Italian Millennials and Subjective Well-Being: the Role of Deprivation on Life Satisfaction

L. Barbieri[1], E. Calegari[1], E. Fabrizi[1]

[1]Università Cattolica del Sacro Cuore, Piacenza
laura.barbieri@unicatt.it

SIS 2021
50th Meeting of Italian Statistical Society
21th-25th June, 2021

Motivation and Contribution

Motivation

- Qualitative Comparative Analysis (QCA) is increasingly applied in social sciences to bridge the gap between qualitative (case oriented) and quantitative (variable oriented) empirical approaches.
- QCA is widely applied in the case of small samples, but it can operate with any number of observations both as an alternative and a complement analysis to conventional regression methods.
- At the best of our knowlege, the QCA has never been applied to study the Subjective Well-Being (SWB). Nevertheless, being case oriented, QCA could allow the researcher to highlight which specific levels of individual characteristics lead to the analysed outcome (presence of SWB).

Contribution

The present work (still in progress) aims to apply QCA in a large sample to highlight how different paths of deprivation could lead to SWB. Morover, we compare the analysis with other statistical approaches usually applied as classifiers (i.e. Bayesian Network and Logistic Regression).

Outline

Data and Variables

We study the determinants of SWB focusing on the relationship between SWB and three deprivation dimensions targeting Italian millennials.

- <u>Data source</u>: EU-SILC data Italian wave 2018 (ad hoc module);
- <u>Units of observation</u>: 3982 Italian Millennials (born between 1981 and 1996) splitted in two subsamples according to the median household disposable income (1637 obs. above and 2345 below);
- <u>Binary outcome variable</u>: *SWB* (0 = presence of SWB, 1=absence of SWB);
- <u>Included variables</u>:
    - *Personal deprivation*, *Material deprivation*, *Financial deprivation* (dicothomous variables obtained with exploratory factor analysis from 13 deprivation items);
    - Dichotomous variable for *geographical location* (1=North, 0=South);
    - *Labour attainment status* (0=Inactive,1=Employed/Student,2=Unemployed)

The Qualitative Comparative Analysis (1)

QCA definition

Analysis which uses combinatorial logic, fuzzy set theory and Boolean minimisation to work out what combinations of case characteristics may be necessary or sufficient to produce an outcome (Kent, 2008).

According to the variables codification, we can distinguish three QCA frameworks: *crisp* (both binary outcome and conditions), *multi-value* (binary outcome and at least one not binary condition), *fuzzy* (either a not binary outcome or at least one not binary condition).

QCA Characteristics

- It admits **equifinality**: different configurations of variables may give rise to the same level of SWB.
- It admits **causal asymmetry**: the occurrence and nonoccurrence of the SWB require separate explanations.
- It has not **omitted variable bias**: in regression analysis, omitting a relevant variable correlated to the included variables leads to an estimation bias. The QCA, being based on Boolean algebra, rather than on correlation, does not suffer from that bias.

Note: in the first attempt here presented, we apply a multi-value QCA.

The Qualitative Comparative Analysis (2)

QCA instruments in multi-value framework

- **Truth table**: list of all the possible combinations of causal conditions (variables' levels) together with the associated value of SWB and their corresponding frequencies in the sample.
- **Analysis of necessity**: identification of combinations that are necessary for the outcome to occur ⤳ A combination of characteristics (variables' levels) that is present in all instances where SWB is present.
- **Analysis of sufficiency**: identification of combinations that are sufficient for the outcome to occur ⤳ A combination of characteristics that is never present in the absence of SWB.
- **Minimization**: set of configurations that meet selected criteria for causal sufficiency, with varying degrees of parsimony.

QCA measures

- **Consistency**: "indicates how closely a perfect subset relation is approximated" (Ragin, 2008, p.44).
- **Coverage**: for each combination of characteristics, it measures the proportion of cases exhibiting the SWB where the combination is observed.

