USC-SIPI REPORT #431

A New Methodology for Calibrating Fuzzy Sets in fsQCA Using Level 2 and Interval Type-2 Fuzzy Sets

by

Jerry M. Mendel and Mohammad M. Korjani

June 2015

Signal and Processing Institute
UNIVERSITY OF SOUTHERN CALIFORNIA
Viterbi School of Engineering
Ming Hsieh Department of Electrical Engineering
3740 McClintock Avenue, Suite 400
Los Angeles, CA 90089-2564 U.S.A.

ABSTRACT

This report provides a new methodology for calibrating the fuzzy sets that are used in fsQCA, one that is based on clearly distinguishing between a linguistic variable and the linguistic terms for that variable. The resulting fuzzy sets are *reduced-information level 2 fuzzy sets* (RI L2 FSs). The major steps for obtaining the RI L2 FSs are: (1) For each linguistic variable, a vocabulary of naturally ordered linguistic terms (words) are chosen; (2) Interval end-point data are collected for each of the linguistic variables, either from a group of subjects or from one expert; (3) The data for each word are mapped into the footprint of uncertainty (FOU) of an interval type-2 fuzzy set (IT2 FS) using the HM approach [10]; (4) An RI L2 FS is created by replacing each word with an uncertainty measure and choosing an appropriate membership grade for it; and, (5) The MF of the RI L2 FS is approximated so that the resulting MF is for $x \in X$. The resulting approximated RI L2 FS MF is for the linguistic variable, and is not the MF of an ordinary FS but instead is the MF of a level 2 FS, one that has an S-shape, the kind of shape that is so widely used by fsQCA scholars, and is so important to fsQCA.

This report also applies its new calibration methodology to Ragin's Breakdown of Democracy example, using new data provided to us by him, and demonstrates that we are able to obtain his earlier solutions using RI L2 FSs in either type-1 or interval-valued (IV) fsQCA, something that should be reassuring to fsQCA scholars. It also studies the robustness of fsQCA to MF breakpoint location uncertainties as well as to membership grade uncertainties. Finally, because the S-shaped MFs are derived from FOUs for all of the linguistic variable's terms, this paper shows how to obtain more precise statements of fsQCA causal combinations for their best instances, something that may be of added value to practitioners of fsQCA.

Table of Contents

Section title **page**

List of Figures

Figure **Page**

List of Tables

Table **Page**

I. INTRODUCTION

As is stated in [34]: Fuzzy Set Qualitative Comparative Analysis (fsQCA), developed by the social scientist Charles Ragin [40–42, 44, Ch. 5], is a methodology for obtaining linguistic summarizations from data that are associated with cases[1]. Unlike more quantitative methods that are based on correlation, fsQCA seeks to establish logical connections between combinations of causal conditions (*conjunctural causation*) and a desired outcome, the result being rules that summarize the sufficiency between subsets of all of the possible combinations of the causal conditions (or their complement) and the desired outcome. The rules are connected by the word OR to the desired outcome. Each rule is a possible path from the causal conditions to the desired outcome and represents *equifinal causation*, i.e. different causal combinations leading to the same desired outcome.

The steps of fsQCA are summarized and explained in [33], and it is not necessary for the reader to understand all of them to read the present paper. The first four steps are very easy to understand and are: 1) Choose a desired outcome and associated cases; 2) Postulate k causal conditions; 3) Treat the desired outcome and causal conditions as fuzzy sets, and determine membership functions (MFs) for all of them; and, 4) Evaluate these MFs for all available cases, obtaining *derived MFs*. This paper focuses mainly on steps 3 and 4, which together we refer to as "calibrating the fuzzy sets." The rest of fsQCA does not change.

The main calibration method that is presently used is Ragin's *direct method* about[2] which Ragin [42, Ch. 5] states:

> Fuzzy sets are calibrated using external criteria, which in turn must follow from and conform to the researcher's conceptualization, definition, and labeling of the set in question. Using the … *direct* method, the researcher specifies the values of an interval scale that corresponds to the three qualitative breakpoints[3] that structure a fuzzy set: full membership, full non-membership, and the crossover point[4]. These three benchmarks are then used to transform the original interval-scale values to fuzzy membership scores. … The end product of this method is the fine-grained calibration of the degree of membership of cases in sets, with scores ranging from 0.0 to 1.0.

In this report we will explain why this calibration method, or any of the other calibration methods that are being used by fsQCA scholars, must be applied with great care in order for their results to actually correspond to fuzzy sets, and that many times (depending, as explained below, upon the wording of the causal condition) they do not lead to fuzzy sets at all, even though users think they do, which calls into question the validity of fsQCA since it is built upon fuzzy sets[5]. In

---

[1] In the rest of this report we use the words: *variable*, *terms*, *causal conditions*, *causal combinations* and *desired outcome*. A *variable* is something that can be measured. Each variable is linguistically described by one or more *terms* that are modeled using fuzzy sets. The terms for each variable that are used as the antecedents of fsQCA rules are called *causal conditions*. In fsQCA the antecedents are combined using the word AND that is modeled using the minimum operation. Such a combination of causal conditions is called a *causal combination*. The term that is used as the consequent of the fsQCA rules is called *desired outcome*.

[2] Of course other methods can be used for calibration, including Ragin's *indirect method* [36], but, regardless of which method is used, the criticism that is given below applies to all of them.

[3] The direct method is not limited to three breakpoints, e.g. [41, p. 156] explains it for three, five and seven breakpoints, [42, p. 88] explains it for seven breakpoints, and [45, p. 29] explains it for ten breakpoints; however, most researchers only use three breakpoints because it becomes more and more difficult for a person to specify five, seven or ten numerical breakpoints.

[4] Equivalent terms for *crossover point* [42, p. 88] are *neither in nor out* [45, p. 29] and *not fully out nor fully in* [41, p. 156].

[5] We were inspired to reach this conclusion by the criticism of Ragin's calibration methods that are in [23].

order to understand this strong (some might even say shocking) criticism better we first return to some history and definitions about fuzzy sets and linguistic variables, much of which is well known to fuzzy set scholars but may not be as well known to scholars in the fsQCA community.

A. Fuzzy Sets and Linguistic Variables

In 1965, when Zadeh introduced a fuzzy set [56] it was a new and novel mathematical construct in which an object could simultaneously reside in more than one set but to different degrees of membership; however, such a set was not yet connected to what he later called a *linguistic variable*, something that occurred only around 1971 in [57, 58] (and then in the more widely referenced [59, 60]). Because the distinction between a linguistic variable and a fuzzy set is crucial to understanding our criticism, a few definitions are presented next.

Definition 1. A *fuzzy set* (FS) A is comprised of a domain X of real numbers (also called the *universe of discourse* of A) together with a *membership function* (MF) $\mu_A : X \to [0,1]$. For each $x \in X$, the value of $\mu_A(x)$ is the *degree of membership*, or *membership grade*, of x in A. If $\mu_A(x) = 1$ or $\mu_A(x) = 0$ for $x \in X$, then the FS A is said to be a *crisp set*.

In this paper we shall refer to A as a *type-1 FS*, because later we shall also use *interval type-2 FSs* (IT2 FSs) for which MF values are no longer single numbers but instead are intervals of numbers that allow for MF uncertainties.

Definition 2. If a variable can take words in natural languages as its values, it is called a *linguistic variable*, where the words are characterized by FSs defined in the universe of discourse in which the variable is defined [49].

Definition 3. Each *linguistic variable* [20, 59, 60] is fully characterized by a quintuple (v, T, X, g, m) in which v is the name of the variable, T is the set of linguistic terms[6] of v that refer to a base variable whose values range over the universal set X, g is a syntactic rule for generating linguistic terms, and m is a semantic rule that assigns each linguistic term $t \in T$ its meaning, $m(t)$, which is a fuzzy set on X, that is, $m : T \to F(X)$, where $F(X)$ denotes the set of fuzzy sets of X, one fuzzy set for each $t \in T$. It is common to refer to v as the linguistic variable.

Example 1. Some examples of linguistic variables, v, are: Developed Country, Urban Country, Industrial Country, Stable Country, Profitable Company, Institutional Veto Points, All-day

[6] Although "term" means one or more words, it is quite common in the FS literature to see "word" used instead of "term," even when a term includes more than one word. In this paper, we also interchange "term" and "word."

School Systems, Horsepower, Acceleration, Production Rate, etc. Some examples of the set of linguistic terms, T, for these linguistic variables are[7]:

1. For Developed, Urban, Industrial, or Stable (Country), and Profitable (Company), $T \triangleq$ {*Barely, Hardly, Somewhat, Moderately, Fully, Extremely*}
2. For Institutional Veto Points and All-day School Systems, $T \triangleq$ {*None to Very Few, Some, A Moderate Amount of, Many, A Large Number of, A Very Large Number of*}
3. For Horsepower, Acceleration and Production Rate, $T \triangleq$ {*Very Low, Low, Moderate, High, Very High*}

Observe that the linguistic terms must make linguistic sense for its linguistic variable, which is where g in Definition 3 comes into play; so, for example, *Somewhat Acceleration* makes no linguistic sense nor does *Very High All-day School Systems*. Note, also, that *it is the elements of T that are treated as fuzzy sets,* and, of course, each of these fuzzy sets is described by a MF.

B. fsQCA Criticism

The concepts of linguistic variable and its linguistic terms seem to be missing explicitly in the fsQCA literature[8]. This is of crucial importance because it is the linguistic terms that must be calibrated and not the linguistic variable. Unfortunately, both of Ragin's calibration methods [42] (the direct and indirect methods) are widely used by many fsQCA adherents to calibrate the linguistic variable.

Fuzzy sets are meant for linguistic terms of a linguistic variable that are *naturally ordered*. Examples of variables that are (are not) naturally ordered are temperature, pressure, height, profit, etc. (beautiful, ill, happy, etc.). It should be clear from the given examples that variables that are "naturally ordered" means variables whose domains are naturally ordered sets, i.e. sets that are equipped with a linear order relation[9]. An example of such a set is the set of real numbers. Ragin's fuzzy-set scoring methods are fuzzy methods only when they are applied to the linguistic terms of naturally ordered linguistic variables.

fsQCA frequently requires a MF for a naturally ordered linguistic variable. In this paper we explain how such a MF can be obtained, even though at this point this sounds contradictory to what we have just explained about MFs being for linguistic terms.

---

[7] Because some of the linguistic terms may be so similar to each other, it may not be necessary to use all of them. One usually chooses the linguistic terms so that their membership functions overlap and cover X.

[8] There is no reference to [57–60] in Ragin's books [40–42, 44] or in the recent book [45]. Interestingly, though, the examples in [45] use one linguistic term as a causal condition for each linguistic variable, whereas the examples in Ragin's just referenced books do not use linguistic terms, but instead use linguistic variables. Recently, however, Ragin [43] stated: "Sets and set labels almost always involve adjectives in some way. For example, it is possible to assess degree of membership in heavy; it is not possible to assess degree of membership in weight. This aspect of sets is fundamental to set-theoretic analysis."

[9] A *linear order relation* is a binary relation that is transitive, anti-symmetric, and total.

C. Size of a Vocabulary and Calibration

Importantly, the size of the vocabulary (i.e., the number of linguistic terms in *T*) for a linguistic variable may affect the calibration of the fuzzy sets. If, for example, only two linguistic terms are used to describe Profitable, namely *Hardly Profitable* and *Fully Profitable*, then their fuzzy sets will look very different from their fuzzy sets when the Example 1 five terms are used for Profitable, because the term *Barely Profitable* now appears before *Hardly Profitable*, and the term *Extremely Profitable* now appears after *Fully profitable*. Except for [36, pp. 41-42], there does not seem to be any discussion about this in the fsQCA literature.

D. When Present Implementations of fsQCA are Correct

As long as the causal conditions and desired outcome in fsQCA are treated properly as linguistic terms (and not as linguistic variables), the size of the vocabulary for a linguistic variable is known during calibration, and it is the linguistic terms that are calibrated, then implementing present versions of fsQCA for an application are okay.

Note that Ragin's direct method may still be used to calibrate the most left- and right-sounding linguistic terms, such as *Very Low* and *Very High*, because those membership functions are[10] S-shaped (Fig. 1) and can therefore be constructed using three (or more) breakpoints; however, it cannot be used to directly calibrate the linguistic terms that lie between *Very Low* and *Very High*, because those terms are not modeled using S-shaped membership functions (see [36, p. 41]).

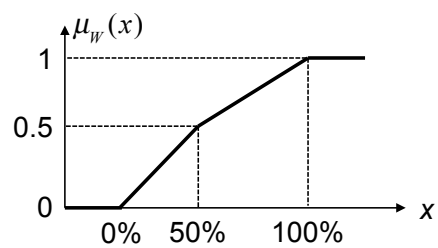

Fig. 1. A representative S-shaped MF. Note that 0%, 50% and 100% denote the locations of *x* that indicate fully-out, neither in nor out (crossover point) and fully-in membership, respectively; and, that sometimes [42] the 0% and 100% breakpoints are located at membership grades of 0.05 and 0.95, respectively, when a log-odds function is used to mathematically describe the MF (see Section 3.2.4 (b) and footnote 22).

E. Robustness to Measurement Errors and Calibration

The robustness of fsQCA to measurement errors and to the entire calibration process are serious concerns to fsQCA practitioners. Presently, fsQCA makes no direct allowances for measurement errors or for preciseness of the choices for MF breakpoints. Some excellent discussions about robustness are in [45, pp. 284-295]. We also address robustness in this report.

[10] The importance of an S-shaped MF for fsQCA is discussed in Section 3.1.

F. Organization of Rest of Report

The rest of this report is organized as follows: Section 2 describes a new calibration methodology, i.e. a new way of collecting data from one or more subjects about the linguistic terms that are associated with a linguistic variable and mapping that data into interval type-2 fuzzy sets; Section 3 explains the importance of the S-shaped MF to fsQCA, introduces a level 2 FS as a FS model for a linguistic variable, and provides a methodology for obtaining a kind of S-shaped (type-1 or interval-valued) level 2 MF for a linguistic variable; Section 4 re-examines Ragin's well-studied Breakdown of Democracy example using new data provided to us by Prof. Ragin; Section 5 examines the robustness of fsQCA to different kinds of uncertainties; Section 6 explains how more precise statements of fsQCA causal combinations can be obtained as a byproduct of our new calibration methodology; Section 7 provides some discussions about how using the S-shaped MFs for the RI L2 FSs obtained in this paper, overcomes past challenges to using fsQCA; and, Section 8 draws conclusions and proposes some future research directions.

II. A New Way to Calibrate Fuzzy Sets for fsQCA

A. Introduction

We require that each linguistic variable used in fsQCA have a vocabulary of linguistic terms assigned to it by one or more experts (practitioners), where the linguistic terms fit a naturally ordered scale, and there can be as many linguistic terms for each linguistic variable, as desired. The size of the vocabulary must be made known to the subjects (or single subject) during the collection of the data from them during the calibration process.

We strongly believe that a FS model for a linguistic term must capture the linguistic uncertainties of that term, so that those uncertainties can flow through fsQCA computations, just as unpredictable uncertainties, that are modeled using probability, flow through probability computations.

There are two kinds of uncertainties associated with a linguistic term (e.g., [28, 37]) (1) *intra-uncertainty*, which is the uncertainty about the linguistic term that each individual has about it; and (2) *inter-uncertainty*, which is the uncertainty about the linguistic term that a group of individuals has about it. These uncertainties are associated with the maxim (e.g., [25-28]) *words mean different things to different people*, which has been the rationale for modeling a linguistic term as an IT2 FS rather than as a T1 FS.

B. IT2 FSs

Because IT2 FSs may be unfamiliar to fsQCA scholars, we first present a few definitions and some discussions about them [27, 29, 37]. Readers who are already familiar with IT2 FSs can skip this section.

Definition 4. An *IT2 FS* denoted $\tilde{A}$, is characterized by an IT2 MF $\mu_{\tilde{A}}: X \to D[0,1]$, where $D[0,1]$ is the set of closed subintervals of $[0,1]$, i.e.[11]

$$\tilde{A} = \{((x,u), \mu_{\tilde{A}}(x,u) = 1) \mid x \in X, u \in [0,1]\} \tag{1}$$

x and u are called the *primary* and *secondary* variables of $\tilde{A}$, respectively, and $\mu_{\tilde{A}}(x,u)$ is a 3D MF. The numerical value of $\mu_{\tilde{A}}(x,u)$ is called the *secondary grade* of $\tilde{A}$; however, because the secondary grade of an IT2 FS is always 1, it plays no role of importance[12] for such a T2 FS.

Definition 5. The *footprint of uncertainty* of IT2 FS $\tilde{A}$, $FOU(\tilde{A})$, is:

$$FOU(\tilde{A}) = \left\{(x,u) \mid x \in X \text{ and } u \in [\underline{\mu}_{\tilde{A}}(x), \overline{\mu}_{\tilde{A}}(x)]\right\} \tag{2}$$

[11] Other kinds of IT2 FSs are described in [4].

[12] In a general T2 FS (GT2 FS) $\mu_{\tilde{A}}(x,u)$ can be different for $x \in X$ and $u \in [0,1]$, in which case the third dimension of a GT2 FS is very important.

It is a closed region that is bounded from below by the *lower MF* of $\tilde{A}$ [$LMF(\tilde{A}) \equiv \underline{\mu}_{\tilde{A}}(x)$] and from above by the *upper MF* of $\tilde{A}$ [$UMF(\tilde{A}) \equiv \overline{\mu}_{\tilde{A}}(x)$], where [1]

$$\underline{\mu}_{\tilde{A}}(x) = \inf\{u \mid u \in [0,1], x \in X, \mu_{\tilde{A}}(x,u) > 0\} \tag{3}$$

$$\overline{\mu}_{\tilde{A}}(x) = \sup\{u \mid u \in [0,1], x \in X, \mu_{\tilde{A}}(x,u) > 0\} \tag{4}$$

Definition 6. At each $x \in X$, J_x is called the *primary membership* of $\tilde{A}$, where

$$J_x = \{(x,u) : u \in [0,1], \mu_{\tilde{A}}(x,u) > 0\} \tag{5}$$

Example 2. Fig. 2 depicts the MFs for crisp sets and T1 FSs, and the FOUs for IT2 FSs for three linguistic terms associated with the variable *Literacy—Low Literacy* (*L*), *Moderate Literacy* (*M*) and *High Literacy* (*H*). The MFs (FOUs) for *L* and *H* are called *shoulder* MFs (FOUs), whereas the MF for *M* is called an *interior* MF (FOU). Sometimes the FOU of each IT2 FS is explained as a blurred version of the MF of each of its respective T1 FSs. Observe in Fig. 2c, for an IT2 FS, that at each *x* its primary memberships may be single points (the flat spots at which the membership grades equal 1), indicating no uncertainty about them, or an interval of values, indicating uncertainty about them. Fig. 3 illustrates $LMF(\tilde{A})$, $UMF(\tilde{A})$, $FOU(\tilde{A})$ and J_x for an interior FOU.

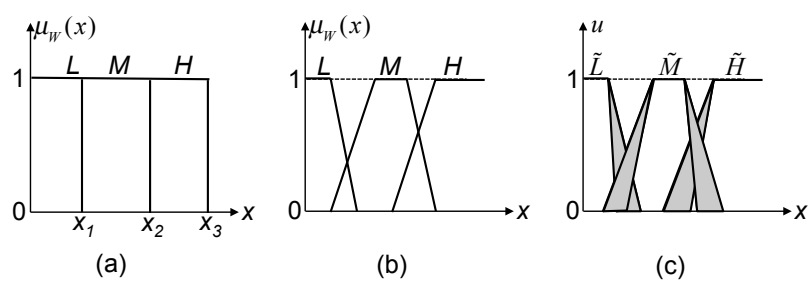

Fig. 2. (a) Crisp MF, (b) type-1 MF and (c) FOUs for the linguistic variable *Literacy* that is described by three terms, *Low Literacy* (*L*), *Moderate Literacy* (*M*) and *High Literacy* (*H*).

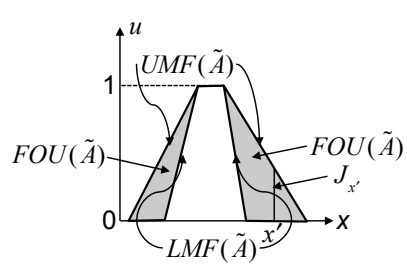

Fig. 3. IT2 FS, $\tilde{A}$, along with its lower and upper MFs, FOU and a primary membership. The flat spot where $u = 1$ is shared by both $LMF(\tilde{A})$ and $UMF(\tilde{A})$.

C. IT2 FS as a Model for a Linguistic Term

It is said (e.g., [28, 37]) that an IT2 FS is a *first-order uncertainty model* for a linguistic term[13] because its grade of membership is the same for the entire FOU. As is explained below, data that are collected from one or more subjects about a linguistic term will be mapped into an IT2 FS, such as any of the ones that are, e.g., shown in Fig. 4 for five generic linguistic terms (words): *Very Low, Low, Moderate, High* and *Very High*. Observe that there are only three kinds of FOUs, Left-shoulder (W_1), Interior (W_2, W_3 and W_4) and Right-shoulder (W_5). The flat spots are associated with the portions of the collected data that are in total agreement across all of the subjects. Note that, in general, there is more linguistic uncertainty about a word that is located near the middle of a scale than at the two ends of the scale, and that uncertainty is usually asymmetric.

