COMPASSS Working Paper 2013 - 72

Fuzzy-Set Coincidence Analysis: the Hidden Asymmetries

Camilla Borgna

Department of Social and Political Studies - University of Milan, Italy

camilla.borgna@unimi.it

Abstract

Despite being conceptually and algebraically straightforward, set coincidence has never received great attention within the framework of set-relational research, which has essentially focused on subset/superset relations. In this article, I advance a novel procedure called fuzzy-set coincidence analysis to systematically assess the degree of overlapping of two or multiple sets. Individual-level analyses, and those in social stratification in particular, are deemed to be a fertile field to apply fuzzy-set coincidence analysis. A main comparative advantage with respect to statistical techniques lies in its ability to uncover hidden asymmetrical patterns. This will be shown with an empirical application that explores patterns of overlapping inequalities for students of different migratory status.

Introduction

In recent years, set-relational research has experienced sophisticated methodological advancements and a sharp increase in empirical applications. However, a noteworthy domain of set relations has been left unexplored, both from the theoretical and the empirical point of view: that of set coincidence or overlapping. To fill this gap, I advance a novel procedure within the framework of set-theoretic methods: the analysis of set coincidence[1], and fuzzy-set coincidence in particular. The focus of fuzzy-set coincidence analysis is the joint subsistence of attributes for given cases and therefore its underlying logic is fundamentally different from that of statistical techniques. Moreover, a major comparative advantage with respect to correlation lies in its capability to uncover hidden asymmetrical patterns, as argued throughout the paper and demonstrated with an empirical application.

The article is structured as follows: first of all, I outline the basics of set-relational research in the social sciences and draw attention on two peculiarities of fuzzy-set coincidence analysis with respect to this tradition; secondly, I describe how coincidence scores can be computed and interpreted in a crisp-set and fuzzy-set framework; I then clarify the differences between fuzzy-set coincidence analysis on the one hand, and correlational analysis and data reduction techniques on the other hand; next, I discuss the usefulness of fuzzy-set coincidence analysis in stratification research; finally, I provide an empirical application to assess overlapping inequalities for second generation migrants in Europe.

1. Set-relational research and set coincidence

The application of fuzzy-sets to social sciences developed from crisp-set Qualitative Comparative Analysis (csQCA) which was intended for Boolean sets, where cases could either assume value 1 (if they belonged to the given set) or 0 (if they belonged to its negation) (Ragin 1987). However, to

[1] To be sure, the analysis of fuzzy-set coincidence that I propose here is not connected to the Coincidence Analysis (CNA) technique proposed by Baumgartner (2013). The latter, indeed, is a technique for the minimization of truth tables alternative to the Quine-McCluskey algorithm.

account for the fact that most phenomena in social sciences exist in degrees, the framework of QCA was later extended to fuzzy logic, under which the degree of membership to a given set may vary, assuming any value from 0 to 1 (Ragin 2000; Ragin 2008). Fuzzy sets do not equate with continuous variables, because their construction relies on a procedure called "calibration" which anchors set membership scores to qualitative thresholds and truncates irrelevant variation. When calibrating raw data, researchers rely on external substantive or theoretical criteria or on meaningful breaks in the internal distribution to determine qualitative thresholds for full membership (1), full non-membership (0) and maximum ambiguity (0.5). Intermediate values are typically assigned using a linear or log-linear function. As Ragin puts it: "Fuzzy sets are simultaneously qualitative and quantitative. They address the varying degree to which different cases belong to a set (including the two qualitative states, full membership and full non membership), not how they differ from one another along quantifiable dimensions of open-ended variation" (Ragin 2000, p.154).

