos de casos conceitualmente não empíricos.

QCA permite ...	Implicação para pesquisa de políticas aplicadas
.. análise de causalidade complexa	Permite que os pesquisadores abordem a causalidade conjuntural e equifinalidade.
	Permite causalidade assimétrica e fácil interpretação dos efeitos combinados das condições.
	A exploração dos "remanescentes lógicos" permite considerar os pressupostos contrafactuais.
	A independência da amostragem aleatória e das distribuições normais permite um melhor tratamento de casos excepcionais e amostragem de casos heterogêneos.
.. análise de contexto específico	A alocação baseada em escores de pertencimento atenua o viés causado por escores discrepantes para casos individuais e melhora a robustez das conclusões para a mudança das amostras de países
	A noção de generalização modesta permite diferentes calibrações de casos e identificação de diferentes caminhos causais, dependendo de diferentes questões de pesquisa, localidades e períodos de tempo.

Fonte: Adaptado de Hudson e Kühner (2013).

Um pouco sobre o estado da arte com QCA

Desde Ragin (1987, 2014), as resenhas sobre QCA são recorrentes na literatura internacional, e um número expressivo de artigos para discussão e publicados em diversos periódicos internacionais podem ser obtidos na própria plataforma da *compass* (<http://www.compass.org/>). Nessa plataforma o pesquisador pode obter outras informações, tais como: eventos, treinamento, livros, visão geral sobre as opções de programas computacionais e sítios eletrônicos úteis. Tendo em vista que compreender as noções e aplicação da QCA leva tempo, Legewie (2013) organizou um apêndice em que subdivide os temas com as respectivas fontes bibliográficas. Esses temas consistem nos conceitos, fundamentos e instrumentos da QCA. Os principais livros são os de Ragin (1987, 2000, 2008, 2014), de Rihoux e Grimm (2006), de Rihoux e Ragin (2009) e de Schneider e Wagemann (2012). Dentre eles, para a calibragem dos conjuntos, o pesquisador pode consultar o capítulo 4 e 5 de Ragin (2008), as páginas 32 a 41 de Schneider e Wagemann (2012) ou o texto de Ragin (2007). Já sobre as propriedades de necessidade e

suficiência nas relações entre os conjuntos, temos novamente o livro de Schneider e Wagemann (2012), especialmente as páginas 56 a 90 e o texto de Cooper e Glaesser (2016). Por sua vez, o assunto sobre as medidas de consistência e cobertura a partir das relações de conjuntos é bem discutido no capítulo 3 de Ragin (2008) e nas páginas 119 a 150 de Schneider e Wagemann (2012). Além dos livros supracitados, a discussão sobre a análise da tabela verdade está bem detalhada em Emmenegger, Schraff e Walter (2014). Para o entendimento das diferenças entre QCA e análise de regressão, o leitor pode consultar os capítulos 10 e 11 de Ragin (2008). Já sobre boas práticas na aplicação de QCA, críticas e discussões, bem como robustez de resultados, podemos listar algumas referências como Schneider e Wagemann (2010), Delreux (2010), Marx (2006), páginas 284 e 295 de Schneider e Wagemann (2012), Lee (2008), Vink e Van Vliet (2009) e Rohlfing (2016).

O leitor pode encontrar uma ampla lista de trabalhos aplicados em Rihoux *et al.* (2013). Podemos listar alguns trabalhos que versam sobre desenvolvimento econômico, bem-estar, análise de política, inovação, crime, eleições ou partidos com a aplicação de QCA. Por exemplo, Ahn e Lee (2012) investigaram os determinantes do desenvolvimento do estado de bem-estar coreano com fsQCA. Seus resultados conclusivos sugerem que, pelo menos até o ano analisado, outros fatores que o desenvolvimento econômico ou industrialização não são robustos o suficiente para fazer uma afirmação de que haveria uma mudança positiva do estado de bem-estar coreano desde o seu estágio de desenvolvimento. Schneider e Makszin (2014) analisaram se o grau de desigualdade política entre grupos sociais é moldado por características do sistema capitalista de bem-estar social. Especificamente, os autores levantaram a hipótese de que mais proteção trabalhista e apoio social diminuem a desigualdade participativa por meio de recursos distribuídos de forma mais equilibrada e o engajamento entre cidadãos com alto e baixo nível educacional. Com fsQCA, o referido trabalho aplicado identificou três tipos funcionalmente equivalentes de capitalismo de bem-estar, que todos produzem a desigualdade participativa por meio de maior proteção, apoio ou ambos, tanto quanto

os mecanismos por trás dessa ligação, que são uma distribuição mais igualitária de recursos e engajamento entre cidadãos de baixa e alta escolaridade. Meuer, Rupietta, Backes-Gellner (2015) utilizaram fsQCA a fim de analisar sistemas de inovação coexistentes considerando 384 empresa suíças, sendo dois sistemas de inovação genéricos, o autárquico e a internalização do conhecimento; um sistema regional de inovação, a hierarquia protegida; e dois de inovação setorial: sistemas, ciências públicas e a aprendizagem organizada. Os resultados conclusivos dos autores sugerem que propensões de sistemas inovativos pareceram ser determinadas por vários tipos de inovação. Nessa constatação, destacam-se as interdependências que existem entre os diferentes tipos inovativos.

Já Olsen e Nomura (2009) avaliaram os resultados de redução da pobreza econômica em países selecionados entre 1992 e 2002 por meio da fsQCA e csQCA. Esse artigo tem um propósito mais metodológico, em que aponta como determinar o equilíbrio de resultados de consistência e cobertura com base em vários pontos de corte como sendo altamente importante para uma avaliação de sensibilidade. A conclusão alcançada é que, usando a redução da pobreza como variável dependente, comparativamente os dois métodos são úteis. Por outro lado, com fsQCA, Alam e Bagchi (2013) analisaram o impacto do ambiente de produção e logístico sobre a entrada do investimento direto estrangeiro (IDE), levando em conta um novo conceito da capacidade da cadeia de suprimentos (SCC) dos países. Os autores se preocuparam em avaliar se essa cadeia é uma causa necessária e suficiente para os fluxos de IDE. O estudo sugere que a atratividade do IDE de um país é mais bem explicada por uma combinação de condições, conforme refletido pela variável composta, SCC, do que por variáveis constituintes individuais. Esse estudo oferece acepções para apontar que gerentes avaliam vários ambientes de países concorrentes, permitindo assim tomar melhores decisões estratégicas sobre investimento direto estrangeiro.

No Brasil, Gurgel (2011) utilizou a fsQCA para identificar quais são as imperfeições de mercado exploradas por empresas brasileiras de capital aberto e de alto desempenho contábil e como essas imperfeições

se configuram. Com a mesma variante de QCA, Sandes-Freitas (2015) analisou a lógica de composição de alianças partidárias nos estados brasileiros, no período de 1986 a 2010, tendo como foco de análise quatro estados brasileiros: Piauí, Ceará, São Paulo e Rondônia. Melo (2015), por seu turno, buscou compreender um eventual paradoxo do Partido dos Trabalhadores (PT) e do Partido Social-Democrata Brasileiro (PSDB) nas eleições de 2014 no Brasil. Xavier (2016) procurou analisar a razão pela qual os países com condições econômicas semelhantes apresentam diferente desenvolvimento financeiro. A pesquisa conclui que a estabilidade política é a única condição necessária, porém não suficiente. Como parte de sua pesquisa, Montenegro (2016) teve o propósito de caracterizar e identificar os possíveis fatores condicionantes da capacidade ambiental dos países desenvolvidos e em desenvolvimento. A pesquisa considerou o desempenho das inovações ambientais nos anos de 1990, 2000 e 2010 por diferentes graus de sistemas nacionais de inovação. A autora mapeou as configurações específicas de quarenta economias mundiais nos períodos de análise, assim como das interações entre áreas científicas e dos campos tecnológicos a partir das patentes ambientais. Já Ferreira (2017) buscou analisar o efeito da dissuasão social como barreira primária às escolhas ilegítimas individuais. Uma vez que os níveis de transgressões podem indiretamente sinalizar o nível dessa dissuasão, o autor desenvolveu uma estratégia metodológica que pudesse mensurá-la. O autor utilizou a fsQCA para avaliar as relações entre as taxas de homicídio e fatores como transgressividade por incapacitação, transgressividade familiar-religiosa e transgressividade educativa adulta. As suas evidências sugerem que alta transgressividade familiar-religiosa implica consistentemente alta taxa de homicídio; inclusive quando combinada com alta dissuasão estatal.

Silva e Betarelli Junior (2016) avaliaram as configurações gerais e específicas das eleições de 2014 nos municípios brasileiros a partir da fsQCA, cujos resultados conclusivos confirmaram a hipótese levantada pela literatura de polarização geográfica dos votos. Ou seja, municípios com características semelhantes, como alto percentual de famílias atendidas pelo Bolsa Família, baixa escolaridade e maior desigualdade de

renda, tiveram um resultado eleitoral próximo, dando a Dilma Rousseff um alto percentual de votos. Eugenio e Betarelli Junior (2016) analisaram as possíveis combinações de condições causais da satisfação com a vida para dados do Brasil no ano de 2013, disponíveis na pesquisa de opinião pública Latinobarómetro. Com a fsQCA, os autores analisaram condições como situação econômica do país (S), renda (R), instituições/democracia (D), escolaridade (E) e emprego (O) para a satisfação individual. Os principais resultados alcançados indicam que há duas possíveis combinações de condições causais, uma delas é boa situação econômica do país, alta renda, boas condições institucionais e democráticas, alta escolaridade e alto nível de emprego, com uma consistência e uma cobertura de, respectivamente, 88,5% e 49,5%. Por fim, Faria, Betarelli Junior e Montenegro (2017) utilizaram a fsQCA para avaliar a multidimensionalidade do desmatamento na Amazônia Legal brasileira. Quatro dimensões características foram criadas: desenvolvimento típico, ambiente familiar, abertura ao comércio e propriedade e escala rural. Os resultados conclusivos dos autores apontaram que a relação entre alto desenvolvimento típico e alta propriedade e escala rural cobre 55% dos municípios analisados e é altamente consistente com o alto desmatamento na Amazônia Legal.

NOÇÕES E RELAÇÕES DE CONJUNTOS

Como mencionado, por definição, a abordagem QCA trata de averiguar a relação entre conceitos, que podem ser entendidos a partir das noções da Teoria de Conjuntos. Os conjuntos contêm informações de natureza qualitativa sobre os casos (unidades de observação ou análise), que podem ser distribuídos em grupos com características distintas e analisados em termos de suas relações entre si, mediante os princípios da álgebra booleana (RAGIN, 1987, 2000, 2014). As técnicas QCA procuram encontrar condições (conjuntos, subconjuntos ou superconjuntos) necessárias e suficientes que engendram o fenômeno da causalidade complexa a partir da Teoria de Conjuntos (SCHNEIDER; WAGEMANN, 2010). Assim, tendo em vista que a QCA emprega as relações teóricas de subconjuntos e superconjuntos e a álgebra booleana para averiguar as relações entre diferentes conjuntos (que são operacionalização de conceitos teóricos), precisamos inicialmente apresentar as noções de conjuntos, as relações entre conjuntos (relações de subconjunto, superconjunto, suficiência e de necessidade), as operações lógicas (união, interseção e negação) e os parâmetros de ajustes (consistência e cobertura). Entender esses conceitos é um pré-requisito para o uso da QCA, além de fornecerem uma base para a interpretação dos resultados (LEGEWIE, 2013).

Noções de conjunto e operações lógicas

O emprego do termo conjunto não é muito difundido nas Ciências Sociais. Não obstante, grande parte do raciocínio conceitual é baseada em uma ideia subjacente de conjuntos, pois muitos conceitos utilizados nas pesquisas científicas sociais podem ser representados por conjuntos e porque um conceito pode ser definido “como uma representação mental de uma propriedade empírica” (SCHNEIDER; WAGEMANN, 2012, p. 24). Por exemplo, o conceito país democrático engendra uma representação mental de um país com características definidoras de uma democracia

(e.g., direito de escolha do chefe do executivo, transição pacífica de governo, liberdade de afiliação partidária). Assim, inúmeros países poderiam pertencer ou não ao conjunto país democrático. A perspectiva baseada em conjuntos utiliza o conceito de pertencimento ao conjunto para estabelecer se um caso ou instância empírica (no exemplo anterior, um país real específico) pode ser descrito por um conceito ou não.

Todavia, as declarações teóricas não são tão simples como o exemplo anteriormente mencionado, país democrático. As mesmas envolvem inúmeros conceitos simultaneamente, como democracia federativa. Nem todos os países são democráticos, nem federativos; alguns são democráticos e federativos. O termo democracia federativa envolve dois conceitos: país democrático e país federativo (dois conjuntos). Expressões mais complexas poderiam surgir, tal como república democrática federativa, e assim sucessivamente. Em suma, as declarações teóricas costumam tratar de expressões conceituais complexas. Ademais, tais expressões complexas podem manter relações de causalidade para com outro conceito: países democráticos e federativos podem estar associados a alto desenvolvimento econômico. Portanto, um pesquisador poderia querer averiguar a força da relação causal entre os conceitos (conjuntos) países democráticos e federativos, e possivelmente outros conjuntos, chamados de **condições causais**, com o conceito alto desenvolvimento econômico, chamado de **resultado**.

Em termos de operacionalização, esses conjuntos são ditos dicotômicos (*crisp set*), ou seja, correspondem a um conjunto binário convencional com apenas duas categorias de informações: score 1 para indicar a presença ou pertencimento pleno dos casos a um conjunto (equivalentemente, adesão total a um conceito) e score 0 ou não pertencimento pleno a um conjunto (não adesão total ao conceito). Dessa maneira, definir um conjunto dicotômico clássico implica tomar uma decisão binária quanto à pertinência de determinado indivíduo (objeto, elemento) numa dada classe (grupo, categoria): aceitar (= 1) ou rejeitar (= 0) tal proposição. Por exemplo, se tivéssemos três países – Brasil, França e EUA –, o conjunto país federativo seria representado por {1, 0, 1}, pois o Brasil e os EUA são países federativos e a França, não.

Quando o QCA foi apresentado inicialmente nos anos 1980 e 1990, tratava-se simplesmente de conjuntos dicotômicos, que exigiam a decisão por parte do pesquisador em classificar um determinado caso como membro ou não de um conjunto. Embora isso corresponda ao modo como os conjuntos, e até mesmo os conceitos, costumam ser percebidos (como caixas em que os casos concretos podem ser “encaixados” ou não), nem sempre a classificação de casos reais é tarefa fácil e clara, especialmente quando se trata de conceitos para os quais estão disponíveis informações mais matizadas e detalhadas. Por exemplo, podem existir variados graus de desenvolvimento econômico. Destarte, os conjuntos discretos ou dicotômicos, embora informativos, limitam informação importante contida na variabilidade dos casos. A diversidade existe não somente em termos dos diferentes tipos de casos nos conjuntos retratados pelos fenômenos, mas também no grau em que cada unidade de análise pertence a tais conjuntos. Por essa razão teve-se a emergência da QCA com conjuntos *fuzzy*, proposta por Ragin (2000).

A utilização de conjuntos *fuzzy* permite capturar a segunda dimensão de diversidade, que é considerar a escala de pertencimento de cada unidade de observação (caso) nos conjuntos de interesse (fenômenos), sem ignorar a primeira dimensão de diversidade, que é a diferença de tipos. Os conjuntos *fuzzy* capturam ambos os aspectos – a distinção quantitativa e a distinção qualitativa – em um único instrumento. Além disso, os conjuntos *fuzzy* são um instrumento importante para avaliar formulações conceituais complexas, comuns nas Ciências Sociais, mediante a aplicação de álgebra booleana (RAGIN, 2000). A grande vantagem do conjunto *fuzzy* reside na possibilidade de escalonar diferentes escores de associação e, por conseguinte, fornecer associações parciais ou completas.

Mais especificamente, os conjuntos *fuzzy* estendem os conjuntos dicotômicos ao assinalar escores no conjunto $[0,1] \subset \mathfrak{R}$; conseqüentemente, possibilitam distinguir elementos com pertencimento parcial dos elementos com pertencimento integral/pleno em um conjunto. Ou melhor, ao usar conjuntos *fuzzy*, a adesão ou pertencimento de cada caso em um conjunto pode assumir qualquer valor entre 0 e 1. O grau

de pertencimento cresce conforme a magnitude do escore se aproxima da unidade (pertencimento pleno) e decresce quando se aproxima de zero (não pertencimento pleno). Próximo a um, indica alto grau de pertencimento; próximo a zero, baixo grau. Se a mediana do conjunto for tomada como o ponto de ambiguidade máxima, no sentido da dificuldade de identificar o grau de pertença do elemento com escore 0,5, então os elementos com escores menores (maiores) de 0,5 indicam que eles estão parcialmente mais fora (dentro) do que dentro (fora) do conjunto. Portanto, três pontos de ancoragem definem um conjunto: a **adesão plena** (indicada por um escore de pertencimento 1); **nenhuma adesão** (escore 0); e um ponto de corte (escore 0,5). Entre os extremos da adesão plena e nula, podem-se refinar os níveis de adesão em um conjunto, gerando quatro níveis (e.g., 0, 0,33, 0,67, e 1) em conjuntos contínuos (sendo o escores-*fuzzy* variando entre 0 e 1) (RAGIN, 2008) . A atribuição dos escores de pertencimento é denominada de método de calibragem. Para atribuir pontuações de pertencimento (adesão) sobre os casos, portanto, são especificadas âncoras qualitativas (LEGEWIE, 2013).

Ou melhor, para conjuntos *fuzzy*, a calibragem é o processo que utiliza informação empírica dos próprios casos para assinalar o seu grau de pertencimento nos seus conjuntos correspondentes. Para isso, requer-se uma definição da população relevante de casos, a definição precisa do significado de todos os conceitos (das condições e dos resultados), a decisão de onde estará o ponto de indiferença máxima entre pertencimento e não pertencimento (que é a âncora 0,5 no conjunto *fuzzy*), a definição de pertencimento pleno, de não pertencimento pleno e sobre a gradação de pertencimento entre as âncoras qualitativas. Ragin (2008) discute detalhadamente vários métodos de calibragem, suas naturezas, vantagens e desvantagens, bem como seu grau de tecnicidade.

A título de ilustração, suponha que dois países (casos) desenvolvidos poderiam pertencer a categorias distintas neste conjunto (alto e baixo desenvolvimento), bem como poderiam diferir entre si quanto ao grau em que pertencem a cada um desses subconjuntos (um determinado caso pode ter um escore maior no conjunto alto desenvolvimento

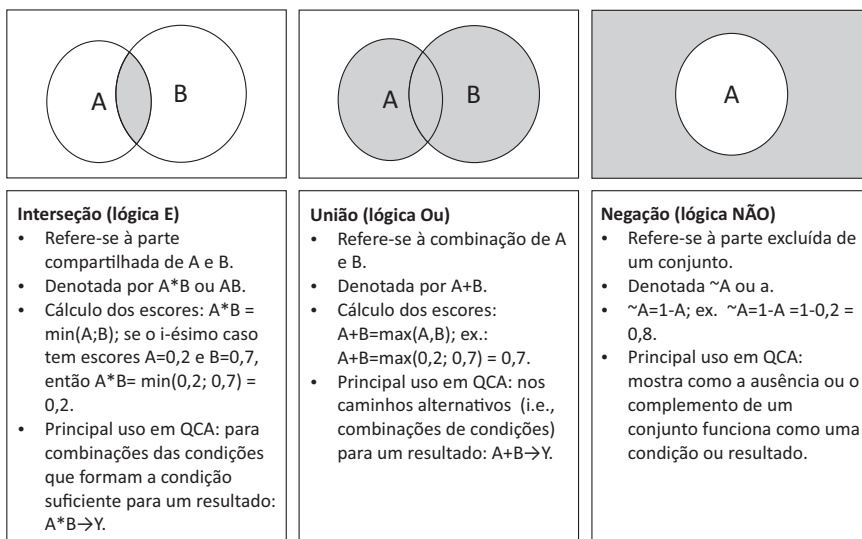
do que outro caso). Quatro países – Brasil, França, EUA e Namíbia – poderiam ser representados pelo conjunto desenvolvimento econômico = $\{0,61; 0,83; 0,94; 0,42\}$. Já a Namíbia está mais dentro do subconjunto baixo desenvolvimento, a França está mais dentro do subconjunto alto desenvolvimento. A transformação de cada variável (conjunto) em conjunto *fuzzy* é um procedimento que transforma os valores originais da variável de tal modo que de cada um deles é subtraído o valor mínimo do conjunto e o resultado é dividido pela diferença entre os valores máximo e mínimo do próprio conjunto. Esse procedimento resulta em um ranqueamento dos valores originais no intervalo $[0,1] \subset \mathfrak{R}$. Os escores *fuzzy*, portanto, assinalam o grau de pertencimento (ou probabilidade condicional) dos casos em cada conjunto da análise.

Independentemente da natureza dos conjuntos, dicotômicos ou *fuzzy*, a complexidade (relações entre conjuntos) pode ser operacionalizada por álgebra de conjuntos de modo a ser reduzida. A própria álgebra booleana foi desenvolvida como linguagem lógica simplificada, por utilizar símbolos, como x, y, z, para representar palavras, frases, ou proposições. A álgebra booleana envolve três operações básicas entre conjuntos: **intersecção** [E (*)], **união** [Ou (+)], e **negação** [Não (~)]. Os conjuntos, por operacionalizarem conceitos teóricos, podem ser rotulados com uma ou mais letras. A Figura 2 apresenta uma visão geral sobre essas operações para dois conjuntos hipotéticos, A e B, em termos de diagramas de Venn, como as mesmas são tratadas na álgebra booleana e as suas utilidades na QCA propriamente dita (as áreas cinza nas ilustrações indicam o conjunto resultante das respectivas operações).

Sabemos, por exemplo, que a declaração país democrático e federativo é uma intersecção dos conjuntos país democrático e país federativo. Se denotarmos país democrático pela letra D e país federativo por F, então a combinação dos conjuntos D e F seria representada por D e F. Esse procedimento também é chamado de **conjunção** ou de conjunção E-lógico. Combinações com E-lógico também são conhecidas na álgebra booleana e na álgebra *fuzzy* como multiplicação booleana ou *fuzzy*. Por outro lado, na Teoria de Conjuntos, uma combinação de conjuntos é chamada

de interseção. Em termos de símbolos, essa operação é representada por um asterisco (*) ou um ponto (·) ou o sinal de interseção (\cap) ou mesmo sem qualquer sinal: $D * F$, $D \cdot F$, $D \cap F$ e DF , cujas formas são equivalentes de expressar a operação interseção/conjunção de conjuntos. Convém não denotar conjuntos com a letra isolada E para não confundirmos com o operador E-lógico. Doravante utilizam-se os símbolos* ou nenhum sinal para denotar a operação de interseção de conjuntos.

Figura 2 – Operações booleanas e QCA



Fonte: Adaptado de Legewie (2013).

Já se uma declaração envolve os conceitos país democrático ou país federativo é uma união entre o conjunto país democrático e o conjunto país federativo. A lógica de proposições utiliza o operador OU-lógico para denotar uniões de conjuntos; também chamadas de **disjunções**. Essa operação é representada por um sinal de mais (+) ou pelo sinal de união (\cup): $D + F$, $D \cup F$. É preciso fazer uma ressalva quanto ao OU-lógico: como as álgebras booleana e *fuzzy* denotam o OU-lógico utilizando um sinal mais (+) (chamado de OU booleano na álgebra booleana e adição *fuzzy* na álgebra *fuzzy*), isso pode gerar confusões. Em termos linguísticos, o sinal + pode ser tomado como a letra e; no entanto, o E-lógico equivale, na lógica das proposições, à multiplicação, representada por asterisco (*), em vez

da adição. Assim, o + booleano (OU-lógico) não deve ser lido como um e no sentido convencional, mas como ou. Outra confusão diz respeito aos resultados da operação adição: nas álgebras convencional e booleana, $0 + 0 = 0$; $1 + 0 = 1$; e $0 + 1 = 1$. Entretanto, na álgebra convencional $1 + 1 = 2$ e na álgebra booleana $1 + 1 = 1$, porque nesta + equivale a ou (SCHNEIDER; WAGEMANN, 2012).

O terceiro operador para a execução dos métodos baseados em conjuntos é o operador **negação** de uma declaração teórica ou conceito; que nada mais é do que o complemento do conjunto. Por exemplo, para representar o grupo de todos os países que não são democracias, pode-se utilizar o símbolo $\sim D$, da teoria de conjuntos, ou $1-D$, da álgebra booleana. Uma forma alternativa de simbolizar o complemento de um conjunto é utilizar a letra minúscula d . Portanto, os símbolos $\sim D$, $1-D$ e d são equivalentes. Doravante utilizam-se os símbolos $\sim D$ ou d para denotar a operação de negação de um conjunto. Além disso, é preciso ter certo cuidado com a operação complemento de conjuntos, pois ela não necessariamente denota sua contrapartida conceitual. Por exemplo, a negação do conceito democracia compreende tipos bastante distintos de regimes políticos, desde teocracias e regimes militares até regimes totalitários e monopartidários (SCHNEIDER; WAGEMANN, 2012).

Sejam os conjuntos hipotéticos A , B e C . Essas três operações básicas também podem ser aplicadas sobre combinações de conjuntos: interseção de uniões, uniões de interseções, negação (complemento) de uniões, negação de interseções, tais como: $(A+B)*(C+D)$, $(A*B)+(C*D)$, $\sim(A+B)$, $\sim(A*B)$, entre outros. A partir disso, emergem algumas propriedades dos operadores: **comutatividade**, **associatividade** e **distributividade**. A **comutatividade** significa que a ordem em que dois ou mais conjuntos estão combinados pelos operadores E e OU lógicos é irrelevante: $A*B$ equivale a $B*A$; $A+B$ equivale a $B+A$. Essa propriedade não é válida para o operador de negação: $1 - A$ não equivale a $A - 1$. Já a **associatividade** estabelece que a sequência em que os conjuntos isolados são combinados é irrelevante. Por exemplo, para a conjunção (interseção) dos conjuntos A , B e C , $(A*B)*C = A*(B*C) = (A*C)*B$; para a disjunção,

$(A+B)+C = A+(B+C) = (A+C)+B$. Por seu turno, a **distributividade** refere-se a fatoração de um conjunto isolado em expressões complexas de conjuntos envolvendo os operadores E e OU lógicos: que são compartilhados pelos vários componentes que podem ser fatorados: $A*B + A*C = A*(B + C)$; equivalentemente $AB + AC = A(B + C)$. Dois conjuntos importantes emergem das três operações básicas: conjuntos universo e vazio. A união de um conjunto com o seu complemento é o conjunto universo, denotado por U, quando se unem todos os elementos de ambos conjuntos ($D \cup \sim D = U$ equivalentemente $D + \sim D = U$). A interseção de um conjunto com o seu complemento é o conjunto vazio, denotado por \emptyset , quando um conjunto e seu complemento não possuem um único elemento em comum ($D \cap \sim D = \emptyset$).