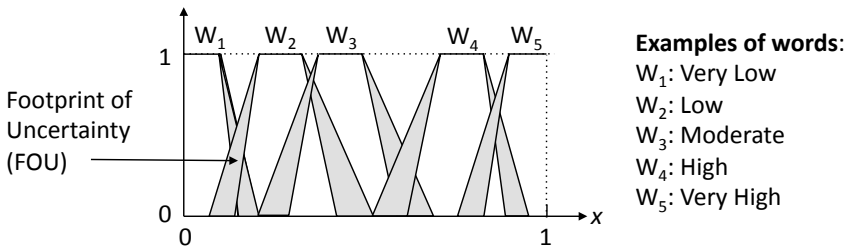

Fig. 4. IT2 FS models for five linguistic terms W_1–W_5. Here x has been normalized to [0, 1], but in general such normalization is unnecessary.

Even though a linguistic variable may reside on a natural scale, sometimes that scale is normalized to [0, 1] or [0, 10] without any loss of generality. Because the normalization constant is known for each linguistic variable, it is straightforward to go back and forth between the normalized and un-normalized scales if one wishes to do this.

D. Mapping Word Data into an FOU

A number of methods have been published on how to map data collected from a group of subjects into the FOU of a word. The Interval Approach (IA) [24], [37, Ch. 3] was the first such method, but it was replaced by the Enhanced Interval Approach (EIA) [53], which makes use of more information from the collected data than does the IA, and also corrects a small problem about the way in which the parameters of the LMF of an FOU are computed. More recently, the HM method was developed [9, 10]; it makes use of even more information from the collected data than does the EIA. The FOUs that are found from the HM method look like the ones in Fig. 4; but, the FOUs that are found from the EIA (or IA) approach do not look like those FOUs, because the EIA does not make explicit use of a data region about which there is total agreement

[13] A *GT2 FS* would be called a *second-order uncertainty model* for a word because its grades of membership vary over its FOU. As of the year 2015, it is not known what data to collect from a group of subjects in order to model a linguistic term as a GT2 FS (e.g., see [3] for a way to model hedged words using GT2 FSs).

across all of the subjects. Consequently, only the HM method is used in this paper[14]. It is reviewed in Appendix A.

The IA, EIA and HM method were all developed assuming that data can be collected from a *group* of subjects, and that such a group is available. Sometimes such a group is not available and only one knowledgeable expert is available, e.g. Prof. Ragin. Mendel and Wu [38] have presented a different way to collect data from one person (it can also be used by a group of subjects), after which that data can be used to generate even more data for what might be called a "virtual group of subjects." The HM method is then directly applied to that larger set of data. The way for going from data that are collected from a single subject to using that data in the HM method is also reviewed in Appendix A.

E. Data Collection and Calibration

It is very important that methods for collecting data from a group of subjects or even from an individual should *not introduce methodological uncertainties* into the data collection procedure. Most people do not know what a fuzzy set is, and so a method that asks an individual to provide a membership function (or FOU) for a linguistic term has methodological uncertainty associated with it that becomes co-mingled with the linguistic term's uncertainty, and the two kinds of uncertainty cannot be separated. Consequently, *we do not advocate asking subjects to provide MFs or FOUs.*

When data are collected from a group of n subjects, they are asked a question like [9, 10, 24, 37, 53]: *Suppose that a word[15] can be located on a scale of l to r, and you want to locate the end-points of the interval that you associate with the word on that scale. On the scale of l to r, what are the endpoints that you associate with the left [right] end-point of the word?* We have administered this kind of survey many times and found that most people have no trouble in answering this question. For each word, the ith subject provides interval endpoints $a^{(i)}$ and $b^{(i)}$, and the group of n subjects provides $[a^{(i)}, b^{(i)}]_{i=1}^{n}$.

Example 3. In [9] the HM method was applied to data collected (on a scale $[l,r]=[0,10]$) from 175 subjects for 32 generic words (i.e., words not attached to any specific context) ranging from *Teeny-weeny* to *Maximum amount*, including the words *Low*, *Moderate*, and *High*. The HM FOUs for *Low*, *Moderate*, and *High* are depicted in Fig. 5.

When data are collected from a single subject, they are asked two similar questions like [38]: *Suppose that a word can be located on a scale of l to r, and you want to locate the end-points of the interval that you associate with the word on that scale, but you are unsure of these two end-points: (Q1)[(Q2)] On the scale of l to r, what are the endpoints of an* ***interval of numbers*** *that*

[14] Another approach for obtaining an IT2 FS word model from the same kind of data that are used by the IA and EIA is [46]; it is based on calculating the median boundaries of the range of membership functions associated with the words.

[15] We do not use "linguistic term" because most people are unsure about what this means.

you associate with the left [right] end-point of the word[16]*?* A single subject provides $[a_L, b_L]$ and $[a_R, b_R]$ for each linguistic term.

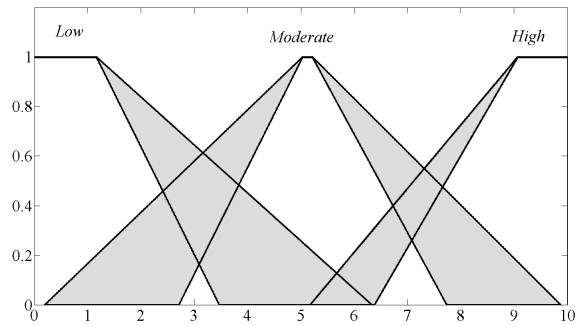

Fig. 5. *Low*, *Moderate*, and *High* FOUs obtained from the HM approach when the data were obtained from a group of 175 subjects.

Example 4. This example shows FOUs obtained from the HM approach when data are available from only one subject. We asked Prof. Charles Ragin to provide us left-hand and right-hand intervals for a three term vocabulary (*Low*, *Moderate*, and *High*) for each variable in the context of his famous Breakdown of Democracy example, for which the desired outcome is *O* = *Breakdown of Democracy* (of 18 European countries between World Wars 1 and 2) and there are five linguistic variables: *A* = *Developed* (country), *B* = *Urban* (country), *C* = *Literate* (country), *D* = *Industrial* (country) and *E* =*Stability* (of a country). In [44, Table 5.2] and [33, Example 1] it is explained that numerical values were initially obtained by Ragin for[17] *A, B,* C, D, *e* and *o*, and that *MF*(*O*) was computed from *MF*(*o*) as 1 - *MF*(*o*), and, *MF*(*E*) was computed from *MF*(*e*) as 1 - *MF*(*e*). Consequently, we asked him to provide end-point intervals also for *A, B,* C, D, *e* and *o*. His endpoint intervals for *e* = *Instablity* (of a country) are presented in Table 1. Fig. 6 depicts three normalized FOUs for *Low*, *Moderate*, and *High Instability* when the HM method was applied to this data. Comparing the sizes of the respective FOUs in Figs. 5 and 6, observe that less linguistic uncertainty (exemplified by smaller FOU area) is present, and FOUs have smaller spans, when data are collected in a context then when they are not.

Table 1

Endpoint intervals for *e* = *Instability* (of a country) provided by Prof. Charles Ragin $(e \in [0,21])$.

	Left-hand Interval		*Right-hand Interval*	
	a_L	b_L	a_R	b_R
Low	0	3	4	5
Moderate	5	7	8	9
High	9	12	15	21

[16] In addition, we instruct the subject that: *If you are absolutely certain about the two end-points then you do not have to provide end-point intervals, you only have to provide the left and right end-points*. We also instruct the subject *not to overlap the left and right end-point intervals*.

[17] A lower case letter denotes the complement of its upper case version.

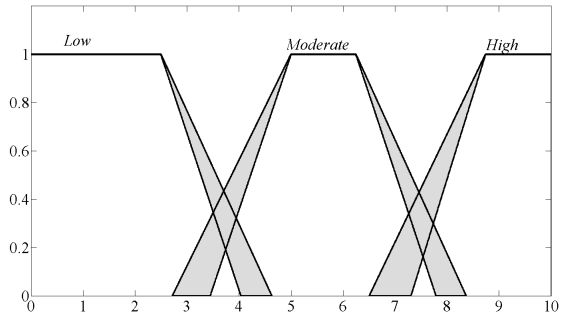

Fig. 6. HM FOUs for *Instablity* (of a country), when data were obtained from one expert, Prof. Charles Ragin. Note that $x = 10/21 \times e$.

F. Recapitulation

So far we have calibrated IT2 FSs for a collection of linguistic terms that are associated with a linguistic variable. To remind the reader, *it is not the linguistic variable that is calibrated; it is the linguistic terms that are associated with the linguistic variable that are calibrated*. The "external criteria" mentioned by Ragin [42, Ch. 5] as a prerequisite to calibration are now provided by collecting data intervals from either a group of subjects or just one subject. These are totally different kinds of external criteria then are presently used in fsQCA. Our approach is now totally in-line with the concepts of linguistic variables and linguistic terms, and so it is not subject to criticism about confusing a linguistic term with a linguistic variable.

III. Membership Functions for Use in fsQCA

At this point, although we now have valid MFs for linguistic terms, they do not conform to the so-called T1 MFs that are presently used in fsQCA, where, as explained above, those T1 MFs may not really be MFs at all for T1 FSs. This section explains how to map our word FOUs into proper S-shaped MFs similar to the ones that are presently used by practitioners of fsQCA.

A. Importance of an S-Shaped MF

One of the most interesting and important features of fsQCA is that it can completely remove a linguistic variable from being considered as a sufficient condition for a desired outcome. For example, Ragin's two final intermediate sufficient conditions for Breakdown of Democracy are [33], [44, Ch. 5]: (*not developed country* and *not urban country* and *not industrial country*) or (*unstable country*). Observe that fsQCA has removed the variable *Literacy* from these sufficient conditions.

Being able to completely remove a variable by fsQCA can only occur when a linguistic variable is described by one linguistic term[18]. So, as soon as a practitioner of fsQCA uses even two linguistic terms for a linguistic variable[19] it is no longer possible to completely remove the linguistic variable from fsQCA. Although it may be possible to remove one of the linguistic terms, e.g. *High IQ* or *Low IQ*, it is not possible to simultaneously remove both *Low IQ* and *High IQ*. So, using only one linguistic term for a linguistic variable is very *crucial/essential* to fsQCA.

As a result of this the MF that has been used by fsQCA scholars and practitioners is the S-shaped MF depicted in Fig. 1. But, as we have explained above, this may not be a MF of a T1 FS, so then what exactly is it? The answer to this question lies in a kind of FS that has not been used too often (even by FS scholars).

B. Level 2 Fuzzy Sets

One kind of generalization [5]-[8], [17, 20, 47] of an ordinary FS (T1 or T2) involves fuzzy sets defined within a universal set whose elements are ordinary FSs. These FSs are known as *level 2 (L2) fuzzy sets*.

Definition 7. A *level 2 FS* (*L2 FS*), $\tilde{\tilde{A}}$ is comprised of a domain of linguistic terms (words) that are associated with the domain X of real numbers for a linguistic variable, together with a MF $\mu_{\tilde{\tilde{A}}}:\mathcal{F}(X)\rightarrow[0,1]$ for a *T1 L2 FS*, or $\tilde{\mu}_{\tilde{\tilde{A}}}:\mathcal{F}(X)\rightarrow D[0,1]$, for an *interval-valued (IV) L2 FS*, where $\mathcal{F}(X)$ denotes the set of all ordinary[20] FSs (or words) of X.

---

[18] This fact is not stated explicitly in the fsQCA literature.

[19] It is customary in fsQCA to treat all of the linguistic terms of a linguistic variable as independent causal conditions; however, there is no theoretical justification for doing this, and doing this can lead to all sorts of problems with remainders that are used in some of the other steps of fsQCA [34] (see, also, Section 7).

[20] Higher-kinds of L2 FSs are possible, but in this paper we are only interested in T1 and IV L2 FSs. $\mathcal{F}(X)$ is also known as the *fuzzy power set* of X. In order to distinguish between a L2 FS and a T1 FS or an IT2 FS, we shall refer to the latter as *ordinary FSs* in the sequel.

A linguistic variable can be treated as a L2 FS by using its linguistic terms that are in its term set as its power set. Because the ordinary FSs of a linguistic variable are naturally ordered, the MF of a L2 FS is also naturally ordered.

Example 5. Referring to Example 1, consider the linguistic variable *Developed Country*, whose term set is $T \triangleq$ {*Barely, Hardly, Somewhat, Moderately, Fully, Extremely*} $\triangleq$ {*B, H, S, M, F, E*}. When *Developed Country* (*D*) is treated as a T1 L2 FS, $\tilde{\tilde{D}}$, then (using the so-called fuzzy set notation for a FS [1], introduced by Zadeh [56])

$$\mu_{\tilde{\tilde{D}}}(T) = \frac{\mu_B}{B} + \frac{\mu_H}{H} + \frac{\mu_S}{S} + \frac{\mu_M}{M} + \frac{\mu_F}{F} + \frac{\mu_E}{E} \tag{6}$$

When *Developed Country* (*D*) is treated as an IV L2 FS, $\tilde{\tilde{D}}$, then

$$\tilde{\mu}_{\tilde{\tilde{D}}}(T) = \frac{[\underline{\mu}_B, \overline{\mu}_B]}{B} + \frac{[\underline{\mu}_H, \overline{\mu}_H]}{H} + \frac{[\underline{\mu}_S, \overline{\mu}_S]}{S} + \frac{[\underline{\mu}_M, \overline{\mu}_M]}{M} + \frac{[\underline{\mu}_F, \overline{\mu}_F]}{F} + \frac{[\underline{\mu}_E, \overline{\mu}_E]}{E} \tag{7}$$

General versions of (6) and (7) are:

$$\text{T1 L2 FS: } \mu_{\tilde{\tilde{v}}}(T) = \sum_{i=1}^{M} \frac{\mu_{t_i}}{t_i} \tag{8}$$

$$\text{IV L2 FS: } \tilde{\mu}_{\tilde{\tilde{v}}}(T) = \sum_{i=1}^{M} \frac{[\underline{\mu}_{t_i}, \overline{\mu}_{t_i}]}{t_i} \tag{9}$$

B.1 On Models for t_i: In order to use (8) or (9) in fsQCA we need to use a mathematical model for t_i. One possible model is its 2D FOU, but then (8) would be a 3D MF (two dimensions are for t_i and one is for μ_{t_i}), and (9) would be a 4D MF (two dimensions are for t_i and the other two are for $[\underline{\mu}_{t_i}, \overline{\mu}_{t_i}]$), neither of which can be connected to the S-shaped MF that is used in fsQCA. The approach that we take is to replace t_i by a *reduced-information version* of it.

Because the FOU of a word conveys both the linguistic uncertainties about the word as well as the natural ordering of the word, we consider three possible *reduced-information versions* of t_i, namely: (1) centroid of the IT2 FS for t_i, C_{t_i}, (2) center of gravity of the centroid, c_{t_i}, and (3) maximum dispersion of the centroid, $\max D_{t_i}$. Appendix B explains these three uncertainty measures and how to compute them. Here we point out that: the centroid of an IT2 FS is an interval of real numbers, $C_{t_i} = [\underline{C}_{t_i}, \overline{C}_{t_i}]$, the center of gravity of the centroid is a single real number, $c_{t_i} = (\underline{C}_{t_i} + \overline{C}_{t_i})/2$, and the maximum dispersion of the centroid is also an interval of real numbers, $\max D_{t_i} = [\underline{\max D_{t_i}}, \overline{\max D_{t_i}}]$, one that is wider than the centroid, and is analogous to the

±standard deviation interval about the mean of a random variable. Each of these uncertainty measures summarizes the linguistic uncertainties about the word and preserves the natural ordering of the word.

B.2 On Models for level 2 membership grades: In order to use (8) or (9) in fsQCA we also need to choose μ_{t_i} for (8) and $[\underline{\mu}_{t_i}, \overline{\mu}_{t_i}]$ for (9). There are no unique ways for making these choices. Some examples of how to choose the grades are: (1) μ_{t_i} may be specified by an expert as a number (e.g., if t_i is a breakpoint term, and there are three such terms, then μ_{t_i} could be chosen as 0, 0.5 or 1; this is discussed in more detail in Section IV.B), or as an interval of numbers (e.g., $[0,0.1]$, $[0.4,0.6]$ and $[0.9,1]$); (2) $[\underline{\mu}_{t_i}, \overline{\mu}_{t_i}]$ may be specified as a normalized[21] $[\underline{C}_{t_i}, \overline{C}_{t_i}]$ accounting for uncertainty about c_{t_i}; or, (3) $[\underline{\mu}_{t_i}, \overline{\mu}_{t_i}]$ may be specified as a normalized $[\underline{\max D_{t_i}}, \overline{\max D_{t_i}}]$ accounting for uncertainty about $[\underline{C}_{t_i}, \overline{C}_{t_i}]$, etc.

B.3 MFs for reduced-information L2 FSs (RI L2 FS): The MFs for the RI L2 FSs and the names for those FSs are given in the third and fourth columns of Table 2, respectively. We use the notations ${}^1\mu_{\tilde{\tilde{v}}}(x)$ and ${}^1\tilde{\mu}_{\tilde{\tilde{v}}}(x)$ instead of $\mu_{\tilde{\tilde{v}}}(x)$ and $\tilde{\mu}_{\tilde{\tilde{v}}}(x)$ because the former MFs are undefined over some $x \in X$, something that is remedied below, e.g. observe, in the first and second rows of Table 2, that ${}^1\mu_{\tilde{\tilde{v}}}(x)$ and ${}^1\tilde{\mu}_{\tilde{\tilde{v}}}(x)$ are defined only at *M* point values of *x*, whereas in the third and fourth rows, ${}^1\tilde{\mu}_{\tilde{\tilde{v}}}(x)$ is defined only over *M* intervals of *x*.

Table 2

Some L2 FSs for term set *T* and their RI L2 MFs.

t_i	Grade	RI L2 FS MF	Name of RI L2 FS[a]
c_{t_i}	μ_{t_i}	$\mu_{\tilde{\tilde{v}}}(T) \to {}^1\mu_{\tilde{\tilde{v}}}(x) = \sum_{i=1}^{M} \frac{\mu_i}{c_{t_i}}$	T1 COG
c_{t_i}	$[\underline{\mu}_{t_i}, \overline{\mu}_{t_i}]$	$\tilde{\mu}_{\tilde{\tilde{v}}}(T) \to {}^1\tilde{\mu}_{\tilde{\tilde{v}}}(x) = \sum_{i=1}^{M} \frac{[\underline{\mu}_{t_i}, \overline{\mu}_{t_i}]}{c_{t_i}}$	IV COG
$[\underline{C}_{t_i}, \overline{C}_{t_i}]$	$[\underline{\mu}_{t_i}, \overline{\mu}_{t_i}]$	$\tilde{\mu}_{\tilde{\tilde{v}}}(T) \to {}^1\tilde{\mu}_{\tilde{\tilde{v}}}(x) = \sum_{i=1}^{M} \frac{[\underline{\mu}_{t_i}, \overline{\mu}_{t_i}]}{[\underline{C}_{t_i}, \overline{C}_{t_i}]}$	IV Centroid
$[\underline{\max D_{t_i}}, \overline{\max D_{t_i}}]$	$[\underline{\mu}_{t_i}, \overline{\mu}_{t_i}]$	$\tilde{\mu}_{\tilde{\tilde{v}}}(T) \to {}^1\tilde{\mu}_{\tilde{\tilde{v}}}(x) = \sum_{i=1}^{M} \frac{[\underline{\mu}_{t_i}, \overline{\mu}_{t_i}]}{[\underline{\max D_{t_i}}, \overline{\max D_{t_i}}]}$	IV Maximum Dispersion

[a] COG is short for Center of Gravity.

[21] Normalization is to [0, 1] to conform to the range of membership grades of a FS.

Fig. 7 depicts the RI MFs for three linguistic terms. Each of its four figures is for one of the rows in Table 2. The shaded rectangles in Fig. 7c and 7d are called *granules* (e.g., [2], [39]). Observe that maximum dispersion granules are wider than centroid granules because they capture more uncertainty about word FOUs.

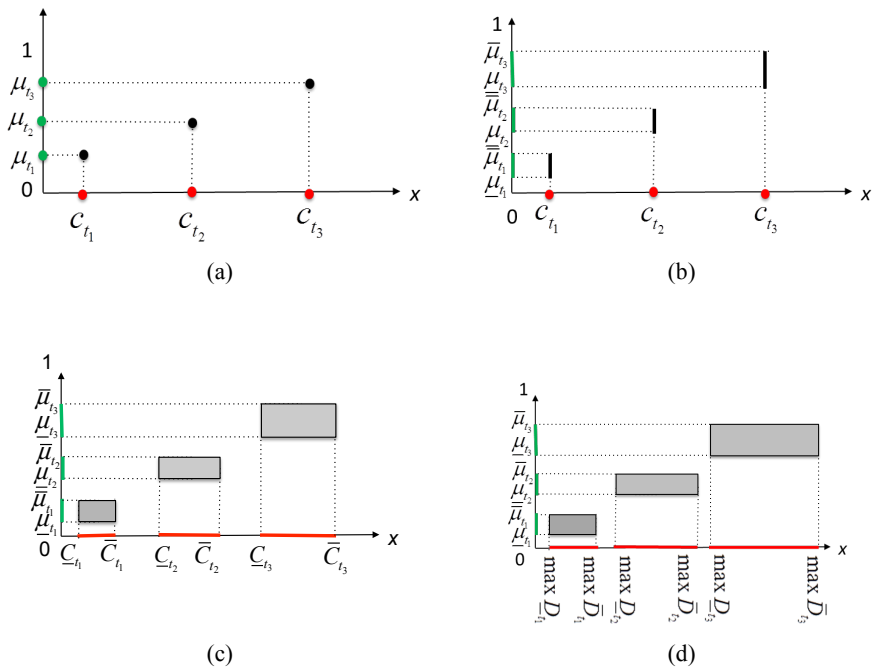

Fig. 7. RI MFs for three linguistic terms: (a) T1 COG, (b) IV COG, (c) IV Centroid, and (d) IV Maximum Dispersion FSs.