Despite being conceptually and algebraically straightforward, the analysis of set coincidence has never been systematized within the framework of set-relational research. There may be two reasons behind the underdevelopment of fuzzy-set coincidence analysis: first, it is best conceived for applications at the meso- or micro-, rather than at the macro-level; second, it is not based on subset/superset relations, but rather on set overlapping. As for the first point, fuzzy-set coincidence analysis can be applied to any level of analysis, but – as will be made clear in section 4 – it is particularly suited for the study of social stratification at the micro level. For example, the empirical application presented in this article is on overlapping individual inequalities. On the contrary, QCA first developed as a "macro-comparative" approach, around "small" or "mid-sized N" research designs (between 10 and 35 cases) and with countries as the typical unit of analysis. Moreover, analyses at the micro-level are problematic for QCA inductive nature, which asks to "go back to the cases" once the analyses have been performed. However, unless data is collected by the researcher within a longitudinal framework, it is only possible to get back to information previously collected. Notwithstanding these limitations, a considerable number of QCA applications is performed at the micro-level, as documented by Rihoux et al. (2013), who also show an increase in their usage since 2010. Vis (2012) has convincingly made the point for the use of QCA in "moderately large N" research designs (between 50 and 100 cases) by assessing its comparative advantages with respect to regression analysis. In this paper, I take a similarly pragmatic approach to argue that, when N is not small, fuzzy-set coincidence analysis shall not be seen as a "second-best" to measures of statistical association, but rather as a useful complement. Indeed, as will result clear from the empirical application, it unveils asymmetrical patterns that are invisible to correlational analysis. Fuzzy-set coincidence analysis is a peculiar technique under another respect: the focus is on the extent to which two or multiple sets may coincide, to the point of being one and the same set. On the contrary, social scientists making use of fuzzy sets are typically concerned with an attribute of a given set of cases being a subset of a broader set. As a matter of fact, subset/superset relations are a natural focus of attention for them because – unlike statistical analysis – they directly target asymmetries. Set-coincidence is a less straightforward focus because it relies on symmetry of attributes and, as such, it might seem very similar to correlation. However, as I argue more explicitly in section 3, set coincidence is fundamentally different from correlation, first of all because it is based on a different logic, and second of all because of its hidden asymmetrical nature.

2. How to assess set coincidence

Set-coincidence[2] can be defined as the degree to which two or more sets overlap, or, in other words, the extent to which they constitute one and the same set. In a crisp-set framework, coincidence scores can be calculated as:

$$\frac{X \cap Y}{X \cup Y} \tag{1}$$

i.e., the number of cases in the intersection of two or more sets over the number of cases in the union of the same set. Given that in fuzzy logic the union of two sets corresponds to taking the

[2] I adopt here the terminology adopted by Ragin (2008), while Smithson and Verkuilen (2006) prefer to talk of "commorbidity" and co-occurrence".

maximum membership score and their intersection corresponds to taking their minimum, fuzzy-set coincidence score can be calculated as:

$$\frac{\sum[\min(x_i, y_i)]}{\sum[\max(x_i, y_i)]} \tag{2}$$

where x_i and y_i are individual membership scores in the sets X and Y (Ragin 2008, 59). Both indexes vary from 0 (absence of set coincidence) to 1 (perfect set coincidence). Readers who are familiar with QCA may have already noticed that expressions (1) and (2) combine formulas for consistency and coverage for sufficiency[3]. Indeed, we can think of set coincidence as a symmetrical relation between two sets, where the first one is contemporarily a subset *and* a superset of the second one[4]. Figure 1 graphically displays this idea for fuzzy sets.

Yet another way to conceptualize set coincidence is referring to relations of necessity and sufficiency between causal conditions and the outcome. While in the upper-left graph B is a sufficient condition for the occurrence of A, and in the lower-left graph it is a necessary condition, in the upper-right graph B is a simultaneously sufficient and necessary condition for the occurrence of A. In other words, A is implied by B *and* A implies B.

