

Measuring Social Capital By Survey Data: A Proposal

Una proposta per misurare il capitale sociale mediante dati campionari

Rita Lima

CIRMET, Università degli studi di Palermo, Edificio 18, Viale delle Scienze, Palermo
90128 , ritalima@hotmail.it

Miranda Cuffaro

Dipartimento di Contabilità Nazionale ed Analisi dei Processi Sociali, Edificio 3,
Università degli studi di Palermo, Viale delle Scienze, Palermo 90128

Riassunto: Lo scopo di questo articolo è quello di proporre una misura del capitale sociale, utilizzando i dati provenienti da un'indagine campionaria. L'indagine fa parte di un progetto di ricerca più ampio, che coinvolge altre Università italiane. Essa ha lo scopo di indagare sulle relazioni tra le caratterizzazioni dell'offerta di lavoro in termini di abilità e conoscenze e i bisogni della domanda di lavoro. Il trattamento statistico dei dati è stato fatto con l'algoritmo CATPCA che usa la metodologia Alternative Least Squares in contesto di Optimal Scaling.

Keywords: Social capital, CATPCA, Optimal Scaling

1. Introduction

The novelty of the paper consists in generating a measure of social capital by survey data, in spite of the majority of recent papers which have proposed synthetic measures by macro data (see for example, Di Giacinto and Nuzzo, 2005). The sociological literature has stressed the role of human and social capital, as physical capital in explaining the different rates of economic growth (see Section 2 for a brief review). Here, according to Coleman (1988), we hypothesize that social capital in the family (family background and social relationship among parents) and social capital in the community (relations among persons outside the family) are very important elements in the formation of human capital. Actually, social capital is strongly related to human capital, as it refers to individuals, to their characteristics and their capacity of relationships. On this basis, we have organized a survey and collected data by a questionnaire made up of two parts: one aimed to evaluate the skills and knowledge acquired by an individual (human capital), the other aimed to evaluate the social relationships of the individual in and outside the family (i.e. bonding and bridging social capital). At our knowledge, the proposal of measuring social capital by survey data is limited only to some experiments (e.g. World Bank, 2004). Here we present, for the sake of illustration, an exercise aimed to synthesize the survey information. The statistical treatment of data by CATPCA algorithm is described in Section 3. Finally, we discuss the results and conclude (Section 4).

2. Social capital: review on definitions, theories and measures

Social capital is an ambiguous concept, which has been defined in various ways, depending on the theoretical context where it is inserted. Indeed, it is a multidimensional concept, comprehensive of different aspects of an individual's or a community's life: sociability, social networks, solidarity, trust, civic engagement. Coleman (1988) distinguishes between social capital within the family and outside the family. As arguably pointed out by Morrow (1999), "*Coleman's notion of social capital (...) provides a conceptual link between the attributes of individual actors and their immediate social contexts, most notably the household, school and neighbourhood*". According to his definition, social capital is a key characteristic of communities rather than of individuals and composed by the following components: networks, people's sense of belonging to the civic community, norms of cooperation, reciprocity and trust. On this basis, different levels of social capital could explain the different rates of economic growth. In the context of our project of research the latter is the aspect more attractive. Indeed, in all recent economic literature, both Italian than foreigner, social capital like physical and human capital, is considered as an input of the growth process. For example, Becattini (2000) indicates it as an important vehicle of the good economic performance of Italian industrial districts in the nineties. Moreover, the World Bank, to understanding the causes and consequences of poverty in some underdeveloped countries, has engaged in extensive conceptual and empirical work on social capital (World Bank, 2004). For these reasons, the interest in the measurement of social capital is becoming more and more pressing. Generally, scholars have provided a synthetic measure of social capital obtained by some multivariate statistical method (i.e. Factor Analysis, Principal component Analysis) applied on macro indicators collected at regional or provincial scale (e.g. number of charitable organizations, number of voluntary members in non-profit organizations etc.). In spite of this, the World Bank has been operating by direct surveys with questionnaire which has been pre-tested in the field in Nigeria and Albania.