B.4 Approximated MFs for RI L2 FS: As has been mentioned above, and is clear from Fig. 7, there are intervals of x where the MF is undefined, something that has to be remedied because a MF should be defined for $x \in X$. In this section we briefly describe two approaches for accomplishing this. The results from both approaches can be summarized as ${}^{1}\mu_{\tilde{\tilde{v}}}(x) \rightarrow \mu_{\tilde{\tilde{v}}}(x)$ or ${}^{1}\tilde{\mu}_{\tilde{\tilde{v}}}(x) \rightarrow \tilde{\mu}_{\tilde{\tilde{v}}}(x)$. It is $\mu_{\tilde{\tilde{v}}}(x)$ or $\tilde{\mu}_{\tilde{\tilde{v}}}(x)$ that will be our MFs for a RI L2 FS model of a linguistic variable for $x \in X$.

A rationale for approximation (that may involve both interpolation and extrapolation) is a thought experiment in which linguistic terms are inserted to the left and to the right of the actually used linguistic terms. Each of these linguistic terms contributes a point, as in Fig. 7a, an interval, as in Fig. 7b, or a granule, as in Figs. 7c or 7d. The more linguistic terms there are the more points, intervals or granules there will be.

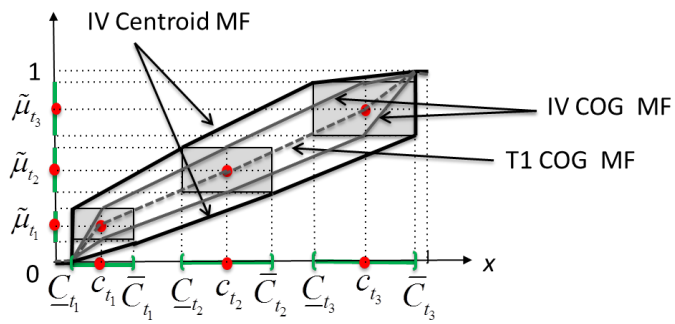

Fig. 8. Piecewise-linear approximation ${}^{1}\mu_{\tilde{\tilde{v}}}(x) \to \mu_{\tilde{\tilde{v}}}(x)$ (dashed) for Fig. 7a T1 COG MF; ${}^{1}\tilde{\mu}_{\tilde{\tilde{v}}}(x) \to \tilde{\mu}_{\tilde{\tilde{v}}}(x)$ (light weight solid lines) for Fig. 7b IV COG MF; and, ${}^{1}\tilde{\mu}_{\tilde{\tilde{v}}}(x) \to \tilde{\mu}_{\tilde{\tilde{v}}}(x)$ (heavy weight solid lines) for Fig. 7c IV Centroid MF.

(a) Piecewise-linear approximation: Examples of piecewise-linear approximations are depicted in Fig. 8. The dashed T1 MF is obtained by using straight lines to connect a Fig. 7a left-dot to its right-dot neighbor. The light weight solid straight lines describe the lower and upper MFs of the IV COG MF, and are obtained by using straight lines to connect the Fig. 7b lower or upper interval end-points (going from left to right), respectively. The heavy weight solid straight lines describe the lower and upper MFs of the IV Centroid MF, and are obtained by using straight lines to connect the Fig. 7c granule lower right corners or upper left corners (going from left to right), respectively. All of these MFs are then extended, by means of vertical and horizontal lines, to 0 and 1 as shown.

(b) Log-odds approximation: Ragin [42, pp. 87-94] prefers to approximate by using a log-odds function, so in this section we explain how to obtain such an approximated MF for $\mu_{\tilde{\tilde{v}}}(x)$. We do this for a three-word term set because we later apply these results to a term set that contains three breakpoint terms.

Let $y_A(x)$ be the odds of membership (the ratio of the membership of being in a generic T1 FS *A* over the membership of not being in *A*), i.e. $y_A(x) = \mu_A(x)/(1-\mu_A(x))$, and $z_A(x)$ be the log odds of membership, i.e. $z_A(x) = \ln y_A(x)$. Then

$$\mu_A(x) = \frac{\exp(z_A(x))}{1+\exp(z_A(x))} \tag{10}$$

Given three points, x_1, x_2, x_3 (to be defined below), whose membership grades are assigned the values[22] 0.05, 0.5 and 0.95, respectively, then [42, pp. 87-94] $z_A(x)$ is found as:

[22] 0.05 and 0.95 are used instead of 0 and 1 because the log-odds transformation is incapable of producing membership grades that are exactly equal to 0 or 1.

$$z_A(x) = \begin{cases} \dfrac{3}{(x_2 - x_1)}(x - x_2) & x < x_2 \\ \dfrac{3}{(x_3 - x_2)}(x - x_2) & x \geq x_2 \end{cases} \tag{11}$$

Once $z_A(x)$ has been computed by using (11), $\mu_A(x)$ is computed by using (10).

Examples of log-odds approximations are depicted in Fig. 9. The dashed T1 MF is obtained by using $x_1 = c_{t_1}$, $x_2 = c_{t_2}$ and $x_3 = c_{t_3}$ in (11). The solid lines describe the lower and upper MFs of the IV Centroid MF, and are obtained by using $x_1 = \underline{C}_{t_1}$, $x_2 = \underline{C}_{t_2}$ and $x_3 = \underline{C}_{t_3}$ to generate the UMF and $x_1 = \overline{C}_{t_1}$, $x_2 = \overline{C}_{t_2}$ and $x_3 = \overline{C}_{t_3}$ to generate the LMF.

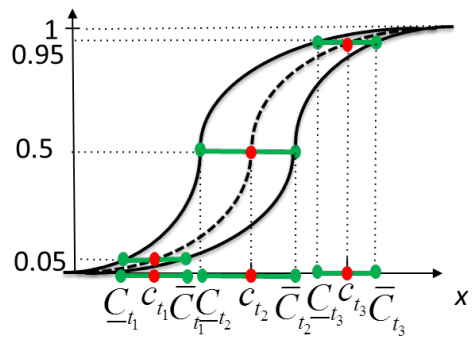

Fig. 9. Log-odds approximations: ${}^1\mu_{\tilde{v}}(x) \to \mu_{\tilde{v}}(x)$ (dashed) for Fig. 7a T1 COG MF and ${}^1\tilde{\mu}_{\tilde{v}}(x) \to \tilde{\mu}_{\tilde{v}}(x)$ (solid lines) for Fig. 7c IV Centroid MF.

C. Recapitulation

It is $\mu_{\tilde{v}}(x)$ or $\tilde{\mu}_{\tilde{v}}(x)$ that is used in fsQCA. Their constructions have made use of: (1) Data collected from a group of subjects or one expert about words in an application dependent vocabulary; (2) HM method to map the data into an FOU for each word in the vocabulary; (3) A L2 FS as the first step in obtaining a MF for the variable; (4) Replacing each term in the domain of linguistic terms of the L2 FS by a reduced-information (RI) version of it, thereby obtaining the MF (T1 or IV) of a RI L2 FS; and, (5) Approximating the MFs for the RI L2 FS, using piecewise-linear or log-odds functions, thereby obtaining approximated MFs for the RI L2 FS. These MFs are S-shaped.

We now have an answer to the question posed at the end of Section 3.1, namely "What exactly is the MF for that is used by the fsQCA community?" Our answer is: *It is the MF of an approximated RI L2 FS for the entire linguistic variable*. Note that it is not the MF of an ordinary FS—T1 or IT2—but it is still the MF of a FS, albeit a L2 FS.

In the sequel, we shall shorten the phrase "approximate MF for a RI L2 FS" to "MF for the linguistic variable."

IV. Breakdown of Democracy Example

In this section we re-examine Ragin's Breakdown of Democracy example. To begin we remind that reader that this example is taken from [44, Chapter 5] for which the desired outcome is O = *Breakdown of Democracy* (of 18 European countries between World Wars 1 and 2) and there are five causal conditions: A = *Developed* (country), B = *Urban* (country), C = *Literate* (country), D = *Industrial* (country) and E = *Stable* (country). O, A, B, C, D and E are linguistic variables (v) and Ragin's term set, $T_R(v)$, for each of these linguistic variable contains three linguistic breakpoint terms (we leave the extensions to more than three breakpoints to the readers):

$$T_R(v) = \{\textit{fully out}, \textit{neither in nor out}, \textit{fully in}\} \tag{12}$$

These linguistic breakpoint terms are linearly ordered and act as adjectives when applied to a linguistic variable (e.g., *fully out* of being a *Developed* country).

A. FOUs

In order to obtain FOUs for the three linguistic breakpoint terms for each of the six linguistic variables, we asked Prof. Ragin to provide us with ranges and uncertainty end-point intervals for each of them. Since he had already chosen breakpoints in his earlier publications, we asked him to provide his minimum and maximum % uncertainties for each of the breakpoints, from which it was easy for us to construct the uncertainty end-point intervals. The uncertainty left and right end-point intervals based on his data are summarized in Table 3. Fig. 10 depicts FOUs for the Table 3 data computed by using the HM method.

Table 3

Three breakpoints (BP), left end-point intervals (LEPI) and right end-point intervals (REPI) for six linguistic variables that are used in the Breakdown of Democracy example, from data provided by Prof. Charles Ragin[a].

Linguistic Variable	Variable's Range	W_1 = *fully out*			W_2 = *neither in nor out*			W_3 = *fully in*		
		BP	*LEPI*	*REPI*	*BP*	*LEPI*	*REPI*	*BP*	*LEPI*	*REPI*
A: Developed	[300, 1200]	400	[320, 380]	[440, 480]	550	[467.5, 550]	[550, 632.5	900	[720, 810]	[990, 1080]
B: Urban	[10, 80]	25	[20, 22.5]	[27.5, 30]	50	[45, 50]	[50, 55]	65	[61.75, 65]	[65, 68.25]
C: Literate	[30, 100]	50	[40, 50]	[50, 60]	75	[63.7, 75]	[75, 86.25]	90	[81, 85.5]	[94.5, 99]
D: Industrial	[10, 50]	20	[17, 20]	[20, 23]	30	25.5, 30]	[30, 34.5]	40	[36, 38]	[42, 44]
E: Stable	[0, 21]	5	[4.5, 4.75]	[5.25, 5.5]	9.5	[8.55, 9.5]	[9.5, 10.45]	15	[12.75, 15]	[15, 17.25]
O: Breakdown of Democracy	[-10, 10]	-9	[-9.45, -9]	[-9, -8.55]	0	[-4, 0]	[0, 4]	10	[9.5, 10]	[10, 10]

[a] The situations where the right end-point of LEPI equals the left end-point of REPI occurred (e.g., all of the W_2 words) as a result of Prof. Ragin providing 0% to ± maximum % uncertainty about his breakpoint.

The FOU centroids and the COGs of those centroids are summarized in Table 4, and the FOU COGs and maximum dispersion intervals are summarized in Table 5. Observe from these tables that the maximum dispersion intervals are wider than their respective centroid intervals. These uncertainty intervals are used to generate MFs for the six linguistic variables, as explained next.

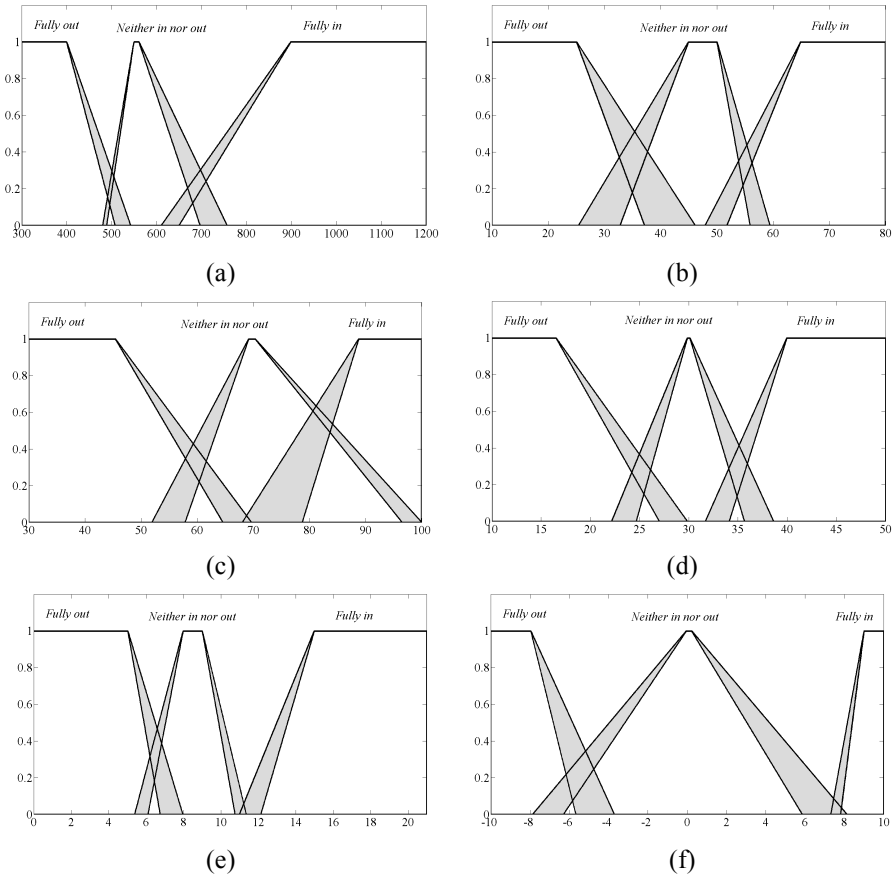

Fig. 10. HM approach FOUs for Breakdown of Democracy example: (a) Developed, (b) Urban, (c) Literate, (d) Industrial, (e) Stable, and (f) Breakdown of Democracy.

Table 4

COG and centroid interval endpoints for $W_1 = fully\ out$, $W_2 = neither\ in\ nor\ out$, and $W_3 = fully\ in$.

Linguistic Variable	$c_{\tilde{W}_1}$	$[\underline{C}_{\tilde{W}_1}, \overline{C}_{\tilde{W}_1}]$	$c_{\tilde{W}_2}$	$[\underline{C}_{\tilde{W}_2}, \overline{C}_{\tilde{W}_2}]$	$c_{\tilde{W}_3}$	$[\underline{C}_{\tilde{W}_3}, \overline{C}_{\tilde{W}3}]$
A: Developed	381.26	[373.84, 388.68]	550.94	[534.21, 567.67]	902.79	[877.41, 928.17]
B: Urban	24.52	[23.72, 25.33]	50.06	[49.15, 50.98]	65.01	[64.52, 65.50]
C: Literate	49.87	[47.95, 51.80]	74.43	[72.00, 76.86]	88.21	[86.91, 89.51]
D: Industrial	19.90	[19.13, 20.67]	30.06	[29.03, 31.10]	39.24	[38.48, 39.99]
E: Stable	5.07	[4.99, 5.15]	9.55	[9.36, 9.74]	14.98	[14.43, 15.54]
O: Breakdown of Democracy	-8.99	[-9.12, -8.87]	0.09	[-0.93, 1.11]	9.72	[9.67, 9.77]

Table 5

COG and maximum dispersion intervals for W_1 = *fully out*, W_2 = *neither in nor out*, and W_3 = *fully in*.

Linguistic Variable	$c_{\tilde{W}_1}$	$[\underline{\max D_{\tilde{W}_1}}, \overline{\max D_{\tilde{W}_1}}]$	$c_{\tilde{W}_2}$	$[\underline{\max D_{\tilde{W}_2}}, \overline{\max D_{\tilde{W}_2}}]$	$c_{\tilde{W}_3}$	$[\underline{\max D_{\tilde{W}_3}}, \overline{\max D_{\tilde{W}_3}}]$
A: Developed	381.26	[332.84, 429.16]	550.94	[461.66, 640.57]	902.79	[800.82, 997.93]
B: Urban	24.52	[20.91, 28.20]	50.06	[43.20, 56.04]	65.01	[61.72, 68.57]
C: Literate	49.87	[38.22, 62.74]	74.43	[63.47, 87.62]	88.21	[82.47, 95.54]
D: Industrial	19.90	[16.98, 23.27]	30.06	[24.56, 35.03]	39.24	[35.21, 43.56]
E: Stable	5.07	[4.64, 5.53]	9.55	[8.22, 10.72]	14.98	[12.15, 18.09]
O: Breakdown of Democracy	-8.99	[-9.52, -8.49]	0.09	[-5.09, 5.63]	9.72	[9.23, 10]

B. MFs for the Six Linguistic Variables

Fig. 11 depicts approximated MFs for RI L2 FSs for the six linguistic variables, obtained by using the log-odds approximation method that is described in Section 3.2.4. Each of the plots contains four items:

1. FOU LMF and UMF obtained by using the maximum dispersion intervals that are given in Table 5, referred to below as a *Maximum Dispersion FOU* (row 4 in Table 2).
2. FOU LMF and UMF obtained by using the centroids that are given in Table 4, referred to below as a *Centroid FOU* (row 3 in Table 2).
3. T1 MF obtained by using the COG of the centroids that are given in Table 4, referred to below as a *COG T1 MF* (row 1 in Table 2).
4. Log-odds T1 MF that was obtained just by using Ragin's three breakpoints that are given in Table 3, referred to below as *Ragin's T1 MF* (row 1 in Table 2).

Observe in Fig. 11 that the Centroid FOUs are contained within the Maximum Dispersion FOUs and are considerably smaller than the latter, and Ragin's T1 MF always falls within the two FOUs and is very close to the COG T1 MF (sometimes they are visually indistinguishable). The closeness of the COG and Ragin's T1 FSs is due to the way in which interval end-point uncertainty bands were created. Recall that they were created about historical breakpoints, which (as can be seen in Table 3) are always the COGs of the interval end-point uncertainty bands. In a new application of our methodology, where historical breakpoints would not be already known, an expert would only specify breakpoint end-point intervals, and there would be no curve that is analogous to Ragin's T1 MF for historical breakpoints.

C. fsQCA Using the MFs for Linguistic Variables

When T1 L2 FSs are used in fsQCA we refer to this as T1 fsQCA, whereas when IV L2 FSs are used in fsQCA we refer to this as IV fsQCA. Because T1 fsQCA has been in the literature for a very long time we do not review it here (see, e.g., [33] for a complete description of all of its steps, and [44, Ch. 5], [45]). Additionally, software is available to perform T1 fsQCA (see Table 11.1 in [45] as well as accompanying discussions about it in its Section 11.1.10). Because IV

fsQCA is not well known, having only appeared in[23] [22], we include a summary of its steps in Appendix B.

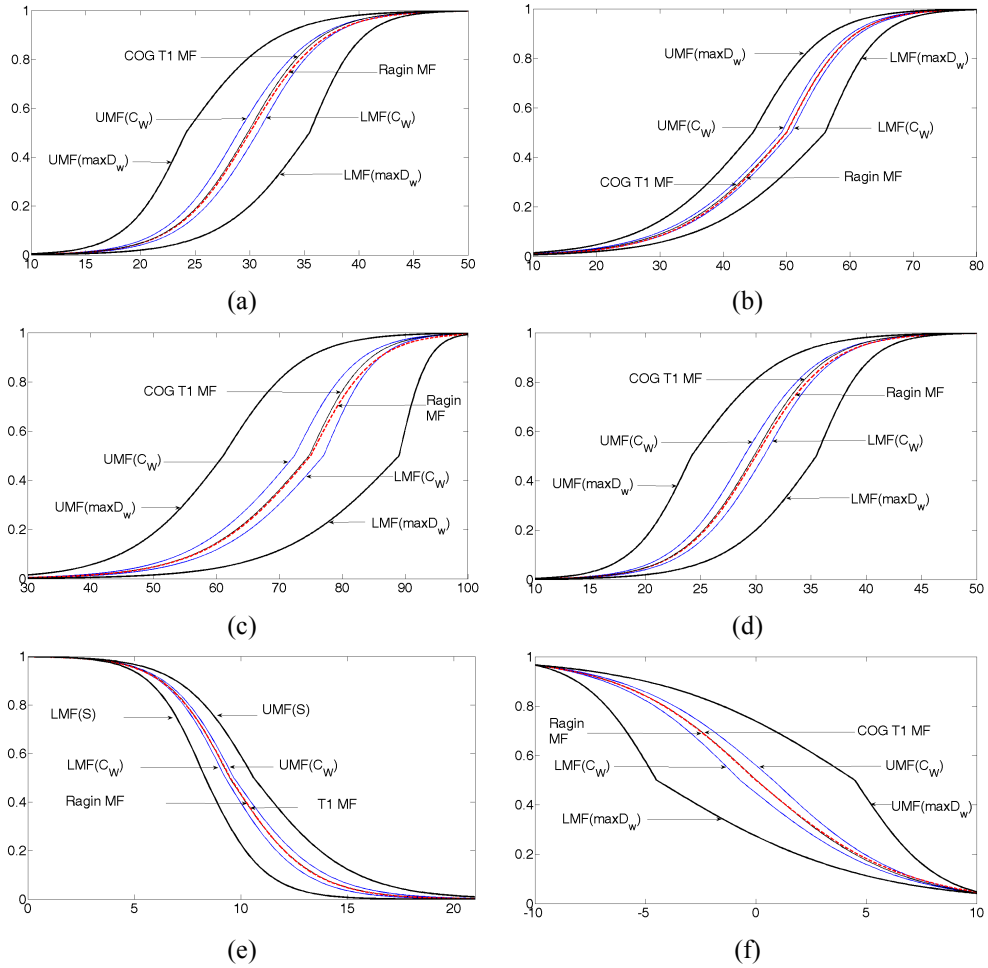

Fig. 11. Maximum Dispersion FOU (black), Centroid FOU (blue), COG T1 MF (black) and Ragin's T1 MF (red dashed) used in Breakdown of Democracy example, for: a) Developed, b) Urban, c) Literate, d) Industrial, e) Stable, and f) Breakdown of Democracy.