Embora um elemento (caso) não possa ser simultaneamente um membro de um conjunto e do seu complemento (princípio conhecido como Regra do Terceiro Excluído), isso não se aplica em conjuntos *fuzzy*, porque um caso particular pode ter pertencimento parcial em um conjunto e no seu complemento. Por exemplo, se um caso tem escore *fuzzy* 0,3 em um conjunto, então tem escore *fuzzy* 0,7 no complemento desse conjunto ($1 - 0,3 = 0,7$). Dessa maneira, é possível aplicar o operador negação em declarações lógicas envolvendo os operadores E e OU lógicos mediante a Lei de DeMorgan, que consiste em duas regras: (i) se uma declaração (ou conceito) for negada, então todos os elementos do conjunto que a representa com *status* de pertencimento (presente) passam a ter *status* de não-pertencimento (ausente) e vice-versa; (ii) os operadores lógicos também devem ser invertidos: o operador OU-lógico (+) torna-se o operador E-lógico (*) e vice-versa. Por exemplo, aplicando-se essas regras aos operadores, teremos: a negação de uma interseção é a união de negações, $\sim(AB) = \sim A + \sim B$; a negação de uma união é a interseção de negações, $\sim(A+B) = \sim A * \sim B = \sim A \sim B$.

Uma vez que os conjuntos são operacionalizados por números reais – conjuntos dicotômicos com escores 1 (presença, pertencimento pleno) ou 0 (ausência, não pertencimento pleno) e conjuntos *fuzzy* com escores pertencentes ao intervalo $[0,1] \subset \mathfrak{R}$ –, é possível calcular os

escores de pertencimento em combinações de conjuntos que emergem das operações supracitadas. Assim como os conjuntos isolados são operacionalizados por valores (escores) no intervalo $[0,1] \subset \mathfrak{R}$, as combinações de conjuntos também podem ser operacionalizadas por valores nesse mesmo intervalo de números reais. No entanto, a forma de calcular os escores de cada caso nas combinações de conjuntos depende da forma da combinação. A **conjunção**, que é a interseção de conjuntos, segue a regra do escore mínimo entre os conjuntos componentes da conjunção; por exemplo, se o *i*-ésimo caso nos conjuntos dicotômicos A e B tem escores 1 e 0, respectivamente, então o escore do *i*-ésimo caso na interseção AB é $\min \{1, 0\} = 0$. Para a **disjunção**, união de conjuntos (OU-lógico), o cálculo segue a regra do escore máximo: o escore do *i*-ésimo caso na união A+B é $\max \{1, 0\} = 1$. Já para a operação **negação**, o escore é igual a um menos o escore do caso: a negação do conjunto A, $\sim A$, implica que o escore do *i*-ésimo caso no conjunto $\sim A$ seria $1 - 1 = 0$.

As regras continuam valendo para conjuntos *fuzzy*: se o *i*-ésimo caso nos conjuntos *fuzzy* A e B tem escores 0,8 e 0,4, respectivamente, então o escore do *i*-ésimo caso na interseção AB é $\min \{0,8; 0,4\} = 0,4$; o escore do *i*-ésimo caso na união A+B é $\max \{0,8; 0,4\} = 0,8$; o escore do *i*-ésimo caso na negação $\sim A$ é $1 - 0,8 = 0,2$. Em suma, essas operações que envolvem combinações de conjuntos – interseção, união, negação – serão úteis quando trabalharmos com as tabelas verdade, pois as mesmas são utilizadas tanto na construção dessas tabelas quanto na sua análise (simplificação da complexidade).

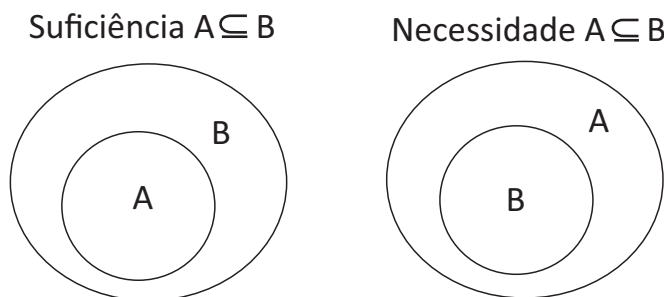
Relações entre conjuntos: suficiência e necessidade

Agora que aprendemos as operações básicas de conjuntos, precisamos dar um passo mais adiante e discutir as relações entre conjuntos. Uma vez que as declarações teóricas costumam estabelecer relações de causalidade entre conceitos (conjuntos) e como as relações de causalidade abarcam as relações de subconjunto ou superconjunto, precisamos discutir os conceitos de suficiência e de necessidade de um

conjunto (ou combinação de conjuntos) para outro conjunto. O principal propósito da QCA é identificar condições (conjuntos) ou combinações de condições que são necessárias e suficientes para o conjunto resultado. A QCA, ao operar com noções da teoria de conjuntos (*i.e.*, superconjuntos e subconjuntos), busca simplificar a análise da complexidade pelas condições necessárias e suficientes (LEGEWIE, 2013).

As propriedades de suficiência e de necessidade nada mais são do que relações entre conjuntos. Suponha hipoteticamente apenas dois conjuntos, A e B. Partindo-se de algum argumento teórico plausível a respeito da relação entre A e B (*e.g.*, A poderia ser subconjunto de B, ou B subconjunto de A ou poderiam ser iguais, $A = B$), podemos averiguar se a relação é válida para todos os casos (instâncias empíricas) de interesse. Se o referencial teórico estabelece que A causa B, então o conjunto (ou condição) A é subconjunto do conjunto (resultado) B, $A \subseteq B$, que é a propriedade de suficiência do conjunto A para o conjunto B. Por outro lado, a propriedade de necessidade define que B é subconjunto de A, $B \subseteq A$. Em termos de diagramas de Venn, a Figura 3 apresenta essas relações de subconjuntos para os conjuntos hipotéticos.

Figura 3 – Relações de suficiência e necessidade entre dois conjuntos hipotéticos



Fonte: Adaptado de Legewie (2013).

O conjunto A é suficiente para B porque todos os casos presentes em A também estão presentes em B: o círculo que representa a condição A é completamente coberto pelo círculo (maior) que corresponde ao resultado B. Para a propriedade de necessidade, o círculo que denota

o resultado B está completamente coberto pelo círculo (maior) que representa a condição A. Há casos incluídos no conjunto A que não estão no conjunto B, mas todos os casos no conjunto B também estão no conjunto A. Caso o referencial teórico estabelecesse que B causa A, então B seria suficiente para A se $A \subset B$ e A seria necessário para B se $A \subset B$. Por isso é importante que estejamos suportados em um referencial teórico adequado para a análise da complexidade de conjuntos (conceitos teóricos). A seguir explicaremos os modos de averiguar a suficiência e a necessidade em conjuntos dicotômicos e *fuzzy* de modo mais detalhado (*vide* Quadro 4).

A propriedade de **suficiência** estabelece que uma determinada condição (conjunto) pode ser considerada suficiente para um resultado se, sempre que ela estiver presente **em todos** os casos, o resultado nesses casos também está presente. Portanto, não pode existir um único caso em que a condição está presente e o resultado, ausente. Para avaliarmos a propriedade de suficiência das condições para o resultado, partimos das condições isoladas uma das outras. Averiguamos se a propriedade é satisfeita para cada condição selecionando os casos em que a condição está presente (casos com escore 1 na condição sob avaliação de suficiência) e conferimos se o correspondente resultado também está presente (escore 1). Se em pelo menos um caso o critério não se aplica à condição, então a mesma não pode ser considerada como suficiente para o resultado. Esse procedimento é executado para cada uma das demais condições isoladamente, inclusive para seus complementos e para combinações de condições. Recordemos que isso é possível porque sabemos como calcular os escores de conjuntos gerados por combinações de conjuntos. Lembremos também que os dados indicam escores de pertencimento às condições, e as condições, por sua vez, podem assumir dois estados qualitativamente distintos: elas podem estar presentes (1) ou ausentes (0), se estamos tratando de conjuntos dicotômicos. Cada condição (A) e o seu complemento ($\sim A$ ou a) denotam duas propriedades qualitativamente distintas. Portanto, a propriedade de suficiência dos complementos das condições também deve ser averiguada. O complemento da condição, por

exemplo, $\sim A$, será suficiente para $B = 1$ se os casos selecionados com $\sim A = 1$ tiverem em correspondência $B = 1$. Se pelo menos um dos casos com $\sim A = 1$ tiver $B = 0$, então $\sim A$ não é suficiente para B (*vide* Quadro 4).

Mesmo que as condições e seus complementos isoladamente não satisfaçam a propriedade de suficiência, ainda é possível que a sua ocorrência simultânea (conjunção de condições e seus complementos) seja suficiente na presença do resultado (resultado com escore 1). Requer-se, portanto, avaliar a propriedade de suficiência para conjunções de condições; para essa etapa calculamos os escores das conjunções. Como uma conjunção de condições é uma interseção de conjuntos, pela álgebra booleana o cálculo de escores, nesse caso, segue a regra do mínimo (o escore da interseção de dois ou mais conjuntos é o escore mínimo de cada um desses conjuntos). Supondo que temos três condições, A , B e C , que teoricamente causam o resultado R , podemos ter 2^k conjunções possíveis (ABC , $AB\sim C$, $A\sim BC$, $A\sim B\sim C$, $\sim ABC$, $\sim AB\sim C$, $\sim A\sim BC$, $\sim A\sim B\sim C$) e devemos averiguar a suficiência de cada uma das oito para o resultado $R = 1$. Assim, seguimos o mesmo procedimento realizado para as condições e seus complementos isolados: selecionamos em cada conjunção os casos com escore 1 (que indica a presença da conjunção) e verificamos se o escore correspondente no resultado é 1. Se pelo menos um deles é nulo ($R = 0$), então a conjunção não é suficiente para o resultado presente. Algumas conjunções podem ser suficientes para R , e outras, não. O procedimento de avaliar se as conjunções são suficientes para o resultado tem uma importância crucial quando analisamos tabelas verdade, conforme será explicado no capítulo 4.

Além da avaliação da propriedade de suficiência dos conjuntos de interesse, o pesquisador também precisa avaliar a propriedade de necessidade. A propriedade de necessidade estabelece que uma determinada condição (ou conjunção de condições) é necessária se, sempre que o resultado está presente, a condição também está presente. Isto é, o resultado não pode ser realizado ($R = 1$) sem a presença da condição (por exemplo, $A = 1$); equivalentemente, nenhum caso com o resultado presente pode exibir uma condição ausente. A identificação das condições que são necessárias para um determinado resultado se

assemelha ao procedimento realizado para a propriedade de suficiência, sendo que, agora, é realizado para um grupo distinto de casos: só averiguamos a necessidade para os casos em que o resultado está presente. Portanto, observar a necessidade consiste em descobrir, caso haja, quais das condições estão presentes em todos esses casos com escore 1 no resultado. Para a propriedade de suficiência, selecionamos os casos em que cada condição estava presente e conferimos se o resultado também estava ou não presente. Inversamente, agora, selecionamos os casos em que o resultado está presente e conferimos se a condição está ou não presente. Se em pelo menos um dos casos a condição (ou seu complemento) estiver ausente, então a condição não é necessária para o resultado ($R = 1$) (*vide* Quadro 4).

O procedimento também se inicia a partir dos conjuntos e seus complementos tomados isoladamente. Caso eles não sejam isoladamente necessários para o resultado presente, não se faz a avaliação da necessidade das conjunções de condições, pois isso não faz sentido. A razão disso é que a regra do mínimo reduz as chances de que as conjunções satisfaçam a propriedade de necessidade. A lógica formal estabelece que: (a) uma conjunção só pode ser necessária quando formada exclusivamente por condições isoladas necessárias; e (b) condições não necessárias não podem formar uma conjunção que seja necessária (SCHNEIDER; WAGEMANN, 2012).

Assim sendo, não é preciso averiguar a necessidade das combinações de condições que não se qualificam como necessárias para o resultado. Por outro lado, as uniões de condições e seus complementos podem ser avaliadas quanto à sua necessidade para o resultado, porque elas são obtidas pelo operador OU-lógico, cujos escores são calculados pela regra do máximo (o maior escore das condições que compõem uma união determina o escore dos casos na união). Essa operação cria conjuntos em que mais casos tendem a ser membros (porque são superconjuntos), conseqüentemente aumentando a chance de que as uniões sejam superconjuntos do resultado. Como base em Schneider e Grofman (2006), elaboramos o Quadro 4 a fim de sintetizar as condições de suficiência e necessidade.

Quadro 4 – Síntese das relações de suficiência e de necessidade

Relação	Descrição
Necessária, mas não suficiente	<p>A condição <i>A</i> é <i>necessária</i> para o resultado <i>Y</i> se a ocorrência de <i>Y</i> não é possível sem a presença de <i>A</i>, mas ela por si só não é <i>suficiente</i> para produzir <i>Y</i>. Ou seja, <i>A</i> é <i>necessária, mas não suficiente</i>, se existir combinações vinculadas com o resultado, que não permite a condição <i>A</i> produzir sozinha.</p> <p>$A * R + A * p = A * (R + p) \rightarrow Y$ em que também: Não <i>A</i> → Não <i>Y</i></p>
Suficiente, mas não necessária	<p>A condição <i>A</i> é <i>suficiente, mas não necessária</i>, se tal condição é capaz de produzir o resultado, mas ao mesmo tempo existem outras combinações também vinculadas com o resultado.</p> <p>$A + R * p \rightarrow Y$</p>
Suficiente e necessária	<p>Condição <i>A</i> é <i>necessária e suficiente</i>.</p> <p>$A \rightarrow Y$ em que também: Não <i>A</i> → Não <i>Y</i></p>
Nem suficiente, nem necessária	<p>A condição <i>A</i> <i>não é suficiente e nem necessária</i> para o resultado, se <i>A</i> produz <i>Y</i> somente se combinado com outras condições.</p> <p>$A * p + R * P + a * R \rightarrow Y$</p>

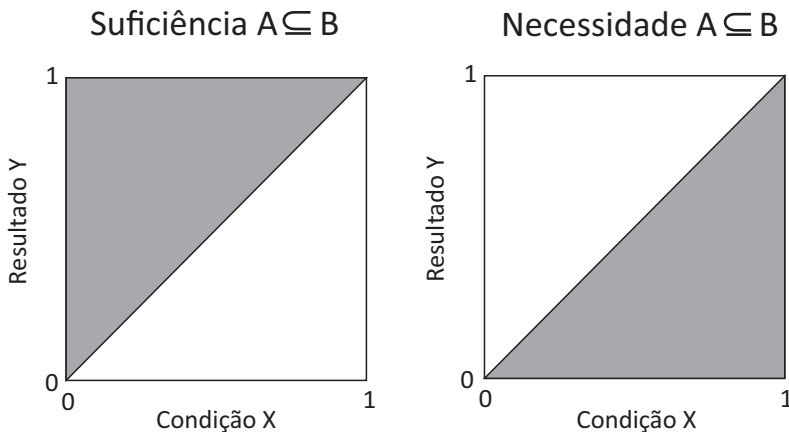
Fonte: Adaptado de Schneider e Grofman (2006).

Nesse Quadro 4, o sinal \rightarrow (juntamente com o seu oposto, \leftarrow) pode ser usado para indicar a lógica de relacionamentos. Tais relações são potencialmente causais, mas elas podem, na verdade, representar simplesmente particulares de concordâncias empíricas observadas de condições e resultado que não são verdadeiramente causais na natureza. A definição se ou não essas relações podem ser lidas como causais é determinada pela teoria (SCHNEIDER; GROFMAN, 2006).

As propriedades de suficiência e necessidade também se aplicam a conjuntos *fuzzy* e não diferem substancialmente das realizadas para conjuntos dicotômicos. A relação aritmética entre os escores de pertencimento em conjuntos *fuzzy* é estabelecida pelo princípio do

subconjunto, de tal modo que um conjunto *fuzzy* A é subconjunto de outro conjunto *fuzzy* B se os escores de pertencimento dos casos em A são menores ou iguais aos escores de pertencimento dos casos correspondentes em B. Formalmente, sejam dois conjuntos *fuzzy*, $A = \{s_1^A, s_2^A, \dots, s_n^A\}$ e $B = \{s_1^B, s_2^B, \dots, s_n^B\}$, e s_i^A, s_i^B os escores do i-ésimo caso em cada conjunto, e $s_i^k \in [0,1] \subset \mathfrak{R}, \forall i, k$; então, $A \subset B$ se $s_i^A \leq s_i^B, \forall i$. Analogamente, B poderia ser subconjunto de A, $B \subset A$ se $s_i^B \leq s_i^A, \forall i$. Esse princípio do subconjunto é uma relação fundamental entre dois conjuntos. Se o conjunto A é subconjunto de B ou vice-versa é uma questão posta pelo referencial teórico que guia o pesquisador. Se os conjuntos A e B forem teoricamente relacionados de tal modo que A é uma condição para o resultado B, então A é condição necessária, mas não suficiente para o resultado B, quando as instâncias do resultado, B, constituem um subconjunto das instâncias da causa, A. Em termos ilustrativos, o **diagrama XY** da Figura 4 mostra as relações entre dois conjuntos *fuzzy* (SCHNEIDER; WAGEMANN, 2012).

Figura 4 – Diagramas XY para conjuntos *fuzzy* hipotéticos



Fonte: Adaptado de Legewie (2013).

Se o eixo Y indica os valores (escores) *fuzzy* dos casos no conjunto resultado, e o eixo X, os escores *fuzzy* dos casos no conjunto condição, então os valores 0 e 1 delimitam respectivamente os níveis mínimo e máximo de pertencimentos (quadrado com lados de igual comprimento), porque os conjuntos *fuzzy* são intervalos de números reais, $[0,1] \subset \mathfrak{R}$. A linha diagonal

representa os casos com os mesmos escores de pertencimento em X e Y. A Figura 4 também descreve as propriedades de suficiência e necessidade para esses dois conjuntos hipotéticos X e Y. Se X é condição e Y é resultado, X é suficiente para Y se $X \leq Y$ (no diagrama XY é a diagonal principal mais a área na parte *superior* à esquerda da diagonal principal; área formada pelos pares de escores de X e Y tais que o escore em X é menor ou igual ao escore em Y). Se X é condição e Y é resultado, X é necessário para Y se $Y \leq X$ (no diagrama XY é a diagonal principal mais a área na parte *inferior* à esquerda da diagonal principal, área formada pelos pares de escores de X e Y tais que o escore em Y é menor ou igual ao escore em X).

Obviamente as propriedades de suficiência (condição subconjunto do resultado) e necessidade (resultado subconjunto da condição) podem ser averiguadas para combinações de conjuntos (conjunções, disjunções e negações). Ressaltemos que alguns pesquisadores, em vez de utilizar os termos propriedade de suficiência e propriedade de necessidade, utilizam os termos condição de suficiência e condição de necessidade, respectivamente. Não existe absolutamente nada de equivocado em utilizar esses termos. Contudo, uma vez que na abordagem QCA o termo condição tem um significado bem definido – é o conjunto estabelecido no referencial teórico como um conceito causal para outro conceito, teoricamente definido como resultado –, preferimos utilizar os termos propriedade de suficiência e propriedade de necessidade para evitar confusão. É claro que após a averiguação dessas propriedades resultarão condições (conjuntos) que são ou não suficientes e/ou necessárias para o conjunto resultado. No capítulo 5, exemplificamos os critérios de suficiência e necessidade em uma aplicação das técnicas csQCA e fsQCA no âmbito de políticas públicas. Lá, algumas das etapas que devem ser executadas para averiguar tais critérios serão detalhadas.

Medidas de consistência e cobertura

Até aqui tratamos de relações teóricas entre conjuntos. Entretanto, quando empiricamente aplicadas às propriedades de suficiência e necessidade são parcialmente satisfeitas, no sentido de que: (a) para

conjuntos dicotômicos, podem existir alguns casos em que a condição está presente e o resultado, ausente sob a verificação da propriedade de suficiência (e vice-versa, quanto à verificação da propriedade de necessidade); e (b) para conjuntos *fuzzy*, nem todos os pares de escores *fuzzy* satisfazem as propriedades $s_i^A \leq s_i^B$ ou $s_i^B \leq s_i^A$, $\forall i$. Portanto, é preciso avaliar a extensão em que os conjuntos (isolados ou combinações deles por conjunção e disjunção) satisfazem cada uma das propriedades. Isso é obtido por meio de medidas de extensão das relações entre conjuntos, denominadas medidas de consistência e de cobertura das condições. Notemos que é preciso avaliar a consistência e a cobertura das condições (isoladas ou combinadas) que importam para o resultado, isto é, das condições que são suficientes e necessárias para o resultado. Queremos saber a extensão da relação de subconjunto.

A **medida de consistência** das condições, que é o principal critério de validação da QCA (RAGIN, 2006), é uma medida descritiva para avaliar a força do suporte empírico para a argumentação teórica, que descreve relações de causalidade entre conjuntos (conceitos). Ela mede o grau de proximidade da relação entre conjuntos; avalia o grau em que os casos que compartilham uma determinada condição (ou combinação de condições) concordam com o resultado. Já a **medida de cobertura**, por outro lado, provê uma quantificação da relevância empírica de uma condição ou combinação causal no conjunto das combinações causais. Ela avalia o grau em que uma condição causal ou combinação de condições causais conta para um resultado. Ela se assemelha ao coeficiente de determinação no contexto econométrico (THIEM, 2010).

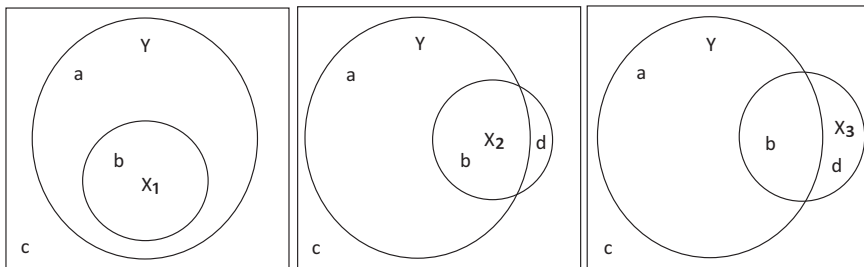
A consistência e cobertura das condições suficientes e necessárias podem ser mais bem compreendidas por **tabelas dois-por-dois** e **diagramas de Venn** quando as condições e resultado são conjuntos dicotômicos, e **diagramas XY** quando são conjuntos *fuzzy*. A seguir apresentamos os modos de calcular a consistência e a cobertura das condições suficientes e necessárias dicotômicas e *fuzzy* utilizando recursos algébricos e diagramáticos, baseando-nos em Schneider e Wagemann (2012). Em todas as análises a seguir, trabalharemos com os conjuntos hipotéticos:

condições X1, X2 e X3 e resultado Y. Primeiramente apresentamos a medida de consistência para conjuntos dicotômicos suficientes e, em seguida, para conjuntos *fuzzy* suficientes.

Consistência e cobertura das condições suficientes

Suponhamos que todos os conjuntos são dicotômicos. A Figura 5 ilustra os diagramas de Venn para três condições X1, X2 e X3 e o resultado Y. Em todos os diagramas os tamanhos dos conjuntos são considerados idênticos e só há mudança quanto à localização dos conjuntos X em relação ao conjunto Y. A condição X1 é subconjunto pleno de Y (isto é, X1 é suficiente para Y) e as condições X2 e X3, não. Contudo, essas duas condições diferem em grau em termos de **violação da propriedade de suficiência**: X3, em comparação à condição X2, está mais fora do que dentro de Y, porque a parcela (número de casos) de X3 que está fora de Y (área d) em relação ao tamanho total de X3 (áreas b e d) é maior do que a parcela (área d) da condição X2. Assim sendo, a condição X2 é relativamente mais consistente do que a condição X3 como condição suficiente para Y.

Figura 5 – Diagramas de Venn de condições suficientes consistentes e inconsistentes



Fonte: Adaptado de Schneider e Wagemann (2012).

Observemos que o termo “está mais fora do que dentro” aqui não pode ser confundido com o mesmo significado quando tratamos de graus de pertencimento em conjuntos *fuzzy* (gradação dentro de um mesmo conjunto). Embora sejam semelhantes, denotam aspectos distintos: “está

mais fora do que dentro” em um conjunto *fuzzy* refere-se à posição dos elementos desse conjunto no próprio conjunto, e “está mais fora do que dentro” no âmbito da consistência (e da cobertura também) representa uma relação entre dois conjuntos distintos, mesmo que também se refira a elementos desses dois conjuntos.

Os diagramas de Venn oferecem uma boa intuição para a compreensão do conceito de consistência de uma condição para um resultado, mas as tabelas dois-por-dois são mais úteis quando queremos compreender a maneira de calcular o parâmetro de ajuste medida de consistência. A Figura 6 exibe as três condições e resultado conforme a Figura 5, em que as células a a d correspondem respectivamente às áreas a a d nos diagramas de Venn. Os números nas células indicam o número de casos que mostram os escores de pertencimento na condição e resultado. Por exemplo, a célula c referente à condição X_1 consiste dos 15 casos dessa condição em que $X_1 = 0$ e $Y = 0$. Notemos que os números de casos para $X_1 = 0$ e $X_1 = 1$ não mudam, são 95 e 100, respectivamente. O que muda é a distribuição dos casos com $X_1 = 1$ quanto ao resultado, isto é, a distribuição dos 100 casos com $X_1 = 1$ entre as células b e d.