Preliminary QCA Results: Solution Formulas

Low Income Sample

M_L: pers[0]*mat[0]*fin[0]*job[0] +
pers[0]*mat[0]*fin[0]*area[0]+
pers[0]*mat[0]*job[1]*area[1]   → swb[1]

	Cons.	Cov.
pers[0]*mat[0]*fin[0]*job[0]	0.95	0.02
pers[0]*mat[0]*fin[0]*area[0]	0.90	0.15
pers[0]*mat[0]*job[1]*area[1]	0.89	0.21
M_L	0.90	0.37

High Income Sample

M_H: pers[0]*mat[0]*fin[0]*job[0] +
pers[0]*mat[0]*fin[0]*job[2]+
pers[0]*mat[0]*fin[0]*area[1]+
pers[1]*mat[0]*fin[1]*job[1]*area[1]→ swb[1]

	Cons.	Cov.
pers[0]*mat[0]*fin[0]*job[0]	1.00	0.02
pers[0]*mat[0]*fin[0]*job[2]	0.92	0.05
pers[0]*mat[0]*fin[0]*area[1]	0.93	0.48
pers[1]*mat[0]*fin[1]*job[1]*area[1]	1.00	0.01
M_H	0.93	0.51

- Symbol * refers to a "logical conjunction" (logical AND operator); symbol + refers to a "logical disjunction" (logical OR operator).
- Solution Formula M_L means that each of its three conjunctions gives a combination of conditions that is sufficient for the presence of SWB (**equifinality**).
- For example, the conjunction pers[0]*mat[0]*fin[0]*job[0], that appears in both M_L and M_H, means that the simultaneous absence of personal, material and financial deprivation for an inactive Millennial implies SWB in both subsample:
    - this condition has a lower **consistency** (Cons.) in the low income sample, suggesting that in this case the configuration produces the outcome with a minor extent;
    - the **coverage** (Cov.) is very low in both cases, reflecting that only a low proportion of cases exhibiting SWB presents the configuration.

QCA vs. Other Models for Binary Outcomes

Logit model

- One of the widely adopted regression models for binary outcomes;
- To model dependencies among variables, interactions are allowed;
- It highlights the magnitude of the effect of each variable on the outcome;
- It can be used for predictive purposes.

Bayesian Network classifier

- Graphical model for binary or categorical outcomes;
- It highlights dependencies among variables;
- It is possible to compute measures of conditional dependence between the outcome and a given variable;
- Widely used for predictive purposes.

QCA

- Useful method to "profile" individuals;
- It takes into account all possible combinations of variables' levels even if results are difficult to interpret when the number of condition increases;
- The relevance of each variable level in term of sufficiency and necessity for an outcome can be evaluated;
- Not focused on "prediction".

Performance Comparison: in Sample Accuracy (High Income Sample)

Logit Model

		Classified	
		0	1
Actual	0	138	58
	1	724	717

Specificity (cutoff=0.9):0.70
Sensitivity(cutoff=0.9):0.49
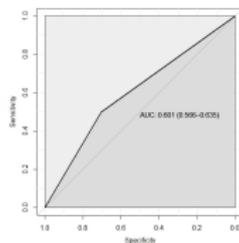
BN Classifier*

		Classified	
		0	1
Actual	0	115	81
	1	503	938

Specificity (cutoff=0.9):0.59
Sensitivity(cutoff=0.9):0.65
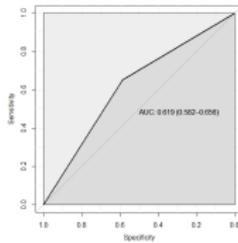
QCA**

		Classified	
		0	1
Actual	0	196	0
	1	1408	33

Specificity (cutoff=0.9):1
Sensitivity(cutoff=0.9):0.02
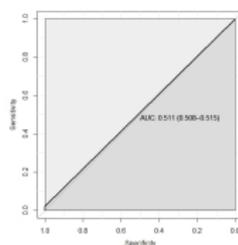
* The BN classifier used here is a Tree Augmented Naive Bayes.
** The analysis refers to the conjunction pers[0]*mat[0]*fin[0]*job[0] of the solution formula M_H, for illustrative purposes only. In the QCA case, the "classified" adjective is intended to distinguish between observations that satisfy (=1) or do not satisfy (=0) the given condition.

Note that the rationale behind the QCA is to look for the best fit only in terms of sufficiency and not the best overall fit, as it happens for the other two models instead. In this perspective, the sufficiency condition is satisfied when a zero appears off the diagonal (upper-right).

Further Developments

In future works, we would like to

- study in depth the model comparison also considering other approaches;
- consider the SWB as an ordinal categorical variable (as it is originally detected) in a fuzzy set QCA framework and compare the results with those of an ordered logistic regression;
- obtain a more precise profile of satisfied people (individuals with a high degree of SWB) in terms of deprivation items.

Some references

- Ragin, C.C. (2008), "Measurement versus calibration: A set-theoretic approach", In *The Oxford handbook of political methodology*.
- Kent, R. (2008), "Using fsQCA - A Brief Guide and Workshop for Fuzzy-Set Qualitative Comparative Analysis", *context, 45*, 489–503.