C.1 T1 fsQCA: Beginning with the case data that are in [44, Table 5.2], and using the COG T1 MFs depicted in Fig. 11, we obtained MF values for each case in the linguistic variables *A–E* (they are the *causal conditions*) as well as in desired outcome *O*. These values are given in Table 6. The one causal combination, out of a possible number of $2^5 = 32$ candidate causal combinations that survives the T1 fsQCA MF>0.5 test for each case is given in the last column of that table[24]. There are only nine uniquely different causal combinations in that column and they appear in the first column of Table 7. Consistency (also known as subsethood) values for each of these nine causal combinations are given in the last column of that table, from which we see that the first six causal combinations have consistency values that are greater than the

[23] [22] calls this IT2 fsQCA; however, we now prefer IV fsQCA and so this is what is used herein.
[24] A proof that only one causal combination survives this test for a case is in [34].

customary threshold value of 0.80. Using these six causal combinations and completing all of the remaining steps of T1 fsQCA, one obtains the complex, parsimonious and intermediate solutions that are given in Table 8. These are exactly the same solutions that were obtained by Ragin [44, pp. 115 & 117].

Table 6

MF values for each case[25] using COG T1 MFs, and surviving causal combination for each case.

Case	Developed (*A*)	Urban (*B*)	Literate (*C*)	Industrial (*D*)	Stable (*E*)	Breakdown of Democracy (*O*)	Surviving Causal Combination
1	0.81	0.13	0.99	0.75	0.43	0.95	*AbCDe*
2	0.99	0.89	0.99	1.00	0.98	0.04	*ABCDE*
3	0.58	0.98	0.99	0.91	0.92	0.10	*ABCDE*
4	0.19	0.08	0.99	0.01	0.92	0.88	*abCdE*
5	0.59	0.04	1.00	0.08	0.59	0.22	*AbCdE*
6	0.98	0.03	0.99	0.83	0.96	0.04	*AbCDE*
7	0.89	0.79	0.99	0.97	0.30	0.95	*ABCDe*
8	0.05	0.10	0.13	0.37	0.43	0.93	*abcde*
9	0.10	0.17	0.91	0.08	0.13	0.57	*abCde*
10	0.72	0.05	0.99	0.01	0.96	0.08	*AbCdE*
11	0.36	0.10	0.42	0.48	0.59	0.95	*abcdE*
12	0.98	1.00	1.00	0.95	0.99	0.04	*ABCDE*
13	0.03	0.18	0.61	0.00	0.00	0.88	*abCde*
14	0.02	0.02	0.01	0.11	0.01	0.95	*abcde*
15	0.02	0.04	0.17	0.00	0.85	0.78	*abcdE*
16	0.04	0.31	0.09	0.21	0.20	0.93	*abcde*
17	0.95	0.14	1.00	0.68	0.92	0.04	*AbCDE*
18	0.98	0.99	1.00	1.00	0.98	0.04	*ABCDE*

C.2 IV fsQCA: Beginning with the case data that are in [44, Table 5.2], and using the Centroid FOUs depicted in Fig. 11, we obtained LMF and UMF values for each case in the linguistic variables *A–E* as well as in desired outcome *O*. These values are given in Table 9. The one causal combination, out of a possible number of $2^5 = 32$ candidate causal combinations that survives the IV fsQCA MF > 0.5 test (see Step 6 in Appendix C, and Appendix D) for each case is given in Table 10. Comparing the $\tilde{F}_j$ column of Table 10 with the last column of Table 6, observe that the surviving causal combinations are the same.

Consistency values for the nine uniquely different causal combinations are given in Table 11. Comparing Tables 11 and 7, observe that they have exactly the same six causal combinations that pass the respective T1 and IT2 consistency tests. Using these six causal combinations and completing all of the remaining steps of IV fsQCA, one obtains the complex, parsimonious and intermediate solutions that are given in Table 12. These are exactly the same solutions that were

[25] The numbered cases correspond to the following countries: 1 – Austria, 2 – Belgium, 3 – Czechoslovakia, 4 – Estonia, 5 – Finland, 6 – France, 7 – Germany, 8– Greece, 9 – Hungary, 10 – Ireland, 11 – Italy, 12 – Netherlands, 13 – Poland, 14 – Portugal, 15 – Romania, 16 – Spain, 17 – Sweden, and 18 – United Kingdom.

obtained in Section 4.3.1 and by Ragin [44, pp. 115 & 117], the same solutions were also obtained by us for the Maximum Dispersion FOUs (we do not show the details because they are so similar to the ones shown for the Centroid FOUs).

Table 7

Distribution of cases and consistency of T1 surviving causal combinations using COG T1 MFs. [a]

COG T1 Surviving Causal Combinations	Frequency	Consistency
abcde	3	1
abcdE	2	0.98
AbCDe	1	0.97
ABCDe	1	0.97
abCdE	1	0.85
abCde	2	0.85
AbCdE	2	0.50
AbCDE	2	0.49
ABCDE	4	0.24

[a] Shaded rows are actual causal combinations with consistencies > 0.8.

Table 8

T1-fsQCA solutions using COG T1 MFs.

Solution	T1 Causal Combination
Complex	*abd* + *ACDe*
Parsimonious	*a* + *e*
Intermediate	*abd* + *e*

Translating the common intermediate solution into words, we have:

- **IF** *Not Developed* and *Not Urban* and *Not Industrial* **THEN** *Breakdown of Democracy*
- **IF** *Unstable* **THEN** *Breakdown of Democracy*

D.4 Recapitulation

We have demonstrated that by our calibration procedure one is able to obtain exactly the same solutions to the Breakdown of Democracy example as obtained by Prof. Ragin. A reader may be wondering at this point "So why was all of this needed? Can't we just continue to use Ragin's breakpoints?" Our answers to these questions are: (1) We have used a very different calibration method, one that overcomes our earlier criticism and therefore strengthens fsQCA as a viable methodology; (2) Using our calibration method we are able to combine information from different experts which is not possible using Ragin's direct method; and, (3) The fact that our two Intermediate solutions, one based on T1 FSs and the other based on IT2 FSs, are the same

(Table 13) demonstrates that fsQCA is robust to the interval end-point uncertainties provided to us by Prof. Ragin. Those uncertainties led to word FOUs whose lower and upper MFs were used in IV fsQCA, as is explained in Appendix C. Because robustness is such an important issue in fsQCA, we examine it more closely next.

Table 9

Lower and Upper MFs for each case using Centroid FOUs.

Case	Causal Conditions										Outcome	
	$\underline{\mu}_{\tilde{A}}$	$\overline{\mu}_{\tilde{A}}$	$\underline{\mu}_{\tilde{B}}$	$\overline{\mu}_{\tilde{B}}$	$\underline{\mu}_{\tilde{C}}$	$\overline{\mu}_{\tilde{C}}$	$\underline{\mu}_{\tilde{D}}$	$\overline{\mu}_{\tilde{D}}$	$\underline{\mu}_{\tilde{E}}$	$\overline{\mu}_{\tilde{E}}$	$\underline{\mu}_{\tilde{O}}$	$\overline{\mu}_{\tilde{O}}$
1	0.79	0.83	0.12	0.14	0.99	1.00	0.70	0.80	0.40	0.46	0.95	0.95
2	0.99	0.99	0.87	0.91	0.99	0.99	1.00	1.00	0.98	0.98	0.04	0.04
3	0.54	0.62	0.97	0.98	0.99	0.99	0.90	0.93	0.91	0.92	0.10	0.11
4	0.16	0.23	0.07	0.09	0.99	0.99	0.01	0.01	0.91	0.92	0.87	0.89
5	0.55	0.62	0.03	0.04	1.00	1.00	0.07	0.10	0.56	0.61	0.20	0.24
6	0.97	0.98	0.03	0.04	0.99	0.99	0.79	0.86	0.95	0.96	0.04	0.04
7	0.88	0.90	0.75	0.82	0.99	1.00	0.96	0.97	0.27	0.34	0.95	0.95
8	0.05	0.06	0.09	0.11	0.10	0.17	0.31	0.43	0.40	0.46	0.93	0.94
9	0.08	0.11	0.15	0.19	0.88	0.93	0.06	0.09	0.10	0.15	0.51	0.62
10	0.69	0.75	0.05	0.06	0.99	0.99	0.01	0.01	0.95	0.96	0.07	0.08
11	0.30	0.44	0.09	0.12	0.36	0.48	0.41	0.55	0.56	0.61	0.95	0.95
12	0.98	0.98	1.00	1.00	1.00	1.00	0.94	0.96	0.99	0.99	0.04	0.04
13	0.03	0.03	0.17	0.20	0.50	0.70	0.00	0.00	0.00	0.00	0.87	0.89
14	0.02	0.02	0.02	0.02	0.01	0.02	0.10	0.14	0.00	0.01	0.95	0.95
15	0.02	0.02	0.03	0.04	0.14	0.21	0.00	0.01	0.84	0.86	0.76	0.81
16	0.03	0.04	0.28	0.34	0.07	0.12	0.18	0.25	0.17	0.23	0.93	0.94
17	0.94	0.95	0.12	0.15	1.00	1.00	0.62	0.74	0.91	0.92	0.04	0.04
18	0.98	0.99	0.99	0.99	1.00	1.00	1.00	1.00	0.98	0.98	0.04	0.04

Table 10

Surviving causal combination, $\tilde{F}_j$, and its LMF and UMF values using Centroid FOUs.

x	$\tilde{F}_j$	$\underline{\mu}_{\tilde{F}_j}$	$\overline{\mu}_{\tilde{F}_j}$
1	$\tilde{A}\tilde{b}\tilde{C}\tilde{D}\tilde{e}$	0.54	0.60
2	$\tilde{A}\tilde{B}\tilde{C}\tilde{D}\tilde{E}$	0.87	0.91
3	$\tilde{A}\tilde{B}\tilde{C}\tilde{D}\tilde{E}$	0.54	0.62
4	$\tilde{a}\tilde{b}\tilde{C}\tilde{d}\tilde{E}$	0.77	0.84
5	$\tilde{A}\tilde{b}\tilde{C}\tilde{d}\tilde{E}$	0.55	0.61
6	$\tilde{A}\tilde{b}\tilde{C}\tilde{D}\tilde{E}$	0.79	0.86
7	$\tilde{A}\tilde{B}\tilde{C}\tilde{D}\tilde{e}$	0.66	0.73
8	$\tilde{a}\tilde{b}\tilde{c}\tilde{d}\tilde{e}$	0.54	0.60
9	$\tilde{a}\tilde{b}\tilde{C}\tilde{d}\tilde{e}$	0.81	0.85
10	$\tilde{A}\tilde{b}\tilde{C}\tilde{d}\tilde{E}$	0.69	0.75
11	$\tilde{a}\tilde{b}\tilde{c}\tilde{d}\tilde{E}$	0.45	0.59
12	$\tilde{A}\tilde{B}\tilde{C}\tilde{D}\tilde{E}$	0.94	0.96
13	$\tilde{a}\tilde{b}\tilde{C}\tilde{d}\tilde{e}$	0.50	0.70
14	$\tilde{a}\tilde{b}\tilde{c}\tilde{d}\tilde{e}$	0.86	0.90
15	$\tilde{a}\tilde{b}\tilde{c}\tilde{d}\tilde{E}$	0.79	0.86
16	$\tilde{a}\tilde{b}\tilde{c}\tilde{d}\tilde{e}$	0.66	0.72
17	$\tilde{A}\tilde{b}\tilde{C}\tilde{D}\tilde{E}$	0.62	0.74
18	$\tilde{A}\tilde{B}\tilde{C}\tilde{D}\tilde{E}$	0.98	0.98

Table 11

Distribution of cases and consistency of IV surviving causal combinations using Centroid FOUs. [a]

Centroid FOU Causal Combinations	Frequency	Consistency
$\tilde{a}\tilde{b}\tilde{c}\tilde{d}\tilde{e}$	3	1.00
$\tilde{a}\tilde{b}\tilde{c}\tilde{d}\tilde{E}$	2	0.98
$\tilde{A}\tilde{b}\tilde{C}\tilde{D}\tilde{e}$	1	0.97
$\tilde{A}\tilde{B}\tilde{C}\tilde{D}\tilde{e}$	1	0.97
$\tilde{a}\tilde{b}\tilde{C}\tilde{d}\tilde{E}$	1	0.86
$\tilde{a}\tilde{b}\tilde{C}\tilde{d}\tilde{e}$	2	0.85
$\tilde{A}\tilde{b}\tilde{C}\tilde{d}\tilde{E}$	2	0.51
$\tilde{A}\tilde{b}\tilde{C}\tilde{D}\tilde{E}$	2	0.49
$\tilde{A}\tilde{B}\tilde{C}\tilde{D}\tilde{E}$	4	0.25

[a] Shaded rows are actual causal combinations with consistencies > 0.8.

Table 12

IV fsQCA Solutions using Centroid FOUs.

Solution	IT2 Causal Combination
Complex	$\tilde{a}\tilde{b}\tilde{d} + \tilde{A}\tilde{C}\tilde{D}\tilde{e}$
Parsimonious	$\tilde{a} + \tilde{e}$
Intermediate	$\tilde{a}\tilde{b}\tilde{d} + \tilde{e}$

Table 13

Comparison of solutions to Breakdown of Democracy example using different kinds of FSs.

Solution	Ragin's T1 MF	COG T1 MF	Centroid FOUs	Maximum dispersion FOUs
Complex	$abd + ACDe$	$abd + ACDe$	$\tilde{a}\tilde{b}\tilde{d} + \tilde{A}\tilde{C}\tilde{D}\tilde{e}$	$\tilde{a}\tilde{b}\tilde{d} + \tilde{A}\tilde{C}\tilde{D}\tilde{e}$
Parsimonious	$a + e$	$a + e$	$\tilde{a} + \tilde{e}$	$\tilde{a} + \tilde{e}$
Intermediate	$abd + e$	$abd + e$	$\tilde{a}\tilde{b}\tilde{d} + \tilde{e}$	$\tilde{a}\tilde{b}\tilde{d} + \tilde{e}$

V. ROBUSTNESS OF FSQCA IN BREAKDOWN OF DEMOCRACY EXAMPLE

In this section we study the robustness of fsQCA to the uncertainty about FOUs as well as to the uncertainty about membership grades assigned to linguistic terms. One way to think about what we are going to do is to examine Fig. 8 (or Figs. 7c and 7d) and ask: "How large can the granules become before the IV fsQCA solutions to the Breakdown of Democracy example change?"

A. Uncertainty About Models for t_i

In this section we examine the robustness of fsQCA to uncertainty about the models for t_i (column 1 in Table 2, the horizontal dimension of the granules in Fig. 8), using log-odds approximated MFs for the RI L2 FSs. The length of the horizontal dimension of the granules in Fig. 8 is determined by the amount of uncertainty there is about the end-point intervals used in the HM method. Hence, we began with Table 3's left and right end-point intervals, and increased each of these interval's end points by 5% increments up to a maximum change of 30% (e.g., $[a,b] \rightarrow [a-0.05a, b+0.05a]$).[26] Doing this leads to word FOUs that look like the ones in Fig. 10, but have larger centroid and maximum dispersion intervals (i.e., the horizontal lengths of the granules shown in Fig. 8 increase). When a change to an fsQCA solution occurred, we then reduced the total % increment at which this occurred by 1% increments, until we reached a word end-point interval size for which the solution did not change.

Doing this, we found that (see Table 14): (1) When Centroid FOUs were used, the results from IV fsQCA are identical to those from T1 fsQCA if less than 22% (i.e., 23% - 1%) uncertainty values are used, and (2) When Maximum Dispersion FOUs were used the results from IV fsQCA are identical to those from T1 fsQCA if less than 17% uncertainty values are used. The robustness of IV fsQCA decreases when Maximum Dispersion FOUs are used because the Maximum Dispersion generates wider intervals as compared to those generated by using the Centroid. Observe also that the intermediate results from IV fsQCA with more uncertainty are simpler than the intermediate results from T1 fsQCA (i.e., $\tilde{a}+\tilde{e}$ versus $\tilde{a}\tilde{b}\tilde{d}+\tilde{e}$). This is due to $\tilde{a}\tilde{B}\tilde{C}\tilde{D}\tilde{E}$ passing the consistency threshold in IV fsQCA and not passing it in T1 fsQCA.

Table 14

Comparisons of fsQCA solutions for different amounts of uncertainty added to data endpoint intervals.

Solution	T1 fsQCA	IV fsQCA using Centroid FOUs		IV fsQCA using Maximum dispersion FOUs	
		10% Uncert.	23% Uncert.	10% Uncert.	18% Uncert.
Complex	$abd + ACDe$	$\tilde{a}\tilde{b}\tilde{d}+\tilde{A}\tilde{C}\tilde{D}\tilde{e}$	$\tilde{a}\tilde{b}\tilde{d}+\tilde{A}\tilde{C}\tilde{D}\tilde{e}+\tilde{a}\tilde{B}\tilde{C}\tilde{D}\tilde{E}$	$\tilde{a}\tilde{b}\tilde{d}+\tilde{A}\tilde{C}\tilde{D}\tilde{e}$	$\tilde{a}\tilde{b}\tilde{d}+\tilde{A}\tilde{C}\tilde{D}\tilde{e}+\tilde{a}\tilde{B}\tilde{C}\tilde{D}\tilde{E}$
Parsimonious	$a + e$	$\tilde{a}+\tilde{e}$	$\tilde{a}+\tilde{e}$	$\tilde{a}+\tilde{e}$	$\tilde{a}+\tilde{e}$
Intermediate	$abd + e$	$\tilde{a}\tilde{b}\tilde{d}+\tilde{e}$	$\tilde{a}+\tilde{e}$	$\tilde{a}\tilde{b}\tilde{d}+\tilde{e}$	$\tilde{a}+\tilde{e}$

[26] If a % change causes a variable to exceed an end-point of its range, we use the end-point.

B. Uncertainty About Grades

In this section we examine the robustness of fsQCA to uncertainty about the grades chosen for a L2 FS (column 2 in Table 2, the vertical length of the granules in Fig. 8 (or Figs. 7c and 7d) using piecewise linear[27] approximated MFs for the RI L2 FSs. The length of the vertical dimension of the granules in Fig. 8 is determined by the amount of uncertainty there is about the breakpoint grades and has nothing to do with using the HM method to obtain FOUs for the breakpoint words (the horizontal length of the granules in Fig. 8); but when the vertical length of a granule changes so do the lower and upper FOU bounding functions.

Recall that Ragin assigned a fixed membership grade of 0.05, 0.5 and 0.95, respectively, to the linguistic terms *fully out*, *neither in nor out*, and *fully out*. Instead of using these fixed grades, we added ±5% to each of them up to a maximum of ±40%, obtaining intervals for each of these breakpoint membership grades (e.g., $0.5 \rightarrow [0.5-0.05\times 0.5, 0.5+0.05\times 0.5]$)[28]. Fig. 12 depicts the resulting piecewise-linear approximations of IV Centroid FOUs when ±10%, ±20%, ±30% and ±40% of the breakpoint grades were added to the above grades for each linguistic term. As before, when a change to an fsQCA solution occurred, we then reduced the ±% change made to the modified breakpoint values at which this occurred by ±1% increments, until we reached a breakpoint interval size for which the solution did not change.

Our results are summarized in see Table 15, from which we see that: (1) (compare columns 2 and 3) using T1 fsQCA results for piecewise-linear approximated MFs for T1 COG (RI L2 FSs) are identical to those obtained by using Ragin's T1 MF; (2) (compare columns 2, 4 and 5) using IV fsQCA with IV COG, we obtain identical results to those from T1 fsQCA if less than ±34% uncertainty is used; (3) (compare columns 2, 6 and 7) using IV fsQCA with IV Centroid, we obtain identical results to those from T1 fsQCA if less than ±31% uncertainty is used; and, (4) (compare columns 2, 8 and 9) using IV fsQCA with Max Dispersion, we obtain identical results to those from T1 fsQCA if less than ±39% uncertainty is used. So, fsQCA seems to be very robust to uncertainty about breakpoint membership grades.

We do not want to read too much into these results because they are all based on replacing 0, and 1 by 0.05 and 0.95, respectively, in order to use the log-odds approximations. It seems to us that a more direct approach is to use a membership grade of 0 for fully out and 1 for fully in and piecewise linear approximations.