Agora podemos não somente averiguar se as condições são consistentes ou não para o resultado, mas também calcular a extensão da suficiência. A diferença entre as condições X_1 , X_2 e X_3 é que a condição X_1 é plenamente consistente, pois todos os elementos de X_1 em que a condição está presente ($X_1 = 1$) estão na célula b (também têm Y presente ($Y = 1$)), e para as condições X_2 e X_3 isso não ocorre. Em alguns casos de X_2 a condição está ausente ($X_2 = 0$) e muitos casos de X_3 , também ausente ($X_3 = 0$). Portanto, as condições X_2 e X_3 são inconsistentes. Lembremos que do, **ponto de vista puramente teórico**, somente X_1 é suficiente, uma vez que não existe um único elemento de X_1 tal que $X_1 = 0$ e $Y = 1$ (não há casos na célula d). As condições X_2 e X_3 não são suficientes até porque têm casos na célula d de suas tabelas dois-por-dois: 10 e 92, respectivamente. Por outro lado, do **ponto de vista empírico**, as condições X_2 e X_3 são suficientes, mas em graus distintos.

Figura 6 – Tabelas dois-por-dois de condições suficientes consistentes e inconsistentes

1	80	100
a		
0	15	0
c		
	0	1
	X ₁	

1	80	90
a		
0	15	10
c		
	0	1
	X ₂	

1	80	8
a		
0	15	92
c		
	0	1
	X ₃	

Fonte: Adaptado de Schneider e Wagemann (2012).

Ragin (2006) sugeriu a seguinte fórmula para mensurar a consistência de uma condição suficiente X para um resultado Y – dividir o número de casos com presença na condição X e (interseção) no resultado Y (X = 1 e Y = 1) pelo número de casos com presença na condição X (X = 1):

$$\text{Consistência da condição suficiente X para o resultado Y} = \frac{\text{Número de casos com X=1 e Y=1}}{\text{Número de casos de X=1}} \quad (1)$$

Em termos de células da tabela dois-por-dois da Figura 6:

$$\text{Consistência da condição suficiente X para o resultado Y} = \frac{\text{Número de casos da célula b}}{\text{Número de casos das células b e d}} \quad (2)$$

Com base nessa fórmula, podemos calcular o **valor de consistência** das condições X1, X2 e X3:

$$I_{X_1,Y} = \frac{100}{(100+0)} = 1,00$$

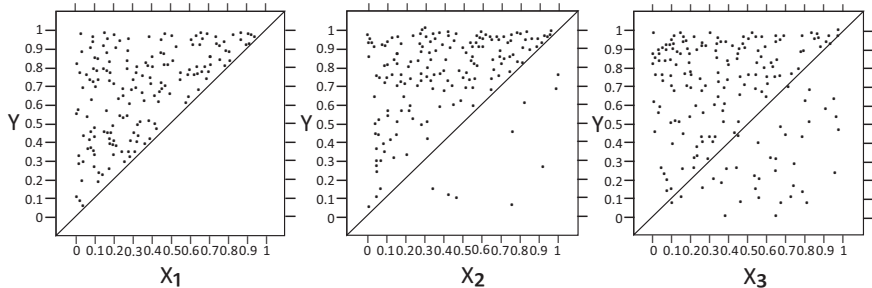
$$I_{X_2,Y} = \frac{90}{(90+10)} = 0,90$$

$$I_{X_3,Y} = \frac{8}{(8+92)} = 0,08$$

Se, entretanto, X1, X2, X3 e Y são conjuntos *fuzzy*, a noção de relações de subconjunto é mais bem representada graficamente pelo diagrama XY. A Figura 7 fornece os diagramas XY de três condições, que

retratam a natureza do grau de consistência para o conjunto resultado semelhantemente às de conjuntos dicotômicos.

Figura 7 – Diagramas XY de condições suficientes consistentes e inconsistentes



Fonte: Adaptado de Schneider e Wagemann (2012).

Poderíamos repetir o mesmo procedimento executado para conjuntos dicotômicos, contando o número de casos que satisfazem a propriedade de suficiência (aqueles que estão acima ou sobre a linha diagonal, porque $X_i \leq Y$) e dividi-lo pelo número de casos que são relevantes para a averiguação dessa propriedade (aqueles com pertencimento na condição X_i superior a 0). Dessa maneira, e conforme a Figura 7 adaptada de Schneider e Wagemann (2012), o valor da consistência da condição X_1 para Y seria 1 ($IX_1, Y = 195/195$), porque nenhum caso com $X_1 > 0$ está abaixo da linha diagonal; para a condição X_2 há 10 casos em 195 com $X_2 > 0$ e abaixo da linha diagonal. Assim, sua consistência é 0,95 ($IX_2, Y = 185/195$); para X_3 , 0,74 ($IX_3, Y = 145/195$).

Não obstante, esse modo de calcular o valor da consistência para conjuntos *fuzzy* é inadequado, porque a distância entre os casos e a linha diagonal é de interesse para o pesquisador. Os **casos mais distantes** da linha diagonal são os que mais se distanciam da alegada relação de subconjunto. Por exemplo, o caso A na Figura 7 tem alto grau de pertencimento (escore em X_2 superior a 0,5) na condição supostamente suficiente X_2 , mas um baixo pertencimento a Y (escore do caso A em Y inferior a 0,5). Portanto, isso não satisfaz a propriedade de suficiência mais do que os casos que estão ligeiramente abaixo da linha diagonal

e/ou que têm baixo grau de pertencimento na condição e no resultado. Lembremos que os casos inconsistentes para a propriedade de suficiência são aqueles localizados abaixo da linha diagonal.

Ragin (2006, 2008; p. 44-68) sugere uma solução para esse problema a partir dos escores de pertencimento dos casos e da regra do mínimo por se tratar justamente de uma conjunção (interseção). Para cada um dos casos, identificamos o valor mínimo entre os escores de pertencimento em X e Y ($s_i = \min(X_i, Y_i)$) é o escore mínimo do i -ésimo caso nos conjuntos X e Y), adicionamos os valores mínimos encontrados ($\sum_{i=1, I}(s_i)$), posteriormente, dividimos pela soma dos escores de pertencimento de todos os casos do X, tal como $\sum_{i=1, I}(X_i)$. Ou melhor:

$$I_{XY}(x_i \leq y) = \frac{\sum \min(x_i, y_i)}{\sum x_i} \quad (3)$$

Quando os valores (escores) de X são todos iguais ou menores do que os seus correspondentes escores em Y, então o valor da medida de consistência é 1 (como é requerido para a consistência plena da suficiência). Na presença de alguns **casos inconsistentes** (são os casos localizados abaixo da linha diagonal do diagrama XY), quando alguns escores de X forem superiores aos de Y, o valor de consistência será menor do que a unidade. Quando há muitos casos inconsistentes, o valor da consistência pode ser menor do que 0,5 e o critério de suficiência não é satisfeito. Note-se que a medida $\sum_{i=1, I} \min(X_i, Y_i)$ é um escore do compartilhamento dos conjuntos, indicando a parte do conjunto X que sobrepõe o conjunto Y. Quanto mais os casos se distanciam da linha principal, maior a diferença entre os seus escores em X e Y, e menor a soma no numerador em relação à soma dos escores de X, no denominador. Portanto, a medida de consistência leva em consideração a extensão em que um determinado caso se localiza abaixo da linha diagonal; ou seja, até que ponto o escore em X excede o de Y.

Em termos práticos, uma relação de suficiência de uma condição para um resultado é estabelecida se significativamente a maioria dos casos satisfaz a propriedade de suficiência – quando mais de 80% dos escores de pertencimento da condição são consistentemente menores

ou iguais aos seus respectivos escores de pertencimento no resultado \rightarrow , então o pesquisador pode declarar que a condição (ou combinação de condições) é “quase sempre” suficiente para o resultado (RAGIN, 2006).

Apesar de ser um modo plausível de expressar numericamente a extensão da relação de subconjunto, a medida de consistência tem uma limitação particular quando aplicada a conjuntos *fuzzy*: ela não leva em consideração se um caso inconsistente está acima ou abaixo da âncora qualitativa de 0,5 em X e/ou em Y. Por exemplo, na Figura 7 no diagrama XY referente à condição X2, as distâncias dos casos A, B e C à linha diagonal são idênticas, contribuindo igualmente para a inconsistência da condição X2 para Y. Contudo, seus escores em X2 são distintos (escore de B < escore de A < escore de C), configurando uma diferença qualitativa entre os casos B e C em relação ao caso A, que é importante para declarar se X2 pode ser interpretada como condição suficiente de Y. Os casos B e C têm escores em X e Y que estão no mesmo lado da âncora qualitativa de 0,5 (*e.g.*, os escores de C em X e Y são superiores a 0,5; significando que o caso C está mais dentro do que fora nesses dois conjuntos). Por outro lado, o caso A tem escores de pertencimento qualitativamente distintos em X e Y: seu escore em X2 é superior a 0,5 (retratando uma boa instância empírica dessa condição), e escore em Y, inferior a 0,5 (retratando uma má instância empírica do resultado). Assim, enquanto os casos B e C são simplesmente casos inconsistentes, o caso A é um **caso logicamente contraditório**.

Em resumo, a medida de consistência indica o grau em que a evidência empírica está em conformidade com a propriedade de suficiência. Quanto mais os casos de uma condição se desviam da definição de subconjunto ($X \leq Y$) de um resultado e quanto mais forte esse desvio, menor o valor de consistência. A medida de consistência pode ser calculada para averiguar a suficiência de qualquer complexidade. Isto é, o conjunto X pode consistir de uma combinação E-lógico (conjunção) e OU-lógico (disjunção) de vários conjuntos. Independentemente do número de conjuntos combinados com os diferentes operadores lógicos, cada caso tem um escore de pertencimento na combinação, definido pela regra do mínimo (para as conjunções) e/ou a regra do máximo (para as disjunções). Consequentemente, os valores de consistência podem ser calculados:

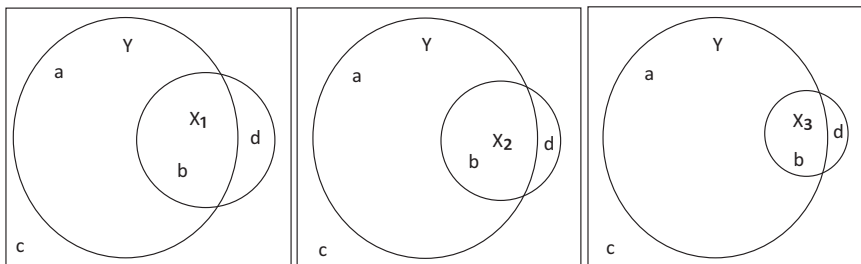
para linhas da tabela verdade (conjunções de condições), para caminhos únicos identificados como suficientes, ou para uma fórmula de solução (que é a solução da aplicação da QCA). A avaliação da consistência das linhas da tabela verdade é chamada de consistência bruta, e o valor de consistência da fórmula de solução é chamado de consistência da solução.

Quando há várias combinações de condições associadas ao resultado (causação conjuntural múltipla), além de medir a suficiência dessas combinações, desejamos quantificar também a importância de cada combinação de condições, na coleção de combinações que implicam para o resultado. Isto é, se uma relação de subconjunto já foi empiricamente estabelecida por meio do parâmetro de consistência, queremos saber qual a relação, em termos de tamanho, entre o subconjunto (X) e o superconjunto (Y). O montante encontrado expressa quanto do resultado Y é coberto (compartilhado) pela condição X; portanto, expressa a importância empírica de X para explicar Y.

Por outro lado, a noção de **cobertura** de um conjunto para outro conjunto pode ser apresentada em termos de diagramas de Venn e de tabelas dois-por-dois, para conjuntos dicotômicos, e diagramas XY, para conjuntos *fuzzy*. Continuemos com as condições X1, X2 e X3 e o resultado Y. Primeiramente apresentamos a medida de cobertura para condições dicotômicas suficientes e, em seguida, para condições *fuzzy* suficientes.

Suponhamos que esses conjuntos sejam dicotômicos. A Figura 8 apresenta as três diferentes condições para o mesmo resultado Y e todas são inconsistentes como condições suficientes para Y.

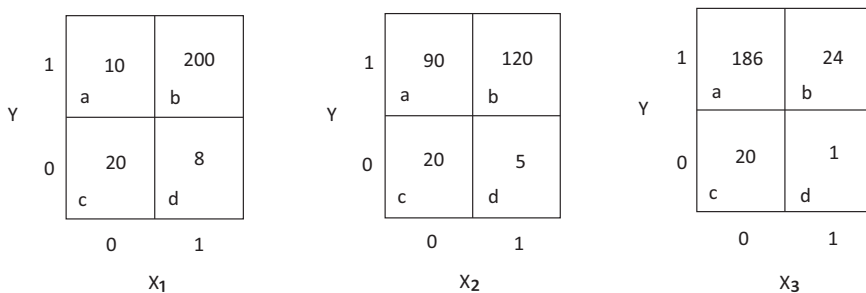
Figura 8 – Diagramas de Venn para diferentes coberturas da suficiência



Fonte: Adaptado de Schneider e Wagemann (2012).

Admitimos também que a razão da área d (parte de X que não é subconjunto de Y) sobre as áreas b e d (tamanho da condição) é igual nos três diagramas. Enquanto o tamanho do conjunto Y não muda, os tamanhos dos conjuntos X em relação ao conjunto Y diferem entre os diagramas: a condição X3 é menor do que a condição X2, que é menor do que a condição X1. A variação do tamanho do conjunto X indica a variação da quantidade de casos com Y = 1 que são cobertos (compartilhados) pelos conjuntos X. Por exemplo, o compartilhamento da condição X1 com o resultado (Y) é maior do que o compartilhamento da condição X3 com o resultado (Y). Portanto, as condições têm coberturas diferentes. A fim de complementar a apresentação da medida de cobertura de condição suficiente, passemos a uma ilustração aritmética por meio de tabelas dois-por-dois. A Figura 9 ilustra tais tabelas, que refletem as propriedades dos conjuntos nos diagramas de Venn da Figura 8. A razão da área b para a área b e d é constante ($200/208 = 120/125 = 24/25 = 0,96$), mas o número de casos de cada condição X compartilhados com o resultado Y se modifica.

Figura 9 – Tabelas dois-por-dois de condições suficientes consistentes e inconsistentes



Fonte: Adaptado de Schneider e Wagemann (2012).

O cálculo da cobertura de X para Y pode ser realizado a partir da fórmula proposta por Ragin (2006):

$$\text{Cobertura de X como condição suficiente para o resultado Y} = \frac{\text{Número de casos com X=1 e Y=1}}{\text{Número de casos de Y=1}} \quad (4)$$

A mesma fórmula, porém em termos das células da tabela dois-por-dois, torna-se:

$$\text{Cobertura de X como condição suficiente para o resultado Y} = \frac{\text{Número de casos da célula b}}{\text{Número de casos das células a e b}} \quad (5)$$

Com base nessa fórmula podemos calcular o **valor de cobertura** das condições X1, X2 e X3, como seguem:

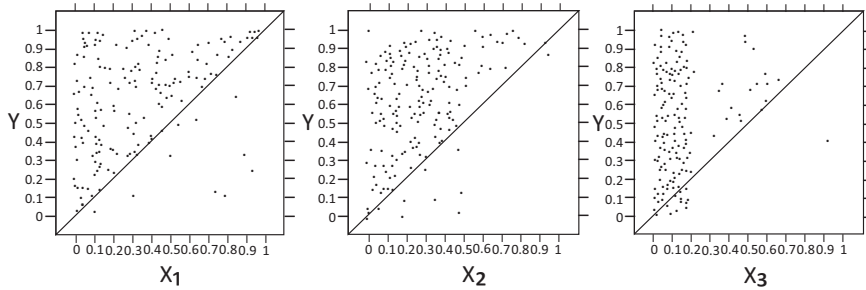
$$C_{X1,Y} = \frac{200}{(10+200)} = 0,95$$

$$C_{X2,Y} = \frac{120}{(90+120)} = 0,57$$

$$C_{X3,Y} = \frac{24}{(186+24)} = 0,11$$

Portanto, a condição X1 compartilha mais casos com Y = 1 do que as demais condições, conforme mostrado pelos diagramas de Venn. Notemos que a cobertura plena de uma condição qualquer só seria alcançada se nenhum caso estivesse na célula a (X = 0 e Y = 1). Se os conjuntos X1, X2 e X3 e Y são *fuzzy*, então recorreremos a diagramas XY para dimensionar a cobertura de uma condição suficiente. Na Figura 10, adaptada de Schneider e Wagemann (2012), os conjuntos X têm valores de consistência idênticos de 0,91, mas diferentes coberturas, cujos coeficientes podem ser observados pela distribuição dos pares de escores de X e Y. Partindo-se do diagrama X1Y para o X3Y, notamos que os casos passam a se localizar mais próximos do eixo-Y, em que os escores de X indicam que os casos estão mais fora do que dentro desse conjunto.

Figura 10 – Diagramas XY de condições suficientes consistentes e inconsistentes



Fonte: Adaptado de Schneider e Wagemann (2012).

Observemos que, quanto mais casos estão localizados no canto superior esquerdo do diagrama, e mais distantes da linha diagonal, menor é o valor de cobertura da condição. Embora esses casos sejam instâncias empíricas adequadas do resultado (alto pertencimento a Y), eles têm instâncias empíricas inadequadas para a condição suficiente (baixo pertencimento a X), exigindo-nos uma explicação adequada para eles. A fórmula de cálculo da cobertura proposta por Ragin (2006) para conjuntos *fuzzy* recorre às informações da gradação dos escores em cada conjunto (e não apenas conta os casos, conforme explicado na fórmula da medida de consistência):

$$C_{XY}(x_i \leq y) = \frac{\sum \min(x_i, y_i)}{\sum y_i} \quad (6)$$

Assim, as coberturas das condições consistentes X1, X2 e X3 para Y são, respectivamente, $C_{X1,Y} = 0,81$; $C_{X2,Y} = 0,6$; $C_{X3,Y} = 0,19$. Portanto, a condição X1 é empiricamente mais importante do que as demais. A medida de cobertura seria análoga ao R^2 em modelos estatísticos¹⁴ (THIEM, 2010). Essa cobertura avalia a parcela dos casos presentes simultaneamente em x e y em relação ao total de casos em y (RAGIN,

¹⁴ Contudo, é importante evitar certas comparações entre as duas metodologias, justamente para não aumentar a “confusão” conceitual da QCA e contribuir para a sua melhor compreensão metodológica (SCHNEIDER; GROFMAN, 2006).

2008), com valores variando entre 0 e 1. Ademais, a cobertura aborda um aspecto diferente do que a consistência. Por exemplo, o conjunto de falhas de paraquedismo seria um subconjunto quase perfeito (*i.e.*, alta consistência) do conjunto de mortes. Entretanto, essa combinação pode não ser muito útil (*i.e.*, baixa cobertura) para determinar os caminhos mais comuns ou significativos para a mortalidade em uma determinada população (LONGEST; VAISEY, 2008). O cálculo da cobertura para conjuntos *fuzzy* leva em conta o quão distante acima da linha diagonal estão localizados os casos e, portanto, o montante de seus escores de pertencimento a Y que não são cobertos pelo seu pertencimento a X. Ou seja, os casos mais próximos do canto superior esquerdo têm pouca contribuição na soma do numerador (baseada na regra do mínimo entre X_i e Y_i), mas relativamente grande contribuição para a soma no denominador (escores elevados de Y).

Convém lembrar que a medida de cobertura apresenta a seguinte desvantagem: uma vez que o cálculo da cobertura também leva em consideração parte do pertencimento dos casos inconsistentes (que não satisfazem a propriedade de suficiência; casos em que o escore de X é maior do que o escore de Y; e que estão abaixo da linha diagonal) em que Y que é coberto por X, o valor de cobertura sofre um viés de alta. Schneider e Wagemann (2012) propõem uma regra prática: a consistência de uma condição suficiente deve ser calculada sempre antes do cálculo de seu grau de cobertura. Já a cobertura só deve ser calculada para condições que passam no teste de consistência (RAGIN, 2006); porque não tem sentido calcular e interpretar o valor de cobertura de uma condição que não é suficiente.

Entretanto, Thiem (2010) adverte que a medida de cobertura é sensível a diferentes funções utilizadas para assinalar pertencimento aos casos, à localização das âncoras qualitativas e a outras restrições, porém, inexistente um critério bem definido para a escolha da função mais apropriada para maximizar a cobertura. Independentemente disso, uma vez que as combinações de condições constituem subconjuntos (ou superconjuntos) de condições causais, podem emergir valores de cobertura para cada

um desses conjuntos, conforme a extensão de compartilhamento que desejamos mensurar. Em suma, a medida de cobertura expressa o grau em que a parte consistente de uma condição suficiente sobrepõe-se a um resultado.

No capítulo 5 aplicamos a medida de cobertura, bem como de consistência, na solução de tabelas verdade. Lá veremos que a minimização lógica das linhas (que são conjunções de condições) desse tipo especial de tabela dá origem a soluções que envolvem operadores lógicos, especialmente uniões de conjunções. Por exemplo, $X1 \sim X3 + X2X3 \rightarrow Y$ (essa disjunção recebe a denominação de fórmula de solução ou solução final; e as conjunções que a definem, de termos de solução). Sob tais circunstâncias podemos estar interessados no valor de cobertura (e até mesmo de consistência) das conjunções $X1 \sim X3$ e $X2X3$ e da disjunção $X1 \sim X3 + X2X3$.

Por conta disso, se o interesse não é só na parte do resultado coberto por qualquer condição suficiente, mas na cobertura total de todas as possíveis combinações suficientes para o resultado, então calcula-se a **cobertura total** de ambas as combinações em uma fórmula de solução. Isso é feito simplesmente ao calcular o escore aderido de cada caso na fórmula de solução. Por sua vez, a **cobertura bruta** mede a proporção do pertencimento no resultado que é explicada por cada termo de solução. Entretanto, pode-se calcular a cobertura de cada combinação suficiente em relação ao resultado (*i.e.*, chamada **cobertura única**). Ou melhor, a **cobertura única** quantifica a proporção do pertencimento no resultado que não é coberta pelo outro termo de solução; ou melhor, a proporção que é unicamente explicada por cada termo de solução. Para tanto, primeiramente computa-se a cobertura solução; em seguida, calcula-se a cobertura das combinações, exceto aquela cobertura específica de interesse; e, posteriormente, subtrai-se esse valor a partir da cobertura da solução. O número se situa entre 0 e 1, e expressa o quanto o resultado é coberto unicamente por uma condição específica – líquido de todas as outras combinações suficientes (SCHNEIDER; GROFMAN, 2006).

A título de ilustração, para a solução final $X_1 \sim X_3 + X_2 X_3 \rightarrow Y$ (com dois termos de solução – causas conjunturais), a cobertura única do primeiro termo é calculada pela diferença entre a cobertura total e cobertura bruta do segundo termo, isto é:

Cobertura total

$$C_{XY}((X_1 \sim X_3 + X_2 X_3)_i \leq Y_i) = \frac{\sum \min((X_1 \sim X_3 + X_2 X_3)_i, Y_i)}{\sum Y_i} \quad (7)$$

Cobertura bruta

$$C_{XY}((X_2 X_3)_i \leq Y_i) = \frac{\sum \min((X_2 X_3)_i, Y_i)}{\sum Y_i} \quad (8)$$

Cobertura única

$$C_{XY}((X_1 \sim X_3)_i \leq Y_i) = C_{XY}((X_1 \sim X_3 + X_2 X_3)_i \leq Y_i) - C_{XY}((X_2 X_3)_i \leq Y_i) \quad (9)$$

Os escores calculados dessas medidas podem revelar que, mesmo que haja muitas combinações causais suficientes que satisfaçam a equifinalidade (diversos caminhos para um único resultado (por exemplo, $X_1 \sim X_3$ e $X_2 X_3$)), poucas delas têm alta cobertura.

Consistência e cobertura das condições necessárias

Os procedimentos para avaliar a consistência e a cobertura de condições suficientes podem ser aplicados em condições necessárias, desde que façamos algumas qualificações. Nós vimos que, se a condição X é suficiente para o resultado Y, então X é um subconjunto de Y. Por outro lado, se X é necessária para Y, então X é um superconjunto de Y. Uma derivação dessa relação espelhada entre a propriedade de suficiência e a propriedade de necessidade é que as fórmulas para calcular os parâmetros de ajuste são intimamente relacionadas. A fórmula para calcular a consistência de suficiência é matematicamente

idêntica à fórmula para calcular a cobertura da necessidade, e a fórmula da cobertura da suficiência é matematicamente idêntica à da consistência da necessidade. Para lembrarmos, a averiguação da propriedade de necessidade – isto é, avaliar se a condição é necessária para o resultado – consiste em selecionarmos os casos em que o resultado está presente e conferirmos se a condição está ou não presente. Portanto, os casos com Y ausente são irrelevantes para avaliar a propriedade de necessidade.

Em vez de partirmos de diagramas de Venn, apresentaremos o procedimento para a quantificação da consistência de condições necessárias partindo de tabelas dois-por-dois. Sejam três condições hipotéticas X4, X5 e X6 que exibem o resultado Y. A Figura 11 ilustra a distribuição dos 210 casos nas tabelas dois-por-dois. Os números nas células continuam indicando o número de casos que mostram os escores de pertencimento na condição e resultado; por exemplo, a célula b referente à condição X5 consiste dos 90 casos dessa condição em que $X5 = 1$ e $Y = 1$. Notemos que os números de casos para $Y = 0$ e $Y = 1$ não mudam, são 110 e 100, respectivamente. O que muda é a distribuição dos casos com $Y = 1$ quanto à condição; isto é a distribuição dos 100 casos com $Y = 1$ entre as células a e b.

Conseguimos, então, não somente averiguar se as condições necessárias são consistentes ou não para o resultado, como calcular a extensão da necessidade. Se uma condição é necessária para o resultado, então nenhum caso pode mostrar o resultado ($Y = 1$) sem a condição. Assim, as células c e d são irrelevantes para a avaliação da necessidade de X para Y. Consequentemente, o grau em que uma condição é consistente com a propriedade de necessidade depende da proporção de casos nas células a e b. Se todos os casos com $Y = 1$ estão localizados na célula b, e a célula a está vazia, a condição é totalmente consistente. Quanto mais casos com resultado presente caem na célula a, menor a consistência da condição que estamos verificando.