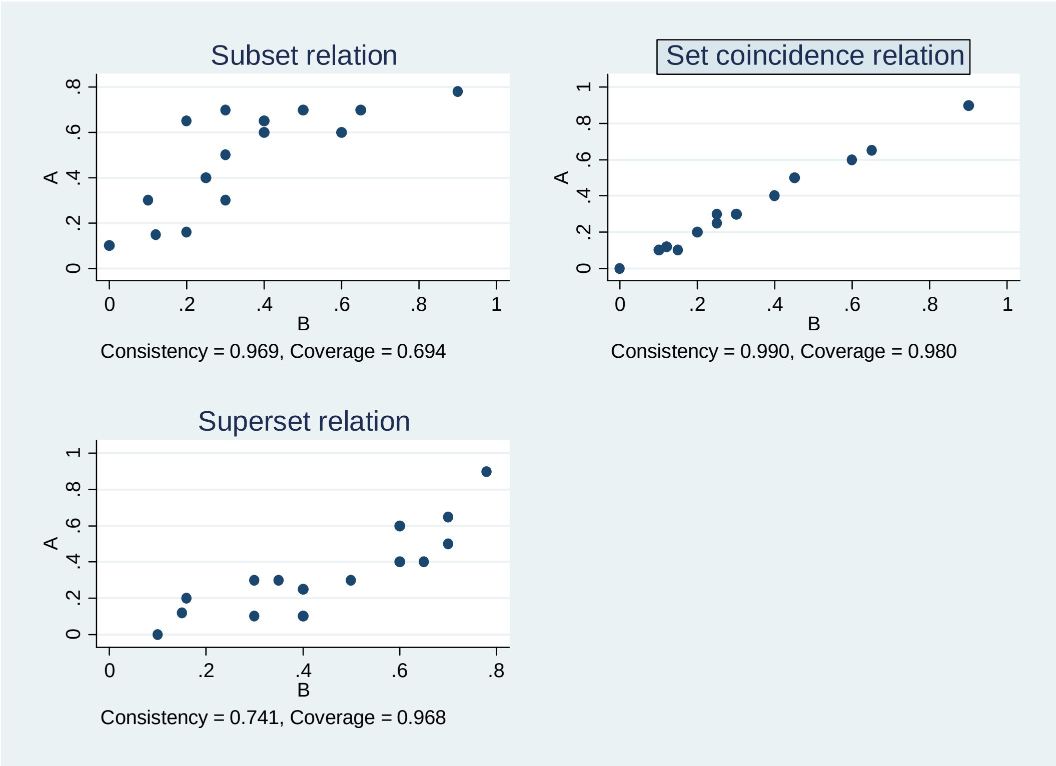
Among the software programs commonly used in set-relational research, version 2.5 of fsQCA features a specific procedure called "Set coincidence", while QCA3 package in R features the command *coincid* whose output is a coincidence matrix; the same output is produced by the option *coincid* of *fuzzy* command in STATA. All these procedures are available for crisp and fuzzy sets. However, only fsQCA makes it possible to assess multiple coincidence, i.e. the degree to which more than two sets overlap. Until software updates make it feasible to compute those coincidence scores directly, it is still possible for R and STATA users to perform coincidence

[3] Consistency for sufficiency assesses the degree to which cases sharing a given condition *X* agree in showing the outcome *Y* and it is computed as $\frac{X \cap Y}{X}$ in a crisp-set framework and as $\frac{\sum[MIN(x_i, y_i)]}{\sum(x_i)}$ in a fuzzy-set framework. Coverage for sufficiency assesses the degree to which instances of the outcome *Y* agree in displaying the causal condition *X* and it is computed as $\frac{X \cap Y}{Y}$ in a crisp-set framework and as $\frac{\sum[MIN(x_i, y_i)]}{\sum(y_i)}$ in a fuzzy-set framework.

[4] Since coincidence reflects the joint subsistence of consistency and coverage, when a researcher is confronted with a low coincidence score, he should always verify whether this situation hides a low consistency, a low coverage or both.

analysis more "manually". Indeed, one can compute the intersection of the sets of interest (logical AND) and divide it by the union of the same sets (logical OR): this, as shown in formulas 1 and 2, returns set coincidence scores.