C. Conclusions

Based on the results presented in Sections 5.1 and 5.2, we conclude that fsQCA is relatively robust in both the horizontal and vertical dimensions of Fig. 8 (or Figs. 7c and 7d). The horizontal dimension lets us incorporate uncertainty about the models for the linguistic terms by moving from point values (Figs. 7a and 7b) to intervals (Figs. 7c and 7d), whereas the vertical dimension lets us incorporate uncertainty about the breakpoint membership grades, also moving from point values (Fig. 7a) to intervals (Figs. 7b, 7c and 7d). Additionally, using piecewise linear

[27] Piecewise-linear approximations are easier to obtain for this part of the robustness study than are the log-odds approximations, because the latter have only been established for the three exact breakpoint values, 0.05, 0.5 and 0.95.

[28] If a % change caused 0.95 to exceed 1, we used 1. Even a -40% change to 0.05 does not cause its lower value to become negative.

approximated RI L2 FSs gives the same results in IV fsQCA as using log-odds approximated MFs in T1 fsQCA.

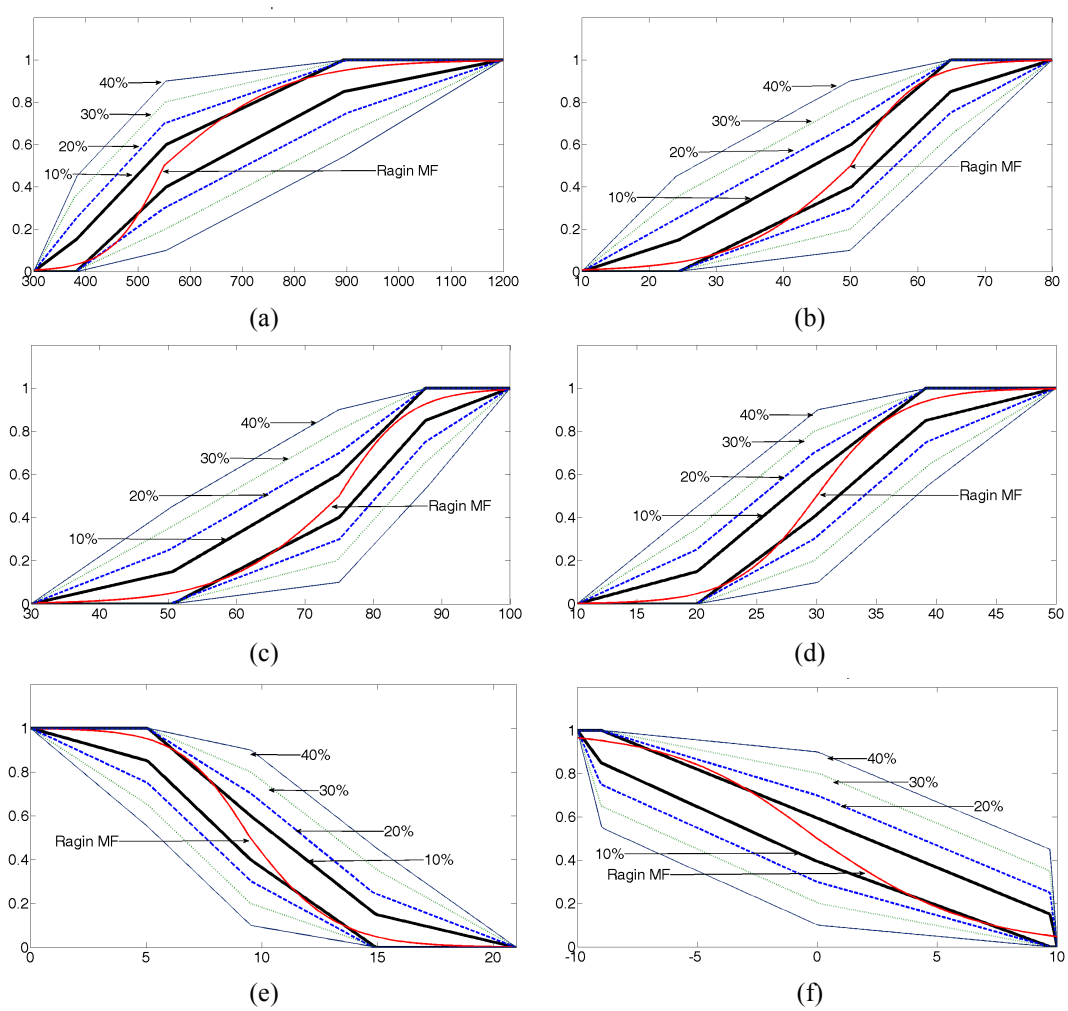

Fig. 12. Piecewise-linear approximations of S-shaped MFs for Breakdown of Democracy example using the centroid for t_i and different amounts of membership grade uncertainty, $\tilde{\mu}_{t_i}$, for: a) Developed, b) Urban, c) Literate, d) Industrial, e) Stable, and f) Breakdown of Democracy. The red curves are Ragin's T1 MFs.

Table 15

Comparison of fsQCA solutions for different amounts of uncertainty (Un) added to breakpoint membership grades.

Solution	Ragin's T1 MF + T1 fsQCA	Piecewise-linear approximations[a]						
		T1 fsQCA using T1 COG	IV fsQCA using IV COG		IV fsQCA using IV centroid		IV fsQCA using Max dispersion	
		0% Un	±30% Un	±35% Un	±30% Un	±32% Un	±30% Un	±40% Un
Complex	$abd + ACDe$	$abd + ACDe$	$\tilde{a}\tilde{b}\tilde{d} + \tilde{A}\tilde{C}\tilde{D}\tilde{e}$	$\tilde{a}\tilde{b}\tilde{d} + \tilde{A}\tilde{B}\tilde{C}\tilde{D}\tilde{e}$	$\tilde{a}\tilde{b}\tilde{d} + \tilde{A}\tilde{C}\tilde{D}\tilde{e}$	$\tilde{a}\tilde{b}\tilde{d} + \tilde{A}\tilde{B}\tilde{C}\tilde{D}\tilde{e}$	$\tilde{a}\tilde{b}\tilde{d} + \tilde{A}\tilde{C}\tilde{D}\tilde{e}$	$\tilde{a}\tilde{b}\tilde{d} + \tilde{A}\tilde{C}\tilde{D}\tilde{e} + \tilde{a}\tilde{b}\tilde{c}\tilde{E}$
Parsimonious	$a + e$	$a + e$	$\tilde{a} + \tilde{e}$	$\tilde{a} + \tilde{B}\tilde{e}$	$\tilde{a} + \tilde{e}$	$\tilde{a} + \tilde{B}\tilde{e}$	$\tilde{a} + \tilde{e}$	$\tilde{a} + \tilde{e}$
Intermediate	$abd + e$	$abd + e$	$\tilde{a}\tilde{b}\tilde{d} + \tilde{e}$	$\tilde{a}\tilde{b}\tilde{d} + \tilde{B}\tilde{e}$	$\tilde{a}\tilde{b}\tilde{d} + \tilde{e}$	$\tilde{a}\tilde{b}\tilde{d} + \tilde{B}\tilde{e}$	$\tilde{a}\tilde{b}\tilde{d} + \tilde{e}$	$\tilde{a}\tilde{b}\tilde{d} + \tilde{e} + \tilde{a}\tilde{b}\tilde{c}$

[a] T1 COG, IV COG, IV Centroid and IV Max Dispersion correspond to the situations that are depicted in Figs. 7a, 7b, 7c and 7d, respectively. Ragin's T1 MF is defined in item 4 of Section 4.2.

VI. On Obtaining More Precise Causal Combinations

The final causal combinations that are obtained from fsQCA, be they the complex, intermediate or parsimonious solutions, will be in terms of linguistic variables, just as they presently are. This is because the S-shaped approximated MFs (FOUs) for RI L2 FSs that are used in T1 or IV fsQCA are for the linguistic variables. Because those MFs (FOUs) were derived from FOUs for all of the linguistic variable's terms, it is (arguably, for the first time) possible to also obtain more precise statements of those causal combinations, e.g. for their best instances, as is explained next. Such more precise statements may be of value to practitioners of fsQCA.

Each case has a numerical value for each of its linguistic variables (e.g., as in [44, Table 5.1]). Our *rule for mapping a number into a word* is: At each measured value of x the winning word is the one with the largest average MF grade, where for an IT2 FS the average MF grade is the average of its LMF and UMF at x.

Example 8. In Fig. 13, at $x = x_2$, the winning word is found by solving $Avg\mu_{W_3}(x_2) \overset{?}{>} Avg\mu_{W_4}(x_2)$. If the inequality is true then x_2 is assigned to the word W_3; otherwise, it is assigned to W_4. At x_2, Fig. 13 reveals that $Avg\mu_{W_3}(x_2) = \frac{1}{2}[\overline{\mu}_{W_3}(x_2) + \underline{\mu}_{W_3}(x_2)]$ and $Avg\mu_{W_4}(x_2) = \frac{1}{2}\overline{\mu}_{W_4}(x_2)$, from which it follows that x_2 is assigned to W_3. At x_1, the winning word is W_2.

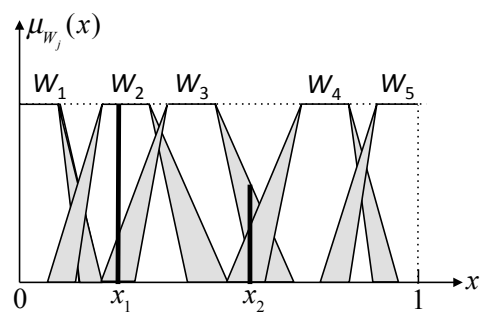

Fig. 13. Mapping a number into a word.

Example 9. In Ragin's Breakdown of Democracy example one intermediate solution is *abd*—*not developed, not urban,* and *not industrial.* Its best instances are Estonia, Greece and Italy. To us *not developed, not urban,* and *not industrial* is somewhat vague, mainly because of the complement; however, because each case can have a linguistic term assigned to its membership in the linguistic variable (by using the method in Example 8), it is now possible to obtain a more precise statement of *abd* for these countries, e.g., for Estonia (Table 16), *abd* can be replaced by *Estonia is fully out of Developed, fully out of Urban,* and *fully out of Industrial*, and, for Italy (Table 16), *abd* can be replaced by *Italy is fully out of Developed, fully out of Urban,* and *neither*

in nor out of Industrial. It is important to observe that no complement of the linguistic terms *Developed, Urban* or *Industrial* have been used.

If we had started with a larger vocabulary (e.g., for 5 or 7 breakpoints) then greater linguistic resolution would have been obtained.

Table 16

On more precise causal combinations for Estonia, Greece and Italy, where the adjectives used for *Developed, Urban, and Industrial* are *fully out*, *neither in nor out*, and *fully in*. The numbers in this table were obtained by using the case data in [44, Table 5.2] and Fig. 10.

Best Instance	Average of LMF and UMF			Precise linguistic term	Average of LMF and UMF			Precise linguistic term	Average of LMF and UMF			Precise linguistic term
	Developed				*Urban*				*Industrial*			
	Fully out	Neither in nor out	Fully in		Fully out	Neither in nor out	Fully in		Fully out	Neither in nor out	Fully in	
Estonia	**0.84**	0.13	0.10	*Fully out*	**0.72**	0.24	0	*Fully out*	**1**	0.01	0	*Fully out*
Greece	**0.67**	0	0	*Fully out*	**0.58**	0.33	0	*Fully out*	0.25	**0.69**	0	*Neither in nor out*
Italy	**0.61**	0.28	0.14	*Fully out*	**0.56**	0.34	0	*Fully out*	0.19	**0.80**	0	*Neither in nor out*

[a] Bold-face numbers are the winning linguistic terms for each case (country).

VII. DISCUSSIONS

Linguistic summarizations for engineering and computer science applications [13-16, 52, 54, 55, 61] begin by establishing term sets that contain more than one linguistic term for one or more linguistic variables and by then creating summarizations using different combinations of the linguistic terms. When we first began working on fsQCA linguistic summarizations (more than six years ago) we were also of this same mindset, but we ran into obstacles that could not be overcome (although we tried many different approaches) [21]. *All obstacles were due to using more than one linguistic term for a linguistic variable and treating each of them as an independent causal condition*. Doing this caused the following problems: (1) A linguistic variable could not be completely removed by means of fsQCA (as explained earlier in Section 3.1); (2) Impossible causal combinations occurred [34][29], and exactly which causal combinations they are and how many of them there are depends on how many linguistic terms each linguistic variable is granulated into (their number becomes very large as the number of linguistic terms increases); (3) Enumeration of the impossible causal combinations must be done ahead of time so that they are not used either as remainders in the QM algorithm that obtains the parsimonious solutions, or during counterfactual analysis; and, (4) Programs for implementing the QM algorithm (e.g., Logic Friday that is available at http://sontrak.com/) often crash because they are limited by the number of remainder terms.

By using the S-shaped MFs (FOUs) for linguistic variables obtained in this paper, all of these challenges disappear, i.e. it is possible to remove a linguistic variable from fsQCA; there are no impossible causal combinations (see [45] for additional discussions); and existing QM algorithm software does not get overwhelmed (unless the number of linguistic variables becomes too large, but, for a linguistic summarization to be understandable to a human there really should not be too many variables in the antecedent of an fsQCA rule).

[29] An *impossible causal combination* is one whose MF can never be greater than 0.5, regardless of the case-based data. For example, when two linguistic terms are used for IQ, namely, *Low IQ* (*L*) and *High IQ* (*H*), it is impossible for the combination *L and H* to occur. Intuitively, this makes sense, because it is not possible for a person to simultaneously be of Low and High IQ (see, also, [45]). In this simple situation it is easy to connect intuition and mathematics; however, this is much more difficult to do when three or more linguistic terms are used to describe a linguistic variable.

VIII. Conclusions and Directions for Further Research

In this paper we have provided a new way for calibrating the fuzzy sets that are used in fsQCA, one that is based on clearly distinguishing between a linguistic variable and the linguistic terms for that variable, and that overcomes our criticism about the MF that is used by fsQCA practitioners. The resulting fuzzy sets are *reduced-information level 2 fuzzy sets* (RI L2 FS) whose MFs are approximated so that they are defined for $x \in X$, and, these MFs for the linguistic variables.

Our new calibration method can be summarized as:

$$\begin{aligned} &\text{Linguistic variable} \rightarrow \text{Linguistic Terms} \rightarrow \text{Data} \rightarrow FOU \\ &\quad \rightarrow RI\ L2\ FS \rightarrow \text{Approximated MF for } RI\ L2\ FS \\ &\quad = \text{MF for the Linguistic variable} \end{aligned} \tag{13}$$

Its major steps are: (1) For each linguistic variable, a vocabulary of linearly ordered linguistic terms (words) are chosen; (2) Interval end-point data are collected for each of the linguistic variables, either from a group of subjects or from one expert; (3) The data for each word are mapped into the footprint of uncertainty (FOU) of an IT2 FS using the HM method; (4) A reduced information L2 FS is created by replacing each word with an uncertainty measure and choosing an appropriate membership grade for it; and, (5) The MF of the RI L2 FS is approximated so that the resulting MF is for $x \in X$.

As mentioned above, the resulting approximated RI L2 FS MF is for the linguistic variable, and, it is not the MF of an ordinary FS (T1 or IT2) but is instead the MF of a L2 FS. This MF has an S-shape, (T1 or IV), which is the kind of MF shape that is so widely used by fsQCA scholars, and (as explained in this paper) is so important to fsQCA.

We have applied our new calibration procedure to Ragin's Breakdown of Democracy example, using new data provided to us by him, and have demonstrated that we are able to obtain his earlier solutions using either T1 or IV fsQCA, something that should be reassuring to fsQCA scholars. By using IV fsQCA we are also able to study the robustness of fsQCA to breakpoint location uncertainties as well as to membership grade uncertainties, and have demonstrated that IV fsQCA is robust to both. Finally, because the S-shaped MFs (FOUs) were derived from FOUs for all of the linguistic variable's terms, we have shown how it is possible to also obtain more precise statements of those causal combinations that do not use complements of the causal conditions (e.g., for their best instances), something that may be of added value to practitioners of fsQCA.

One of the big challenges to using our new calibration procedure is to either modify existing fsQCA software or to create new software for both T1 fsQCA and IV fsQCA.

We would like to remind the reader that as one goes from linguistic-term FOUs to their centroids (or maximum dispersion intervals) to the COGs of those centroids information is lost. How to use the entire FOU in fsQCA computations for such a L2 FS remains to be studied.

Finally, although our new calibration procedure has been motivated by fsQCA and has led to the S-shaped interpolated MF for a RI L2 FS, we suggest that perhaps this mapping from a set of FOUs for the linguistic terms of a linguistic variable into a L2 MF for the linguistic variable may also be useful outside of fsQCA, e.g. in computing with words [37].

APPENDIX A. THE HM APPROACH

The material in this appendix is taken from [9] and [10] and is included here for completeness.

A.1 Calibration for Data Collected From a Group of Subjects

Going from n data intervals $[a^{(i)}, b^{(i)}]_{i=1}^{n}$ for a word to an IT2 FS for the word using the HM Method is done in two parts, one that focuses on removing uninformative intervals—the Data Part—and the other that focuses on mapping the remaining set of intervals into the FOU—the Fuzzy Set Part.

A.2 Data Part

The *Data Part* uses statistics and probability, starting with the n intervals, to (see [53] for the details of each of these steps[30]):

1) Remove bad data (i.e., intervals that fall outside of [*l, r*]—not everyone takes a survey seriously)— $n \to n'$.
2) Remove outliers (using Box and Whisker tests [50]) in two steps, $n' \to n'' \to m'$.
3) Keep only the data intervals that are within an acceptable two-sided tolerance limit (using statistical tolerance limits [50]), in two steps, $m' \to m^{+} \to m''$.
4) Remove data intervals that have no overlap or too little overlap with other data intervals (using a statistical test that enforces the maxim [24, 37] *words must mean similar things to different people* (for effective communication to occur)), $m'' \to m$.

At the end of the Data Part, the original n data intervals have been reduced to a set of m (surviving) data intervals $\{[a^{(i)}, b^{(i)}]\}_{i=1}^{m}$, where $m \le n$.

A.3 Fuzzy Set Part

The *Fuzzy Set Part* uses these m data intervals to (see [10] for the details of each of these steps):

5) *Establish the nature of the FOU as either a Left-shoulder, Interior or a Right-shoulder FOU* (*the data speaks*, i.e. we do not choose the nature of the FOU ahead of time) by using one-sided tolerance intervals for the end-points $[a^{(i)}, b^{(i)}]$ of the m data intervals, $\underline{a}$ for $\{a^{(i)}\}_{i=1}^{m}$ and $\overline{b}$ for $\{b^{(i)}\}_{i=1}^{m}$. The first (last) word $[i = 1\ (m)]$ is always modeled as a Left (Right)-shoulder FOU. For $i = 2,...,m-1$, if $\underline{a} < 0$, the word is modeled as (Fig. A.1a) a Left-shoulder FOU; or, if $\overline{b} > r$, the word is modeled as (Fig. A.1c) a Right-shoulder FOU; or, otherwise the word is modeled as (Fig. A.1b) an Interior FOU.

[30] The processing steps in the Data Part for the HM method are the same as in the Data Part of the EIA.

6) *Compute the overlap* $[o_l, o_r]$ *of the m intervals*. For a Left-shoulder FOU, $[o_l, o_r] = [0, \min_i b^{(i)}]_{i=1}^m$; for an Interior FOU, $[o_l, o_r] = [\max_i a^{(i)}, \min_i b^{(i)}]_{i=1}^m$; and, for a Right-shoulder FOU, $[o_l, o_r] = [\max_i a^{(i)}, 10]_{i=1}^m$.

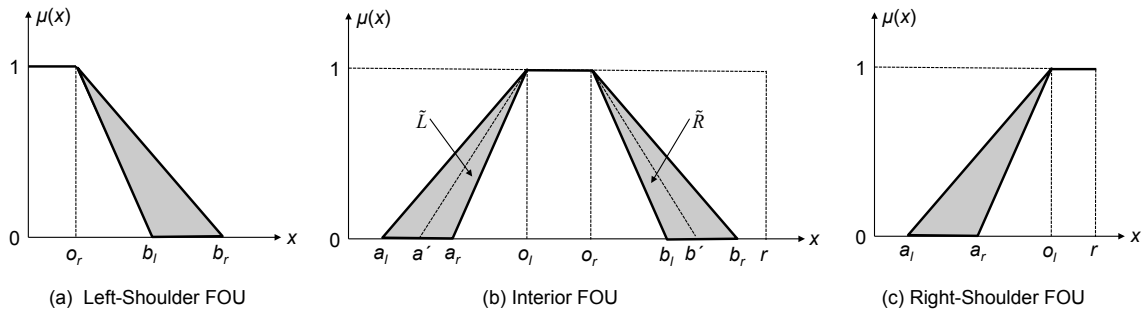

Fig. A.1. Three kinds of FOUs and their parameters: (a) Left-shoulder FOU, (b) Interior FOU, and (c) Right-shoulder FOU. In these figures $l = 0$ and $r = 1$.