Figura 11 – Tabelas dois-por-dois de condições necessárias consistentes e inconsistentes

	1	0	100
	a	b	
Y	0	10	100
	c	d	
		0	1
		X ₁	

	1	10	90
	a	b	
Y	0	10	100
	c	d	
		0	1
		X ₂	

	1	80	20
	a	b	
Y	0	10	100
	c	d	
		0	1
		X ₃	

Fonte: Adaptado de Schneider e Wagemann (2012).

A diferença entre as condições X4, X5 e X6 é que a condição X4 é plenamente consistente para a necessidade de Y, porque todos os elementos de Y em que o resultado está presente (Y = 1) estão na célula b (também têm X4 presente (X4 = 1)). Para as condições X5 e X6 isso não ocorre porque alguns casos com Y = 1 tem X5 = 0 (célula a com 10 casos); idem para X6 (célula a com 80 casos). Portanto, as condições X5 e X6 não são condições necessárias totalmente consistentes para Y. Assim como alertamos para o caso da suficiência, do **ponto de vista puramente teórico**, somente X4 é necessária, pois não existe um único elemento de Y tal que X4 = 0 e Y = 1 (não há casos na célula a). As condições X5 e X6 não são necessárias porque têm casos na célula a de suas tabelas dois-por-dois: 10 e 80, respectivamente. Por outro lado, do **ponto de vista empírico**, as condições X5 e X6 são necessárias, porém em distintos graus.

Ragin (2006) sugere a seguinte fórmula para calcular a consistência de uma condição necessária:

$$\text{Consistência de X como uma condição necessária para Y} = \frac{\text{Número de casos com X=1 e Y=1}}{\text{Número de casos de Y=1}} \quad (10)$$

No numerador somamos todos os casos que são membros de resultado e da condição necessária, e no denominador somamos todos os casos que são membros apenas do resultado. Em termos da tabela dois-por-dois:

$$\text{Consistência de X como uma condição necessária para Y} = \frac{\text{Número de casos da célula b}}{\text{Número de casos das células a e b}} \quad (11)$$

A partir das fórmulas podemos calcular a consistência da necessidade das condições X4, X5 e X6:

$$IN_{X4,Y} = \frac{100}{(0+100)} = 1,00$$

$$IN_{X5,Y} = \frac{90}{(10+90)} = 0,90$$

$$IN_{X6,Y} = \frac{20}{(20+80)} = 0,08$$

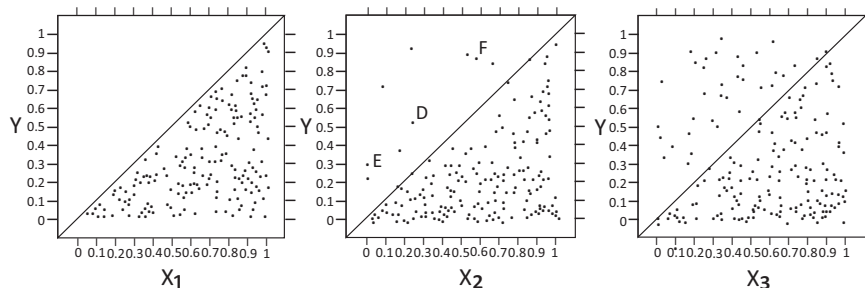
Para conjuntos *fuzzy*, na verificação da consistência da suficiência fizemos uso da informação contida nos escores de pertencimento a conjuntos *fuzzy*. O mesmo procedimento se aplica na verificação da consistência da necessidade de uma condição para um resultado, que é definida pelo grau em que o pertencimento de cada caso em X é igual ou maior do que o pertencimento em Y. Portanto, para o cálculo da consistência da necessidade, relacionamos o pertencimento de cada caso em X que satisfaz a propriedade de necessidade (é consistente com a declaração de necessidade: $X_i \geq Y$) à soma dos escores de pertencimento em Y. De acordo com Ragin (2006), para cada um dos casos, identificamos o valor mínimo entre os escores de pertencimento em X e Y ($s_i = \min(X_i, Y_i)$ é o escore mínimo do i-ésimo caso nos conjuntos X e Y), adicionamos os valores mínimos encontrados ($\sum_{i=1, I}(s_i)$) e, em seguida, dividimos essa soma pela soma dos escores de pertencimento de todos os casos do Y, tal como ($\sum_{i=1, I}(Y)$):

$$IN_{XY}(x_i \geq y) = \frac{\sum \min(x_i, y_i)}{\sum y_i} \quad (12)$$

Em termos diagramáticos, se, para todos os casos, os seus escores em X são iguais ou superiores aos seus correspondentes escores em Y, então todos os casos estão localizados abaixo ou na linha diagonal do diagrama

XY. Desse modo, a medida de consistência da necessidade assume valor 1. Na Figura 12, a condição X4 é uma condição necessária plenamente consistente, porque nenhum caso está localizado acima da linha diagonal. Quanto mais casos exibem um escore de pertencimento em Y maior do que seu escore em X (e quanto maior a magnitude da diferença entre esses escores), mais esses casos se localizarão acima da linha diagonal (e mais distantes estarão da linha diagonal). Nessa situação, o valor da consistência da necessidade de X se desviará do valor 1, uma vez que a soma no numerador será menor do que a soma no denominador (escores menores vão para o numerador ao invés de ir para o denominador).

Figura 12 – Diagramas XY de condições necessárias consistentes e inconsistentes



Fonte: Adaptado de Schneider e Wagemann (2012).

A condição X6 tem mais casos acima da linha diagonal, bem como mais distantes dela, do que a condição X5; por isso são condições inconsistentes, mas em distintos graus. Particularmente, os casos D, E e F são inconsistentes em termo de necessidade da condição X5 para o resultado R, pois os escores na condição são inferiores aos correspondentes escores no resultado Y. Para que esses três casos fossem consistentes para a propriedade de necessidade deveria ocorrer o contrário, seus escores na condição deveriam ser superiores aos correspondentes escores no resultado. Em termos práticos, uma relação de necessidade de uma condição para um resultado é estabelecida se significativamente a maioria dos casos satisfaz a propriedade de necessidade – quando mais de 90% (valor de consistência da necessidade maior ou igual a 0,90) dos escores

de pertencimento da condição são consistentemente maiores ou iguais aos seus respectivos escores de pertencimento no resultado –, então o pesquisador pode declarar que a condição (ou combinação de condições) é “quase sempre” necessária para o resultado (RAGIN, 2006).

Já a cobertura de uma condição necessária (Y subconjunto de X) indica quão pequeno o conjunto resultado Y é em relação ao conjunto X. Aqui, não faremos uso de diagramas de Venn, tabelas dois-por-dois e diagramas XY porque o princípio explicativo para a noção de cobertura de condições necessárias se assemelha aos anteriormente apresentados. Trabalharemos supondo condições e resultado *fuzzy*. O que precisamos lembrar é que estamos interessados agora na relação de Y como subconjunto de X. Se X satisfaz o teste de consistência como condição necessária para Y, então, por definição, X é um superconjunto de Y. Portanto, X cobre plenamente Y. A fórmula de cálculo do nível de cobertura de uma condição necessária é idêntica à fórmula da consistência de uma condição suficiente (RAGIN, 2006):

$$CN_{XY}(x_i \geq y) = \frac{\sum \min(x_i, y_i)}{\sum x_i} \quad (13)$$

De acordo com essa fórmula, se X e Y têm tamanhos aproximadamente iguais, a cobertura de X, como condição necessária para Y, é elevada. Equivalentemente, quanto mais o tamanho de X supera o tamanho de Y, menor a cobertura de X como condição necessária.

Observemos que a fórmula da cobertura da condição necessária (equação 13) é idêntica à fórmula da consistência da condição suficiente (equação 6). Não obstante a equivalência dessas medidas (ou simetria desses parâmetros), Ragin (2008) e Goertz (2006) salientam que a interpretação do valor de cobertura de uma condição necessária é distinta da interpretação do valor de cobertura de uma condição suficiente, porque para o valor de cobertura da condição necessária estão em questão as noções de relevância e trivialidade. Essa é uma questão que está além do escopo deste livro, que procura apresentar os elementos essenciais da abordagem QCA, sem se aprofundar em questões de fronteira.

Sugerimos Schneider e Wagemann (2012) para aqueles que pretendem obter uma visão mais detalhada desse problema envolvendo a cobertura de condições necessárias.

Exemplo numérico para consistência e cobertura

Para ilustrar o cálculo dos parâmetros de consistência e cobertura propomos um exercício hipotético. Suponhamos que temos uma matriz de dados com conjuntos fuzzy A, B e C e resultado R (Tabela 1); porém estendida com os conjuntos complementares ~A, ~B e ~C, as conjunções (interseções) ABC, A~BC, AB e ~A~C, e a disjunção AB + ~A~C (união das conjunções AB e ~A~C).

Tabela 1 – Matriz de dados com conjuntos fuzzy hipotéticos

Casos	Condições			Resultado	Negações			Conjunções				Disjunções	
	A	B	C		R	~A	~B	~C	ABC	A~BC	AB	~A~C	AB + ~A~C
1	0,67	0,6	0,2	0,7	0,33	0,4	0,8	0,2	0,2	0,6	0,33	0,6	
2	0,33	0,9	0,7	0,9	0,67	0,1	0,3	0,33	0,1	0,33	0,3	0,33	
3	1	0,2	0,7	0,4	0	0,8	0,3	0,2	0,7	0,2	0	0,2	
4	0	0,8	0,3	0,3	1	0,2	0,7	0	0	0	0,7	0,7	
5	0,67	0,7	0,1	0,7	0,33	0,3	0,9	0,1	0,1	0,67	0,33	0,67	

Fonte: Elaboração própria.

Recordemos que os escores das conjunções são calculados com a regra do mínimo e os das disjunções, com a regra do máximo. Queremos saber a consistência da suficiência e da necessidade de algumas condições isoladas, conjunções e da disjunção ~A, B, C, ~A, ABC e A~BC. Começamos calculando os **valores da consistência** da suficiência a partir da equação (3):

$$I_{XY}(x_i \leq y) = \frac{\sum \min(x_i, y_i)}{\sum x_i}$$

então:

$$I_{BY} = \frac{\sum \min(B_i, Y_i)}{\sum B_i} = \frac{[\min(0,6;0,7) + \min(0,9;0,9) + \min(0,2;0,4) + \min(0,7;0,7)]}{(0,6 + 0,9 + 0,2 + 0,8 + 0,7)} = \frac{(0,6 + 0,9 + 0,2 + 0,7)}{(0,6 + 0,9 + 0,2 + 0,8 + 0,7)} = 0,75$$

Cabe destacar que o $\min(0,8;0,3)$ não entra no numerador, pois não satisfaz a propriedade de suficiência. Similarmente para as condições $\sim A$, $A\sim BC$, $\sim A\sim C$ e $AB + \sim A\sim C$, têm-se:

$$I_{\sim A, Y} = \frac{(0,33 + 0,67 + 0 + 0,33)}{(0,33 + 0,67 + 0 + 1 + 0,33)} = 0,86$$

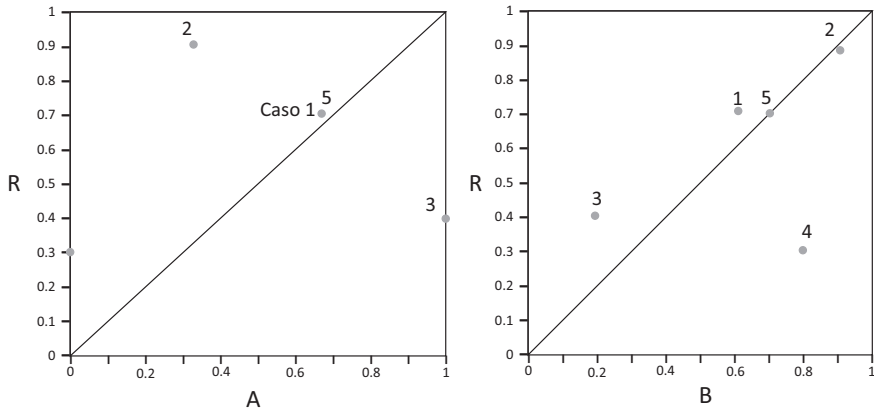
$$I_{\sim A\sim BC, Y} = \frac{(0,2 + 0,1 + 0 + 0,1)}{(0,2 + 0,1 + 0,7 + 0 + 0,1)} = 0,36$$

$$I_{\sim A\sim C, Y} = \frac{(0,33 + 0,3 + 0 + 0,33)}{(0,33 + 0,3 + 0 + 0,7 + 0,33)} = 0,58$$

$$I_{AB + \sim A\sim C, Y} = \frac{(0,6 + 0,33 + 0,2 + 0,67)}{(0,6 + 0,33 + 0,2 + 0,7 + 0,67)} = 0,72$$

A partir desses cálculos vemos, por exemplo, que a negação de A é mais suficiente para o resultado do que a condição B, uma que o seu valor de consistência (0,86) é maior do que o de B (0,75). Procedendo-se semelhantemente aos cálculos acima realizados, as consistências da suficiência das condições A e C são 0,63 e 0,65. Portanto, se considerássemos apenas os conjuntos A, B e C como condições para R, o maior valor de consistência da condição B oferece evidência para interpretarmos B como uma condição suficiente para a presença do resultado. A Figura 13 reporta os diagramas XY para os pares condição-resultado A-R e B-R. A maioria dos casos se localiza acima da linha diagonal. O caso 3 não satisfaz a propriedade de suficiência porque o seu escore em A (1,0) é superior ao seu escore em R (0,4). Portanto, é um caso inconsistente. O mesmo raciocínio vale para o caso 4 para a condição B.

Figura 13 – Diagramas XY para os pares condição-resultado A-R e B-R



Fonte: Elaboração própria.

Já a **consistência da necessidade** dos conjuntos, cuja definição é obtida pela equação (12), $IN_{XY}(x_i \geq y_i) = \frac{\sum \min(x_i, y_i)}{\sum y_i}$, podemos obter para as condições A, B e $AB + \sim A \sim C$:

$$IN_{A,Y} = \frac{(1)}{(0,7 + 0,9 + 0,4 + 0,3 + 0,7)} = 0,33$$

$$IN_{B,Y} = \frac{(0,9 + 0,8 + 0,7)}{(0,7 + 0,9 + 0,4 + 0,3 + 0,7)} = 0,80$$

$$IN_{AB + \sim A \sim C, Y} = \frac{(0,7)}{(0,7 + 0,9 + 0,4 + 0,3 + 0,7)} = 0,23$$

Nesse modo a consistência da necessidade da condição B é 0,80 e a da condição A, 0,33. Portanto, a condição B é relativamente mais necessária para o resultado R do que a condição A. Note-se que os casos estão agrupados próximos à linha diagonal e no canto superior direito (região de alto pertencimento em ambos os conjuntos).

Aplicadas as medidas de consistências, calcularemos, então, os **valores de cobertura**. Enquanto as medidas de consistência da suficiência

e da necessidade indicam o grau em que a evidência empírica está em conformidade com as propriedades de suficiência e de necessidade, respectivamente, a medida de cobertura mensura a extensão do compartilhamento dos casos de uma condição com o resultado. Com base na equação (06), $C_{XY}(x_i \leq y) = \frac{\sum \min(x_i, y_i)}{\sum y_i}$, podemos ter as coberturas para alguns conjuntos da Tabela 1. Como que $C_{XY}(x_i \leq y) = IN_{XY}(x_i \geq y)$, então, por exemplo, $C_{B,Y} = IN_{B,Y} = 0,80$. Por outro lado, sabemos que $CN_{B,Y} = I_{B,Y}$, logo $CN_{B,Y} = 0,75$. Os cálculos das consistências da suficiência e da necessidade e das coberturas de outros conjuntos (conjunções, disjunções e seus complementares) poderiam ser executados, inclusive referentes ao conjunto complemento de R, $\sim R$.

Convém enfatizarmos que os exemplos das conjunções ABC e $A \sim BC$ são úteis porque, quando analisamos as tabelas verdade, como veremos no capítulo 4, as linhas dessas tabelas são exatamente conjunções de condições. Ademais, o exemplo da disjunção $AB + \sim A \sim C$, por outro lado, é útil porque, em geral, a fórmula de solução de uma análise QCA costuma, mas nem sempre, ser expressa em termos de disjunção. Notemos ainda que aqui aplicamos as fórmulas de consistência e de cobertura nos conjuntos de uma matriz de dados, cujo procedimento também é executado no âmbito das tabelas verdade, que é o que realmente importa na abordagem QCA.

Trade off da consistência e cobertura de condições suficientes e necessárias

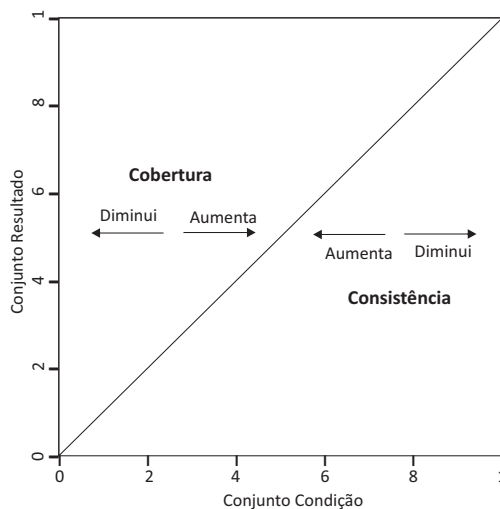
As medidas de consistência mensuram o grau ao qual o pertencimento em cada conjunto é subconjunto do resultado. Já as medidas de cobertura evidenciam a importância empírica de cada conjunto, informando a proporção do pertencimento no resultado que é explicada por uma condição (RAGIN, 2006). A consistência e a cobertura podem ser pensadas como números-sumário que descrevem o padrão de dados da matriz de dados (SCHNEIDER; WAGEMANN, 2012). As fórmulas dessas medidas equivalem entre si, tal como:

$$I_{XY}(x_i \leq y) = CN_{XY}(x_i \geq y) \quad (14)$$

$$C_{XY}(x_i \leq y) = IN_{XY}(x_i \geq y) \quad (15)$$

Essas duas classes de medidas de qualidade das relações entre conjuntos são distintas e às vezes produzem escores que são inversamente relacionados, quando alta consistência produz baixa cobertura. Portanto, pode existir um *trade off* entre as medidas de consistência e cobertura; particularmente quando a base teórica envolve uma causalidade complexa com a interseção de muitas condições causais (pequeno número de casos) para um mesmo resultado. O conjunto interseção pode ser altamente consistente, mas com uma cobertura muito baixa (RAGIN, 2006). A Figura 14 ilustra esse *trade off*.

Figura 14 – Diagrama XY



Fonte: Adaptado de Legewie (2013).

Se todos os casos estiverem localizados acima da linha diagonal, então a propriedade de suficiência é satisfeita e a condição é dita suficiente para o resultado; se estiverem abaixo dessa linha, satisfaz-se a propriedade de necessidade, e a condição é dita necessária para

o resultado. Os casos sobre a linha diagonal satisfazem ambas as propriedades. Portanto, aumenta-se a consistência quando os casos se concentram à esquerda da linha diagonal, porém diminui a cobertura (indicada pela direção das setas). Inversamente, aumenta-se a cobertura quando os casos se concentram à direita da linha diagonal, porém diminui a consistência.

Do ponto de vista prático da pesquisa, o cálculo da consistência deve sempre preceder o cálculo da cobertura. Primeiro, porque não tem sentido interpretar o valor de cobertura de uma condição suficiente (ou necessária) que não é consistente. Segundo, porque agindo assim evitamos confusões na interpretação dos valores obtidos das fórmulas de consistência e cobertura para suficiência e para necessidade de uma condição (SCHNEIDER; WAGEMANN, 2012). Conhecidas a terminologia e as medidas de qualidade das relações entre conjuntos do programa de pesquisa do método comparativo configuracional, é possível discutir a natureza da aplicação desse método. Por isso, no próximo capítulo dar-se-á atenção ao conceito de tabela verdade, álgebra e minimização booleana e a influência das configurações causais sem instâncias empíricas, denominadas remanescentes lógicos.

Os métodos baseados em conjuntos, em essência, tratam do modo de desvendar padrões de relações (subconjunto) entre conjuntos. Veremos no capítulo 4 que o algoritmo tabela verdade (RAGIN, 2008) nos permite descobrir tais padrões. Além disso, as soluções finais derivadas da análise de tabelas verdade têm fórmulas que relacionam (em termos de operadores lógicos) condições interpretadas em termos de propriedades de necessidade e de suficiência, bem como de suas derivações, como as condições INUS e SUIN. Uma condição INUS é uma condição que é parte insuficiente, mas necessária, de uma condição (conjunção) que é ela mesma desnecessária, mas suficiente para um resultado. Já uma condição SUIN significa a parte suficiente, mas não necessária de uma condição que é insuficiente, mas necessária para um resultado. Essas duas formas de condições causais representam formas avançadas de complexidade causal e referem-se a componentes

que não contam como condições necessárias ou suficientes quando tomados isoladamente, mas que desempenham um papel causal crucial em combinações complexas com outras condições. Para ilustrar, no exemplo apresentado anteriormente, temos a condição $AB + \sim A \sim C$, que é uma disjunção (união) de duas conjunções (AB e $\sim A \sim C$) suficientes para o resultado R . Esse é um exemplo de condição INUS, pois a condição isolada A é parte insuficiente, mas necessária da conjunção AB que, por sua vez, é não necessária, mas suficiente para o resultado. Ou seja, a condição A é **necessária** para o resultado R se a ocorrência de R não é possível sem a presença de A , mas A por si só não é **suficiente** para produzir R . Semelhantemente, isso vale para a condição isolada B em AB e as condições isoladas $\sim A$ e $\sim C$ em $\sim A \sim C$.

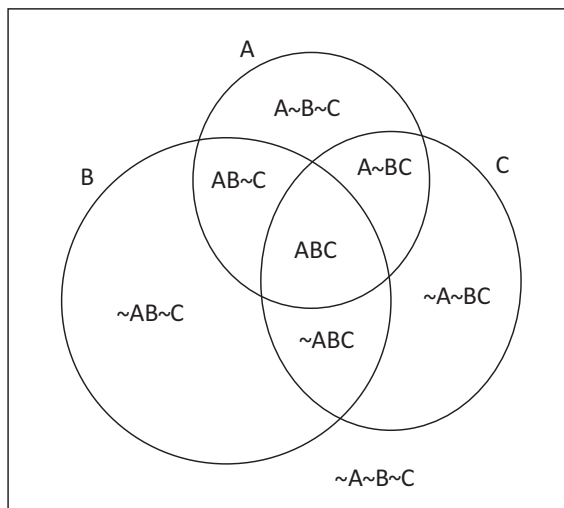
TABELAS VERDADE

A tabela verdade constitui o núcleo de qualquer técnica de QCA. À primeira vista, ela se parece com a matriz dos dados em estudo, porque, assim como na matriz, cada coluna da tabela verdade denota uma variável, ou melhor, um conjunto. A diferença específica está no significado das linhas; enquanto, em uma matriz de dados, cada linha representa um caso diferente (as linhas são as unidades de observação); na tabela verdade cada linha representa uma combinação (ou configuração) de condições logicamente possíveis (isto é, conjunções E-lógico). Uma vez que cada condição isolada pode denotar dois estados (presença ou ausência de um determinado conceito), o número total de linhas da tabela verdade é 2^k combinações de condições; em que a letra k indica o número de condições em pesquisa e o número 2 indica os dois estados distintos (presença ou ausência) pelas quais cada uma das condições pode ocorrer. Notemos que cada linha da tabela verdade descreve uma combinação qualitativamente distinta de condições das demais linhas. Portanto, a diferença entre os casos em diferentes linhas é uma diferença de tipo, em vez de ser uma diferença de grau.

Uma vez que o número de linhas reflete o número de combinações logicamente possíveis de condições, quanto maior o número de condições, maior será o número de linhas. Ademais, o número de linhas da tabela verdade aumenta exponencialmente conforme o número de condições: para $k = 3$, há oito configurações possíveis; $k = 4$, 16 configurações; $k = 5$, 32 combinações, e assim sucessivamente. Entretanto, na prática das pesquisas aplicadas de QCA, nem todas as configurações logicamente possíveis se materializam empiricamente. Isto é, não necessariamente existem instâncias empíricas (casos concretos) para cada uma das combinações. Quando isso ocorre, estamos diante do fenômeno da diversidade limitada: quando a diversidade das combinações logicamente ou teoricamente possíveis de condições é limitada na realidade a um determinado número de casos reais. Agora concentraremos atenção na análise de uma tabela verdade em que todas as combinações possíveis têm instâncias empíricas.

As conjunções E-lógico (ou combinações de condições) poderiam ser representadas também por diagramas de Venn. Para três condições, A, B e C, haveria oito combinações possíveis (oito áreas distintas), e o diagrama de Venn se pareceria com a Figura 15, a seguir. Cada área no diagrama de Venn corresponde a uma linha da tabela verdade, e cada área descreve uma expressão booleana. Por exemplo, a área no topo do diagrama, que é o conjunto $A\sim B\sim C$, compreende os casos em que a condição A está presente ($A = 1$) e as condições B e C estão ausentes ($B = 0, C = 0$), e assim sucessivamente. Embora bastante sugestiva e intuitiva, a representação de todas as combinações de condições torna-se difícil de desenhar e interpretar quando o número de condições é superior a três condições. Além disso, o diagrama de Venn, como o mostrado na Figura 15, exibe apenas os conjuntos A, B e C e suas interseções, mas nada indica sobre as relações de subconjuntos de suficiência e necessidade entre condições e resultado. Claro, os diagramas de Venn podem fazer ambos simultaneamente, ou seja, mostrar a relação de subconjunto de uma interseção de condições e resultado.

Figura 15 – Diagrama de Venn com todas as combinações de três condições (A, B e C)



Fonte: Adaptado de Schneider e Wagemann (2012).