Figure 1 Graphical meaning of fuzzy-set coincidence: hypothetical set relations



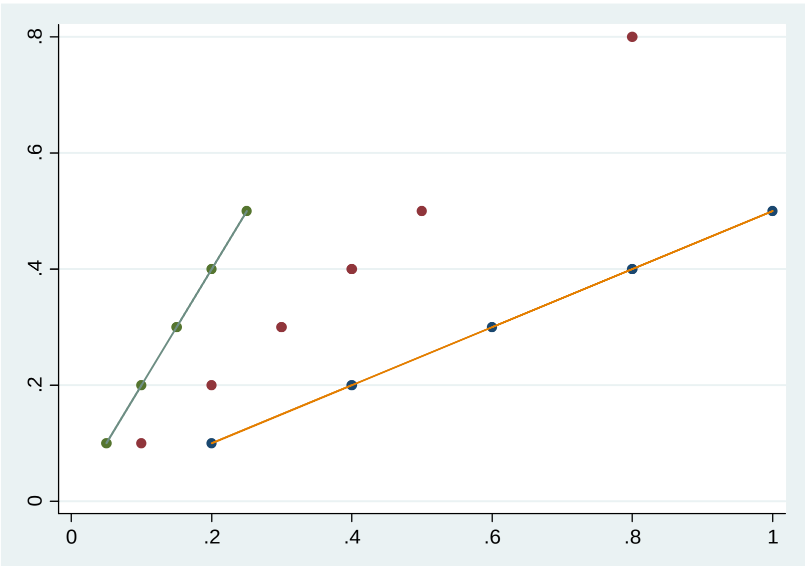
3. Differences from mainstream statistical techniques

The symmetrical nature of set coincidence might induce more than one reader to wonder if there is any difference with correlation, and why we should even bother to define it and apply it. However, set coincidence differs from correlation in at least two respects. The first one pertains to its underlying logic, while the second one derives from more pragmatic considerations regarding the "hidden" asymmetrical nature of set coincidence.

To start with, a correlation coefficient is an estimate of a common probabilistic trend of two variables among our observations. It concerns the functional form of the relation between two variables, be it linear in the case of strictly defined correlation coefficients – such as Pearson's R –

or more complex in the case of other measures of association – such as odds-ratios. Instead, a set coincidence score close to 1 indicates that most of our cases share exactly the same degree of membership in two sets. It is about the joint subsistence of attributes in given cases. In Ragin's words: (fuzzy-set coincidence) "is a special case of correlation. In a plot of two fuzzy sets, any straight line that is neither vertical nor horizontal yields a perfect correlation coefficient. However, perfect set coincidence occurs only when all the cases plot exactly on the main diagonal of the fuzzy plot" (*ibidem*). As shown in Figure 2, the brown dots represent cases with equal membership on the two attributes, therefore a situation of perfect fuzzy-set coincidence (*and* perfect correlation). On the contrary, the green and red lines represent a positive linear relation between the two variables: whenever *x* increases by one point, *y* increases by two points in the case of the green line, and by half a point in the case of the red line. Both lines correspond to a correlation coefficient of 1, i.e., there is no error. However, a perfect correlation does not necessarily imply a perfect set coincidence, which is therefore worth to investigate separately.

Figure 2. Fuzzy-set coincidence *vs.* correlation



A second aspect which differentiates coincidence from correlation is that, despite appearances, fuzzy-set coincidence is not completely symmetrical in nature. Indeed, as I have mentioned, any fuzzy-set analysis relies on calibration, which anchors set membership scores to qualitative thresholds and truncates irrelevant variation. Calibration can – and in many cases should – be asymmetrical (or "dual") rather than symmetrical (or "unipolar"). For instance, if a researcher moves from an interval scale variable to define two different sets, referring to the two extremes of this source variable, when imputing membership scores he may decide to leave space for a "grey" area, where cases do not hold membership in one set, but neither do they in the other set. Put differently, one set is not the negation of the other. Note that, even if calibration usually refers to fuzzy-set analyses, asymmetrical set memberships can apply to crisp sets as well. It is easy to grasp "hidden" asymmetrical nature of set coincidence referring to the empirical application object of this paper: I assess the degree to which individuals cumulate socio-economic advantages and disadvantages. If I find an indicator of economic wealth to be strongly correlated with one of cultural status, I can conclude that, in the given sample, the two simultaneously increase or decrease. However, I do not know if this is true because some individuals are able to accrue their position by cumulating richness with a high cultural status, or rather if some other individuals are trapped in the lowest strata by their poorness and their low cultural status. On the contrary, coincidence analysis is able to uncover these different patterns, because it requires the researcher to carefully calibrate the two indexes into precise sets defining e.g. "richness" and "poorness". Since many individuals are neither "rich" nor "poor", it is reasonable to perform an asymmetrical calibration: in my case, these individuals will have a fuzzy membership score lower than 0.5 in both sets.