7) *Remove the overlap* $[o_l, o_r]$ *from each of the original m intervals,* $\{[a^{(i)}, b^{(i)}]\}_{i=1}^m$. For a Left-shoulder FOU, this leaves one new set of smaller intervals, $\{[o_r, b^{(i)}]\}_{i=1}^m$; for a Right-shoulder FOU, this leaves one new set of smaller intervals, $\{[a^{(i)}, o_l]\}_{i=1}^m$; and, for an Interior FOU, this leaves two new sets of smaller intervals, $\{[a^{(i)}, o_l]\}_{i=1}^m$ and $\{[o_r, b^{(i)}]\}_{i=1}^m$.

8) *Map the set (s) of smaller intervals into the two parameters that define the respective FOU*. For a left- or right-shoulder FOU (Figs. A.1a and A.1c) there are exactly two such parameters, b_l and b_r, or a_l and a_r; but, for an interior FOU (Fig. A.1b) there are four such parameters, a_l, a_r, b_l and b_r. For the interior FOU, $\{[a^{(i)}, o_l]\}_{i=1}^m$ are mapped into a_l and a_r, and $\{[o_r, b^{(i)}]\}_{i=1}^m$ are mapped into b_l and b_r.

In Step 8, the mappings are done so that two measures of uncertainty about the *m* smaller-length data intervals are mapped into two comparable measures of uncertainty about the FOU. As an illustration of this, consider the left-hand portion of the FOU for an Interior FOU, $\tilde{L}$ (Fig. A.1b), and the mapping of $\{[a^{(i)}, o_l]\}_{i=1}^m$ into a_l and a_r. This mapping is obtained in four steps:

i. The centroid of $\tilde{L}$, $C_{\tilde{L}}$, and the average centroid, $c(\tilde{L})$, [19, 51] are computed as [10]:

$$C_{\tilde{L}} = [c(UMF(\tilde{L}), c(LMF(\tilde{L}))] = \left[\frac{2o_l + a_l}{3}, \frac{2o_l + a_r}{3}\right] \tag{A.1}$$

$$c(\tilde{L}) = \tfrac{1}{2}[c(LMF(\tilde{L})) + c(UMF(\tilde{L}))] = \tfrac{2}{3}o_l + \tfrac{1}{6}(a_l + a_r) \tag{A.2}$$

ii. The standard deviation of $\tilde{L}$, $STD_{\tilde{L}}$, [19, 51] is computed as [10]:

$$STD_{\tilde{L}} = [s(LMF(\tilde{L})), s(UMF(\tilde{L}))] = \left[\frac{o_l - a_r}{3\sqrt{2}}, \frac{o_l - a_l}{3\sqrt{2}} \right] \tag{A.3}$$

iii. The sample mean and standard deviation of the m intervals $[a^{(i)}, o_l]_{i=1}^{m}$ are computed, and are denoted $\hat{m}_{LH}$ and $\hat{s}_{LH}$, respectively.

iv. a_l and a_r are solved for from the following two equations:

$$\begin{cases} c(\tilde{L}) = \hat{m}_{LH} \\ s(UMF(\tilde{L})) = \hat{s}_{LH} \end{cases} \tag{A.4}$$

$s(UMF(\tilde{L}))$ is used in (A.4) because it is larger than $s(LMF(\tilde{L}))$ and because the average of $s(LMF(\tilde{L}))$ and $s(UMF(\tilde{L}))$ contains $a_l + a_r$ as does $c(\tilde{L})$, so that it would be impossible to unravel both a_l and a_r by using both $c(\tilde{L})$ and the average standard deviation.

The solutions to (A.4) are:

$$\begin{cases} a_l = \max(0, o_l - 3\sqrt{2}\hat{s}_{LH}) \\ a_r = \min(o_l, 6\hat{m}_{LH} + 3\sqrt{2}\hat{s}_{LH} - 3o_l) \end{cases} \tag{A.5}$$

For a_l, the max guards against $o_l - 3\sqrt{2}\hat{s}_{LH}$ being negative, and for a_r, the min guards against $6\hat{m}_{LH} + 3\sqrt{2}\hat{s}_{LH} - 3o_l$ being larger than o_l.

Comparable results for the right-hand portion of the FOU for an Interior FOU, $\tilde{R}$ (Fig. A.1b), are:

$$\begin{cases} b_l = \max(o_r, 6\hat{m}_{RH} - 3\sqrt{2}\hat{s}_{RH} - 5o_r) \\ b_r = \min(r, o_r + 3\sqrt{2}\hat{s}_{RH}) \end{cases} \tag{A.6}$$

where $\hat{m}_{RH}$ and $\hat{s}_{RH}$ are the sample mean and standard deviation, respectively, of $\{[o_r, b^{(i)}]\}_{i=1}^{m}$.

Compare Figs. A.1a and A.1b to observe that the Left-shoulder FOU is the same as $\tilde{R}$, so (A.6) is also used to determine the two parameters of a Left-shoulder FOU. Similarly, the Right-shoulder FOU is the same as $\tilde{L}$, so (A.5) is also used to determine the two parameters of a Right-shoulder FOU.

A.4 Calibration for Data Collected From One Subject

Starting with the left-hand and right-hand intervals, $[a_L, b_L]$ and $[a_R, b_R]$ [38]:

1) Assume that each of the end-point intervals is uniformly distributed, and then compute the mean and variance for both of them.
2) Assign the mean and variance of the left and right intervals from Step 1 to uniform probability distributions and generate 100 random numbers $(L_1, L_2, ..., L_{50}; R_1, R_2, ..., R_{50})$.

Form 50 end-point pairs from these random numbers[31] $\{(L_1,R_1),...,(L_{50},R_{50})\}$.

3) Assume each pair of end-points has been collected from a different subject (or the same subject who is sampled 50 times, where the spacing of the samples is long enough so that the subject does not remember his/her past responses), a virtual group of subjects.
4) Apply the HM method to the 50 intervals to obtain the (Person) FOU for the word. This FOU only accounts for a person's intra-uncertainty about the word.

Using the additional data that are provided by a single subject, this four-step procedure reduces to the HM method for collecting data from a group of subjects[32]

[31] We chose 50 intervals because the convergence results given in [53] have demonstrated that FOUs converge in a mean-square sense when around 30 or more intervals are used. Any number ≥ 30 should be adequate.

[32] Of course, the interval end-point data could also be obtained from a group of subjects, in which case the four-step procedure of this section could be applied to each of the subject's data, leading to a larger set of *n* starting data intervals than in Section II. Our experience with extracting data from subjects by means of questionnaires is that they like simple questions. The Section II questions are simpler than the ones in Section III, so we do not advocate doing what we have just suggested.

APPENDIX B. UNCERTAINTY MEASURES FOR IT2 FSS

In this appendix we provide background information about the centroid of an IT2 FS, (2) center of gravity of the centroid, and (3) maximum dispersion of the centroid.

B.1 Centroid of an IT2 FS

It is always possible to cover the FOU of an IT2 FS, $\tilde{A}$, by a collection of T1 FSs (Fig. B.1) that are called *embedded T1 FSs*, so that $\tilde{A}$ can be expressed as the set-theoretic union of those T1 FSs [31, 37 (Ch. 2)]. Conceptually, it is then possible to compute the center of gravity (COG) for each of the embedded T1 FSs. Regardless of how many embedded T1 FSs it takes to cover $FOU(\tilde{A})$[33] the total number of COGs has both a smallest value, $\underline{C}_{\tilde{A}}$, and largest value, $\overline{C}_{\tilde{A}}$. The collection of these COGs is called the *centroid* of $\tilde{A}$, $C_{\tilde{A}}$, where[34]

$$C_{\tilde{A}} = \left\{ x \mid x \in \{\underline{C}_{\tilde{A}}, ..., \overline{C}_{\tilde{A}}\} \right\} \equiv \left\{ x \mid x \in [\underline{C}_{\tilde{A}}, \overline{C}_{\tilde{A}}] \right\} \tag{B.1}$$

$C_{\tilde{A}}$ is a measure of the uncertainty in $\tilde{A}$.

The larger (smaller) in length that $C_{\tilde{A}}$ is the larger (smaller) is the uncertainty in $\tilde{A}$. Geometrically, larger (smaller) uncertainty in $\tilde{A}$ is manifested by fatter (thinner) FOUs. In the limiting case when both $LMF(\tilde{A})$ and $UMF(\tilde{A})$ become the same T1 MF $C_{\tilde{A}}$ reduces to one number, the COG, $c_{\tilde{A}}$, of that T1 MF.

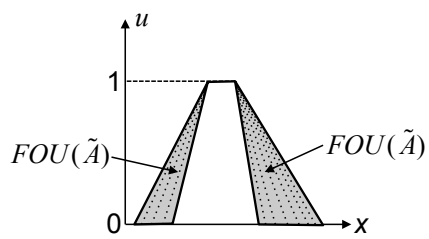

Fig. B.1. Covering an FOU by T1 FSs. Each T1 FS (called an *embedded T1 FS*) contains one dashed line from the left-portion of the FOU, the common flat top, and one dashed line from the right-portion of the FOU. $LMF(\tilde{A})$ and $UMF(\tilde{A})$ are also embedded T1 FSs. $FOU(\tilde{A})$ is the union of all of the embedded T1 FSs. In general, embedded T1 FSs do not have to be straight lines.

Generally speaking, closed-form formulas do not exist for computing $\underline{C}_{\tilde{A}}$ and $\overline{C}_{\tilde{A}}$, even for the seemingly simple FOU in Fig. B.1. Fortunately, there are many simple algorithms for computing $\underline{C}_{\tilde{A}}$ and $\overline{C}_{\tilde{A}}$ [30], and software is available on-line for doing this (e.g., Enhanced KM Algorithms

[33] For continuous universes of discourse, this number is uncountably infinite; however, for sampled continuous universes of discourse (as is the case when a digital computer is used to solve problems involving IT2 FSs), there are a large but countable number of such T1 FSs.

[34] Strictly speaking, the centroid is also a FS; however, because its membership grade is 1 for all of its elements (due to the membership grades of an T2 FS all being equal to 1), it is customary not to show its membership grades.

http://sipi.usc.edu/mendel/software/,
https://sites.google.com/site/drwu09/publications/completepubs)

but none are not needed for reading the rest of this paper.

For linguistic terms, whose FOUs are located in a natural ordering on the variable's numerical axis, the centroid is also located in a natural ordering on the variable's numerical axis. This is shown in Fig. B.2 for the three linguistic terms, *W1–W3*.

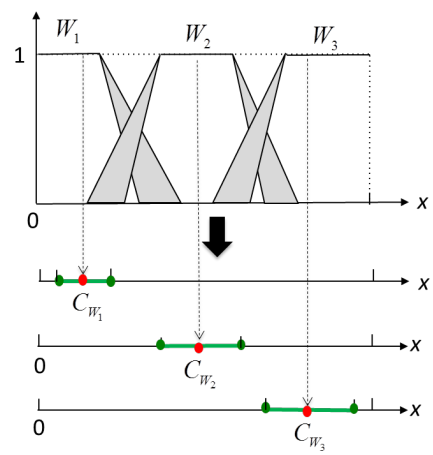

Fig. B.2. Mapping each word's FOU into an interval of numbers, its Centroid C_{W_i}, or a single number, its COG c_{W_i}, the dot in the middle of the Centroid.

B.2 COG of Centroid of an IT2 FS

The COG, $c_{\tilde{A}_i}$, of $C_{\tilde{A}_i}$ is a number that expresses the average uncertainty about each linguistic term. We shall refer to $c_{\tilde{A}_i}$ as the *COG of the word*. COGs are also located in a natural ordering on the variable's numerical axis, as is indicated in Fig. B.2. The formula for $c_{\tilde{A}_i}$ is:

$$c_{\tilde{A}_i} = \tfrac{1}{2}[\underline{C}_{\tilde{A}_i} + \overline{C}_{\tilde{A}_i}] \tag{B.2}$$

It lies along the *x*-axis.

B.3 Maximum Dispersion of the Centroid

Just as the standard deviation provides a useful measure of dispersion about the mean in probability, it also provides a useful measure of dispersion about the centroid end-points for an IT2 FS. Wu and Mendel [51] show that the variance of an IT2 FS $\tilde{A}_i$, $V_{\tilde{A}_i}$, is the union of relative variance, $v_{\tilde{A}_i}(A_e)$, of all its embedded T1 FSs A_e, i.e.,

$$V_{\tilde{A}_i} = \bigcup_{\forall A_e} v_{\tilde{A}_i}(A_e) = [v_l(\tilde{A}_i), v_r(\tilde{A}_i)] = \left[\min_{\forall A_e} v_{\tilde{A}_i}(A_e), \max_{\forall A_e} v_{\tilde{A}_i}(A_e)\right] \tag{B.3}$$

where (for sampled values of x)

$$v_{\tilde{A}_i}(A_e) = \frac{\sum_{i=1}^{N}\left[x_i - c_{\tilde{A}_i}\right]^2 \mu_{A_e}(x_i)}{\sum_{i=1}^{N}\mu_{A_e}(x_i)} \tag{B.4}$$

in which $c_{\tilde{A}_{ii}}$ is computed by (B.2). How to compute $V_{\tilde{A}_i}$ (by using EKM algorithms) is explained in [51].

The standard deviation of $\tilde{A}_i$, $S_{\tilde{A}_i}$, is

$$S_{\tilde{A}_i} = \left[\underline{S}_{\tilde{A}_i}, \overline{S}_{\tilde{A}_i}\right] = \left[\sqrt{v_l(\tilde{A}_i)}, \sqrt{v_r(\tilde{A}_i)}\right] \tag{B.5}$$

One way that $S_{\tilde{A}_i}$ has been used [51] is to combine it with $c_{\tilde{A}_i}$ to provide *left and right end point intervals of dispersion* $D_{\tilde{A}_i}$ about $c_{\tilde{A}_i}$, as

$$D_{\tilde{A}_i} = \left[[c_{\tilde{A}_i} - \sqrt{v_r(\tilde{A}_i)}, c_{\tilde{A}_i} - \sqrt{v_l(\tilde{A}_i)}], [c_{\tilde{A}_i} + \sqrt{v_l(\tilde{A}_i)}, c_{\tilde{A}_i} + \sqrt{v_r(\tilde{A}_i)}]\right] \tag{B.6}$$

We refer to the outer interval of $D_{\tilde{A}}$ as $\max D_{\tilde{A}_i}$, where

$$\max D_{\tilde{A}_i} = [\underline{\max D_{\tilde{A}_i}}, \overline{\max D_{\tilde{A}_i}}] = [c_{\tilde{A}_i} - \sqrt{v_r(\tilde{A}_i)}, c_{\tilde{A}_i} + \sqrt{v_r(\tilde{A}_i)}] \tag{B.7}$$

It is an interval of real numbers that lies along the x-axis and is of length $2\sqrt{v_r(\tilde{A}_i)}$.

A figure showing $\max D_{\tilde{A}_i}$ for the three FOUs depicted in Fig. B.1 would look very similar to the centroid intervals that are shown on that figure. Each of the $\max D_{\tilde{A}_i}$ intervals would also be centered around the COG of each centroid, but the $\max D_{W_i}$ interval would be of length $2\sqrt{v_r(W_i)}$ whereas C_{W_i} is of length $\overline{C}_{W_i} - \underline{C}_{W_i}$.

APPENDIX C. IV FSQCA

In this appendix, the T1 FSs that are used in the steps of T1 fsQCA are replaced by the new S-shaped approximated MFs for RI L2 FS, $\tilde{\mu}_{\tilde{\tilde{v}}}(x)$, i.e. the MF of the linguistic variable $\tilde{\tilde{v}}$. We use double-tilde over-bars over a letter to remind the reader that these quantities are modeled as L2 FSs. Some of the T1 fsQCA steps do not change when $\tilde{\mu}_{\tilde{\tilde{v}}}(x)$ are used; however, for completeness, we show all of the steps of IV fsQCA. Our wording of the steps follows the wording of the T1 fsQCA steps given in [33] and [34].

Step 1. *Choose a desired outcome,* $\tilde{\tilde{O}}$*, and its appropriate cases*: Let $S_{\tilde{\tilde{O}}}$ be the finite space of possible outcomes, $\tilde{\tilde{O}}^w$, for a specific application, i.e.

$$S_{\tilde{\tilde{O}}} = \{\tilde{\tilde{O}}^w, w = 1,...,n_O\} \tag{C.1}$$

The desired outcome, which is application dependent, is $\tilde{\tilde{O}}$, where $\tilde{\tilde{O}} \in S_{\tilde{\tilde{O}}}$. IV fsQCA, just as T1 fsQCA, focuses on one outcome at a time, and each IV fsQCA is independent of the others.

Step 1 for IV fsQCA is exactly the same as it is for T1 fsQCA.

Let S_{Cases} be the finite space of all appropriate cases (x) that have been labeled 1, 2, …, N, i.e.

$$S_{Cases} = \{1,2,...,N\} \tag{C.2}$$

Step 2. *Choose k causal conditions (variables*[35]*),* $\tilde{\tilde{C}}_i$ (i =1,…, k). Let $S_{\tilde{\tilde{C}}'}$ be the finite space of all possible causal conditions, $\tilde{\tilde{C}}'_{i'}$, for the specific application, i.e. $S_{\tilde{\tilde{C}}'} = \{\tilde{\tilde{C}}'_{i'},\ i' = 1,...,n_{C'}\}$. A subset of the $\tilde{\tilde{C}}'_{i'}$ possible causal conditions, $S_{\tilde{\tilde{C}}}$, is chosen whose elements are re-numbered $1,2,...,k$, i.e.

$$S_{\tilde{\tilde{C}}} = \{\tilde{\tilde{C}}_i,\ i = 1,...,k\} \ni \forall \tilde{\tilde{C}}_i \in S_{\tilde{\tilde{C}}'} \tag{C.3}$$

Step 2 for IV fsQCA is exactly the same as it is for T1 fsQCA.

Step 3. *Treat the desired outcome and causal conditions as RI L2 FSs and determine* ***MFs*** *for them*: Obtaining an S-shaped approximated MF for a RI L2 FS (i.e., the MF of the linguistic variable) is explained in Sections 2 and 3. Let $\tilde{\mu}_{\tilde{\tilde{O}}}$ be the MF of the desired outcome, $\tilde{\tilde{O}}$, and $\tilde{\mu}_{\tilde{\tilde{C}}_i}$ be the MF of causal condition i, $\tilde{\tilde{C}}_i$, i.e.[36]

[35] Because we are now using a MF for the entire linguistic variable, there will only be one causal condition per variable, and so *causal condition* and *variable* are synonymous. If, instead, one chooses to use one of the linguistic terms as a causal condition (e.g., Low profit, or High profit), then we advocate using only one of them at a time in fsQCA, in order to avoid the impossible causal combinations problems that are explained in Section 7.

[36] The notation $1/[a,b]$ indicates that membership grade is 1 for all elements in [a, b].

$$\left.\begin{array}{r}\tilde{\mu}_{\tilde{O}}:\Omega\subseteq\mathbb{R}\to[0,1]\\ \omega\mapsto\mu_{\tilde{O}}(\omega)\\ \tilde{\mu}_{\tilde{O}}(\omega)=1/[\underline{\tilde{\mu}}_{\tilde{O}}(\omega),\overline{\tilde{\mu}}_{\tilde{O}}(\omega)]\end{array}\right\} \tag{C.4}$$

$$\left.\begin{array}{r}\tilde{\mu}_{\tilde{C}_i}:\Xi_i\subseteq\mathbb{R}\to[0,1]\\ \xi_i\mapsto\mu_{\tilde{C}_i}(\xi_i)\\ \tilde{\mu}_{\tilde{C}_i}(\xi_i)=1/[\underline{\tilde{\mu}}_{\tilde{C}_i}(\xi_j),\overline{\tilde{\mu}}_{\tilde{C}_i}(\xi_i)]\end{array}\right\} i=1,...,k \tag{C.5}$$

This step is different from T1 fsQCA because it uses S-shaped approximated MFs.

Step 4. *Evaluate the MFs for all N appropriate cases the results being the* ***IV derived membership functions****, i.e.* ($x=1,...,N$ and $i=1,...,k$)

$$\left.\begin{array}{r}\tilde{\mu}^D_{\tilde{O}}:(S_{Cases},S_{\tilde{O}})\to[0,1]\\ x\mapsto\omega(x)\mapsto\tilde{\mu}_{\tilde{O}}(\omega(x))\equiv\tilde{\mu}^D_{\tilde{O}}(x)=1/[\underline{\tilde{\mu}}_{\tilde{O}}(x),\overline{\tilde{\mu}}_{\tilde{O}}(x)]\end{array}\right\} \tag{C.6}$$

$$\left.\begin{array}{r}\tilde{\mu}^D_{\tilde{C}_i}:(S_{Cases},S_{\tilde{C}})\to[0,1]\\ x\mapsto\xi_i(x)\mapsto\tilde{\mu}_{\tilde{C}_i}(\xi_i(x))\equiv\tilde{\mu}^D_{\tilde{C}_i}(x)=1/[\underline{\tilde{\mu}}_{\tilde{C}_i}(x),\overline{\tilde{\mu}}_{\tilde{C}_i}(x)]\end{array}\right\} \tag{C.7}$$

Step 5. *Conceptually, create 2^k* ***candidate causal combinations*** *(rules) and view each as a corner in a 2^k-dimensional vector space.* Let $S_{\tilde{\tilde{F}}}$ be the finite space of 2^k candidate causal combinations, called (by us) ***firing interval fuzzy sets***, $\tilde{\tilde{F}}_j$, i.e. ($i=1,...,k$ and $j=1,...,2^k$)

$$\left.\begin{array}{r}S_{\tilde{\tilde{F}}}=\{\tilde{\tilde{F}}_1,...,\tilde{\tilde{F}}_{2^k}\}\ \ni\tilde{\tilde{F}}_j=\tilde{\tilde{A}}_1^j\wedge\tilde{\tilde{A}}_2^j\wedge...\wedge\tilde{\tilde{A}}_k^j\\ \tilde{\tilde{A}}_i^j=\tilde{\tilde{C}}_i\text{ or }\tilde{\tilde{c}}_i\end{array}\right\} \tag{C.8}$$

where $\tilde{\tilde{c}}_i$ denotes the complement of $\tilde{\tilde{C}}_i$. This step is taken from Fast fsQCA [32, 35], and is actually the same as in Fast fsQCA, except that it now uses RI L2 FSs.