Independentemente do tipo de conjuntos, dicotômicos ou *fuzzy*, a construção da tabela verdade para os conjuntos condições e resultado e para o conjunto de casos selecionados pelo pesquisador segue um mesmo procedimento. Embora a maioria dos programas computacionais para operar as técnicas QCA seja capaz de produzir tabelas verdade a partir de uma matriz de dados que representa escores de pertencimento, é muito útil detalhar os três passos necessários para construí-la. Primeiro, listamos todas as 2^k combinações logicamente possíveis das k condições, deixando vazias as células da coluna dos escores do resultado. Segundo, atribuímos cada caso de nossa matriz de dados para a linha da tabela verdade que corresponde a seus escores nas k condições. Notemos que cada caso só pode pertencer a uma única linha da tabela verdade. Contudo, as linhas individuais dessa tabela podem conter mais do que um caso (até mesmo nenhum caso, conforme falamos do fenômeno da diversidade limitada. Por ora, não consideramos essa possibilidade). Terceiro, devemos atribuir um escore de resultado a todas as linhas da tabela verdade. Isso é determinado pelos escores do resultado das instâncias empíricas que se enquadram na respectiva linha.

É importante compreender a informação contida na coluna Resultado. Do ponto de vista do caso, o escore 1 nessa coluna indica que os casos com determinadas características também mostram o resultado de interesse ($R = 1$). Por outro lado, sob a perspectiva de configurações (conjunções E-lógico), podemos dizer que uma conjunção é suficiente para o resultado presente ($R = 1$). Uma linha da tabela verdade com escore 1 está explicitamente ligada a esse resultado (RIHOUX; RAGIN, 2009). Em essência, cada linha de tabela verdade é uma declaração de suficiência (RAGIN, 2008).

Tabela verdade em csQCA

Suponhamos um conjunto de doze casos selecionados, três condições e um resultado, a matriz (hipotética) de dados dicotômicos

(Tabela 2). Por exemplo, o oitavo caso tem as condições A e C presentes e a condição B ausente, configuração (A = 1, B = 0, C = 1), e resultado presente (R = 1). O décimo primeiro caso, por outro lado, tem todas as condições e resultado ausentes (A = 0, B = 0, C = 0 e Y = 0). E assim, sucessivamente.

Tabela 2 – Matriz de dados dicotômicos

Casos	Condições			Resultado
	A	B	C	R
1	1	1	1	1
2	0	0	1	0
3	1	0	0	1
4	1	0	0	1
5	1	1	0	0
6	0	0	1	0
7	1	1	1	1
8	1	0	1	1
9	1	0	0	1
10	0	1	0	0
11	0	0	0	0
12	0	1	1	1

Fonte: Adaptado de Schneider e Wagemann (2012).

Agora é possível construir a tabela verdade para essa matriz de dados, com base nos três passos supracitados. Com $k = 3$, há oito configurações logicamente possíveis (tabela verdade com oito linhas); alocamos os casos para a linha conforme sua configuração (lembrando que cada caso só pode estar em uma única linha da tabela verdade, mas cada linha dessa tabela pode conter vários casos, por exemplo, a primeira linha tem dois casos (1 e 7) e a quinta linha tem três casos (3, 4 e 9). Posteriormente, atribuímos um escore de resultado a todas as linhas da tabela verdade (escores determinados pelo escore de resultado das instâncias empíricas que se enquadram na respectiva linha – poderíamos estar diante de uma situação em que dois casos poderiam ter a mesma configuração mas resultados distintos. Isso caracteriza uma **contradição**. Todavia, isso não ocorre neste exemplo hipotético) (Tabela 3).

Tabela 3 – Tabela verdade a partir da matriz de dados dicotômicos

Combinações	Condições			Resultado	Casos
	A	B	C	R	
ABC	1	1	1	1	caso 1, 7
ABc	1	1	0	0	caso 5
AbC	1	0	1	1	caso 8
aBC	0	1	1	1	caso 12
Abc	1	0	0	1	casos 3, 4, 9
aBc	0	1	0	0	caso 10
abC	0	0	1	0	casos 2, 6
abc	0	0	0	0	caso 11

Fonte: Adaptado de Schneider e Wagemann (2012).

Tabela verdade em fsQCA

Com conjunto *fuzzy* também se aplicam os três passos para converter uma matriz de dados em uma tabela verdade, conforme feito para conjuntos dicotômicos. Primeiro criamos a tabela verdade, depois alocamos cada caso em cada uma das linhas da tabela e, em seguida, determinamos o escore do resultado para cada linha, conforme os escores de resultado dos casos associados a cada linha. Suponhamos agora um conjunto de doze casos selecionados, três condições e um resultado, a matriz (hipotética) de dados *fuzzy*¹⁵ (Tabela 4).

Uma pergunta que surgirá é: já que os conjuntos fuzzy permitem qualquer escore de pertencimento entre os escores extremos 0 e 1, como se parece a tabela verdade para tais conjuntos? Assim como se dá para conjuntos dicotômicos, o número de linhas de tabela verdade para conjuntos fuzzy continua sendo determinado pela fórmula 2^k combinações de condições. Isso ocorre porque, exatamente como em conjuntos dicotômicos, os conjuntos fuzzy também estabelecem uma diferença qualitativa entre casos acima da âncora qualitativa 0,5 (mais dentro do que fora do conjunto) e casos abaixo dessa âncora (mais fora do que dentro do conjunto).

¹⁵ Vamos momentaneamente supor que o pesquisador já fez a calibragem de cada um dos quatro conjuntos *fuzzy* (três condições e um resultado).

Tabela 4 – Matriz de dados fuzzy

Casos	Condições			Resultado
	A	B	C	R
1	1	0,8	0,1	0,8
2	0,67	0,6	0,2	0,7
3	0,33	0,7	0,3	0,4
4	0,67	0,3	0,9	0,3
5	0,33	0,9	0,9	0,4
6	1	0,2	0,7	0,4
7	0	0,8	0,3	0,3
8	0,33	0,4	0,8	0,6
9	0,67	0,1	0,4	0,7
10	0	0,1	0,2	0,3
11	0,67	0,9	0,7	0,9
12	0,67	0,7	0,1	0,7

Fonte: Adaptado de Schneider e Wagemann (2012).

Por conta disso as k condições *fuzzy* produzem 2^k linhas na tabela de verdade. Embora siga os mesmos três passos, aqui se requer uma explicação mais detalhada, devido às nuances próprias de um conjunto *fuzzy*. Com conjuntos dicotômicos, a identificação da linha da tabela verdade para a qual um determinado caso deve ser alocado requer encontrar a correspondência exata entre os escores dicotômicos de pertencimento dos casos e as linhas da tabela verdade. Por exemplo, o terceiro caso da matriz de dados (Tabela 2) tem configuração (A = 1, B = 0, C = 0) e é alocado para a quinta linha da tabela verdade de dados dicotômicos (Tabela 3). Para conjuntos *fuzzy*, porém, os casos com escores *fuzzy* de pertencimento nas condições k **não correspondem** exatamente a qualquer uma das linhas da tabela verdade. Por exemplo, em que linha da tabela verdade (Tabela 4) o caso 4 da matriz de dados *fuzzy* (Tabela 5) deveria ser alocado, uma vez que seus escores de pertencimento nos conjuntos *fuzzy* são A = 0,67, B = 0,3 e C = 0,9?

Para facilitar o modo de responder esse tipo de questão, Ragin (2008, cap. 7) sugeriu o conceito de espaço de propriedade (*property space*), em que cada conjunto constitui uma dimensão do espaço de propriedade.

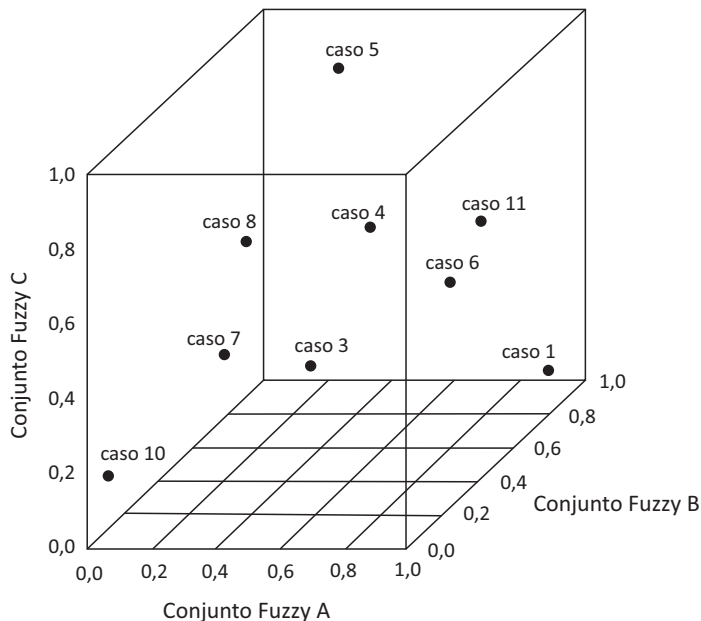
Tabela 5 – Tabela verdade a partir da matriz de dados *fuzzy*

Combinações	Condições			Resultado	Casos
	A	B	C		
ABC	1	1	1	1	caso 11
ABc	1	1	0	1	casos 1, 2, 12
AbC	1	0	1	0	casos 4, 6
aBC	0	1	1	0	caso 5
Abc	1	0	0	1	caso 9
aBc	0	1	0	0	casos 3, 7
abC	0	0	1	1	caso 8
abc	0	0	0	0	caso 10

Fonte: Adaptado de Schneider e Wagemann (2012).

Em nosso exemplo, as três condições *fuzzy* A, B e C, produzem um espaço tridimensional como exibido na Figura 16 (as configurações *fuzzy* de alguns casos não estão nessa figura ilustrativa).

Figura 16 – Espaço de propriedade tridimensional dos conjuntos *fuzzy* A, B e C.



Fonte: Adaptado de Schneider e Wagemann (2012).

O espaço de propriedade apresenta algumas características importantes:

- a) independentemente dos valores dos escores *fuzzy* de um caso qualquer nas condições A, B e C, a localização do caso será, por definição, dentro do espaço da propriedade, pois os escores e as dimensões do espaço da propriedade têm um valor mínimo 0 e máximo de 1;
- b) por conta dos escores de pertencimento em A, B e C, cada caso tem uma localização precisa dentro desse espaço (cubo);
- c) cada canto do espaço corresponde a uma combinação específica dos dois valores extremos dos conjuntos *fuzzy* (pertencimento pleno ou adesão total ao conjunto (escore 1) e não-pertencimento pleno ao conjunto (escore 0)). Por exemplo, o canto na parte inferior esquerda da Figura 16 indica a configuração em que os três conjuntos *fuzzy* assumem o escore zero e pode ser rotulado como $\sim A \sim B \sim C$ ou a^*b^*c e escores “0, 0, 0”. Por extensão, o canto superior direito indica os escores “1,0,1” rotulado por $A \sim BC$ ou A^*b^*C , e assim sucessivamente;
- d) uma vez que os cantos do espaço de propriedade indicam as combinações específicas com escores extremos de pertencimento nas condições *fuzzy*, podemos tomá-los como situações ideais-típicas. Assim, os casos que caem exatamente em um desses cantos são instâncias empíricas do tipo-ideal (combinação ideal-tipo). Observemos que, a menos que um caso tenha escores de (não) pertencimento pleno em todas as condições que constituem o espaço da propriedade (que o caracterizaria como um caso dicotômico), ele não estará localizado em um dos cantos. Portanto, na maioria das aplicações fsQCA, muitos, se não todos, casos se aproximam dos tipos-ideais apenas em algum grau. É possível calcular a distância das configurações dos casos empíricos em relação aos tipos-ideais.

- e) um espaço de propriedade com três dimensões tem oito cantos, assim como a tabela verdade baseada em três condições tem oito linhas. Essa propriedade segue diretamente do fato de que os cantos de um espaço de propriedade de conjuntos *fuzzy* equivalem às linhas de uma tabela verdade; porque os cantos desse espaço denotam as configurações em que as condições *fuzzy* assumem escores extremos 0 ou 1. Ou seja, os cantos são onde os lugares em que os conjuntos *fuzzy* mostram escores de pertencimento de conjuntos dicotômicos. Por isso que se diz que o csQCA pode ser tomado como um caso particular do fsQCA.

Resumindo o que foi dito até aqui, com as k condições *fuzzy*, criamos um espaço de propriedade com 2^k cantos que correspondem às 2^k combinações ideais, às 2^k linhas da tabela verdade e às 2^k conjunções E-lógicas entre as k condições. Convém lembrar que a Figura 16 é apenas ilustrativa e para fins didáticos, pois nas aplicações de fsQCA, que muitas vezes envolvem mais do que três condições, o espaço de propriedade não constitui uma das ferramentas das técnicas QCA para simplificar a complexidade dos casos, mas sim a tabela verdade. Tendo em vista que os escores *fuzzy* são valores entre os valores extremos 1 e 0, e as instâncias empíricas, a partir de suas configurações complexas, podem estar localizadas em qualquer lugar do espaço de propriedade, conforme o exemplo da Figura 16, surgem duas questões: A qual canto um determinado caso mais provavelmente pertence? E qual a extensão desse caso em relação ao tipo ideal (isto é, à linha da tabela verdade)?

Para responder a essas perguntas, tomemos o caso 11, cuja configuração é $A = 0,67$, $B = 0,9$ e $C = 0,7$. Todos os escores são superiores à âncora qualitativa 0,5. Este não é o caso ideal “1, 1, 1” ($A*B*C$; canto superior direito; primeira linha da tabela verdade), mas as magnitudes dos seus escores *fuzzy* o aproximam dessa configuração ideal. Por outro lado, o caso 10 se aproxima da configuração ideal “0, 0, 0” ($\sim A \sim B \sim C$; canto inferior esquerdo; última linha da tabela verdade), porque sua configuração empírica é $A = 0$, $B = 0,1$ e $C = 0,2$.

Com base nisso, existe alguma maneira de definir (alocar) precisamente os casos nas linhas da tabela verdade? Sim, porque é fácil calcular o escore da sua conjunção. Uma vez que cada um dos 2^k cantos corresponde a uma das 2^k combinações logicamente possíveis e que o escore de pertencimento de uma interseção (conjunção) é calculado pela regra do mínimo (escore mínimo das condições isoladas referentes ao i -ésimo caso, isto é $\min(c_i) = \{A_i, B_i, C_i\}$), então é possível determinarmos o escore de pertencimento dos casos em todas as combinações logicamente possíveis de condições (nos cantos do espaço de propriedade). Portanto, para os exemplos acima: o caso 10 tem escore *fuzzy* de pertencimento de 0,8 no tipo-ideal $\sim A \sim B \sim C$, porque o mínimo entre 1,0, 0,9 e 0,8 é 0,8; o caso 11 tem escore *fuzzy* de 0,67 no tipo-ideal $A * B * C$ ($\min(c_{11}) = \{0,67; 0,9; 0,7\} = 0,67$).

Observemos que a tabela verdade (Tabela 5) não apresenta os escores das negações dos conjuntos *fuzzy*, por isso para o caso 10 eles precisam ser calculados antes de calcularmos o escore mínimo da conjunção empírica desse caso. Procedimento semelhante deve ser realizado para o cálculo dos escores dos casos nos tipos-ideais. Notemos também que os dois casos não são exemplos plenos de seus respectivos tipos-ideais, porque os escores da conjunção são inferiores a 1.

Os casos não têm apenas seus escores de pertencimento quanto ao seu próprio tipo-ideal, mas também em relação aos demais tipos-ideais (Tabela 6). Contudo, esses escores são baixos, em comparação ao escore de seu tipo ideal, uma vez que foram calculados com a regra do mínimo. Um aspecto relevante a enfatizar é que os casos têm escores de pertencimento parcial em todas as linhas da tabela verdade, porém existe uma única linha em que seu escore de pertencimento é superior à âncora qualitativa 0,5. Portanto, não interessa quantos conjuntos *fuzzy* são combinados, cada caso terá um escore de pertencimento superior a 0,5 em uma e somente uma das 2^k combinações logicamente possíveis.

Tabela 6 – Escores de pertencimento dos casos 10 e 11 nos tipos-ideais

Casos	Condições			Negações			Conjunções tipos-ideais							
	A	B	C	$\sim A$	$\sim B$	$\sim C$	ABC	$AB \sim C$	$A \sim B C$	$A \sim B \sim C$	$\sim A B C$	$\sim A \sim B C$	$\sim A \sim B \sim C$	
caso 10	0	0,1	0,2	1	0,9	0,8	0	0	0	0	0,1	0,1	0,2	0,8
caso 11	0,67	0,9	0,7	0,33	0,1	0,3	0,67	0,3	0,1	0,1	0,33	0,3	0,1	0,1

Fonte: Elaboração própria.

Essa propriedade dos conjuntos *fuzzy* é importante para a nossa tarefa de identificação da linha da tabela verdade para a qual um caso melhor se ajusta, que acaba sendo a linha em que seu escore de pertencimento parcial é superior a 0,5. Falta calcular o escore do resultado de cada uma das 2^k linhas da tabela verdade. Uma vez que cada linha da tabela verdade é uma declaração de suficiência, isso significa que cada linha da tabela deveria ser considerada uma conjunção suficiente para o resultado se o escore de pertencimento de cada caso na linha é menor ou igual ao seu pertencimento no conjunto resultado. A Tabela 7 lista os escores *fuzzy* de pertencimento dos doze casos hipotéticos, as oito linhas da tabela verdade (da oitava à décima quinta coluna) e a coluna de resultado.

Tabela 7 – Escores de pertencimento *fuzzy* nas linhas e resultado

Casos	Condições			Negações			Conjunções								Resultado
	A	B	C	$\sim A$	$\sim B$	$\sim C$	ABC	AB $\sim C$	A $\sim B$ C	A $\sim B$ $\sim C$	$\sim A$ BC	$\sim A$ B $\sim C$	$\sim A$ $\sim B$ C	$\sim A$ $\sim B$ $\sim C$	R
1	1	0,8	0,1	0	0,2	0,9	0,1	0,8	0,1	0,2	0	0	0	0	0,8
2	0,67	0,6	0,2	0,33	0,4	0,8	0,2	0,6	0,2	0,4	0,2	0,33	0,2	0,33	0,7
3	0,33	0,7	0,3	0,67	0,3	0,7	0,3	0,33	0,3	0,3	0,3	0,67	0,3	0,3	0,4
4	0,67	0,3	0,9	0,33	0,7	0,1	0,3	0,1	0,67	0,1	0,3	0,1	0,33	0,1	0,3
5	0,33	0,9	0,9	0,67	0,1	0,1	0,33	0,1	0,1	0,1	0,67	0,1	0,1	0,1	0,4
6	1	0,2	0,7	0	0,8	0,3	0,2	0,2	0,7	0,3	0	0	0	0	0,4
7	0	0,8	0,3	1	0,2	0,7	0	0	0	0	0,3	0,7	0,2	0,2	0,3
8	0,33	0,4	0,8	0,67	0,6	0,2	0,33	0,2	0,33	0,2	0,4	0,2	0,6	0,3	0,6
9	0,67	0,1	0,4	0,33	0,9	0,6	0,1	0,1	0,4	0,6	0,1	0,1	0,33	0,33	0,7
10	0	0,1	0,2	1	0,9	0,8	0	0	0	0	0,1	0,1	0,2	0,8	0,3
11	0,67	0,9	0,7	0,33	0,1	0,3	0,67	0,3	0,1	0,1	0,33	0,3	0,1	0,1	0,9
12	0,67	0,7	0,1	0,33	0,3	0,9	0,1	0,67	0,1	0,3	0,1	0,33	0,1	0,3	0,7

Score na linha \leq que escore em R 1 1 0 1 0 0 0 0

Fonte: Elaboração própria.

Agora é possível averiguarmos a propriedade de suficiência de cada linha da tabela verdade: comparamos o escore da linha ao escore do resultado na linha correspondente; se o escore da linha for menor ou igual ao escore do resultado, então a linha da tabela verdade satisfaz o critério de suficiência. É preciso que todos os escores de cada linha da tabela verdade sejam menores ou iguais aos escores do resultado; caso contrário, o critério de suficiência não é satisfeito. Por exemplo, as linhas

ABC, $AB\sim C$ e $A\sim B\sim C$ (oitava, nona e décima primeira colunas da Tabela 7) satisfazem esse critério e são subconjuntos perfeitos do resultado. As demais linhas não são suficientes para o resultado. Por isso, na última linha da Tabela 7, as colunas ABC, $AB\sim C$ e $A\sim B\sim C$ receberam valor 1, para indicar que são suficientes para o resultado, e as demais colunas, valor zero, porque não são suficientes para o resultado.

As informações da Tabela 7 nos permitem representar a matriz de dados dos conjuntos *fuzzy* no formato de uma tabela verdade para conjuntos dicotômicos: para cada uma das 2^k linhas, sabemos quais delas são subconjuntos do resultado e conhecemos também os casos correspondentes a cada linha. Assim, a Tabela 8 apresenta a tabela verdade resultante dos dados *fuzzy* da Tabela 7.

Tabela 8 – Tabela verdade derivada dos dados *fuzzy* hipotéticos

Condições			Suficiente para	Casos com
A	B	C	R	escores $\geq 0,5$ na linha *
1	1	1	1	caso 11 (0,67)
1	1	0	1	casos 1 (0,8), 2 (0,6), 12 (0,67)
1	0	1	0	casos 4 (0,67), 6 (0,7)
0	1	1	0	caso 5 (0,67)
1	0	0	1	caso 9 (0,6)
0	1	0	0	casos 3 (0,67), 7 (0,7)
0	0	1	0	caso 8 (0,6)
0	0	0	0	caso 10 (0,8)

* Números em parêntesis = escore fuzzy de pertencimento na linha (conjunção)

Fonte: Elaboração própria.

Tabelas verdade com “ruídos”

Vimos que as tabelas verdade são o núcleo da abordagem QCA porque elas resumiam as informações contidas nas matrizes de dados e permitem que façamos a averiguação da consistência e da cobertura das condições suficientes e que apliquemos a minimização booleana de modo a reduzir a complexidade original dos casos quanto às suas configurações de condições (ou interseções fundamentais). Além disso, vimos também

que, mesmo tendo uma matriz de dados *fuzzy*, ainda assim podemos convertê-la em uma tabela verdade dicotômica e proceder à análise de consistência e cobertura da suficiência conforme executada para uma tabela verdade derivada de uma matriz de dados dicotômicos.

As tabelas verdade examinadas e reduzidas, pelo algoritmo de Quine-McCluskey, não apresentaram problemas em suas linhas. Ou melhor, as conjunções, que constituem todas as combinações de condições logicamente possíveis, tinham instâncias empíricas para o conjunto resultado e/ou nenhuma conjunção cujos casos associados se subdividiam em dois grupos: um grupo de casos com presença do resultado e outro grupo, com ausência do resultado. Todavia, na prática de pesquisas surgem situações em que nem todas as linhas de uma tabela verdade possuem instâncias empíricas – quando isso ocorre estaremos diante da chamada diversidade limitada e tais linhas são chamadas de remanescentes lógicos – ou podem existir tais instâncias, mas com resultado contraditório. A Tabela 9 exibe uma tabela verdade em que esses dois fenômenos estão presentes.

Tabela 9 – Tabela verdade com diversidade limitada e linhas contraditórias

Combinações	Condições			Resultado	Casos
	A	B	C	R	
ABC	1	1	1	1	caso 11
ABc	1	1	0	1	casos 1, 2, 12
AbC	1	0	1	0	casos 4, 6
aBC *	0	1	1	?	caso 5
Abc	1	0	0	1	caso 9
aBc **	0	1	0	0	casos 3 (1), 7 (0)
abC	0	0	1	1	caso 8
abc *	0	0	0	?	caso 10

Nota: * linha remanescente lógico; ** linha contraditória.

Fonte: Elaboração própria.

Isso não quer dizer que na prática da abordagem QCA eles necessariamente ocorrem ou ocorrerão. Esta tabela nada mais é do que a Tabela 5, em que as linhas 4 (conjunção aBC) e 8 (conjunção abc) são remanescentes lógicos, simbolizados pela interrogação (?), porque não há

instâncias empíricas para o resultado (não sabemos se elas têm presença ou ausência do resultado),; e a linha 6 (conjunção aBc), uma contradição lógica, porque enquanto o caso 3 tem o resultado presente, o caso 7 tem ausência do resultado. Examinamos separadamente os fenômenos da diversidade limitada e das linhas contraditórias, suas naturezas, suas fontes e soluções possíveis para esses problemas.

Diversidade limitada e remanescentes lógicos

Vimos que as linhas da tabela verdade podem ou não ser suficientes para um resultado. Utilizamos os escores de pertencimento dos casos em cada condição e comparamos com os escores de pertencimento do resultado a fim de verificarmos a consistência da suficiência de cada uma das condições ou de combinações de condições. Se não houver informação em algumas linhas da tabela verdade para o resultado, como avaliaremos a consistência dessas linhas? Isto é, se uma linha consistir numa conjunção de condições que é logicamente possível, mas que não é empiricamente observável?

Quando isso ocorre, estamos diante de uma situação em que algumas, ou mesmo muitas, linhas da tabela verdade são chamadas de **remanescentes lógicos**. Isto é, estamos diante da chamada **diversidade limitada**, porque a diversidade de combinações de condições logicamente (teoricamente) possíveis para um resultado (presença ou ausência), que poderiam ter instâncias empíricas que as representassem, é limitada, já que apenas algumas delas exibem tais instâncias. Na presença de remanescentes lógicos, a avaliação da consistência da suficiência é inviabilizada; conseqüentemente não é possível selecionar as linhas suficientes para aplicar a minimização booleana e, portanto, reduzir a complexidade causal para um determinado resultado.

Em suma, embora se possam listar todas as combinações possíveis de condições, existem fenômenos sociais com diversidade limitada, no sentido de que não apresentam instâncias empíricas para todas as configurações causais que implicam um determinado resultado (RAGIN; SONNETT, 2005). A presença de remanescentes lógicos compromete a

solução final de uma aplicação da QCA. Por isso, precisamos de alguma forma de abordar esse problema de modo que a aplicação seja bem-sucedida. Antes de tratar das soluções possíveis para o problema da diversidade limitada, convém discutirmos as fontes desse problema. A diversidade limitada surge a partir das seguintes situações: (a) quando o número de combinações logicamente possíveis (quantidade de linhas da tabela verdade) é superior ao número de casos sob exame; (b) quando o tipo de caso não existe na realidade como a conhecemos, porque essa realidade é estruturada por processos históricos, sociais, culturais etc; (c) quando uma configuração particular é impossível à luz do que conhecemos sobre o mundo (*e.g.*, um país tropical na Antártida) (SCHNEIDER; WAGEMANN, 2012).