Another analogy that might come to mind is that between fuzzy-set coincidence analysis and data reduction techniques, such as Principal Component Analysis (PCA) and Factor Analysis (FA). Indeed, like these procedures, fuzzy-set coincidence analysis produces a synthetic indicator that refers to several dimensions, which are found to be strictly connected, to the point of representing a single dimension. In particular, in confirmatory factor analysis, researchers move from their theoretical knowledge to define *a priori* expectations about which latent factor is to be associated with a specified subset of observed variables. Consequently, they test their hypotheses by modeling such variables as linear combinations of the potential factors. Fuzzy-set coincidence analysis requires researchers to use theoretical and substantive information in order to choose which sets are expected to overlap in the population of cases of interest. Then, they test their hypothesis by computing the degree to which these sets coincide to the point of being one and the same set.

However, these apparent similarities should not misguide the reader to the point of thinking that fuzzy-set coincidence analysis is just another data reduction technique. Its goal is not the search for a latent structure underlying the data, which would allow the researcher to reduce dimensions of variation. Rather, the analysis of coincidence is case-oriented. Just like other set-theoretic procedures, its goal is to systematically assess how cases of interest are positioned with respect to what Lazarsfeld (1937) called "property space". More specifically, fuzzy-set coincidence analysis verifies the joint subsistence of attributes in given cases. As a consequence, findings of coincidence analysis are always context-dependant and there is no broad generalization purpose. Moreover, its deterministic framework is another sharp difference *vis-à-vis* factor analysis' inferential approach. Therefore, one can easily conclude that coincidence scores provide qualitative different information with respect to "loadings" issued by factor analysis.

It should be noticed that, given its focus on cases, fuzzy-set coincidence analysis is also different from set-theoretic procedures aimed at reducing dimensionality, and in particular, the use of "macro-variables" proposed in chapter 11 of Ragin (2000). Ragin puts forward this "macro-variables approach" as a strategy to streamline as much as possible the explicatory model and therefore to reduce limited diversity in the truth table. To do so, a researcher can look for conditions that he expects to be substitutable and connect them with a logical "OR". The newly built condition takes the name of "macro-variable". Clearly, coincidence analysis can be used as a preliminary step for the construction of a truth table: if two causal conditions are found to coincide, then it is superfluous to include them both in the model. However, the heuristic value of coincidence analysis is broader. As I argued throughout this article, one could be interested in coincidence scores *per se*, just like she is interested in subset/superset relations. Moreover, the analysis of set coincidence is case-oriented and can be seen as starting point toward more qualitative and in-depth research. For instance, one can compute the degree to which the sets coincide for different groups of cases and use the information in order to further investigate similarities and differences among such cases.

4. Social stratification and fuzzy-set coincidence analysis

The empirical application presented in Section 5 concerns overlapping inequalities for second generation migrants in Europe. Indeed, I argue that social stratification is a particularly fertile field for the analysis of set coincidence. Vertical stratification of individuals typically occurs along a number of different dimensions, as profusely recognized by sociologists since Max Weber. When these dimensions coincide to a great extent, social structures are said to be “crystallized” (Freedman et al. 1951). From a set-theoretical perspective, one can think of individuals as simultaneously belonging to multiple sets, according to the dimension considered: social class (élite, bourgeoisie, proletariat…), cultural or social status (high, low…), education level (primary, secondary, tertiary…). If, in a given society, factors of advantage or factors of disadvantage are polarized to the point of coinciding, then only one dimension of vertical differentiation exist.

In order to assess the crystallization of social structures, sociologists have relied on correlation measures (Lenski 1954; Landecker 1981). However, fuzzy-set coincidence analysis can be a more appropriate technique to assess the degree of crystallization. Indeed, as I have argued in the previous section, set coincidence is fundamentally different from correlation, first of all because its focus is on the joint subsistence of attributes rather than on a common probabilistic trend. This is particular important when studying stratification systems, which – as Landecker himself maintains – are “constellations of rank systems" (Landecker 1981, 40), that is, constellations of "qualitatively different" status hierarchies” (*ivi*, 19). Secondly, the usefulness of fuzzy-set coincidence analysis in this field derives from its capability to account for asymmetrical patterns. Indeed, a social structure can be crystallized either because some individuals are able to accrue their position by cumulating several assets, or because some other individuals are trapped in the lowest strata by multiple deprivations, or, still, because both these processes are in place. Correlational analysis is not able to disentangle these processes and could fail to detect patterns of crystallization if only one of the two is in place. On the contrary, fuzzy-set coincidence analysis requires careful calibration of set membership for factors of advantage and factors of disadvantage separately, and the two do not need to be symmetrical: in this way, crystallization processes "from the top" and "from the bottom" can be distinguished.

