Data Analytics: Challenges in Using the Fuzzy Sets Qualitative Comparative Analysis (FsQCA)

Dr. Dimitris K. Kardaras
Athens University of Economics and Business

Structure of presentation…

- Introduction to the FsQCA method
- Applications
- Example of Applying FsQCA
- The Challenges and Future Research

Introduction to FsQCA

Social scientist Charles Ragin introduced the foundations of Qualitative Comparative Analysis (QCA) in three major books (1987, 2000, and 2008)

Many cases/phenomena in data analysis exhibit causal complexity, i.e.

causal factors combine with each other to lead to the occurrence of an event or phenomenon.

different combinations of causal factors can lead to the occurrence of a given type of event or phenomenon.

causal factors can have opposing effects depending on the combinations with other factors.

Aims of FsQCA-I

- Assume an event or phenomenon (Y) and a number of factors (X_1, X_2, X_3,…, X_i)

- The FsQCA is particularly effective in investigating intertwined relationships between multiple factors that contribute to the realisation of certain outcome.

Aims of FsQCA-II

- FsQCA may detect multiple paths, i.e. alternative causal combinations that can lead to high levels of the same outcome.
- The FsQCA models allow a detailed analysis of how alternative conditions of causes combine and contribute to high membership scores of the outcome.
- Which cases/How many cases are represented by each causal combination?

Examples of FsQCA Applications

Hotel Management

- An FsQCA model that links the causal conditions (human capital, social capital, and contingency factors)
- to the outcome (high performance), measured by occupancy rate (commercial effectiveness), profit, and profit per employee (efficiency).

Environmental Issues

- An FsQCA model that identifies the Multi-causal pathways of public opposition to dam projects in Asia.
- e.g. lack of social safeguards in combination with rampant corruption and environmental risk induce opposition.

Business Model for IoT smart cities

- An FsQCA model that identifies the Multi-causal pathways of value propositions.

Identify alternative pathways to utilizing customer knowledge

- Identify pathways such as that a combination
- of high levels of Customer Relationship Management Technologies and Key Account Management teams, and etc., etc., lead to
- high degrees of **Customer Knowledge Utilization**.

Applying FsQCA

Terminology in FsQCA

1. The term “condition” is used, not “independent variable”
2. phenomenon to be explained is called “outcome,” not “dependent variable;”
3. The results of a FsQCA are called “solution formula” or “solution term,” not “equation.”

FsQCA steps I

1. **Identify** the factor that will represent the outcome set (Y).
2. **Identify the factors (Xi)** that will be used to form the causal combinations that produce the outcome set (Y).
3. **Select Data**.

FsQCA steps II

4. **Produce the truth table** of all possible permutations of the terms considered. Each permutation is a possible causal combination.

5. **Calculate membership degrees for each combination**. Its calculation is performed drawing on the fuzzy sets operations theory (Union, Intersection, Complement).

FsQCA steps III

6. **Calculate the consistency and the coverage** of the solutions.

7. **Identify best combinations**, by selecting the combinations that exhibit a consistently rate above a threshold (threshold often set at 0.8).

$$Consistency(X \prec Y) = \frac{\sum \min(X, Y)}{\sum X}$$

$$Coverage = \frac{\sum \min(X, Y)}{\sum Y}$$

Data Analysis-Input Sets

- Assume the following five (5) **Input Sets (**factors):

Quietness	Sea View	Staff Friendliness	Cultural Activities	Restaurant

Data Analysis-Output Set

- The **Outcome Set** is the *large amount of money spent* by each user during his/her hotel stay.

Data Analysis-Output set TFN

- The **Outcome Set** is modelled as a TFN with the following linguistic scales, that indicate the extent a customer is included to the set of those who spend large amount of money during their hotel stay.

Linguistic scale	Triangular fuzzy scale			Mean of fuzzy numbers
Very High	(0.75,	1.00,	1.00)	1.00
High	(0.50,	0.75,	1.00)	0.75
Medium	(0.25,	0.50,	0.75)	0.50
Low	(0.00,	0.25,	0.50)	0.25
Very Low	(0.00,	0.00,	0.25)	0.00

Data Analysis-Input Dataset

Large Amount Spent membership degree Outcome Set (Y)	Customer	Quietness	Sea View	Staff Friendliness	Cultural Activities	Restaurant
0.50	1	0.30	0.50	0.40	0.70	0.70
0.70	2	0.30	0.70	0.60	0.70	0.90
0.1	3	0.10	0.30	0.20	0.60	0.50
0.7	4	0.50	0.70	0.40	0.50	0.70
0.9	5	0.30	0.70	0.60	0.70	0.70

Data Analysis-Truth Table (FsQCA)

Developing the truth table is developed.

Since there are 5 terms to consider the number of permutations is

$$2^5 = 32$$

The cells in the truth table take the value (1) or (0) representing true or false.