Antes de tratarmos do efeito dos remanescentes lógicos na solução final da análise de uma tabela verdade, vamos analisar um exemplo hipotético em que inexistem tais remanescentes. Suponhamos que quatro fenômenos sejam descritos por conjuntos representados pelas letras A, B, C e R. A título de simplificação, para todos esses conjuntos, a presença do fenômeno é simbolizada pela letra maiúscula e a ausência, pela letra minúscula, em vez de utilizarmos o sinal \sim . Assim teremos as seguintes possibilidades quanto à natureza qualitativa de cada fenômeno: A, a, B, b, C, c, R e r; em que A significa “presença do fenômeno A”, a significa “ausência do fenômeno A” e assim por diante.

Sendo conjuntos, podemos, mediante os operadores lógicos, produzir outros conjuntos denominados conjunções, que são interseções de conjuntos, e disjunções, que são uniões de conjuntos isolados e/ou de conjunções de conjuntos. Portanto, os conjuntos A, B e C permitem estabelecer oito combinações logicamente possíveis: ABC, ABc, AbC, aBC, Abc, aBc, abC e abc. Essas interseções, simplificada e representadas sem o sinal \cap , também são conhecidas como “interseções fundamentais” (THIEM; DUSA, 2012a). Poderíamos ampliar o conjunto das combinações logicamente possíveis ao introduzir o conjunto R, mas ele tem um significado específico, conforme veremos a seguir. Suponhamos também que uma determinada teoria estabeleça que os fenômenos A, B e C são

condições causais para o fenômeno R (denominado resultado). Assim sendo, há, no máximo, dezesseis implicações teóricas possíveis: oito para R (presença do fenômeno R) e oito para r (ausência do fenômeno R). Contudo, implicações teóricas não necessariamente terão instâncias empíricas que as suportem.

A título de ilustração, vamos supor que as combinações ABC, ABc, aBC e aBc estejam associadas à presença de R, que é o resultado de interesse. Se as instâncias empíricas evidenciam que essas conjunções são consistentemente suficientes para o resultado R (isto é, cada uma delas é um subconjunto consistente do resultado R), então a presença simultânea das condições “A e B e C” ou “A e B e c” ou “a e B e C” ou “a e B e c” implicam (são subconjuntos de) “Y”. A equação booleana é $ABC + ABc + aBC + aBc \rightarrow R$.

A regra de minimização booleana estabelece que, se duas interseções fundamentais para um mesmo resultado diferem na valência de uma única condição, então essa condição pode ser eliminada de modo a resultar um termo mais simples (THIEM; DUSA, 2012a). Por exemplo, as interseções fundamentais aBC e aBc diferem quanto à valência do fenômeno C. Portanto, ele é redundante para o resultado R, pois sua presença, C, ou ausência, c, não influencia a valência do fenômeno R, que continua presente. Isto é, a mudança qualitativa na condição C não está associada à mudança qualitativa no resultado. Essas interseções podem ser simplificadas para: AB implica Y ($AB \subseteq Y$) ou $AB \rightarrow Y$.

Essa regra de simplificação pode ser executada para outros pares de interseções fundamentais. A Figura 17 apresenta o processo de redução da complexidade das interseções fundamentais em duas etapas: na primeira etapa (figura à esquerda) produzimos as combinações AB, BC, Bc e aB; e na segunda etapa (figura à direita), quando as combinações simplificadas resultantes da primeira etapa são comparadas para identificar aquelas que satisfazem a regra de minimização, gerando, neste caso, a combinação B. Essa **solução final**, em que a condição B (presença) isoladamente implica R (presença) é denominada o *prime implicant* de R, e a sua **fórmula de solução** é $B \rightarrow R$.

Figura 17 – Etapas da minimização booleana

	ABC	ABc	aBC	aBc			AB	BC	Bc	aB	
ABC	-	AB	BC	-		AB	-	-	-	B	
ABc	AB	-	-	Bc	\Rightarrow	BC	-	-	B	-	\Rightarrow B \rightarrow R
aBC	BC	-	-	aB		Bc	-	B	-	-	
aBc	-	Bc	aB	-		aB	B	-	-	-	

Fonte: Elaboração própria.

Nem sempre as soluções finais de aplicações da abordagem QCA têm fórmulas de solução contendo condições causais isoladas, como no exemplo descrito na Figura 17. Elas podem conter uma ou várias combinações de condições causais. Nesse caso, estamos diante de uma causalção conjuntural múltipla para a presença do resultado. Poderíamos ter como fórmulas de solução “ $aB + Ac \rightarrow R$ ” ou “ $AB + aC \rightarrow$ ” ou “ $aB \rightarrow R$ ” ou outra conjunção e/ou disjunção de condições para R. Por exemplo, se a solução final da minimização de uma tabela verdade gerasse a fórmula de solução “ $aB + Ac \rightarrow R$ ”, então “ausência do fenômeno A e presença do B” OU “presença de A e ausência de C” é consistente para a “presença do fenômeno R”. Portanto, os dois caminhos causais sugerem que duas condições alternativas suficientes, aB e Ac, existem e, mutuamente exclusivas, acarretam o resultado R.

Esse exemplo hipotético das relações entre condições e resultado compreende as duas dimensões da causalidade complexa, conforme definida no capítulo 2: a **equifinalidade**, porque a conjunção aB ou a conjunção Ac conduzem a um mesmo resultado, R; e a **causalidade conjuntural**, porque as condições a e B combinadas, mas não isoladas, levam a R, bem como condições A e c combinadas. Se a solução final é tal que sua fórmula de solução é, por exemplo, “ $Bc + ac \rightarrow r$ ”, então, além da equifinalidade e causalidade conjuntural, teríamos também a dimensão **causalidade assimétrica**, quando a análise empírica pode revelar não apenas associação de combinações de condições para a presença do resultado, R, mas a análise de outra tabela verdade, porém para a ausência do resultado, r.

As combinações de condições aB e Ac, nesse exemplo, são denominadas condições INUS – insuficiente, mas parte necessária de uma

condição que em si é não-necessária, mas suficiente para o resultado. A condição a isolada não é suficiente, mas é uma componente necessária da condição combinada aB, que por si mesma não é necessária, mas somente suficiente para R; bem como a condição B não é suficiente, mas é uma componente necessária da condição combinada aB. Analogamente, para a conjunção Ac.

Uma regra, quando se analisam dados com a QCA, é que uma condição isolada não é necessária nem suficiente, mas tem um papel crucial para o conjunto resultado (SCHNEIDER; WAGEMANN, 2010). Agora podemos avaliar o efeito da presença de remanescentes lógicos na solução final da análise de uma tabela verdade. A diversidade limitada constitui um problema para a análise de causalidade complexa, pois o conjunto das combinações para um determinado resultado é truncado pela inexistência de casos empíricos que respaldem a avaliação da relação entre essas combinações e o resultado, não sendo factível a análise da relação (de subconjunto) entre essas combinações ausentes e o conjunto resultado.

Uma solução para esse problema é imaginar experimentos mentais como contrafactuais. Isto é, supor presença do resultado, R, e ausência do resultado, r, para as combinações empiricamente ausentes e executar as análises das relações entre todas as combinações possíveis de condições, inclusive as contrafactuais, e o conjunto resultado suposto, mediante as operações de minimização booleana. Como desejamos encontrar as configurações causais mínimas para um determinado resultado, a presença ou ausência dos remanescentes lógicos nas operações de minimização altera a solução final.

A análise de cada contrafactual (resultado suposto), mesmo sendo um experimento mental, nos permite fazer uma declaração de se a combinação de condições que está sendo examinada é ou não uma instância provável do resultado. Assim, podemos averiguar que, se ela existisse, implicaria no resultado (presença do fenômeno). A plausibilidade da solução final, supondo o remanescente lógico como consistente ao resultado, depende do estado dos nossos conhecimentos teórico e empírico. Um remanescente lógico (linha da tabela verdade) é dito *easy* quando temos informação sobre a relação entre cada condição causal e o resultado, de modo que possamos utilizá-la no experimento

mental, e certificarmos-nos de que a combinação remanescente lógico é subconjunto consistente do resultado. Se a informação sobre o efeito qualitativo da condição que está sendo examinada sobre o resultado indica que ela é redundante para o resultado, então ela pode ser descartada como subconjunto consistente do resultado (RAGIN; SONNETT, 2005).

Supondo que, para as quatro conjunções anteriormente mencionadas ABC, ABc, aBC e aBc, a conjunção aBc é um remanescente lógico – isto é, não apresenta instância empírica para o seu correspondente resultado – então a solução final, via minimização booleana, utilizando apenas as combinações consistentes para o resultado (ABC, ABc e aBC), seria “ $AB + BC \rightarrow R$ ”, alcançada conforme as operações de minimização representadas na Figura 18. Essa solução é denominada **solução complexa**. Portanto, a solução complexa é alcançada quando consideramos os remanescentes lógicos como inexistentes; eles são excluídos do processo de minimização.

Por outro lado, se considerássemos todas as conjunções para o resultado R – nesse caso estamos supondo que a conjunção aBc é consistente para o resultado R –, alcançaríamos a solução “ $B \rightarrow R$ ”, conforme representada na Figura 18. Quando consideramos no processo de minimização os remanescentes lógicos que contribuem para alcançar a solução final mais simples quanto possível, estamos tentando solucionar o problema da diversidade limitada pela chamada **solução parcimoniosa**.

Figura 18 – Etapas da minimização booleana (sem remanescente lógico)

	ABC	ABc	aBC		AB	BC		
ABC	-	AB	BC		AB	-		
ABc	AB	-	-	\Rightarrow	BC	-		
aBC	BC	-	-				\Rightarrow	$AB + BC \rightarrow R$

Fonte: Elaboração própria.

Quando consideramos apenas os remanescentes lógicos ditos *easy*, alcançamos uma solução denominada **solução intermediária**. Observemos que a declaração da combinação aBc como consistente para R é possível porque temos instâncias empíricas que nos permitem esta conclusão: as combinações ABC, aBC e ABc, como sabemos por hipótese, são consistentes para R e não são remanescentes lógicos. Portanto, a

comparação das combinações consistentes ABC e aBC mostra que a presença/ausência da condição A é redundante para R; semelhantemente, a comparação das combinações ABC e ABc indica que a presença/ausência da condição C é redundante para o resultado R. Assim, nosso conhecimento existente torna a asserção “aBc \rightarrow R”, um contrafactual fácil (*easy*). Notemos que, nesse exemplo hipotético, só existe um único remanescente lógico, aBc, conseqüentemente, a solução intermediária é idêntica à solução parcimoniosa.

Contudo, isso nem sempre é possível. Em algumas situações, podemos estar diante de remanescentes lógicos tais que a análise do efeito qualitativo de algumas condições não nos permite declará-las como redundantes para o resultado R. Por exemplo, se as combinações aBc e ABc fossem remanescentes lógicos, a comparação ABc e ABC não seria possível, pois não sabemos se ABC implica R ou r (princípio da multifinalidade). Portanto, o efeito da mudança qualitativa na condição C não é conhecido. Esses contrafactuais são ditos *difficult*, não porque são difíceis de lidarmos com eles ou devemos evitá-los, mas porque exigem uma justificativa cuidadosa para declará-los como subconjuntos consistentes do resultado R (RAGIN; SONNETT, 2005).

Em suma, essas distinções entre os remanescentes lógicos impõem especificidades às soluções finais de uma aplicação QCA, pois a maneira pela qual eles são considerados na operação de minimização, partindo das interseções fundamentais, conduz a três tipos de solução final: complexa, parcimoniosa e intermediária. A solução final complexa deriva da operação de minimização booleana considerando inicialmente somente as combinações de condições consistentes com o resultado, ou seja, ignoramos totalmente os remanescentes lógicos. As demais soluções consideram, além das combinações consistentes empregadas na solução complexa, os remanescentes lógicos, porém de modo diferenciado: enquanto na solução parcimoniosa utilizamos os remanescentes lógicos que conduzem à solução mais simples possível, na solução intermediária empregamos na operação de minimização somente os remanescentes que são contrafactuais *easy*.

Observemos que essas três soluções, embora sejam operadas distintamente, são conjuntos relacionados entre si. A solução complexa

privilegia a complexidade (interseções fundamentais) e é subconjunto da solução intermediária que, por sua vez, é subconjunto da solução parcimoniosa, que consiste em alguma condição causal isolada (superconjunto). Essa relação entre as soluções ocorre porque ambas devem cobrir as linhas da tabela verdade para o resultado (presença do fenômeno representado pelo conjunto resultado). Equivalentemente, podemos dizer que a solução parcimoniosa é superconjunto da solução intermediária, que, por sua vez, é superconjunto da solução complexa. A título de ilustração, tomemos a Tabela 7 e vamos supor resultados possíveis para os remanescentes lógicos aBC e abc . As tabelas verdade geradas por esse experimento mental estão disponibilizadas na Tabela 10. As diferenças entre elas ocorrem na coluna resultado das linhas que são remanescentes lógicos. Na Tabela 10-A, supomos que aBC e abc são suficientes para R; na B, aBC é suficiente; na C, abc é suficiente; e na D, nenhum deles é suficiente para R.

Tabela 10 – Tabelas verdade para o resultado dos remanescentes lógicos

Condições			Resultado	Condições			Resultado
A	B	C	R	A	B	C	R
1	1	1	1	1	1	1	1
1	1	0	1	1	1	0	1
1	0	1	0	1	0	1	0
0	1	1	1	0	1	1	1
1	0	0	1	1	0	0	1
0	1	0	0	0	1	0	0
0	0	1	1	0	0	1	1
0	0	0	1	0	0	0	0

Tabela C				Tabela D			
Condições			Resultado	Condições			Resultado
A	B	C	R	A	B	C	R
1	1	1	1	1	1	1	1
1	1	0	1	1	1	0	1
1	0	1	0	1	0	1	0
0	1	1	0	0	1	1	0
1	0	0	1	1	0	0	1
0	1	0	0	0	1	0	0
0	0	1	1	0	0	1	1
0	0	0	1	0	0	0	0

Fonte: Elaboração própria.

Figura 19 – Minimizações das tabelas verdade para resultado dos remanescentes lógicos

Minimização da tabela verdade A

	ABC	ABc	aBC	Abc	abC	abc
ABC	-	AB	BC	-	-	-
ABc	AB	-	-	Ac	-	-
aBC	BC	-	-	-	aC	-
Abc	-	Ac	-	-	-	bc
abC	-	-	aC	-	-	ab
abc	-	-	-	bc	ab	-

Minimização da tabela verdade B

	ABC	ABc	aBC	Abc	abC	abc
ABC	-	AB	BC	-	-	-
ABc	AB	-	-	Ac	-	-
aBC	BC	-	-	-	aC	-
Abc	-	Ac	-	-	-	-
abC	-	-	aC	-	-	-
abc	-	-	-	-	-	-

Minimização da tabela verdade C

	ABC	ABc	Abc	abC	abc
ABC	-	AB	-	-	-
ABc	AB	-	Ac	-	-
Abc	-	Ac	-	-	bc
abC	-	-	-	-	ab
abc	-	-	bc	ab	-

Minimização da tabela verdade D

	ABC	ABc	Abc	abC
ABC	-	AB	-	-
ABc	AB	-	Ac	-
Abc	-	Ac	-	-
abC	-	-	-	-

Fonte: Elaboração própria.

Supondo todos os resultados possíveis para as linhas remanescentes lógicos aBC e abc, e procedendo a minimização booleana (Figura 19), encontramos quatro fórmulas de solução distintas. Como nesse exemplo só há remanescentes lógicos *easy*, porque há instâncias empíricas nas linhas passíveis de comparação com os remanescentes lógicos (aBC pode ser minimizada com ABC e se torna BC ou com aBc e se torna aB; abc pode ser minimizada com abC e se torna ab e com aBc, para ac), então as soluções parcimoniosas também são intermediárias:

(a) Soluções parcimoniosas e intermediárias

$$AB + BC + Ac + aC + ab + bc \rightarrow R$$

$$AB + BC + Ac + aC \rightarrow R$$

$$AB + Ac + ab + bc \rightarrow R$$

(b) Solução complexa, que ignora os remanescentes lógicos no processo de minimização da tabela verdade: $AB + Ac \rightarrow R$.

Caso ABC também fosse um remanescente lógico, não haveria instância empírica de modo que ele pudesse ser comparado com aBC. Raciocínio semelhante para abc. Nesse caso, seriam contrafactuais

difficult, porque exigiriam uma justificativa adequada para declará-los como subconjuntos consistentes do resultado R.

Linhas contraditórias da tabela verdade

A natureza das linhas contraditórias na tabela verdade é mais bem compreendida mediante o uso de condições dicotômicas. Consistem naquelas linhas em que os casos a elas pertencentes não compartilham o mesmo escore de pertencimento no conjunto resultado. Ou seja, a mesma linha pode conduzir à presença (ocorrência) e à ausência (não ocorrência) do resultado. No exemplo dado na Tabela 7, o caso 3 tem o resultado presente, R, e o caso 7 tem ausência do resultado, r (ou $\sim R$). Por isso, a conjunção aBc é uma linha contraditória. Uma vez que as linhas da tabela verdade são declarações de suficiência, a presença de contradições lógicas significa que a mesma combinação de condições implica tanto R quanto r. Assim, o problema analítico que temos em mão é que, com base na evidência empírica, não é possível decidir se a linha contraditória é suficiente para R, para r, ou nenhum dos dois; conseqüentemente, nem se ela deve ou não entrar no processo de minimização booleana para o resultado R, r ou nenhum deles (remanescente lógico).

De acordo com Schneider e Wagemann (2012), há algumas estratégias, não mutuamente exclusivas, para tratar a presença das linhas contraditórias em tabelas verdades em csQCA e fsQCA: estratégias implementáveis **antes** da minimização lógica e estratégias implementáveis **durante** o processo de minimização. O primeiro conjunto de estratégia, que pertence aos padrões de boas práticas da QCA, consiste de:

- (a) Adicionar uma nova condição na tabela verdade.

A contradição é superada porque a introdução de uma nova condição faz com que a linha contraditória seja “dividida” em duas linhas, distribuindo, portanto, os casos com distintos escores de pertencimento no resultado nas duas novas linhas. A desvantagem dessa solução está no fato de que ela duplica o número de linhas da tabela verdade, podendo conduzir à emergência de remanescentes lógicos.

(b) Re-especificar a definição da população de interesse.

Esse procedimento implicaria que alguns casos poderiam ser incluídos e outros, excluídos, mas isso, para ser executado, precisa estar baseado em argumentos teóricos. Afinal, alguns casos não podem ser excluídos arbitrariamente do conjunto de dados simplesmente por não satisfazerem a propriedade de suficiência. Ademais, isso poderia influenciar a seleção das condições e resultado e suas respectivas calibrações, o que, por sua vez, poderia gerar novas linhas contraditórias.

(c) Re-especificar a definição, a conceituação e/ou a mensuração das condições ou resultado.

A observação das semelhanças e dessemelhanças nos casos contraditórios de uma determinada linha poderia revelar que a especificação de uma (ou algumas) das condições e/ou do resultado pode ter sido feita de modo impreciso ou mesmo equivocado. Mas essa solução também precisa estar suportada em uma boa argumentação teórica.

Quanto às estratégias implementáveis durante o processo de minimização, alguns pesquisadores conduzem a operação de minimização lógica incluindo a(s) linha(s) contraditória(s). Os tratamentos, mutuamente excludentes, durante a minimização são:

(a) A exclusão de todas as linhas contraditórias da minimização.

Isso significa que apenas as relações perfeitas de subconjunto se qualificam como condições suficientes para o resultado. A desvantagem de agirmos assim é que qualquer caso que é membro do resultado (*i.e.*, presença do resultado) e está na linha contraditória não será explicado ou coberto pelo termo de solução produzido por essa estratégia.

(b) A inclusão de todas as linhas contraditórias na minimização.

Essa estratégia baseia-se no argumento de que uma linha contraditória pelo menos faz com que seja possível a ocorrência do resultado; e a fórmula da solução alcançada representará as conjunções de condições que tornam o resultado possível. Todos os

casos que são membros do resultado serão explicados, ou cobertos, por esse termo de solução. Contudo, a desvantagem é que o termo de solução cobrirá alguns casos que não são membros do resultado.

(c) A inclusão das linhas contraditórias fica “a cargo” do computador. O próprio programa de computação seleciona quais linhas contraditórias incluir e não incluir no processo de minimização, a partir de conjecturas geradas pelo computador a respeito do escore de pertencimento no conjunto resultado. A *rationale* subjacente é que as linhas contraditórias são incluídas, desde que tornem mais parcimonioso o termo de solução resultante. Essa estratégia é, entre as três, a menos justificável e menos encontrada nas aplicações de QCA.

Essas estratégias enfatizam um fato importante: os métodos baseados em conjuntos, em geral, e a QCA, em particular, são uma abordagem que especifica requisitos para o processo de pesquisa anterior e posterior à análise dos dados, e não são apenas um conjunto de técnicas para analisar dados. Portanto, refletem natureza dual da QCA como técnica de análise de dados e como abordagem de pesquisa.

EXEMPLO DE APLICAÇÕES

Este capítulo tem o propósito de apresentar duas aplicações, permitindo ao leitor se familiarizar de uma maneira prática com todos os conceitos e fundamentos tratados nos capítulos antecedentes, especialmente a principal ferramenta de análise dos casos, ou seja, a tabela verdade. O leitor notará também como essas aplicações podem contribuir nas avaliações de políticas. A primeira aplicação utiliza parte das variáveis constantes na base de dados do artigo de Betarelli Junior e Simões (2011). Tendo em vista que *fsQCA* nos fornece a configuração específica de cada caso, uma análise da configuração que levou ao alto grau de industrialização torna-se uma importante caracterização regional, contribuindo para os planejadores públicos de políticas regionais. Por exemplo, pode-se identificar as regiões que detêm um alto poder aquisitivo no mercado local (economia de urbanização), uma grande concentração de oferta de serviços produtivos e de força de trabalho qualificado (economia de urbanização), e alto nível de analfabetismo (deseconomias urbanas). Já a segunda aplicação utiliza a *QCA* para certos países a fim de avaliar o sucesso da implementação de política pública (qualidade ambiental) a partir de um conjunto de medidas de capacidade institucional dos referidos países. Uma das principais contribuições deste estudo empírico é prover às autoridades específicas informações úteis ao processo decisório relacionado às condições institucionais.

Aplicação ilustrativa

Aqui, nosso objetivo é puramente técnico: mostrar os passos básicos de uma aplicação da *QCA* com conjuntos fuzzy (*fsQCA*). Procuramos identificar as configurações consistentes para alto grau de industrialização nas microrregiões do estado de São Paulo. Suponhamos que um determinado estudo qualitativo comparativo quer averiguar a causalidade conjuntural múltipla de cinco condições causais para o resultado alto grau de industrialização (sigla *GINDUST*). O Quadro 5 reporta os seis conjuntos (conceitos), com suas siglas, descrição, fórmula de cálculo dos valores dos casos em cada conjunto

(posteriormente transformados em escores fuzzy) e interpretação de cada conceito.¹⁶ Além disso, quanto à calibragem dos conjuntos *fuzzy*, utilizaremos o método do ranqueamento para todos os seis conjuntos, deixando que o valor do escore estabeleça o grau de pertencimento no conjunto. Portanto, não recorreremos à informação externa aos dados para estabelecer as âncoras qualitativas. Escores entre 0 e 0,5 indicam casos de não-ocorrência do conceito, e escores entre 0,5 e 1, de ocorrência do conceito. Por exemplo, um caso com um escore 0,32 no conjunto GINDUST indica seu baixo pertencimento nesse conjunto (baixo grau de industrialização).

Quadro 5 – Lista de conjuntos, suas descrições e interpretações.

Sigla	Conjuntos	Interpretação
GINDUST	Grau de industrialização	Está associada com a economia de urbanização (no sentido que apresenta um mercado local significativo de compradores e de vendedores), de forma que provoca maiores efeitos encadeamento (para frente e para trás).
MERCADO	Mercado	O propósito é tentar captar o poder de compra do mercado local (fator aglomerativo), contudo também pode ser tratado o custo da força de trabalho urbano. Nesse trabalho, considerou como um fator aglomerativo de urbanização.
FESPEC	Força de trabalho especializado	Procura-se obter o nível de qualificação da força de trabalho, o que representa uma vantagem potencial de uma cidade, ou seja, economias de urbanização.
SERVIÇOS	Oferta de Serviços produtivos	Esta variável envolve-se com o conceito de economias de urbanização de modo que mede o nível de oferta de serviços.
CESP	Coefficiente de Especialização	Compara a estrutura produtiva da microrregião em relação à estadual. Quando se aproxima de 1, a região apresenta um elevado grau de especialização em um dado setor ou está diversa da estrutura do emprego estadual (HADDAD, 1989).
ANALFAB	Analfabetismo	Representa a baixa qualificação da força de trabalho da região e atua como um fator desaglomerativo para as atividades intensivas em trabalho qualificado. Contudo, para setores que não são demandantes de qualificação, essa variável pode representar uma fonte de atração.

Fonte: Betarelli e Simões (2011).

¹⁶ Os autores podem disponibilizar a base de dados, do-file e arquivo log para executar essa aplicação no programa Stata.