5. An empirical application

By using data provided by the 2009 wave of the *Programme for International Student Assessment* (PISA), I assess overlapping inequalities for 15-year-old students according to their migratory status in 14 West European countries. In particular, aim of this exercise is to explore patterns of fuzzy-set coincidence in ascriptive characteristics that may influence educational achievement. Four dimensions of inequality are considered: occupational status, institutionalized cultural capital, objectified cultural capital and wealth. For each of these dimensions, I calibrate two fuzzy sets: one as a factor of advantage (e.g. "high occupational status") and one as a factor of disadvantage (e.g. "low occupational status")[5]. Dual unipolar calibrations are performed, so that one set is not the negation of the other. As a consequence, some individuals may fail to display the factor of advantage *and* the factor of disadvantage and they will have membership score lower than 0.5 in both sets.

I calculate fuzzy-set coincidence scores separately by country and by migratory status (natives and second generation migrants) assessing the degree to which factors of advantage and disadvantage coincide. As displayed in Table 1, fuzzy-set coincidence scores are definitely closer to zero than to one, supporting the argument for a multidimensional assessment of ascriptive factors that may have an effect on educational achievement. However, Table 1 also shows that factors of advantage systematically coincide more for natives than for second generation migrants, while the opposite is

[5] For information on the source variables and the calibration thresholds chosen, refer to Appendix 1.

true for factors of disadvantage. In other words, it is more likely for households which are privileged under one aspect to enjoy other assets if parents are native-born than if parents are foreign-born. On the contrary, unprivileged households are more likely to suffer from a multiplicity of deprivations if parents are foreign-born than if they are native-born. From a substantive point of view, this finding is interesting because compounded disadvantages may point to a situation of social exclusion where only highly resilient students can escape educational failure. On the contrary, empirical literature has shown that simultaneously enjoying several assets is not decisive to be among top-performers. Therefore, if a researcher is interested in the effects of ascriptive characteristics on educational achievement, her focus would be on the native/migrant discrepancy between coincidence scores for disadvantages rather than on that for advantages.

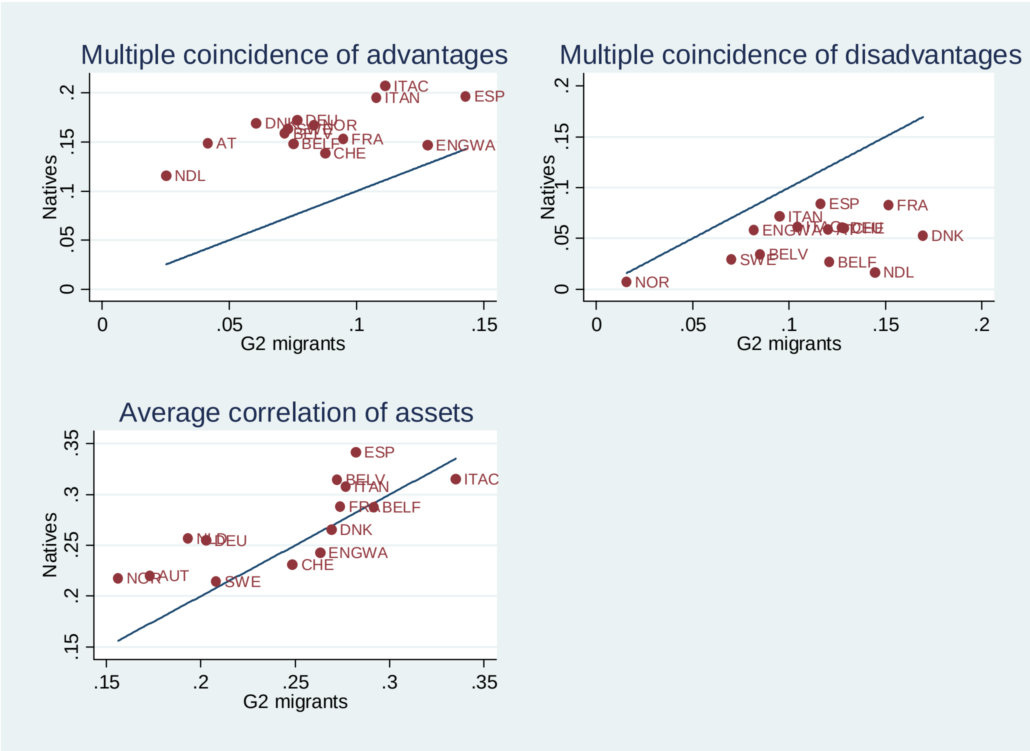
Table 1. Advantages and disadvantages coincidence scores, by country and migratory status