Step 6. *Compute the R_S* ***surviving causal combinations*** *(firing interval surviving rules)*

$$\tilde{\tilde{F}}_{j^*(x)}(x)=\tilde{\tilde{A}}_1^{j^*(x)}\wedge\tilde{\tilde{A}}_2^{j^*(x)}\wedge...\wedge\tilde{\tilde{A}}_k^{j^*(x)} \tag{C.9}$$

where $(\forall x\in S_{Cases})$

$$\tilde{\tilde{A}}_i^{j^*(x)}\triangleq arg\max\left(\tilde{\mu}_{\tilde{\tilde{C}}}(x),\tilde{\mu}_{\tilde{\tilde{c}}}(x)\right)=\left\{\begin{array}{ll}\tilde{\tilde{C}} & \overline{\tilde{\mu}}_{\tilde{\tilde{C}}}(x)\geq\overline{\tilde{\mu}}_{\tilde{\tilde{c}}}(x)\\ \tilde{\tilde{c}} & \overline{\tilde{\mu}}_{\tilde{\tilde{c}}}(x)\geq\overline{\tilde{\mu}}_{\tilde{\tilde{C}}}(x)\end{array}\right. \tag{C.10}$$

in which

$$\tilde{\mu}_{\tilde{c}_j}(x) = 1/[1-\bar{\tilde{\mu}}_{\tilde{C}_j}(x), 1-\underline{\tilde{\mu}}_{\tilde{C}_j}(x)] \tag{C.11}$$

Then, keep the adequately different R_S *surviving causal combinations that occur for more than f cases.* This is a mapping from $\{S_{\tilde{C}}, S_{Cases}\}$ into $S_{\tilde{F}^S}$ that makes use of $\tilde{\mu}_{\tilde{C}_i}(x)$ and $\tilde{\mu}_{\tilde{c}_i}(x)$, where $S_{\tilde{\tilde{F}}^S}$ is a subset of $S_{\tilde{\tilde{F}}}$, with R_S elements, i.e. $\tilde{\tilde{F}}_{j*(x)}(x)$ is first computed by using (C.9) and (C.10), after which the R_S *uniquely different* $\tilde{\tilde{F}}_{j*(x)}(x)$ are relabeled $\tilde{\tilde{F}}_l$ ($l = 1,...,R_S$, and $S_{\tilde{\tilde{F}}_l} = \{\tilde{\tilde{F}}_1,...,\tilde{\tilde{F}}_{R_S}\}$). The detailed computations are ($l = 1,...,R_S$ and $x = 1,...,N$):

$$\begin{aligned} &\tilde{\mu}_{\tilde{\tilde{F}}_{j*(x)}} : (S_{\tilde{\tilde{F}}}, S_{Cases}) \to [0,1] \\ &\quad x \mapsto \tilde{\mu}_{\tilde{\tilde{F}}_{j*(x)}}(x) = 1/[\underline{\tilde{\mu}}_{\tilde{\tilde{F}}_{j*(x)}}(x), \bar{\tilde{\mu}}_{\tilde{\tilde{F}}_{j*(x)}}(x)] \end{aligned} \tag{C.12}$$

$$\left.\begin{aligned} \underline{\tilde{\mu}}_{\tilde{\tilde{F}}_{j*(x)}}(x) &= \min\left\{\underline{\tilde{\mu}}_{\tilde{\tilde{A}}_1^{j*(x)}}(x), \underline{\tilde{\mu}}_{\tilde{\tilde{A}}_2^{j*(x)}}(x), ..., \underline{\tilde{\mu}}_{\tilde{\tilde{A}}_k^{j*(x)}}(x)\right\} \\ \bar{\tilde{\mu}}_{\tilde{\tilde{F}}_{j*(x)}}(x) &= \min\left\{\bar{\tilde{\mu}}_{\tilde{\tilde{A}}_1^{j*(x)}}(x), \bar{\tilde{\mu}}_{\tilde{\tilde{A}}_2^{j*(x)}}(x), ..., \bar{\tilde{\mu}}_{\tilde{\tilde{A}}_k^{j*(x)}}(x)\right\} \end{aligned}\right\} \tag{C.13}$$

$$\left.\begin{aligned} &t_{\tilde{\tilde{F}}_{j*}} : (S_{\tilde{\tilde{F}}_{j*}}, S_{\tilde{\tilde{F}}_{j'}}, S_{Cases}) \to \{0,1\} \\ &\quad x \mapsto t_{\tilde{\tilde{F}}_{j*}}(x) = \begin{cases} 1 & \text{if } \tilde{\tilde{F}}_{j*} = \tilde{\tilde{F}}_{j*(x)}(x) \\ 0 & \text{otherwise} \end{cases} \end{aligned}\right\} \tag{C.14}$$

$$\left.\begin{aligned} &N_{\tilde{\tilde{F}}_{j*}} : \{0,1\} \to I \\ &\quad t_{\tilde{\tilde{F}}_{j*}} \to N_{\tilde{\tilde{F}}_{j*}} = \sum_{x=1}^{N} t_{\tilde{\tilde{F}}_{j*}}(x) \end{aligned}\right\} \tag{C.15}$$

$$\left.\begin{aligned} &\tilde{\tilde{F}}_l^S : (S_{\tilde{\tilde{F}}}, I) \to S_{\tilde{\tilde{F}}^S} \\ &\quad \tilde{\tilde{F}}_{j*} \mapsto \tilde{\tilde{F}}_l^S = \begin{cases} \tilde{\tilde{F}}_{j*}(j* \to l) & \text{if } N_{\tilde{\tilde{F}}_{j*}} \geq f \\ 0 & \text{if } N_{\tilde{\tilde{F}}_{j*}} < f \end{cases} \end{aligned}\right\} \tag{C.16}$$

In (C.16) *f* is an integer frequency threshold that must be set by the user. The firing intervals for these R_S *surviving rules* are denoted $\tilde{\tilde{F}}_l^S$ with associated re-numbered membership functions $\mu_{\tilde{\tilde{F}}_l^S}(x)$, $l = 1,...,R_S$.

This step is very different from T1 fsQCA because the winning causal condition in (C.9) is computed by using the IT2 min-max theorem (see Appendix D), which is an extension of the T1 min-max Theorem [34] (that is used in T1 fsQCA) to IT2 FSs [22].

Step 7. *Compute the consistencies (subsethoods) of the R_S surviving causal combinations, and keep only those R_A causal combinations—the* ***actual causal combinations*** *(actual rules)—whose consistencies are greater than 0.80.* This is a mapping from $(S_{\tilde{\tilde{F}}^S}, \tilde{\tilde{O}}, S_{Cases}\}$ into $S_{\tilde{\tilde{F}}^A}$, where $S_{\tilde{\tilde{F}}^A}$ is a subset of $S_{\tilde{\tilde{F}}^S}$, with R_A elements, i.e. ($l = 1,...,R_S$ and $m = 1,...,R_A$):

$$\left.\begin{array}{l} ss_{VS}(\tilde{\tilde{F}}_l^S, \tilde{\tilde{O}}) : S_{\tilde{\tilde{F}}^S}, \tilde{\tilde{O}}, S_{Cases}\} \to [0,1] \\ \left\{\tilde{\mu}_{\tilde{\tilde{F}}_l^S}(x), \tilde{\mu}_{\tilde{\tilde{O}}}^D(x)\right\}_{x=1}^N \mapsto ss_{VS}(\tilde{\tilde{F}}_l^S, \tilde{\tilde{O}}) \\ \qquad = \dfrac{\sum_{x=1}^N \min\left(\underline{\tilde{\mu}}_{\tilde{\tilde{F}}_l^S}(x), \underline{\tilde{\mu}}_{\tilde{\tilde{O}}}(x)\right) + \sum_{x=1}^N \min\left(\overline{\tilde{\mu}}_{\tilde{\tilde{F}}_l^S}(x), \overline{\tilde{\mu}}_{\tilde{\tilde{O}}}(x)\right)}{\sum_{x=1}^N \underline{\tilde{\mu}}_{\tilde{\tilde{F}}_l^S}(x) + \sum_{x=1}^N \overline{\tilde{\mu}}_{\tilde{\tilde{F}}_l^S}(x)} \end{array}\right\} \tag{C.17}$$

$$\left.\begin{array}{l} \tilde{\tilde{F}}_m^A : [0,1] \to S_{\tilde{\tilde{F}}^A} \\ ss_{VS}(\tilde{\tilde{F}}_l^S, \tilde{\tilde{O}}) \mapsto \tilde{\tilde{F}}_m^A = \left\{\tilde{\tilde{F}}_l^S (l \to m) \middle| ss_{VS}(\tilde{\tilde{F}}_l^S, \tilde{\tilde{O}}) \geq 0.80,\ l = 1,...,R_S\right\} \end{array}\right\} \tag{C.18}$$

The firing intervals for these actual rules, that are denoted $\tilde{\tilde{F}}_m^A$, have associated re-numbered MFs $\tilde{\mu}_{\tilde{\tilde{F}}_m^A}(x)$, $m = 1,...,R_A$, and can be expressed (the superscript A in each $\tilde{\tilde{A}}_i^{A,m}$ denotes "actual"), as:

$$\tilde{\tilde{F}}_m^A = \tilde{\tilde{A}}_1^{A,m} \wedge \tilde{\tilde{A}}_2^{A,m} \wedge ... \wedge \tilde{\tilde{A}}_k^{A,m} \tag{C.19}$$

This step is different from T1 fsQCA because it uses S-shaped approximated MFs and a subsethood formula in (C.17) that is valid for such FSs that is due originally to Vlachos and Sergiadis [48], because (C.17) reduces to the same subsethood formula that is used in T1 fsQCA when all IV L2 FSs reduce to T1 L2 FSs.

The rest of IV fsQCA simplifies the actual causal combinations, as in T1 fsQCA.

Step 8. *Use the Quine-McCluskey (QM) algorithm to obtain* ***complex solutions*** (*prime implicants*) *and* ***parsimonious solutions*** (*minimal prime implicants*). A mapping from $S_{\tilde{\tilde{F}}^A}$, $S_{\tilde{\tilde{F}}} - S_{\tilde{\tilde{F}}^S}$ and $S_{\tilde{\tilde{F}}^S} - S_{\tilde{\tilde{F}}^A}$ into $S_{\tilde{\tilde{F}}^{PI}}$ is used to obtain the complex solution, i.e.

$$\left.\begin{array}{l} \tilde{\tilde{F}}^{PI} : \{S_{\tilde{\tilde{F}}^A}, S_{\tilde{\tilde{F}}^S} - S_{\tilde{\tilde{F}}^A}, S_{\tilde{\tilde{F}}} - S_{\tilde{\tilde{F}}^S}\} \to S_{\tilde{\tilde{F}}^{PI}} \\ \left\{\tilde{\tilde{F}}_j\right\}_{j=1}^{2^k} \mapsto \left\{\tilde{\tilde{F}}_n^{PI}\right\}_{n=1}^{R_C}, \tilde{\tilde{F}}_n^{PI} = QM_{PI}\begin{pmatrix} S_{\tilde{\tilde{F}}^A} & \text{present} \\ S_{\tilde{\tilde{F}}} - S_{\tilde{\tilde{F}}^S} & \text{absent} \\ S_{\tilde{\tilde{F}}^S} - S_{\tilde{\tilde{F}}^A} & \text{absent} \end{pmatrix} \end{array}\right\} \tag{C.20}$$

The parsimonious solution, $S_{\tilde{\tilde{F}}^{MPI}}$, is obtained by a different application of the QM algorithm, i.e.

$$\left.\begin{array}{l} \tilde{\tilde{F}}^{MPI} : \{S_{\tilde{\tilde{F}}^{A}}, S_{\tilde{\tilde{F}}^{S}} - S_{\tilde{\tilde{F}}^{A}}, S_{\tilde{\tilde{F}}} - S_{\tilde{\tilde{F}}^{S}}, S_{\tilde{F}}\} \rightarrow S_{\tilde{\tilde{F}}^{MPI}} \\ \left\{\tilde{\tilde{F}}_j\right\}_{j=1}^{2^k} \mapsto \left\{\tilde{\tilde{F}}_p^{MPI}\right\}_{p=1}^{R_P}, \tilde{\tilde{F}}_p^{MPI} = QM_{MPI}\begin{pmatrix} S_{\tilde{\tilde{F}}^{A}} & \text{present} \\ S_{\tilde{\tilde{F}}} - S_{\tilde{\tilde{F}}^{S}} & \text{don't care} \\ S_{\tilde{\tilde{F}}^{S}} - S_{\tilde{\tilde{F}}^{A}} & \text{absent} \end{pmatrix} \end{array}\right\} \quad \text{(C.21)}$$

This step of IV fsQCA is performed in the crisp domain so it is the same as in T1 fsQCA.

Step 9. *Use additional substantive knowledge, obtained from an expert, to perform* ***Counterfactual Analysis*** *(CA) on each term of the complex solutions (one at a time), but constrained by each term of the parsimonious solutions (one at a time), to obtain* ***intermediate solutions***. To begin, this substantive knowledge is used (in thought experiments) to establish the presence or the absence of each causal condition or its complement on the desired outcome. This is a transformation of $S_{\tilde{\tilde{C}}}$ into $S_{K(\tilde{\tilde{C}})}$, where $K(\tilde{\tilde{C}})$ denotes knowledge applied to $\tilde{\tilde{C}}$. Then $S_{K(\tilde{\tilde{C}})}$ is used to map ($S_{\tilde{\tilde{F}}^{PI}}, S_{\tilde{\tilde{F}}^{MPI}}$), by means of CA into $S_{\tilde{\tilde{F}}^{I}}$, the space of *intermediate solutions* that contains R_I elements, i.e.

$$\left.\begin{array}{l} \tilde{\tilde{C}} : S_{\tilde{\tilde{C}}} \rightarrow S_{K(\tilde{\tilde{C}})} \\ \tilde{\tilde{C}}_i \overset{K}{\mapsto} \left\{\text{Either } \tilde{\tilde{C}}_i, \text{ or } \tilde{\tilde{c}}_i, \text{ or } \textit{unknown}\right\} \end{array}\right\} \quad \text{(C.22)}$$

$$\left.\begin{array}{l} \tilde{\tilde{F}}^{I} : (S_{K(\tilde{\tilde{C}})}, S_{\tilde{\tilde{F}}^{MPI}}, S_{\tilde{\tilde{F}}^{PI}}) \rightarrow S_{\tilde{\tilde{F}}^{I}} \\ \left(K(\tilde{\tilde{C}}), \{\tilde{\tilde{F}}_n^{PI}\}_{n=1}^{R_C}, \{\tilde{\tilde{F}}_p^{MPI}\}_{p=1}^{R_P}\right) \overset{CA}{\mapsto} \{\tilde{\tilde{F}}_q^{I}\}_{q=1}^{R_I} \end{array}\right\} \quad \text{(C.23)}$$

In the thought experiments one asks: Based on my expert knowledge, (1) Do I believe that $\tilde{\tilde{C}}_i$ strongly influences the desired output? If the answer is YES, then stop, and $\tilde{\tilde{C}}_i$ is put on the list of substantive knowledge. On the other hand, if the answer is NO or DON'T KNOW, then one asks: (2) Is it, instead, $\tilde{\tilde{c}}_i$ that strongly influences the desired output? If the answer is YES, then $\tilde{\tilde{c}}_i$ is put on the list of substantive knowledge. If the answer is NO or DON'T KNOW, then neither $\tilde{\tilde{C}}_i$ nor $\tilde{\tilde{c}}_i$ are put on the list of substantive knowledge, i.e. the substantive knowledge is silent about the causal condition or its complement.

This step is the same as in T1 fsQCA and is also performed in the crisp domain.

Step 10. *Perform QM on the intermediate solutions to obtain the* ***simplified intermediate solutions***. It is quite possible that the union of the intermediate rules can be further simplified;

this step does this, and is similar to Step 8. This is a mapping from $S_{\tilde{\tilde{F}}^{I}}$ and $S_{\tilde{\tilde{F}}} - S_{\tilde{\tilde{F}}^{I}}$, but into $S_{\tilde{\tilde{F}}^{SI}}$ by yet another application of the QM algorithm, i.e.

$$\left.\begin{array}{c} \tilde{\tilde{F}}^{SI} : \{S_{\tilde{\tilde{F}}^{I}}, S_{\tilde{\tilde{F}}} - S_{\tilde{\tilde{F}}^{I}}\} \rightarrow S_{\tilde{\tilde{F}}^{SI}} \\ \left\{\tilde{\tilde{F}}_j\right\}_{j=1}^{2^k} \mapsto \left\{\tilde{\tilde{F}}_r^{SI}\right\}_{r=1}^{R_{SI}}, \tilde{\tilde{F}}_r^{SI} = QM_{SI}\begin{pmatrix} S_{\tilde{\tilde{F}}^{I}} & \text{present} \\ S_{\tilde{\tilde{F}}} - S_{\tilde{\tilde{F}}^{I}} & \text{absent} \end{pmatrix} \end{array}\right\} \tag{C.24}$$

This step is the same as in T1 fsQCA and is also in the crisp domain.

Step 11. *Retain only those simplified intermediate solutions whose consistencies are approximately greater than 0.80, the* ***believable simplified intermediate solutions.*** This is a mapping from $\{S_{\tilde{\tilde{F}}^{SI}}, \tilde{\tilde{O}}, S_{Cases}\}$ into $S_{\tilde{\tilde{F}}^{BSI}}$, where $S_{\tilde{\tilde{F}}^{BSI}}$ is a subset of $S_{\tilde{\tilde{F}}^{SI}}$ with R_{BSI} elements, i.e. $(r = 1,...,R_{SI}$ and $s = 1,...,R_{BSI})$:

$$\left.\begin{array}{c} ss_{VS}(\tilde{\tilde{F}}_r^{SI}, \tilde{\tilde{O}}) : \{S_{\tilde{\tilde{F}}^{SI}}, \tilde{\tilde{O}}, S_{Cases}\} \rightarrow [0,1] \\ \left\{\tilde{\mu}_{\tilde{\tilde{F}}_r^{SI}}(x), \tilde{\mu}_{\tilde{\tilde{O}}}^{D}(x)\right\}_{x=1}^{N} \mapsto ss_{VS}(\tilde{\tilde{F}}_r^{SI}, \tilde{O}) = \dfrac{\sum_{x=1}^{N} \min\left(\underline{\tilde{\mu}}_{\tilde{\tilde{F}}_r^{SI}}(x), \underline{\tilde{\mu}}_{\tilde{\tilde{O}}}(x)\right) + \sum_{x=1}^{N} \min\left(\overline{\tilde{\mu}}_{\tilde{\tilde{F}}_r^{SI}}(x), \overline{\tilde{\mu}}_{\tilde{\tilde{O}}}(x)\right)}{\sum_{x=1}^{N} \underline{\tilde{\mu}}_{\tilde{\tilde{F}}_r^{SI}}(x) + \sum_{x=1}^{N} \overline{\tilde{\mu}}_{\tilde{\tilde{F}}_r^{SI}}(x)} \end{array}\right\} \tag{C.25}$$

$$\left.\begin{array}{c} \tilde{\tilde{F}}_s^{BSI} : [0,1] \rightarrow S_{\tilde{\tilde{F}}^{BSI}} \\ ss_{VS}(\tilde{\tilde{F}}_r^{SI}, \tilde{\tilde{O}}) \mapsto \tilde{\tilde{F}}_s^{BSI} = \left\{\begin{array}{c} \tilde{\tilde{F}}_r^{SI}(r \rightarrow s) \,\middle|\, ss_K(\tilde{\tilde{F}}_r^{SI}, \tilde{\tilde{O}}) \geq 0.80, \\ r = 1,...,R_{SI} \end{array}\right\} \end{array}\right\} \tag{C.26}$$

This step is different from the one in T1 fsQCA because it uses Vlachos and Sergiadis's IT2 FS subsethood measure [48, 51, 37], as in Step 7.

APPENDIX D. IT2 MIN-MAX THEOREM

We use interval ranking [18] in order to find the winning causal conditions between an IT2 causal condition and its complement.