A avaliação da consistência da suficiência das configurações iniciais é executada pelo comando “fuzzy varlist settest (...)”, que exibe, para todas as combinações logicamente possíveis de condições (linhas da tabela verdade), o número de casos que são membros dessas combinações (conjunções) e informa se cada uma delas é (ocorrência) ou não é (não ocorrência) um subconjunto do resultado¹⁷. A matriz de coincidência fornece as proporções de casos (mesorregiões) contidos em cada uma das cinco condições que coincidem com alto grau de industrialização (I) e com cada uma das demais condições (Tabela 11). Os elevados compartilhamentos dos conjuntos alto grau de mão de obra especializada (F) e amplo mercado (M) com alto grau de industrialização (I) são 0,706 e 0,700, respectivamente; cerca de 71% dos casos são compartilhados pelos conjuntos F e I. Lembremos que os coeficientes da matriz de coincidência não delimitam as condições de suficiência e de necessidade; ou seja, não indicam qual desses conjuntos está contido no outro, e vice-versa. Para isso, precisamos averiguar, entre todas as combinações lógicas possíveis, aquelas com relações de suficiência com alto grau de industrialização.

Tabela 11 – Matrizes de coincidência e de suficiência e necessidade

Matriz de Coincidência							Matriz de Suficiência e Necessidade						
	I	F	S	C	M	A		I	F	S	C	M	A
I	1						A	0,667	0,617	0,572	0,788	0,595	1
F	0,706	1					M	0,7	0,857	0,84	0,532	1	0,595
S	0,693	0,826	1				C	0,65	0,562	0,541	1	0,532	0,788
C	0,65	0,562	0,541	1			S	0,693	0,826	1	0,541	0,84	0,572
M	0,7	0,857	0,84	0,532	1		F	0,706	1	0,826	0,562	0,857	0,617
A	0,667	0,617	0,572	0,788	0,595	1	I	1	0,706	0,693	0,65	0,7	0,667

Fonte: Resultados do exercício de aplicação da fsQCA.

Podemos também dimensionar a relação de compartilhamento de casos entre as condições. Por exemplo, amplo mercado (M) e alto grau de mão de obra especializada (F) compartilham cerca de 86% dos casos. As

¹⁷ O Quadro 6 apresenta os principais comandos no Stata.

tabelas de consistência para o resultado (e para o não resultado) listam todas as interseções primitivas e seus valores de consistência. Elas não são colocadas aqui, por conta do tamanho (cinco condições implicam em 32 linhas na tabela verdade). Essas tabelas indicam que somente as combinações mfsCA e MFsCa têm valores de consistência de 0,718 e 0,764, inferiores ao limiar (*threshold*) de consistência da suficiência para o resultado alto grau de industrialização, estabelecido em 0,800. As combinações mais consistentes para o resultado são MFsCa (0,970) e mFsca (0,963).

As tabelas anteriormente mencionadas nos permitem identificar a presença da diversidade limitada, porque há 14 remanescentes lógicos (combinações de condições sem instâncias empíricas). Não há o problema de linhas contraditórias. A etapa de minimização da complexidade do algoritmo tabela verdade é executada no Stata pelo algoritmo Quine-McCluskey (*reduce*), com opções para o tratamento dos remanescentes lógicos. Para tanto, precisamos recorrer às linhas consistentes da tabela verdade (Tabela 12).

Tabela 12 – Tabela verdade

M	F	S	C	A	I	Consistência
0	1	1	1	0	1	0,948
1	1	0	0	1	1	0,948
1	1	0	1	0	1	0,970
1	1	1	1	0	1	0,957

Fonte: Elaboração própria.

Já a Tabela 13 exibe as soluções finais, seus termos de solução e medidas de consistência e coberturas. A **solução complexa**, “mFSCa → I”, que não envolve num remanescente lógico, indica que mercado com baixo poder de compra (m), alta mão de obra especializada (F), altos serviços produtivos (S), alto grau de especialização produtiva (C), baixo analfabetismo (a) tem consistência de 95% para conduzir a alto grau de industrialização (I). Como só há um termo, só há uma medida de cobertura, de 0,297. Ou seja, 29,7% das mesorregiões com configuração mFSCa compartilham casos com I.

Tabela 13 – Soluções finais, termos de solução e seus valores de consistência e coberturas

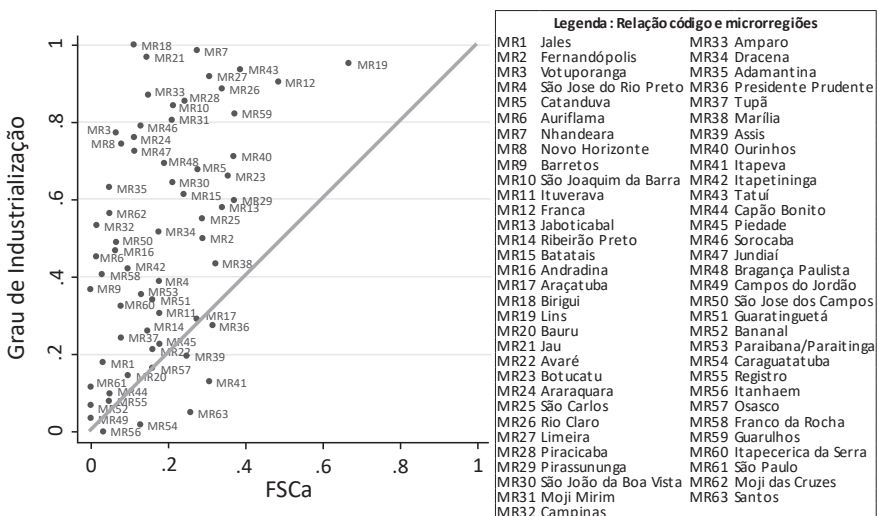
Tipo de solução	Configurações	Consistências	Coberturas	
			Bruta	Única
Solução final	MFscA	0,948	0,324	0,070
	FSCa	0,945	0,344	0,028
	MFCa	0,959	0,334	0,010
Solução complexa	MFSCa	0,948	0,297	0,297
Solução parcimoniosa	mFSC	0,901	0,334	0,037
	FCa	0,942	0,366	0,069
Solução				
	Final	Complexa	Parcimoniosa	
Consistência	0,93	0,948	0,904	
Cobertura	0,43	0,297	0,404	

A **solução parcimoniosa**, “mFSC + FCa \rightarrow I”, que considera os remanescentes lógicos *easy*, é uma disjunção; portanto, ela mostra que há dois caminhos conjunturais alternativos para alcançar alto grau de industrialização (I): mFSC (mercado com baixo poder de compra (m), alta mão de obra especializada (F), altos serviços produtivos (S), alto grau de especialização produtiva (C)) ou FCa (alta mão de obra especializada (F), alto grau de especialização produtiva (C), e baixo analfabetismo (a)). Ambos os caminhos são altamente consistentes para I. Os dois caminhos causais compartilham 40,4% dos casos com alto grau de industrialização e suas coberturas únicas são muito pequenas, mas a cobertura bruta do caminho FCa é ligeiramente superior à do caminho mFSC. Isso quer dizer que a conjunção FCa é a relativamente mais representativa para o resultado do que a outra.

A **solução intermediária**, “MFscA + FSCa + MFCa → I”, produzida quando diferentes conjuntos de remanescentes lógicos, que são utilizados para gerar a solução parcimoniosa, são incorporados na minimização booleana, oferece três caminhos alternativos consistentes para I (sua descrição se assemelha à descrição da solução parcimoniosa, no sentido de que são disjunções). Como as demais, tanto a solução quanto os termos têm altos valores de consistência. As coberturas dos três termos de solução indicam que essas conjunções sobrepõem entre 32,4% e 34,4% dos casos no conjunto de alto grau de industrialização.

A título de ilustração, utilizamos um diagrama XY para visualizarmos o padrão de suficiência causal entre o termo da solução intermediária mais representativa (FSCa – alta mão de obra especializada, altos serviços produtivos, alto grau de especialização produtiva, e baixo analfabetismo) para alto grau de industrialização, na Figura 20. A maioria das mesorregiões tem escores que satisfazem a propriedade de suficiência, porém alguns não, como MR36, MR39, MR41, MR44, MR54, MR63 e outros; e os mais consistentes são, por exemplo, MR19, MR12, MR43, MR59, entre outros.

Figura 20 – Diagrama XY de suficiência de FSCa para I



Fonte: Elaboração própria.

Por fim, o Quadro 6 fornece os principais comandos executados nesta aplicação. Para maiores detalhes dos comandos e atualizações das versões no programa Stata, o leitor deve consultar Longest e Vaisey (2008).

Quadro 6 – Relação dos comandos principais executados

Método de calibragem	foreach var of varlista gindust mercado fespec cesp analfab { setgen st`var' = stdrank(`var')}
Executar <i>fsQCA</i>	fuzzy stgindust stmercado stfespec stservicos stcesp stanalfab, label (I M F S C A)
Matrizes da Tabela 11	fuzzy I M F S C A, matx(coincid suffnec) standardized altdisplay
Minimização Quine- McCluskey	
Convencional	fuzzy I M F S C A, settest(yvv yvn) sigonly greater(col1) conval(.800) common reduce
Com remanescentes	fuzzy I M F S C A, settest(yvv yvn) sigonly greater(col1) conval(.800) common reduce remainders(0)
Tabela Verdade	truthTAB
Escore	cnfgen MFscA FSCa MFCa
Diagrama XY	gen U = FSCa fzplot I U, mlabel(tipo)

Fonte: Elaboração própria.

Estudo selecionado

Para ilustrar a abordagem e sua aplicação no âmbito de políticas públicas, descreveremos o estudo de Korhonen-Kurki *et al.* (2014), que utilizam os dados do *Global Comparative Study on REDD+* (GCS-REDD) para analisar o processo de implementação do REDD+ em um grupo selecionado de países. O **REDD+** (*Reducing Emissions from Deforestation and Forest Degradation*, ou Redução das Emissões por Desmatamento e Degradação florestal) é um conjunto de incentivos econômicos, com o intuito de reduzir as emissões de gases de efeito estufa resultantes da degradação florestal e desmatamento; mudanças que prejudicam as florestas e limitam sua provisão de serviços ambientais (BRASIL, 2018).

O objetivo do GCS-REDD é prover os formuladores de políticas e profissionais associados com conhecimentos relevantes para assegurar uma redução efetiva, custo-eficiente e equitativa das emissões de carbono decorrentes do desmatamento e da degradação florestal (isso forma os critérios 3E+), bem como prover cobenefícios, tais como alívio da pobreza e a conservação da biodiversidade nos países participantes. Os países¹⁸ selecionados são países em desenvolvimento ou emergentes, tropicais ricos em florestas com um compromisso político de implementar REDD+, mas também com potentes motivadores/condutores de desmatamento, governança multinível limitada, baixa coordenação transetorial horizontal e falta de capacidade. Esses aspectos formam o contexto comum dos casos (países).

O objetivo do estudo foi empregar a QCA para identificar, mediante uma comparação sistemática, sob que condições esses países podem implementar com sucesso o REDD+, e desenvolver generalizações e recomendações políticas para eles e para países que compartilham contexto semelhante. Uma vez que as análises ocorreram simultaneamente aos esforços para começar a implementação do REDD+, os pesquisadores procuraram identificar também **o que funciona e o que não funciona nesse processo**. Para realizar isso, Korhonen-Kurki *et al.* (2014) empregaram um processo iterativo para selecionar os fatores (condições) mais relevantes (etapa de especificação) que conduziriam ao sucesso ou ao fracasso do estabelecimento de uma estrutura política adequada ao REDD+. Esse processo identificou seis **fatores** ou **condições**. O **resultado** bem-sucedido foi definido como o estabelecimento de uma política integrada promotora da mudança no domínio da política REDD+ que é susceptível de conduzir ao 3E-REDD+ (redução efetiva, custo-eficiente e equitativa das emissões de carbono decorrentes do desmatamento e da degradação florestal). Doravante o

¹⁸ África: Burkina Fuzzy, Camarões, República Democrática do Congo, Moçambique e Tanzânia; América do Sul: Bolívia, Brasil e Peru; Ásia e Oceania: Indonésia, Nepal, Papua Nova Guiné e Vietnam.

resultado será denotado pela sigla R. A apresentação desse estudo será dividida em duas partes, de modo a enfatizar as distintas utilidades da QCA como abordagem metodológica: na estruturação dos dados a serem analisados por algumas técnicas da QCA, e na extração de inferências.

QCA na estruturação de dados

Um dos aspectos da abordagem QCA pode ocorrer antes do início da análise propriamente dita, que é a identificação e a verificação da coerência dos dados. Os participantes do estudo exploraram minuciosamente os fatores que potencialmente afetariam a implementação bem-sucedida do REDD+. A lista de tais fatores posteriormente foi reduzida a um número gerenciável de fatores relevantes, e estes, por sua vez, foram operacionalizados, atribuindo-lhes indicadores que mensuram tais fatores. Korhonen-Kurki *et al.* (2014) usaram a QCA de forma descritiva para sumariar as informações adquiridas durante os dois anos de implementação do projeto por dezenas de pesquisadores. A lista preliminar de fatores potenciais foi desenvolvida em um *workshop* com participantes de várias equipes dos países selecionados. Embora a lista não tenha sido formada sob pesquisa representativa, forneceu informações sobre os fatores relevantes para comparações entre países.

Etapas subsequentes, inclusive a revisão dos estudos de contexto específico dos países produzida pelas equipes dos próprios países, possibilitaram a simplificação do número inicial desses fatores, mediante um processo de priorização e consolidação; foram selecionados oito fatores para inclusão na análise e operacionalizados por indicadores. Esses indicadores foram desenvolvidos após a primeira avaliação das discrepâncias reveladas, de modo a assegurar transparência e comparabilidade das avaliações. Esses dados forneceram uma base confiável e válida para dar início à QCA propriamente dita, durante a qual os fatores foram discutidos novamente e alterados – que é uma das características típicas da QCA, referida como “diálogo com o caso” (BERG-SCHLOSSER *et al.*, 2008). Ao cabo desse processo foram selecionados seis

fatores como chave para explicar o sucesso ou o fracasso da execução do mecanismo 3E-REDD+. A análise foi conduzida usando o *software* TOSMANA.

O processo levou vários meses, mas assegurou que as informações fossem intersubjetivamente verificáveis e comparáveis, respeitando as especificidades dos casos. A QCA ajudou a garantir que todas as equipes dos países compartilhassem os mesmos entendimento e definição dos fatores. Desse modo, as informações puderam ser tabuladas numa tabela verdade, permitindo a discussão em conjunto da identificação da causação conjuntural múltipla para 3E-REDD+. As aplicações da QCA ofereceram aos participantes a possibilidade de obterem uma visão geral do projeto e reduzirem a complexidade das informações derivadas de numerosos e extensos estudos de caso. Além disso, as aplicações mostraram também as discrepâncias nas avaliações, que puderam ser esclarecidas. A partir da definição dos indicadores, o processo assegurou que todos tivessem um mesmo entendimento e que os valores atribuídos aos fatores fossem comparáveis, mas suficientemente sensíveis ao contexto conforme a realidade. Embora o estudo original tenha considerado seis fatores para a implementação bem-sucedida do 3E-REDD+, o exemplo de aplicações da QCA a seguir considera apenas três fatores institucionais: (a) pressão por escassez de recursos florestais (sigla P): grande parte da área florestal do país está sob pressão da atividade humana, devido a padrões institucionalizados de uso da floresta, que são incapazes de atender às necessidades ou cumprir interesses de uso; (b) características principais da legislação, política e governança florestais efetivas (E): há uma estrutura legal que define a posse, uso e gestão de direitos, incluindo regras formais e consuetudinárias. Leis e políticas sobre gestão e participação florestal sustentável são aplicadas pelas autoridades nacionais e locais e cumpridas pelos usuários da floresta; e (c) processo de mudança política já iniciado (C): mudança política direcionada a mudanças climáticas desenvolvidas independentemente do REDD+ já está em andamento – por exemplo, programas antidesmatamento, estratégias de desenvolvimento de baixo teor de carbono, esquemas de pagamento por serviços ambientais, entre outros.

QCA para extrair inferências

O QCA é útil para estruturar dados e validar avaliações, conforme descrito anteriormente, porém serve principalmente para analisar o material resultante dessa etapa inicial, de modo a produzir respostas tentativas para o estudo que se está realizando. Muitas agências oficiais têm como função principal formular recomendações aos gestores públicos sobre estratégias e desenho institucional que alcancem, de modo eficiente, efetivo e equitativo, os objetivos de políticas previamente estabelecidos. Essas recomendações podem ser formuladas especificamente para determinados domínios das políticas (subnacionais, nacionais ou supranacionais), mas podem ser estendidas, no sentido de que se tornem recomendações gerais aplicáveis em outros domínios. Assim sendo, é importante que haja uma base confiável para se fazerem generalizações a partir dos resultados encontrados em um determinado conjunto de casos. A QCA permite realizar inferências sem negligenciar as circunstâncias específicas de cada caso e os diferentes caminhos que as unidades macrossociais podem optar para realizar a mudança política (ou institucional) em estudo. O estudo de Korhonen-Kurki *et al.* (2014) empregou variantes da QCA para identificar estratégias e desenhos institucionais que conduzam eficiente, efetiva e equitativamente à política REDD+.

Aplicação da csQCA aos dados do GSC-REDD

O elemento central da QCA é a **tabela verdade**; para cada uma das condições, o limiar entre ausência e presença é definido teoricamente e avaliado com base no conhecimento específico de cada caso¹⁹. A Tabela 14 (tabela verdade da aplicação csQCA) apresenta a aplicação desse processo de atribuição de escores 0 ou 1 aos países escolhidos para o estudo. Ela mostra também todas as combinações possíveis de fatores institucionais cobertas pelos doze casos (países). Em nosso exemplo com três fatores (ou condições causais), são possíveis oito combinações de condições. Notemos

¹⁹ Descritos no Apêndice 1 em Korhonen-Kurki *et al.* (2014).

que todas as oito condições estão representadas no conjunto dos doze casos empíricos, e isso é uma situação nem sempre comum. Na ausência de instâncias empíricas, há as linhas remanescentes lógicas e estaremos diante da diversidade empírica limitada (ver capítulo 4).

Na configuração final, as letras maiúsculas mostram a presença de uma condição, e letras minúsculas indicam ausência. Por exemplo, a configuração para a Indonésia (quinta linha da Tabela 14), em termos de letras representando presença/ausência das condições, é P^*e^*C , com a presença do resultado ($R = 1$). Isso é interpretado da seguinte forma: pressão por escassez de recursos florestais E ausência de características legislativas, políticas e de governança florestais efetivas E presença de mudança já iniciada de política pró-clima leva à presença de uma política abrangente promotora de mudança no domínio da política REDD +.

Tabela 14 – Tabela verdade da csQCA

Fator/Condição			Resultado	Caso/País
P	E	C	R	
0	0	0	0	Papua Nova Guiné
1	0	0	0	Burkina Faso, Moçambique e Tanzânia
1	1	0	0	Camarões
0	0	1	0	República Democ. do Congo, Peru
1	0	1	1	Indonésia
0	1	0	0	Nepal
0	1	1	1	Vietnam
1	1	1	C	Brasil (1), Bolívia (0)

Fonte: Adaptado de Sehring, Kurki e Brockhaus (2013).

Notemos que a presença do resultado se dá apenas no Brasil, Indonésia e Vietnam. Por outro lado, todas as condições (última linha da tabela) estão presentes no Brasil e Bolívia, mas o resultado não está presente em ambos os países: embora presente no Brasil, na Bolívia está ausente (letra C para indicar uma contradição no resultado em que todas as condições estão presentes). As configurações que conduzem à presença do resultado são P^*e^*C (Indonésia), p^*E^*C (Vietnam) e P^*E^*C (Brasil). A complexidade dessas fórmulas pode ser reduzida: tanto p^*E^*C (Vietnam) quanto P^*E^*C (Brasil) conduzem ao resultado $R = 1$. Portanto, é irrelevante

se a condição P está ou não presente nas combinações específicas do Vietnã e do Brasil. E , pela álgebra booleana, as combinações p^*E^*C e P^*E^*C podem ser reduzidas para E^*C . Portanto, a presença de estrutura legal-política e de governança do uso florestal E em um processo já iniciado de mudança institucional pró-clima levam consistentemente ao REDD+ bem-sucedido. Contudo, a configuração P^*e^*C (Indonésia) também leva a $R = 1$. A redução da complexidade não pode ser feita quando se compara P^*e^*C com p^*E^*C (Vietnã), mas sim com P^*E^*C (Brasil), pois nesse caso a condição E seria irrelevante. Pela álgebra booleana, as combinações P^*e^*C e P^*E^*C podem ser reduzidas para P^*C . Assim, a presença de escassez de recursos florestais E em um processo já iniciado de mudança institucional pró-clima também levam ao REDD+ bem-sucedido.

A aplicação da QCA a conjuntos dicotômicos (csQCA) identificou dois caminhos que conduzem ao sucesso da implementação do REDD+: E^*C ou P^*C ; isto é $R = E^*C + P^*C$. Portanto, conforme a estrutura de condições dos casos bem-sucedidos da política REDD+, a presença de estrutura legal-política e de governança do uso florestal E em um processo já iniciado de mudança institucional pró-clima *OU* presença de escassez de recursos florestais E em um processo já iniciado de mudança institucional pró-clima podem levar ao resultado de interesse. Não importa qual a frequência que uma determinada combinação de condições é encontrada – se é observada uma única vez, ela é causal (no sentido da abordagem QCA). Quanto às condições necessárias e suficientes, pode-se concluir que, no contexto conjunto dos casos selecionados, o fato de que são países que já iniciaram mudanças de política (condição C) é uma condição necessária, embora não suficiente, para a presença de políticas pró-REDD+. A condição E , características da política, governança e legislação florestal efetiva, é uma condição do tipo INUS: é insuficiente para o resultado, mas é parte necessária da configuração E^*C , que, por sua vez, não é em si necessária, mas é suficiente para o resultado REDD+. Se esses resultados relacionados à Indonésia, ao Vietnã e ao Brasil puderem ser generalizados, é possível que Camarões e Bolívia sejam bem-sucedidos na implementação da política REDD+.

Por outro lado, a condição de suficiência (capítulo 3) estabelece que, partindo-se de algum argumento teórico plausível, uma condição pode ser considerada suficiente para um resultado se, sempre que ela estiver presente em todos os casos, o resultado também está presente nesses casos. Ou seja, não pode existir um único caso em que a condição está presente, mas o resultado, ausente. Para avaliarmos a suficiência das condições para o resultado, partimos das condições isoladas umas das outras. Avaliemos a condição P; neste caso estamos interessados somente nos casos em que $P = 1$ (Burkina Faso, Moçambique, Tanzânia, Camarões, Indonésia, Brasil e Bolívia). Se P é condição suficiente para R, sempre que P ocorre ou está presente, R também deve ocorrer ou estar presente. Além disso, nenhum caso em que P está presente poderia ter ausência no resultado ($R = 0$). Uma vez que há pelo menos um caso em que a condição P está presente e R ausente (por exemplo, Camarões), então esta condição não é suficiente isoladamente para REDD+ bem-sucedido.

Esse mesmo procedimento é feito para cada uma das demais condições isoladamente. A condição E está presente em Camarões, Nepal, Vietnam, Brasil e Bolívia. Contudo, embora no Vietnam e no Brasil o resultado esteja presente, o Nepal apresentou ausência do resultado. Portanto, a condição E não é isoladamente suficiente para o resultado REDD+. Semelhantemente ocorre para a condição C. Portanto, nenhuma das condições é isoladamente suficiente para um resultado bem-sucedido. Todavia, até agora só avaliamos apenas a condição de suficiência das condições nos casos em que elas estavam presentes. Temos que lembrar que os dados indicam escores de pertencimento às condições (a conjuntos), e as condições podem, portanto, assumir dois estados qualitativamente diferentes (quando se aplica a QCA a conjuntos dicotômicos) elas podem estar presentes (1) ou ausentes (0). Cada conjunto condição (P, E e C) e os seus complementos (“p”, “e” e “c”) denotam duas propriedades qualitativamente distintas. Portanto, a propriedade de suficiência dos complementos das condições também precisa ser avaliada.

A Tabela 15 se assemelha à Tabela 14 quanto às colunas de resultado e de casos, mas substitui as condições pelos seus complementos: p (ausência

de escassez de recursos florestais), e (ausência de estrutura legal-política e de governança do uso florestal) e c (processo não iniciado de mudança institucional pró-clima). Notemos que ambas as tabelas de dados contêm exatamente a mesma informação, mas ligeiramente distintas na forma de apresentação dos escores das condições. Avaliemos a condição p; neste caso estamos interessados somente nos casos em que $p = 1$ (Papua Nova Guiné, Congo, Peru, Nepal e Vietnam). Se p é condição suficiente para R, sempre que p ocorre ou está presente, R também deve ocorrer ou estar presente ($R = 1$). Além disso, nenhum caso em que p está presente poderia ter ausência no resultado ($R = 0$). Uma vez que há pelo menos um caso em que a condição “p” está presente e R ausente (por exemplo, Papua Nova Guiné), então essa condição não é suficiente isoladamente para REDD+ bem-sucedido. O mesmo procedimento avaliativo é aplicado para “e” e “c”. E nenhuma delas é isoladamente suficiente para o resultado.

Tabela 15 – tabela verdade da csQCA

Negação da Condição			Resultado	Caso/País
p	e	c	R	
1	1	1	0	Papua Nova Guiné
0	1	1	0	Burkina Faso, Moçambique e Tanzânia
0	0	1	0	Camarões
1	1	0	0	República Democ. do Congo, Peru
0	1	0	1	Indonésia
1	0	1	0	Nepal
1	0	0	1	Vietnam
0	0	0	C	Brasil (1), Bolívia (0)

Fonte: Elaboração própria a partir da tabela verdade original de Korhonen-Kurki *et al.* (2014).

Em suma, nenhuma das condições, e seus complementos, é suficiente para um resultado bem sucedido na implementação da 3E-REDD+. Contudo, mesmo que as condições e seus complementos isoladamente não satisfaçam a propriedade de suficiência, ainda é possível que a sua ocorrência simultânea (ou seja, a conjunção ou combinação de condições e seus complementos) seja suficiente para $R = 1$. Portanto temos de avaliar a propriedade de suficiência para conjunções de condições. Para essa tarefa,

precisamos calcular os escores das combinações de condições. Uma vez que uma combinação de condições é uma interseção de conjuntos (condições) e, pela álgebra booleana, o cálculo de escores, nesse caso, segue a regra do mínimo (o escore da interseção de dois ou mais conjuntos é o escore mínimo de cada um desses conjuntos), então utilizamos os escores de P, E, C (Tabela 14), “p”, “e” e “c” (Tabela 14) para calcular os escores das oito combinações P^*E^*C , P^*E^*c e assim sucessivamente.