	Advantages			Disadvantages		
	Nat	**G2**	**Nat - G2**	**Nat**	**G2**	**Nat - G2**
Austria	0.223	0.084	0.139	0.146	0.269	-0.123
Belgium Fla	0.207	0.103	0.104	0.119	0.239	-0.120
Belgium Wal	0.237	0.121	0.116	0.128	0.174	-0.046
Switzerland	0.212	0.124	0.088	0.140	0.249	-0.109
Germany	0.259	0.116	0.143	0.142	0.259	-0.117
Denmark	0.268	0.078	0.190	0.138	0.306	-0.168
England Wal	0.213	0.198	0.015	0.151	0.179	-0.028
Spain	0.271	0.185	0.086	0.134	0.197	-0.063
France	0.243	0.149	0.094	0.166	0.258	-0.092
Italy Center	0.273	0.123	0.149	0.131	0.192	-0.061
Italy North	0.275	0.180	0.095	0.148	0.199	-0.051
Netherlands	0.229	0.071	0.158	0.118	0.266	-0.148
Norway	0.242	0.124	0.117	0.090	0.120	-0.029
Sweden	0.238	0.117	0.122	0.101	0.164	-0.063

Multiple fuzzy-set coincidence scores of HISTATUS, HIEDU, HICULT, RICH (Advantages) and LOSTATUS, LOEDU, LOCULT, POOR (Disadvantages).

It is worthwhile to point out that we could not have unveiled these patterns by performing mainstream correlational analysis, as can be clearly seen in Figure 3[6]. The average correlation of uncalibrated assets is not systematically higher for either of the two student categories defined by migratory status. In effect, it is exactly the asymmetry of coincidence patterns for factors of advantage and disadvantage that make it invisible to a symmetrical technique like correlation: the higher coincidence of disadvantages for second generation migrants compensates the higher coincidence of advantages for natives. What might seem a "zero pattern" from a correlational point of view is revealed to be a double pattern of inequality by fuzzy-set coincidence analysis.

Figure 3. Hidden asymmetrical patterns: set coincidence *vs.* correlation



---

[6] Average correlations of assets are computed as the simple average of the components of the correlation matrix between HISEI, HISCED, CULTPOSS and WEALTH. Fuzzy-set coincidence scores are instead computed as *multiple* coincidence scores in the first place, as shown in Table 1. However, paired coincidence scores display the same pattern, as shown in Appendix 2.

Conclusion

In this article, I proposed a novel set-theoretic procedure to assess the degree of overlapping of two or multiple sets. After clarifying its peculiarities within the framework of set-relational research, I have presented more in detail its logic and its operationalization in a crisp- and fuzzy-set framework. I have argued in favor of its specificity with respect to correlational analysis and data reduction techniques, both statistical and set-theoretic. Conceptual and pragmatic gains resulting from the use of fuzzy-set coincidence analysis have been considered, and a point has been made in favor of its use in social stratification research. Fuzzy-set coincidence analysis appears to be a more appropriate technique than correlation analysis to assess the degree of crystallization of social structures. Through an empirical application, I have practically demonstrated how fuzzy-set coincidence analysis is able to uncover hidden asymmetrical patterns that would have been invisible to correlation analysis.

References

Baumgartner, Michael. 2013. “Detecting Causal Chains in Small-N Data.” *Field Methods* 25 (1): 3–24.