Data Analysis-The Truth Table

Causal Combination	Quietness	Sea View	Staff Friendliness	Cultural Activities	Restaurant
1	0	0	0	0	0
2	0	0	0	0	1
3	0	0	0	1	0
4	0	0	0	1	1
5	0	0	1	0	0
6	0	0	1	0	1
7	0	0	1	1	0
8	0	0	1	1	1
9	0	1	0	0	0
10	0	1	0	0	1
11	0	1	0	1	0
12	0	1	0	1	1
13	0	1	1	0	0
14	0	1	1	0	1
15	0	1	1	1	0
16	0	1	1	1	1
17	1	0	0	0	0

Data Analysis- Calculate Membership degrees

Calculate membership degrees for each causal combination, using fuzzy operations (union, intersection, complement)

Data Analysis- Calculate Membership degrees II

For example, combination number 3 for customer-1:

μ=(*Quietness=false* ∩*Sea View=false* ∩ *Staff Friendliness=false* ∩ *Cultural Activities=true* ∩ *Restaurant=false)* = (not (*Quietness), not (Sea View), not (Staff Friendliness), Cultural Activities, not (Restaurant)).*

The μ(Quietness=false) = ((1- (Quietness)) = (1- 0.3)=0.7

Data Analysis

- The Membership degrees for combinations for each customer

Causal Combination	Customer 1	Customer 2	Customer 3	Customer 4	Customer 5
1	0.3	0.1	0.4	0.3	0.3
2	0.3	0.3	0.4	0.3	0.3
3	0.3	0.1	0.5	0.3	0.3
4	0.5	0.3	0.5	0.3	0.3
5	0.3	0.1	0.2	0.3	0.3
6	0.3	0.3	0.2	0.3	0.3
7	0.3	0.1	0.2	0.3	0.3
8	0.4	0.3	0.2	0.3	0.3
9	0.3	0.1	0.3	0.3	0.3
10	0.3	0.3	0.3	0.5	0.3
11	0.3	0.1	0.3	0.3	0.3
12	0.5	0.4	0.3	0.5	0.4
13	0.3	0.1	0.2	0.3	0.3
14	0.3	0.3	0.2	0.4	0.3
15	0.3	0.1	0.2	0.3	0.3
16	0.4	0.6	0.2	0.4	0.6
17	0.3	0.1	0.1	0.3	0.3

Data Analysis

- Calculate Consistency and Coverage for each Causal combination.

Causal Combination	Consistency	Coverage
1	0.785714286	0.379310345
2	0.8125	0.448275862
3	0.733333333	0.379310345
4	0.789473684	0.517241379
5	0.916666667	0.379310345
6	0.928571429	0.448275862
7	0.916666667	0.379310345
8	0.933333333	0.482758621
9	0.846153846	0.379310345
10	0.882352941	0.517241379
11	0.846153846	0.379310345
12	0.904761905	0.655172414
13	0.916666667	0.379310345
14	0.933333333	0.482758621
15	0.916666667	0.379310345
16	0.954545455	0.724137931
17	1	0.379310345

Data Analysis- Necessary and Sufficient Causal Combinations

- Assuming a threshold value of 0.8 for the consistency firstly,
- and then looking for the higher possible coverage,

- **the analysis results into** two causal combinations; the combinations number 12 and 16.

Data Analysis

Causal Combination	Quietness	Sea View	Staff Friendliness	Cultural Activities	Restaurant
16	0	1	1	1	1
12	0	1	0	1	1

Data Analysis Causal Combinations Final set

Customers who spend a large amount of money, show interest in

- (Sea View) AND (Staff friendliness) AND (Cultural activities) AND (Restaurant) OR
- (Sea View) AND (Cultural activities) AND (Restaurant).

Challenges-Future Research

FsQCA Should Be Applied together with Other Data Analysis Techniques

- Due to its focus on complex causal structures, FsQCA provides more precise information about the analytically relevant **similarities** and **differences** between cases, by clustering them into different causal paths towards an outcome.

FsQCA Should Be Applied together with Other Data Analysis Techniques

- The identified clusters of causal paths, can be a useful starting point for selecting cases for subsequent (comparative) case studies.

The Membership Degrees Should be chosen carefully

- In FsQCA values are used to indicates fuzzy sets (e.g. 0, 0.5, 1).
- There is need for set of arguments in order to determine which empirical evidence qualifies for set membership degrees above and/or below these pre-set values.

Number of Conditions: Truth tables may become very complex

- The Number of Conditions should be kept at a moderate level;
- Many conditions, FsQCA produces very complex results, making interpretations a daunting task.
- At the same time, using more conditions help to raise the consistency values.

Choosing thresholds for consistency and coverage

- Earlier on in the example we assumed a consistency level of 0.8;
- The appropriate levels for consistency and coverage are research-specific.
- They vary with the number of cases studied, the knowledge the researcher has about the cases, the quality of data gathered, the specificity of theories and hypotheses at hand.

Identify possible temporal relations among causal combinations

- In some cases there may be a temporal order in which conditions occurred.
- If temporal relationships exist then causal combinations are interconnected;
- Identifying the temporal order of conditions needs be addressed methodologically.

Conclusions…more research is needed…suggestions

- To specify some rules of good practice, broadly defined.
- To specify the level of importance-necessity and sufficiency of conditions in producing the outcome (Y).
- Define fuzzy sets on conditions.
- Devise strategies for combining FsQCA with othert data analysis techniques.

References (indicative list)

- Rihoux, B., Ragin, C.C. (2008). Configurational comparative methods: Qualitative comparative analysis (QCA) and related techniques. Sage Publications.
- Skarmeas, D., Leonidou, C.N., Saridakis, C. (2014). Examining the role of CSR skepticism using fuzzy-set qualitative comparative analysis. Journal of Business Research, 67, 1796–1805.
- Chari, S., Tarkiainen, A., Salojärvi, H. (2016). Alternative pathways to utilizing customer knowledge: A fuzzy-set qualitative comparative analysis. Journal of Business Research, 69, 5494–5499.
- Schneider, C., and Wagemann, C. (2010). Standards of Good Practice in Qualitative Comparative Analysis (QCA) and Fuzzy-Sets, Comparative Sociology 9, 1–22.

Thank you!!!