3. Data Handling

Statistical analyses were carried out by the CATPCA (the Categorical Principal Components Analysis or PRINCALS; see Gifi, 1991). CATPCA include the multiple correspondence analysis and all nonmetric PCA methods. This procedure simultaneously quantifies categorical m variables while reducing the dimensionality of the data. By reducing the dimensionality, one can interpret few components rather than a large number of variables. Moreover, CATPCA, with respect to ordinary principal components analysis (PCA), allows to treat variables not only as numeric, but as ordinal or nominal variables as well. The technique consists of finding object scores X of order $n \times p$ (i.e. n =number of case-objects, p =number of dimensions) and a set of multiple category quantifications Y_j of order $k_j \times p$ (i.e. k_j =number of categories of each variable and $j = 1, \dots, m$) so that the loss function:

$$\sigma(X; Y) = \frac{1}{m} \sum_j \text{tr} \left\{ (X - G_j Y_j)' M_j (X - G_j Y_j) \right\} \quad (1)$$

is minimal, under the normalization restriction $X'M_*X = mnI$, and the condition $M_* = \sum_j M_j$, where: a) M_j is a diagonal $n \times n$ matrix in which the diagonal elements $m_{j(ii)}$ are 0 when the i -th observation is within the range $[l, k_j]$ and are 1 when the i -th observation is outside the range $[l, k_j]$; b) G_j is an indicator matrix for variable j , of order $n \times k_j$. The elements of $G_j g_{(j)ir}$ are defined as $i=1, \dots, n$; $r=1, \dots, k_j$, and are 0 when the i -th object is in the r -th category of variable j and 1 when the i -th object is not in the r -th category of variable j ; c) I is the $p \times p$ identity matrix. The algorithm uses Alternating Least Squares to minimize the loss function. It consists of two phases, a model estimation phase and an optimal scaling phase, iteratively alternated until convergence is reached. Actually, in the search for an optimal mean correlation between the recoded variables and the components, both the component loadings and the category quantifications are varied until the optimum is found. Relating to our data, different measurement level have been assigned for each variable independently (nominal, ordinal and numerical). Finally among these we have chosen the ordinal one, as it explains the highest percentage of variance (41%) after only 23 iterations¹.