Freedman, Ronald, Amos Henry Hawley, Werner S. Landecker, Horace Mitchell Miner, and Guy E. Swanson. 1951. *Principles of Sociology: a Text and a Reader*. Overbeck.

Landecker, Werner S. 1981. *Class Crystallization*. Rutgers Univ Pr.

Lazarsfeld, Paul. 1937. “Some Remarks on the Typological Procedures in Social Research.” *Zeitschrift Für Sozialforschung* 6.

Lenski, Gerhard E. 1954. “Status Crystallization: A Non-Vertical Dimension of Social Status.” *American Sociological Review* 19 (4) (August): 405. doi:10.2307/2087459.

Ragin, Charles C. 1987. *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*. University of California Press.

———. 2000. *Fuzzy-Set Social Science*. University of Chicago Press.

———. 2008. *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. University of Chicago Press.

Rihoux, Benoit, Priscilla Alamos-Concha, Damien Bol, Axel Marx, and Ilona Rezsöhazy. 2013. “From Niche to Mainstream Method? A Comprehensive Mapping of QCA Applications in Journal Articles from 1984 to 2011.” *Political Research Quaterly* 66 (1): 9–18.

Smithson, Michael, and Jay Verkuilen. 2006. *Fuzzy Set Theory: Applications in the Social Sciences*. SAGE Publications. Vol. 147. Qualitative Applications in the Social Sciences.

Vis, Barbara. 2012. “The Comparative Advantages of fsQCA and Regression Analysis for Moderately Large-N Analyses.” *Sociological Methods & Research* 41 (1) (February 1): 168–198.

Appendix 1. Thresholds for the calibration of fuzzy sets

Source variable	Fuzzy set	Qualitative thresholds	Criteria
HISEI *Highest occupational level among parents according to Ganzeboom socio-economic index*	High occupational status: HISTATUS	Country-specific	Internal distribution (80,60,40 percentiles)
	Low occupational status: LOSTATUS	//	Internal distribution (20,40,60 percentiles)
HISCED *Highest educational level among parents according to ISCED 1997scale*	Highly educated parents: HIEDU	1 if HISCED >=5 0.5 if HISCED = 4 0 if HISCED <=3	Substantive knowledge: 0-2 =None, primary, lower-secondary 3=Vocational secondary 4-5=Generalist secondary, Post-secondary non tertiary 6=Tertiary
	Low educated parents: LOEDU	1 if HISCED <=2 0.5 if =3 0 >=4	//
CULTPOS *Home possessions of classical literature, books of poetry and pieces of art*	High cultural capital: HICULT	Country specific	Internal distribution (first 3 peaks out of 4)
	Low cultural capital: LOCULT	//	Internal distribution (last 3 peaks out of 4)
WEALTH *Home possessions of comfort items*	Rich household: RICH		Internal distribution (80,60,40 percentiles)
	Poor household: POOR		Internal distribution (20,40,60 percentiles)

Appendix 2. Average paired advantage and disadvantage coincidence scores, by country

	Advantages			Disadvantages		
	Natives	**G2**	**Natives - G2**	**Natives**	**G2**	**Natives - G2**
Austria	0.41	0.25	0.16	0.26	0.37	-0.11
Belgium Fla	0.40	0.29	0.11	0.22	0.39	-0.17
Belgium Wal	0.30	0.20	0.10	0.22	0.28	-0.06
Switzerland	0.39	0.31	0.09	0.27	0.39	-0.12
Germany	0.43	0.27	0.16	0.27	0.42	-0.14
Denmark	0.42	0.28	0.15	0.25	0.43	-0.18
England Wal	0.40	0.38	0.02	0.27	0.31	-0.04
Spain	0.45	0.37	0.07	0.32	0.38	-0.07
France	0.40	0.30	0.11	0.31	0.43	-0.12
Italy North	0.46	0.34	0.13	0.27	0.36	-0.08
Netherlands	0.45	0.34	0.11	0.30	0.36	-0.06
Norway	0.36	0.20	0.16	0.22	0.43	-0.21
Sweden	0.45	0.39	0.05	0.42	0.47	-0.04

Simple country averages of 6 advantage-advantage and 6 disadvantage-disadvantage paired fuzzy-set coincidence scores.