Como ilustração, o escore da combinação P^*e^*C para a Indonésia é 1, pois é o escore mínimo de $P = 1$, $e = 1$ e $C = 1$. Por outro lado, para o mesmo país a combinação P^*E^*C tem escore nulo, pois $P = 1$, $E = 0$ e $C = 1$ e o valor mínimo entre os três escores é 0. A Tabela 16 apresenta os escores para todas as combinações possíveis de presença/ausência de combinações. Essa tabela não é uma tabela verdade, pois nas linhas estão os casos e não as combinações teoricamente possíveis.

Tabela 16 – Tabela dos escores das combinações de condições

Caso/País	Escore das Conjunções								Resultado R
	P^*E^*C	P^*E^*c	P^*e^*C	P^*e^*c	p^*E^*C	p^*E^*c	p^*e^*C	p^*e^*c	
Brasil	1	0	0	0	0	0	0	0	1
Bolívia	1	0	0	0	0	0	0	0	0
Burkina Faso	0	0	0	1	0	0	0	0	0
Camarões	0	0	0	0	0	0	0	0	0
Congo	0	0	0	0	0	0	1	0	0
Indonésia	0	0	1	0	0	0	0	0	1
Moçambique	0	0	0	1	0	0	0	0	0
Nepal	0	0	0	0	0	1	0	0	0
Peru	0	0	0	0	0	0	1	0	0
Papua Nova Guiné	0	0	0	0	0	0	0	1	0
Tanzânia	0	0	0	1	0	0	0	0	0
Vietnam	0	0	0	0	1	0	0	0	1

Fonte: Elaboração própria a partir da tabela verdade original de Korhonen-Kurki *et al.*(2014).

Agora podemos avaliar a propriedade de suficiência para os casos de interesse. Notemos que poucos países são membros de cada combinação. Tal observação é uma consequência direta da regra do valor mínimo ditada pelo operador E-lógico. A predominância de escores

zero nas conjunções propicia uma avaliação mais rápida da condição de suficiência, por conta do menor número de linhas a examinar. Identificamos os casos em que as condições combinadas estão presentes (com escore 1) e verificamos se o resultado também está presente ($R = 1$). Lembramos que não pode existir um único caso em que a condição combinada esteja presente, mas o resultado, ausente. Por exemplo, para a conjunção P^*E^*C há duas instâncias empíricas (Brasil e Bolívia), mas o resultado só está presente para o caso brasileiro. Portanto, esta conjunção P^*E^*C não é suficiente para o resultado bem-sucedido da política REDD+. Somente as condições combinadas P^*e^*C (Indonésia) e p^*E^*C (Vietnam) satisfazem a propriedade de suficiência.

Quando passamos da avaliação de conjuntos amplos, como as condições P , E e C , para conjuntos mais restritos, como as suas conjunções, refinamos as características de cada país (combinando mais conjuntos mediante o operador E-lógico); conseqüentemente, menos casos exibirão a conjunção de condições, aumentando a chance de que eles satisfaçam a propriedade de suficiência para o resultado. Enquanto a complexidade máxima (as conjunções específicas de condições), além de ser empírica e teoricamente menos interessante, não necessariamente permite as generalizações para outros casos; a complexidade mínima (condição causal isolada) em geral nem sempre é suficiente para um determinado resultado de interesse.

Podemos agora avaliar quais condições e combinações de condições são necessárias para o resultado REDD+ bem-sucedido. A condição de necessidade (capítulo 3) estabelece que uma determinada condição (ou conjunção de condições) é necessária se, sempre que o resultado está presente, a condição também está presente. Em outras palavras, o resultado não pode ser alcançado sem a presença da condição; nenhum caso com o resultado presente pode exibir uma condição ausente.

O procedimento para identificar quais são as condições necessárias para um determinado resultado é semelhante ao realizado para a condição de suficiência, sendo que, agora, é realizado para um grupo diferente de casos. Isso implica que só averiguamos a condição de necessidade para os

casos em que o resultado está presente. Avaliar a necessidade consiste em descobrir, caso haja, quais das condições estão presentes em todos esses casos. Para a condição de suficiência identificamos os casos em que cada condição estava presente e conferimos se o resultado estava ou não presente. Agora, identificamos os casos em que o resultado está presente e conferimos se a condição está ou não presente.

Para avaliarmos a necessidade das condições para o resultado, partimos das condições isoladas umas das outras (Tabela 14). Avaliemos a necessidade da condição P para o resultado R = 1. Portanto, só nos interessam os casos em que o resultado está presente (isto é, Indonésia, Vietnam e Brasil). A condição P está presente para a Indonésia, mas está ausente para o Vietnam. Portanto, a condição P não é necessária para o resultado. Para a condição E, ela está presente no Vietnam e no Brasil, mas ausente na Indonésia; por isso não é uma condição necessária para o resultado. Quanto à condição C, ela é necessária para o resultado, pois está presente em todos os casos em que o resultado está presente. Observemos que, na Bolívia, embora tenha a mesma configuração do Brasil, o resultado está ausente. Portanto, esse país é um caso que não entra no conjunto dos casos passíveis de averiguação da condição de necessidade.

Procedimento semelhante executamos para os complementos das condições (Tabela 15). Os casos de interesse continuam sendo Indonésia, Vietnam e Brasil. Averiguamos a presença das condições isoladas “p”, “e” e “c”. A condição “p” está ausente na Indonésia, logo ela não é uma condição necessária. O mesmo desfecho se dá para as condições “e” e “c”. Portanto, ambas não são necessárias. Assim sendo, apenas a condição C é necessária para a presença do resultado. Interessante aqui é que, para a implementação bem sucedida do REDD+, é necessário que o processo de mudança institucional pró-clima já tenha sido iniciado. Conforme dissemos no capítulo 3, a avaliação da condição de necessidade das conjunções de condições não faz sentido, porque a regra do mínimo reduz as chances de que a conjunção passe pelo critério de necessidade. Além disso, a lógica formal pura estabelece que nenhuma conjunção pode passar no critério de necessidade, desde que seja exclusivamente formada por condições

isoladas necessárias; e condições não necessárias não podem formar uma conjunção que seja necessária (SCHNEIDER; WAGEMANN, 2012).

Aqui somente a condição C é necessária. Portanto, não faz sentido investigar se as conjunções de condições se qualificam como necessárias. Contudo, o uso do operador OU-lógico (união de conjuntos) pode ser adequado para a averiguação da necessidade, já que ele segue a regra do máximo (o maior escore das condições que fazem parte de uma expressão OU-lógico determina o escore dos casos nessa união). Esse procedimento cria conjuntos a que mais casos tendem a pertencer (porque são superconjuntos). Isso aumenta a chance de que as uniões sejam superconjuntos do resultado. A Tabela 17 reporta os escores das uniões das condições isoladas e de seus complementos. Por exemplo, a união P+C (terceira coluna da tabela) para Moçambique tem escore 1 porque é o escore máximo para as condições isoladas P e C desse país ($P = 1$ e $C = 0$). Notemos que agora aparecem muitos escores 1.

Tabela 17 – Tabela dos escores das uniões de condições isoladas

Caso/País	Escore das Uniões									Resultado
	P+E	P+C	P+e	P+c	E+C	E+c	p+e	p+c	e+c	R
Brasil	1	1	1	1	1	1	0	0	0	1
Bolívia	1	1	1	1	1	1	0	0	0	0
Burkina Faso	1	1	1	1	0	1	1	1	1	0
Camarões	1	1	1	1	1	1	0	1	1	0
Congo	0	1	1	0	1	0	1	1	1	0
Indonésia	1	1	1	1	1	0	1	0	1	1
Moçambique	1	1	1	1	0	1	1	1	1	0
Nepal	1	1	0	1	1	1	1	1	1	0
Peru	0	1	1	0	1	0	1	1	1	0
Papua Nova Guiné	0	0	1	1	0	1	1	1	1	0
Tanzânia	1	1	1	1	0	1	1	1	1	0
Vietnam	1	1	0	0	1	1	1	1	0	1

Fonte: Elaboração própria da tabela verdade original de Korhonen-Kurki *et al.* (2014).

Aplicando-se o critério de necessidade podemos averiguar quais uniões são necessárias para o resultado. Basta verificar se as uniões estão presentes para os casos com $R = 1$. O critério é satisfeito pelas uniões P+E, P+C e E+C. Portanto, para uma bem sucedida implementação da

política REDD+ é necessário que haja escassez de recursos florestais OU presença de estrutura legal-política e de governança do uso florestal OU ambos; haja escassez de recursos florestais OU processo já iniciado de mudança institucional pró-clima OU ambos; haja presença de estrutura legal-política e de governança do uso florestal OU processo já iniciado de mudança institucional pró-clima OU ambos.

Em suma nenhuma das condições, e seus complementos, é suficiente para um resultado bem-sucedido na implementação da 3E-REDD+. Somente as condições combinadas P^*e^*C (Indonésia) e p^*E^*C (Vietnam) satisfazem a propriedade de suficiência. Quanto à propriedade de necessidade, somente a condição C é necessária; ao se fazer as uniões de condições isoladas, somente as uniões P+E, P+C e E+C podem ser ditas necessárias.

Lembramos que o objetivo desta seção foi de apenas apresentar de maneira prática a relação de suficiência e necessidade na variante *csQCA*. Essa análise foi realizada no âmbito da importância de política pública, especialmente para a temática aqui tratada. As medidas de consistência e cobertura não foram o foco, evidenciando uma forma alternativa na utilização dessa abordagem.

Referências bibliográficas

- AHN, S.; LEE, S. S. Explaining Korean welfare state development with new empirical data and methods. *Asian Social Work and Policy Review*, v. 6, n. 2, p. 67–85, 2012.
- ALAM, A.; BAGCHI, P. K. Supply chain capability of countries and its effect on foreign direct investment: a fuzzy-set analysis. *International Journal of Logistics Economics and Globalisation*, v. 5, n. 1, p. 59–80, 2013.
- BERG-SCHLOSSER, D. *et al.* Qualitative Comparative Analysis (QCA) as an Approach. *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques*, p. 1–18, 2008.
- BETARELLI JUNIOR, A. A.; SIMÕES, R. F. A dinâmica setorial e os determinantes locais das microrregiões paulistas. *Economia Aplicada*, v. 15, n. 4, p. 641–670, 2011. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-80502011000400006&lng=pt&nrm=iso&tIng=en>.
- BRADY, H. E. Causation and Explanation in Social Science. In: BOX-STEFFENSMEIER, J. M.; BRADY, H. E.; COLLIER, D. (Org.). *The Oxford Handbook of Political Methodology*. New York: The Oxford Handbook of Political Science, 2008. .
- BRASIL. MINISTÉRIO DO MEIO AMBIENTE (MMA). *O que é REDD+*. Disponível em: <<http://redd.mma.gov.br/pt/pub-apresentacoes/item/82-o-que-e-redd>>. Acesso em: 8 fev. 2018.

- BRAUMOELLER, B. F. Causal Complexity and the Study of Politics. *Political Analysis*, v. 11, n. 3, p. 209–233, 2003. Disponível em: <https://www.cambridge.org/core/product/identifier/S1047198700010317/type/journal_article>.
- BYRNE, D. Complexity, Institutions and Public Policy: Agile Decision-making in a Turbulent World - By Graham Room. *Social Policy & Administration*, v. 46, n. 7, p. 845–847, 2012. Disponível em: <<http://doi.wiley.com/10.1111/j.1467-9515.2012.00872.x>>.
- CARAMANI, D. *Introduction to the comparative method with Boolean algebra*. Los Angeles, United States: Sage publications, 2009.
- CLARK, W. R.; GILLIGAN, M. J.; GOLDBERGER, M. A simple multivariate test for asymmetric hypotheses. *Political Analysis*, v. 14, n. 3, p. 311–331, 2006.
- COOPER, B.; GLAESER, J. Analysing necessity and sufficiency with Qualitative Comparative Analysis: how do results vary as case weights change? *Quality and Quantity*, v. 50, n. 1, p. 327–346, 2016.
- CRONQVIST, L. *Tosmana: Tool for Small-N Analysis, Version 1.54*. Trier: University of Trier, 2017. Disponível em: <<http://www.tosmana.net>>.
- DELREUX, T. *Solving contradictory simplifying assumptions in QCA: representation of a new best practice*. Analysis, COMPASS Working Papers. Houston: Department of Sociology, University of Houston-Downtown, 2010. Disponível em: <https://dial.uclouvain.be/downloader/downloader.php?pid=boreal:87885&datastream=PDF_01>.
- EBBINGHAUS, B. Comparing Welfare State Regimes: Are Typologies an Ideal or Realistic Strategy? *ESPAAN, Edinburg, UK*, p. 1–20, 2012. Disponível em: <http://www.espanet2012.info/___data/assets/pdf_file/0005/89033/Ebbinghaus_-_Stream_2.pdf>.
- EBBINGHAUS, B. When less is more selection problems in large-N and small-N cross-national comparisons. *International Sociology*, v. 20, n. 2, 2005.
- EMMENEGGER, P.; SCHAFF, D.; WALTER, A. *QCA, the Truth Table Analysis and Large-N Survey Data: The Benefits of Calibration and the Importance of Robustness Test*. COMPASS Working Paper 2014-79, nº 79. Houston: Department of Sociology, University of Houston-Downtown, 2014. Disponível em: <<http://www.compass.org/wpseries/EmmeneggerSchraffWalter2014.pdf>>.
- EUGÊNIO, A. S.; BETARELLI JUNIOR, A. A. Impactos na satisfação com a vida: Uma abordagem por Análise Comparativa Qualitativa. 2016, Aracajú: XIV Encontro Nacional da Associação Brasileira de Estudos Regionais e Urbanos, 2016.
- FARIA, W. R.; BETARELLI JUNIOR, A. A.; MONTENEGRO, R. L. G. Multidimensional characteristics and deforestation: an analysis for the Brazilian Legal Amazon. 2017, Natal: 45º Encontro Nacional de Economia (ANPEC), 2017.
- FERREIRA, S. DE F. *Transgressão, norma social e crime: o papel da dissuasão social nas áreas mínimas comparáveis no Brasil (1991-2010)*. 2017. Tese (Doutorado em Economia) – Programa de Pós-graduação em Economia (PPGE), Universidade Federal de Juiz de Fora (UFJF), 2017.
- GANTER, A.; HECKER, A. Configurational paths to organizational innovation: Qualitative comparative analyses of antecedents and contingencies. *Journal of Business Research*, v. 67, n. 6, p. 1285–1292, 2014.
- GOERTZ, G. Assessing the trivialness, relevance and relative importance of necessary or sufficient conditions in social science. *Studies in Comparative International Development*, v. 41, n. 2, p. 88–109, 2006.

- GOERTZ, G.; MAHONEY, J. Two-level theories and fuzzy-set analysis. *Sociological Methods and Research*, v. 33, n. 4, p. 497–538, 2005.
- GURGEL, M. C. B. DO A. *Configurações estratégicas de empresas de alto desempenho: análise qualitativa comparativa das imperfeições de mercado por elas exploradas*. 2011. Tese (Doutorado em Economia) - Fundação Getúlio Vargas, Escola de Administração de Empresas de São Paulo, 2011.
- HUANG, R. *QCA3: Yet another Package for Qualitative Comparative Analysis*. . [S.l.]: R Package Version 0.0-7, 2014. Disponível em: <<http://cran.r-project.org/package=QCA3>>.
- HUDSON, J.; KÜHNER, S. Qualitative comparative analysis and applied public policy analysis: New applications of innovative methods. *Policy and Society*, v. 32, n. 4, p. 279–287, 2013.
- KORHONEN-KURKI, K. et al. Enabling factors for establishing REDD+ in a context of weak governance. *Climate Policy*, v. 14, n. 2, p. 167–186, 2014.
- LEE, S. *A Critique of the Fuzzy-set Methods in Comparative Social Policy*. , COMPASS Working Paper., nº 53. Houston: Department of Sociology, University of Houston-Downtown, 2008. Disponível em: <<http://www.compass.org/wpseries/Lee2008.pdf>>.
- LEE, S. S. Fuzzy-set method in comparative social policy: a critical introduction and review of the applications of the fuzzy-set method. *Quality & Quantity*, v. 47, n. 4, p. 1905–1922, 2013. Disponível em: <<http://link.springer.com/10.1007/s11135-011-9633-8>>.
- LEGEWIE, N. An Introduction to Applied Data Analysis with Qualitative Comparative Analysis (QCA). *Forum: Qualitative Social Research*, v. 14, n. 3, p. 1–45, 2013. Disponível em: <<http://www.qualitative-research.net/index.php/fqs/article/view/1961/3594>>.
- LONGEST, K. C.; VAISEY, S. fuzzy : A program for performing qualitative comparative analyses (QCA) in Stata. *Stata Journal*, n. 1, p. 79–104, 2008a.
- LONGEST, K. C.; VAISEY, S. Fuzzy: A program for performing qualitative comparative analyses (QCA) in Stata. *Stata Journal*, v. 8, n. 1, p. 79–104, 2008b.
- MACKIE, J. L. CAUSES AND CONDITIONS. *American Philosophical Quarterly*, v. 2, n. 4, p. 245–264, 1965. Disponível em: <<http://www.jstor.org/stable/20009173>>.
- MAHONEY, J.; KIMBALL, E.; KOIVU, K. L. The logic of historical explanation in the social sciences. *Comparative Political Studies*, v. 42, n. 1, p. 114–146, 2009.
- MARX, A. Towards more robust model specification in QCA: Results from a methodological experiment. *COMPASS-working paper*, n. 2005, p. 1–25, 2006.
- MELO, C. R. The 2014 elections and the Brazilian party system. *Brazilian Political Science Review*, v. 9, n. 1, p. 93–114, 2015.
- MEUER, J.; RUPIETTA, C.; BACKES-GELLNER, U. Layers of co-existing innovation systems. *Research Policy*, v. 44, n. 4, p. 888–910, 2015.
- MIETHE, T. D.; DRASS, K. A. Exploring the Social Context of Instrumental and Expressive Homicides: An Application of Qualitative Comparative Analysis. *Journal of Quantitative Criminology*, v. 15, n. 1, p. 1–21, 1999.
- MILL, J. S. *1843 A System of Logic Ratiocinative and Inductive*. Hawaii: University Press of the Pacific, 2002.
- MONTENEGRO, R. L. G. *Inovações ambientais e sistemas nacionais de inovação: caracterizações para o período 1990-2010*. 2016. Tese (Doutorado em Economia) – Centro de Desenvolvimento e Planejamento Regional da Universidade Federal de Minas Gerais (CEDEPLAR/UFMG), 2016.

- OLSEN, W.; NOMURA, H. Poverty Reduction. *Sociological Theory and Methods*, v. 24, n. 2, p. 1–31, 2009. Disponível em: <https://www.jstage.jst.go.jp/article/ojams/24/2/24_2_219/_pdf/-char/ja>.
- RAGIN, C. Fuzzy sets: calibration versus measurement. *Methodology volume of Oxford handbooks of ...*, p. 1–31, 2007. Disponível em: <[http://toktok.persiangu.com/other/Calibration fuzzy measurement.pdf](http://toktok.persiangu.com/other/Calibration%20fuzzy%20measurement.pdf)>.
- RAGIN, C. *Redisigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago: University of Chicago Press, 2008.
- RAGIN, C. C. *Fuzzy-Set Social Science*. Chicago: University of Chicago Press., 2000.
- RAGIN, C. C. Measurement Versus Calibration: A Set-Theoretic Approach. In: BOX-STEFFENSMEIER, J. M.; BRADY, H. E.; COLLIER, D. (Org.). *The Oxford Handbook of Political Methodology*. Oxford: Oxford University Press, 2008. .
- RAGIN, C. C. Set relations in social research: Evaluating their consistency and coverage. *Political Analysis*, v. 14, n. 3, p. 291–310, 2006.
- RAGIN, C. C. *The Comparative Method: moving Beyond Qualitative and Quantitative Strategies*. Oakland: University of California Press, 2014.
- RAGIN, C. C. *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*. Berkeley, CA: University of California Press, 1987. Disponível em: <<http://www.jstor.org/stable/2579563>>.
- RAGIN, C. C.; DRASS, K. A.; DAVEY, S. *Fuzzy-Set/Qualitative Comparative Analysis, fs/QCA [Computer Programme], Version 3.0*. Irvine, CA: University of California., 2017. Disponível em: <<http://www.fsqca.com>>.
- RAGIN, C. C.; SONNETT, J. Between Complexity and Parsimony: Limited Diversity, Counterfactual Cases, and Comparative Analysis. *Vergleichen in der Politikwissenschaft*. [S.l: s.n.], 2005. p. 180–197. Disponível em: <http://link.springer.com/10.1007/978-3-322-80441-9_9>.
- RIHOUX, B. *et al.* From Niche to Mainstream Method? A comprehensive mapping of QCA applications in journal articles from 1984 to 2011. *Political Research Quarterly*, v. 66, n. 1, p. 175–184, 2013.
- RIHOUX, B. Qualitative Comparative Analysis (QCA) and related techniques: recent advances and challenges. In: PICKEL, S. *et al.* (Org.). *Methoden der vergleichenden Politik- und Sozialwissenschaft*. Wiesbaden: VS Verlag, 2009. p. 365–385. Disponível em: <http://dx.doi.org/10.1007/978-3-531-91826-6_18>.
- RIHOUX, B.; GRIMM, H. *Innovative comparative methods for policy analysis: Beyond the quantitative-qualitative divide*. New York: Springer, 2006.
- RIHOUX, B.; RAGIN, C. C. *Configurational Comparative Methods: QCA and Related Techniques*. Thousand Oaks, CA and London: Sage, 2009.
- ROHLFING, I. Why simulations are appropriate for evaluating Qualitative Comparative Analysis. *Quality and Quantity*, v. 50, n. 5, p. 2073–2084, 2016.
- SANDES-FREITAS, V.; BIZZARRO-NETO, F. Qualitative Comparative Analysis (QCA): usos e aplicações do método. *Revista Política Hoje*, v. 24, n. 2, p. 103–118, 2015.
- SANDES-FREITAS, V. E. *Alianças partidárias nos estados brasileiros: das coligações às coalizões de governo (1986-2010)*. 2015. Tese de doutorado em ciência política, Campinas, IFCH, Unicamp., 2015.
- SCHNEIDER, C. Q.; GROFMAN, B. It might look like a regression equation... but it's not! An intuitive approach to the presentation of QCA and FS/QCA results. 2006, [S.l: s.n.], 2006. p. 15–17.

- SCHNEIDER, C. Q.; MAKSZIN, K. Forms of welfare capitalism and education-based participatory inequality. *Socio-Economic Review*, v. 12, n. 2, p. 437–462, 2014.
- SCHNEIDER, C. Q.; WAGEMANN, C. *Set-Theoretic Methods for the Social Sciences: A Guide to Qualitative Comparative Analysis*. United Kingdom: Cambridge University Press, 2012.
- SCHNEIDER, C. Q.; WAGEMANN, C. Standards of Good Practice in Qualitative Comparative Analysis (QCA) and Fuzzy-Sets. *Comparative Sociology*, v. 9, n. 3, p. 397–418, 2010.
- SILVA, M. V. A. E; BETARELLI JUNIOR, A. A. Configurações gerais e específicas das eleições-2014 nos municípios brasileiros. 2016, Curitiba: XIII Encontro Nacional da Associação Brasileira de Estudos Regionais e Urbanos - XIII ENABER, 2016. Disponível em: <<http://abep.org.br/xxencontro/files/paper/602-688.pdf>>.
- SMITHSON, M.; VERKUILEN, J. *Fuzzy set theory: applications in the social sciences*. [S.l.: s.n.], 2006. Disponível em: <<http://www.loc.gov/catdir/toc/ecip061/2005028893.html%5Cnhttp://www.loc.gov/catdir/enhancements/fy0658/2005028893-d.html>>.
- THIEM, A. *A ranking of the most QCA-friendly journals: some observations*. Disponível em: <<http://alrik-thiem.net/blog/ranking-most-qca-friendly-journals-some-observations/>>. Acesso em: 7 out. 2015.
- THIEM, A. *Set-relational fit and the formulation of transformational rules in fsQCA*. , COMPASS Working Paper., nº 2010–61. Houston: Department of Sociology, University of Houston-Downtown, 2010. Disponível em: <<http://www.compass.org/wpseries/Thiem2010.pdf>>.
- THIEM, A.; DUSA, A. QCA : A Package for Qualitative Comparative Analysis. *The R Journal*, v. X, p. 1–7, 2012a.
- THIEM, A.; DUSA, A. *Qualitative Comparative Analysis with R: A User's Guide*. New York: Springer, 2012b.
- VARONE, F.; RIHOUX, B.; MARX, A. A new method for policy evaluation? longstanding challenges and the possibilities of Qualitative Comparative Analysis (QCA). *Innovative Comparative Methods for Policy Analysis: Beyond the Quantitative-Qualitative Divide*. [S.l.: s.n.], 2006. p. 213–236.
- VINK, M. P.; VAN VLIET, O. Not quite crisp, not yet fuzzy? assessing the potentials and pitfalls of multi-value QCA. *Field Methods*, v. 21, n. 3, p. 265–289, 2009.
- VIS, B. The Comparative Advantages of fsQCA and Regression Analysis for Moderately Large-N Analyses. *Sociological Methods and Research*, v. 40, n. 1, p. 168–198, 2012.
- WAGEMANN, C.; SCHNEIDER, C. Q. Qualitative comparative analysis (QCA) and fuzzy-sets: Agenda for a research approach and a data analysis technique. *Comparative Sociology*, v. 9, n. 3, p. 376–396, 2010.
- XAVIER, C. D. DE O. *Os determinantes político-institucionais do desenvolvimento financeiro: uma análise QCA dos países emergentes de renda média alta*. 2016. (Dissertação de mestrado) Instituto de Relações Internacionais, São Paulo, Universidade de São Paulo, São Paulo, 2016.

Esta obra foi impressa pela Imprensa Nacional
SIG, Quadra 6, Lote 800
70610-460 – Brasília-DF