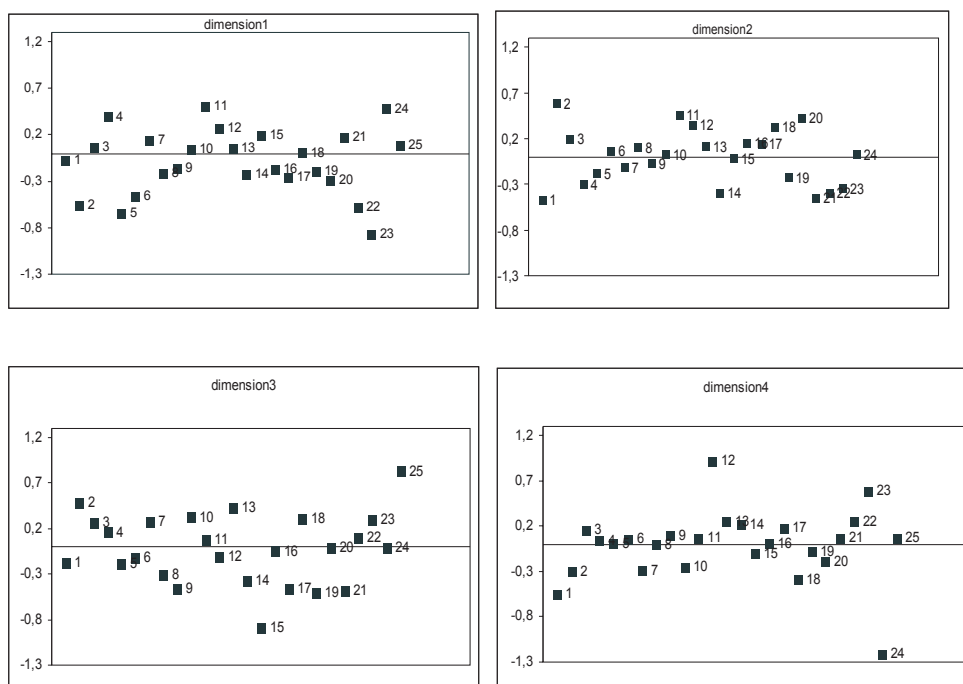
4. Empirical findings

The CATCPA on the data of 420 questionnaires highlights the role of some variables in defining social capital. The structure of data is well represented by four dimensions, as indicated by the Cronbach's alpha test (0,94) and by the percentage of total variance explained by the model (41%). About 14% of the total variance was explained by the first dimension while only the 10% by the second dimension. To interpret the four dimensions we considered those variables with component loading greater than 0,35 on each dimension. Dimension 1 could represent the "*structural social capital*" as it is mainly composed by those factors related to trust and solidarity. It measures how much the respondents trust in their teachers or in social and political institutions (judges, monks, policemen, local police officers, union). Dimension 2 is mainly explained by social participation in voluntary and religious associations or in artistic and cultural groups (or activities) which could have long distance network with other groups out of the city. For this reason dimension 2 could be related to the personal improvement and to the social environment ("*civicness*"). In fact, it represents the availability of students to belong to groups and calls attention to the fact that "*civic virtue is most powerful when embedded in a sense network of reciprocal social relations*" (Putnam, 1995). Dimension 3 is defined by the availability of individuals to meet friends, to be satisfied with their friendship, to consider their friends trustworthy. Such dimension could indicate the "*communication and sociability factor*". Dimension 4, as is explained by to go to the stadium, to go to the disco, to do sport, could represent the free-time activities in public spaces ("*social network of short distance*"). The object scores for the four dimensions solution are plotted in Figure 1, in which x axis is used only to ranking the 25 districts of the city of Palermo. Along dimension 1 we can observe a wider dispersion of zones with respect to the median value. Districts 2, 5, 6, 22 and 23 are the lowest structural social capital areas. Along dimension 2, zones 2, 11, 12, 18 and 20 are

¹ By this way, in order to reduce missing data, we replaced missing values with the most frequent category. When there are multiple modes, the one with the smallest category indicator is used. In fact CATPCA treats the missing values by setting weights in the loss function (1) equal to 0.

the most actively engaged in groups and networks. Looking the shape of zones, according to dimension 3, we see that more than 24% of districts were below the median while only district 25 is very far from it. Finally, according to dimension 4 the most districts are concentrated around the median.

Figure 1: *Object Scores by dimension and zone code*



References

- Becattini G. (2000) *Dal distretto industriale allo sviluppo locale.*, Bollati Boringhieri, Torino.
- Chiesi A. (2003) Problemi di rilevazione empirica del capitale sociale. *Inchiesta*, XXXIII, 139, 86-97.
- Di Giacinto V., Nuzzo G. (2005) Measuring social capital: evidence from Italy, *Proceedings of "Local Economies and Internationalization in Italy"*, 159–183. Bologna .
- Colemann J.S. (1988) Social Capital in the creation of Human Capital., *American Journal of Sociology* .,94, 95-120.
- Morrow V.(1999) Conceptualising social capital in relation to the well-being of children and young people: a critical review, *American Sociological Review*, 744-765.
- Gifi A. (1991) *Nonlinear Multivariate Analysis* (Reprint with corrections), Chichester, England: John Wiley & Sons.
- Putnam R. D. (1995) Bowling Alone: America's Declining Social Capital, *Journal of Democracy*, 6:1, Jan, 65-78.
- Word Bank (2004), Measuring Social Capital, *Working paper*,18, 1-19.