Theorem 1 (IT2 Min-max Theorem): Given k causal conditions, $\tilde{\tilde{C}}_1$, $\tilde{\tilde{C}}_2$,…, $\tilde{\tilde{C}}_k$ and their respective complements, $\tilde{\tilde{c}}_1$, $\tilde{\tilde{c}}_2$, …, $\tilde{\tilde{c}}_k$ where

$$\tilde{\mu}_{\tilde{\tilde{C}}_i}(x)=\left[\underline{\tilde{\mu}}_{\tilde{\tilde{C}}_i}(x),\overline{\tilde{\mu}}_{\tilde{\tilde{C}}_i}(x)\right]=\left[1-\overline{\tilde{\mu}}_{\tilde{\tilde{c}}_i}(x),1-\underline{\tilde{\mu}}_{\tilde{\tilde{c}}_i}(x)\right] \tag{D.1}$$

Consider the 2^k candidate causal combinations ($j=1,...,2^k$) $\tilde{\tilde{F}}_j=\tilde{\tilde{A}}_1^j\wedge\tilde{\tilde{A}}_2^j\wedge...\wedge\tilde{\tilde{A}}_k^j$ where $\tilde{\tilde{A}}_i^j=\tilde{\tilde{C}}_i$ or $\tilde{\tilde{c}}_i$ and $i=1,...,k$. Let $\tilde{\mu}_{\tilde{\tilde{F}}_j}(x)=\left[\underline{\tilde{\mu}}_{\tilde{\tilde{F}}_j}(x),\overline{\tilde{\mu}}_{\tilde{\tilde{F}}_j}(x)\right]$ where ($x=1,2,...,N$):

$$\left.\begin{aligned}\underline{\tilde{\mu}}_{\tilde{\tilde{F}}_j}(x)&=\min\left\{\underline{\tilde{\mu}}_{\tilde{\tilde{A}}_1^j}(x),\underline{\tilde{\mu}}_{\tilde{\tilde{A}}_2^j}(x),...,\underline{\tilde{\mu}}_{\tilde{\tilde{A}}_k^j}(x)\right\}\\ \overline{\tilde{\mu}}_{\tilde{\tilde{F}}_j}(x)&=\min\left\{\overline{\tilde{\mu}}_{\tilde{\tilde{A}}_1^j}(x),\overline{\tilde{\mu}}_{\tilde{\tilde{A}}_2^j}(x),...,\overline{\tilde{\mu}}_{\tilde{\tilde{A}}_k^j}(x)\right\}\end{aligned}\right\} \tag{D.2}$$

$\tilde{\tilde{F}}_{j^*(x)}(x)$ is determined by:

$$\tilde{\tilde{F}}_{j^*(x)}(x)=\tilde{\tilde{A}}_1^{j^*(x)}\wedge\tilde{\tilde{A}}_2^{j^*(x)}\wedge...\wedge\tilde{\tilde{A}}_k^{j^*(x)} \tag{D.3}$$

where

$$\tilde{\tilde{A}}_i^{j^*(x)}\triangleq arg\max\left(\tilde{\mu}_{\tilde{\tilde{C}}}(x),\tilde{\mu}_{\tilde{\tilde{c}}}(x)\right)=\begin{cases}\tilde{\tilde{C}} & \overline{\tilde{\mu}}_{\tilde{\tilde{C}}}(x)\geq\overline{\tilde{\mu}}_{\tilde{\tilde{c}}}(x)\\ \tilde{\tilde{c}} & \overline{\tilde{\mu}}_{\tilde{\tilde{c}}}(x)\geq\overline{\tilde{\mu}}_{\tilde{\tilde{C}}}(x)\end{cases} \tag{D.4}$$

Then for each x (case) there is only one j, $j^*(x)$, for which $\tilde{\mu}_{\tilde{\tilde{F}}_{j^*(x)}}(x)=\left[\underline{\tilde{\mu}}_{\tilde{\tilde{F}}_{j^*(x)}}(x),\overline{\tilde{\mu}}_{\tilde{\tilde{F}}_{j^*(x)}}(x)\right]$ has the highest interval ranking and $\tilde{\mu}_{\tilde{\tilde{F}}_{j^*(x)}}(x)$ can be computed as $\left[\underline{\tilde{\mu}}_{\tilde{\tilde{F}}_{j^*(x)}}(x),\overline{\tilde{\mu}}_{\tilde{\tilde{F}}_{j^*(x)}}(x)\right]$, where:

$$\left.\begin{aligned}\underline{\tilde{\mu}}_{\tilde{\tilde{F}}_{j^*(x)}}(x)&=\min\left\{\underline{\tilde{\mu}}_{\tilde{\tilde{A}}_1^{j^*(x)}}(x),\underline{\tilde{\mu}}_{\tilde{\tilde{A}}_2^{j^*(x)}}(x),...,\underline{\tilde{\mu}}_{\tilde{\tilde{A}}_k^{j^*(x)}}(x)\right\}\\ \overline{\mu}_{\tilde{F}_{j^*(x)}}(x)&=\min\left\{\overline{\tilde{\mu}}_{\tilde{\tilde{A}}_1^{j^*(x)}}(x),\overline{\tilde{\mu}}_{\tilde{\tilde{A}}_2^{j^*(x)}}(x),...,\overline{\tilde{\mu}}_{\tilde{\tilde{A}}_k^{j^*(x)}}(x)\right\}\end{aligned}\right\} \tag{D.5}$$

In (D.5), $arg\max\left(\mu_{\tilde{\tilde{C}}}(x),\mu_{\tilde{\tilde{c}}}(x)\right)$ denotes the winner of $\max\left(\mu_{\tilde{\tilde{C}}}(x),\mu_{\tilde{\tilde{c}}}(x)\right)$, namely $\tilde{\tilde{C}}_i$ or $\tilde{\tilde{c}}_i$.

Proof: We used Ishibuchi and Tanaka's [12] or Hu and Wang's [11] approaches for maximization problems of intervals. Intervals can be classified into non-overlapping, partially overlapping and completely overlapping intervals. Only two types of intervals occur when a causal condition is compared to its complement, as is depicted in Fig. D.1 for four different situations.

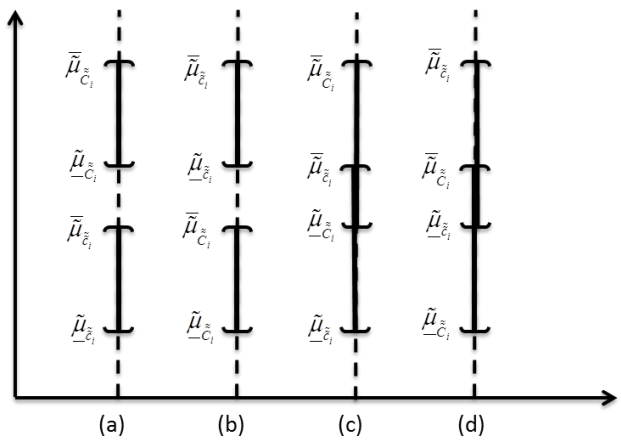

Fig. D.1. Four different situations of the IT2 FSs for a case: (a), (b) non-overlapping intervals, and (c), (d) overlapping intervals.

If $A=\left[a_L,a_R\right]$ and $B=\left[b_L,b_R\right]$ are two intervals, then Ishibuchi and Tanaka's [12] approach defines the order relation for maximization problems as

$$A \geq B \quad iff \quad a_L \geq b_L \quad and \quad a_R \geq b_R \tag{D.6}$$

Applying (D.6) to $\tilde{\mu}_{\tilde{C}_i}(x)=\left[\underline{\tilde{\mu}}_{\tilde{C}_i}(x),\overline{\tilde{\mu}}_{\tilde{C}_i}(x)\right]$ and $\tilde{\mu}_{\tilde{c}_i}(x)=\left[\underline{\tilde{\mu}}_{\tilde{c}_i}(x),\overline{\tilde{\mu}}_{\tilde{c}_i}(x)\right]$ for the four cases in Fig. D.1, we have:

$$\begin{cases} (a): \quad \overline{\tilde{\mu}}_{\tilde{C}_i} \geq \overline{\tilde{\mu}}_{\tilde{c}_i} \;\; and \;\; \underline{\tilde{\mu}}_{\tilde{C}_i} \geq \underline{\tilde{\mu}}_{\tilde{c}_i} \Rightarrow \quad \tilde{\tilde{C}}_i \geq \tilde{\tilde{c}}_i \\ (b): \quad \overline{\tilde{\mu}}_{\tilde{c}_i} \geq \overline{\tilde{\mu}}_{\tilde{C}_i} \;\; and \;\; \underline{\tilde{\mu}}_{\tilde{c}_i} \geq \underline{\tilde{\mu}}_{\tilde{C}_i} \Rightarrow \quad \tilde{\tilde{c}}_i \geq \tilde{\tilde{C}}_i \\ (c): \quad \overline{\tilde{\mu}}_{\tilde{C}_i} \geq \overline{\tilde{\mu}}_{\tilde{c}_i} \;\; and \;\; \underline{\tilde{\mu}}_{\tilde{C}_i} \geq \underline{\tilde{\mu}}_{\tilde{c}_i} \Rightarrow \quad \tilde{\tilde{C}}_i \geq \tilde{\tilde{c}}_i \\ (d): \quad \overline{\tilde{\mu}}_{\tilde{c}_i} \geq \overline{\tilde{\mu}}_{\tilde{C}_i} \;\; and \;\; \underline{\tilde{\mu}}_{\tilde{c}_i} \geq \underline{\tilde{\mu}}_{\tilde{C}_i} \Rightarrow \quad \tilde{\tilde{c}}_i \geq \tilde{\tilde{C}}_i \end{cases} \tag{D.7}$$

Notice that in (D.7) the winning causal condition is the one whose interval has the higher upper bound interval, as expressed in (D.4).

ACKNOWLEDGEMENT

The authors are pleased to acknowledge the data that was provided to them by Prof. Charles C. Ragin without which it would not have been possible for us to provide the results for the Breakdown of Democracy examples, and to Mohammad Reza Rajati for suggesting L2 fuzzy sets to us.

REFERENCES

[1] J. Aisbett, J. T. Rickard and D. G. Morgenthaler, "Type-2 fuzzy sets as functions on spaces," *IEEE Trans. on Fuzzy Systems*, vol. 18, pp.841-844, June 2010.

[2] A. Bargiela and W. Pedrycz, *Granular Computing: An Introduction*, Kluwer Academic Publishers, Dordrecht, 2003.

[3] A. Bilgin, H. Hagras, A. Malibari, M. Alhaddad and D. Alghazzawi, "Towards A General Type-2 Fuzzy Logic Approach for Computing With Words Using Linear Adjectives," *Proc. FUZZ-IEEE 2012*, pp. 1130-1137, Brisbane, AU, June 2012.

[4] H. Bustince, et al., "Interval type-2 fuzzy sets are generalization of interval-valued fuzzy sets: towards a wide view on their relationship," *IEEE Trans. on Fuzzy Systems,* vol. 23, pp. 1876-1882, October 2015.

[5] G. de Tre and R. de Caluwe, "Level-2 fuzzy sets and their usefulness in object-oriented database modeling," *Fuzzy Sets and Systems*, vol. 140, pp. 29-49, 2003.

[6] D. Dubois and H. Prade, *Fuzzy Sets and Systems: Theory and Applications*, Chapter 2, Academic Press, NY, 1980.

[7] M. Ganesh, *Introduction to Fuzzy Sets and Fuzzy Logic*, Prentice-Hall of India, June 2006.

[8] S. Gottwald, "Set theory for fuzzy sets of higher level," *Fuzzy Sets and Systems,* vol. 2, pp. 125-151, 1979.

[9] M. Hao and J. M. Mendel, "Modeling words by normal interval type-2 fuzzy sets," *Proc. of NAFIPS 2014*, Paper #89, Boston, MA, June 2014.

[10] M. Hao and J. M. Mendel, "Encoding words into normal interval type-2 fuzzy sets: HM approach," *IEEE Trans. on Fizzy Systems*, vol. 24, no. 4, pp. 865-879, 2016.

[11] B. Q. Hu and S. Wang, "A novel approach in uncertain programming part 1: New arithmetic and order relation for interval numbers," *J. of Industrial and Management Optimization*, vol. 2(4), pp. 351-371, 2006.

[12] H. Ishibuchi and H. Tanaka, "Multiobjective programming in optimization of the interval objective function," *European J. of Operational Research*, Vol. 48, pp. 219-225, 1990.

[13] J. Kacprzyk and A. Wilbik, "A comprehensive comparison of time series described by linguistic summaries and its application to the comparison performance of a mutual fund and its benchmark," in *Proc. FUZZ-IEEE'10*, Barcelona, Spain, pp. 25-32, 2010.

[14] J. Kacprzyk and R. R. Yager, "Linguistic summaries of data using fuzzy logic," *Int'l. J. of General Systems*, vol. 30, pp. 133–154, 2001.

[15] J. Kacprzyk and S. Zadrożny, "Computing with words is an implementable paradigm: fuzzy queries, linguistic data summaries and natural language generation," *IEEE Trans. on Fuzzy Systems*, vol. 18, pp.461-472, June 2010.

[16] J. Kacprzyk J., R. R. Yager, and S. Zadrożny, "A fuzzy logic based approach to linguistic summaries of databases," *Int'l. J.of Applied Mathematics and Computer Science*, vol. 10, pp. 813-834, 2000.

[17] G. S. Kanhaiya and S. B. Nimse, "A comparative study of level II fuzzy sets and type-II fuzzy sets," *Int'l. J. of Advanced Research in Computer Engineering & Technology*, vol. 1, no. 4, pp. 215-219, 2012.

[18] S. Karmakar and A.K. Bhunia, "A comparative study of different order relations of intervals," *Reliable Computing*, vol. 16, pp. 38-72, 2012.

[19] N. N. Karnik and J. M. Mendel, "Centroid of a type-2 fuzzy set," *Information Sciences*, vol. 132, pp. 195-220, 2001.

[20] G. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*, Prentice-Hall, Upper Saddle River, NJ, 1995.

[21] M. M. Korjani and J. M. Mendel, "Fuzzy set qualitative comparative analysis (fsQCA): challenges and applications," *Proc. of 2012 NAFIPS Conference*, 6 pages, Berkeley, CA, August 2012.

[22] M. M. Korjani and J. M. Mendel, "Interval type-2 fuzzy set qualitative comparative analysis (IT2-fsQCA)," in *Proc. of NSFIPS 2014*, Paper #93, Boston, MA, June 2014.

[23] G. Lakoff, "Set theory and fuzzy sets: their relationship to natural language," in *Qualitative & Multi-Method Research*, vol. 12, no. 1, pp. 9-14, Spring 2014.

[24] F. Liu and J. M. Mendel, "Encoding words into interval type-2 fuzzy sets using an *Interval Approach*," *IEEE Trans. on Fuzzy Systems*, vol. 16, pp. 1503-1521, December 2008.

[25] J. M. Mendel, "Computing with words, when words can mean different things to different people," Int'l. ICSC Congress on Computer Intelligence: Methods & Applications, *Third Int'l. ICSC Symposium on FL and Appl's.*, Rochester Institute of Technology, Rochester, NY, June, 1999, pp. 158-164.

[26] J. M. Mendel, "The Perceptual Computer: an architecture for computing with words," Modeling With Words Workshop, Proc. of FUZZ-IEEE 2001, Melbourne, AU, Dec. 2001, pp. 35-38.

[27] J. M. Mendel, *Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions*, Prentice-Hall, Upper-Saddle River, NJ, 2001.

[28] J. M. Mendel, "Computing with words and its relationships with fuzzistics," *Information Sciences*, vol. 177, pp. 998-1006, 2007.

[29] J. M. Mendel, "Type-2 fuzzy sets and systems: an overview," *IEEE Computational Intelligence Magazine*, vol. 2, pp. 20-29, Feb. 2007.

[30] J. M. Mendel, "On KM algorithms for solving type-2 fuzzy set problems," *IEEE Trans. on Fuzzy Systems*, vol. 21, pp. 426-446, June 2013.

[31] J. M. Mendel and R. I. John, "Type-2 fuzzy sets made simple," *IEEE Trans. on Fuzzy Systems*, vol. 10, pp. 117-127, April 2002.

[32] J. M. Mendel and M. Korjani, "Fast fuzzy set qualitative comparative analysis (Fast fsQCA)," *Proc. NAFIPS 2012*, Berkeley, CA, Aug. 2012, 6 pages.

[33] J. M. Mendel and M. Korjani, "Charles Ragin's Fuzzy Set Qualitative Comparative Analysis (fsQCA) used for linguistic summarizations," *Information Sciences*, vol. 202, pp. 1-23, 2012.

[34] J. M. Mendel and M. Korjani, "Theoretical aspects of fuzzy set qualitative comparative analysis (fsQCA)," *Information Sciences*, vol. 237, pp. 137-161, 2013.

[35] J. M. Mendel and M. Korjani, "On establishing nonlinear combinations of variables from small to big data for use in later processing," *Information Sciences*, vol. 280, pp. 98-110, 2014.

[36] J. M. Mendel and C. C. Ragin, *fsQCA: Dialog Between Jerry M. Mendel and Charles C. Ragin*, USC-SIPI Report #411 (2nd Ed.) Jan. 2012.

[37] J. M. Mendel and D. Wu, *Perceptual Computing: Aiding People in Making Subjective Judgments*, John Wiley and IEEE Press, Hoboken, NJ, 2010.

[38] J. M. Mendel and D. Wu, "Determining interval type-2 fuzzy set models for words using only data collected from one subject: *Person FOUs*," *Proc. of 2014 IEEE Int'l. Conf. on Fuzzy Systems* (FUZZ-IEEE), June 2014, Beijing, China, pp. 768-775.

[39] W. Pedrycz, "Granular computing: concepts and algorithmic developments," *Appl. Comput. Math.,* vol. 10, no. 1, Special Issue, pp. 175-194, 2011.

[40] C. C. Ragin, *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*, University of California Press, Berkeley, Los Angeles and London, 1987.

[41] C. C. Ragin, *Fuzzy-set Social Science*, Univ. of Chicago Press, Chicago, IL, 2000.

[42] C. C. Ragin, *Redesigning Social Inquiry: Fuzzy Sets and Beyond*, Univ. of Chicago Press, Chicago, IL, 2008.

[43] C. C. Ragin, "Comment: Lucas and Szatrowski in critical perspective," *Sociological Methodology*, Vol. 44, pp. 80-94, 2014.

[44] B. Rihoux and C. C. Ragin (eds.), *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques*, SAGE, Los Angeles, CA 2009.

[45] C. Q. Schneider and C. Wagemann, *Set-Theoretic Methods for the Social Sciences: a Guide to Qualitative Comparative Analysis*, Cambridge Univ. Press, 2012, Cambridge, UK.

[46] H. Tahayori and A. Sadeghian, "Median interval approach to model words with interval type-2 fuzzy sets," *International Journal of Advanced Intelligence Paradigms*, vol. 4, no. 3, pp. 313-336, 2012.

[47] G. de Tre and R. de Caluwe, "Level-2 fuzzy sets and their usefulness in object-oriented database modeling," *Fuzzy Sets and Systems*, vol. 140, pp. 29-49, 2003.

[48] I. Vlachos and G. Sergiadis, "Subsethood, entropy, and cardinality for interval-valued fuzzy stes-an algebraic derivation," *Fuzzy Sets and Systems*, vol. 158, pp. 1384-1396, 2007.

[49] L.-X. Wang, *A Course in Fuzzy Systems and Control*, Prentice-Hall, PTR, Upper Saddle River, NJ, 1997.

[50] R. W. Walpole, R. H. Myers, A.L. Myers and K. Ye, *Probability & Statistics for Engineers and Scientists*," 8th Edition, Prentice-Hall, Upper Saddleback River, NJ, 2007.

[51] D. Wu and J. M. Mendel, "Uncertainty measures for interval type-2 fuzzy sets," *Information Sciences*, vol. 177, pp. 5378-5393, 2007.

[52] D. Wu and J. M. Mendel, "Linguistic summarizations using IF-THEN rules and interval type-2 fuzzy sets," *IEEE Trans. on Fuzzy Systems*, vol. 19, pp. 136-151, February 2011.

[53] D. Wu, J. M. Mendel and S. Coupland, "Enhanced interval approach for encoding words into interval type-2 fuzzy sets and its convergence analysis," *IEEE Trans. on Fuzzy Systems*, vol. 20, pp. 499-513, June 2012.

[54] R. R. Yager, "A new approach to the summarization of data," *Information Sciences*, vol. 28, pp. 69–86, 1982.

[55] R. R. Yager, "On linguistic summaries of data," in Piatetsky-Shapiro, G. and Frawley, B. (Eds.): *Knowledge Discovery in*

Databases, MIT Press, Cambridge, USA, pp. 347-363, 1991.

[56] L. A. Zadeh, “Fuzzy sets,” *Information and Control*, vol. 8, no. 3, pp.338-353, 1965.

[57] L. A. Zadeh, “Fuzzy languages and their relation to human and machine intelligence,” in *Proc. Conf. Man and Computer*, 1970; also, Electron. Res. Lab., Univ. California, Berkeley, Memo. M-302, 1971.

[58] L. A. Zadeh, “Quantitative fuzzy semantics,” *Information Sciences*, vol. 3, pp. 159-176, 1971.

[59] L. A. Zadeh, “Outline of a new approach to the analysis of complex systems and decision processes,” *IEEE Trans. on Systems, Man, and Cybernetics*, vol. SMC-3, no. 1, pp. 28-44, 1973.

[60] L. A. Zadeh, “The concept of a linguistic variable and its application to approximate reasoning-2,” *Information Sciences*, vol. 8, pp. 301-357, 1975.

[61] S. Zadrożny, and Kacprzyk J., “Fuzzy querying using the 'query-by example' option in a Windows-based DBMS,” in *Proc. Third European Congress on Intelligent Techniques and Soft Computing - EUFIT'95,* Aachen, Germany, vol. 2, pp. 733-736, 1995.