Identification of relationship between housing difficulty and property values in urban areas

Silvestro Montrone and Paola Perchinunno and Carmelo M. Torre

Università di Bari, Università di Bari, Politecnico di Bari

October 2008

Identification of relationship between housing difficulty and property Values in Urban Areas

Silvestro Montrone, Paola Perchinunno[1], and Carmelo Maria Torre[2]

[1] 1 Department of Statistical Science, University of Bari,

Via C. Rosalba 53, 70100 Bari, Italy

p.perchinunno@dss.uniba.it

2 Department of Architecture and Urban Planning, Polytechnic of Bari,

Via Orabona 4, 70125 Bari, Italy

torre@poliba.it

Abstract. The objective of the present work is to use statistical data to identify territorial zones characterized by the correlation between urban access to services and quality of housing and the value of property ownership. While poverty is widely accepted to be an inherently multi-dimensional concept, it has proved very difficult to develop measures that both capture this multidimensionality and make comparisons over time and space easy. With this in mind, we attempt to apply a Total Fuzzy and Relative (TFR) approach, based on a fuzzy measure of the degree of association of an individual to the totality of the poor and an approach of semantic distance (Munda, 1995), based on the definition of a “fuzzy distance” as a discriminating reference to rank the availability to property in real estate market, as complement of urban poverty, in a specific case (the Italian City of Bari).

JEL Classification: C1, C5, R2, R31,

1 Introduction

The question of housing in Italy is of interest to academics, politicians and social operators. Some estimates of the Bank of Italy show that more than 28 % of Italian families (six million households) were already in a state of residential discomfort in 2002. The Italian housing problems come after other countries experienced the same difficulties, due to high housing costs heavily affecting household budgets (Poterba 1984), or due to the inadequacy of housing, of urban facilities and of high quality dwellings. New social groups that suffer housing problems are identified in the lower middle class and working class in areas characterized by high urban

density. In the light of the actual conditions it becomes necessary and urgent to deal with the policies that should be adopted to manage the housing emergency in metropolitan areas: the lack of available housing or the offer of acceptable rent levels, the lack of social services and household overcrowding, or the inadequacy of economic resources to achieve a better living standard are all conditions which, cause housing problems and that means urban poverty.
Poverty does not only mean scarcity of financial self-funding: it embraces a plurality of social and cultural, issues, such as education, health and housing. In other words it represent a measure of the inequality of access to various basic goods and services.
This work aims to determine territorial zones (hot spots) characterized by the presence of urban poverty in the City of Bari, based on the most recent available data released by ISTAT (Italian National Statistics Institute) in the last Population and Housing Census (2001).
Such multi-dimensionality suggests an exploration of a new definition of indicators useful to describe housing discomfort, related to the poverty of small areas and to real estate values (Montrone et al. 2007).
The presence of a varied range of definitions on the theme of poverty shows the necessity of no longer relying on a single indicator but on a group of indicators which are useful in the definition of living conditions of various subjects.
The approach chosen in order to arrive at the synthesis and measurement of the incidence of relative poverty in the population in question is the so-called "*Total Fuzzy and Relative*" method, "which uses the techniques of the *Fuzzy Set* in order to measure the incidence of relative poverty within a population, beginning from statistical information gathered from a plurality of indicators" (Lemmi e Pannuzi, 1995).

2 Identifying indicators of social and housing difficulty

2.1 Introduction

Since the end of the 1970s, numerous studies have been based on a variety of approaches, each of which adopted an attentive definition and conceptualization of the phenomena. Townsend (1979), defines poor families those that “lack the resources for a quality of alimentation, participation in activities and enjoyment of the living conditions which are standard, or at least widely accepted, in the society in which they are living”. The reference is, therefore, towards a concept of poverty as relative privation, which takes into account the particular historical, economic, social, geographical, cultural and institutional context under examination. Within this study, twelve principal dimensions of poverty were identified which are:, diet, clothing, housing costs, costs within the household, living conditions, working conditions, health, education, the environment, family activities, recreational activities and social relations. It may be noted that three of the twelve areas considered are connected to housing conditions. The twelve categories described above have been used in many later studies based on the concept of so-called multidimensional poverty, carried out amongst others by Gailly and Hausman (1984), Mack and Lansley (1985) and Desai and Shah (1988).

2.2 Measuring indicators of social and housing difficulty

The subject of this study derives from the necessity to identify geographical areas characterized by situations of residential deprivation or urban poverty in the City of Bari (hot spots). In order to analyse the phenomena of residential poverty on a geographical basis, this work uses data from the most recent *Population and Housing Census 2001* carried out by ISTAT; this information allows geographical analysis in sections according to the census, albeit hindereded by the lack of more recent data. Specifically, the geographical units of the survey are the 1,421 census sections within the City of Bari, of which 109 are uninhabited areas or are designated for other uses (for example parks or universities).
The choice of those indexes to take into consideration has been made according to the analysis of some socio-cultural aspects of the resident population in the City of Bari considered useful in defining urban poverty: these are education levels, working conditions and living conditions.
It emerges from the study of "Relative Poverty in Italy" conducted annually by ISTAT that low levels of education, exclusion from the employment market and precarious residential standards are closely linked to conditions of poverty.
The various indices were classified into *two sets*:

- *Social difficulty*, related to the conditions of the resident population within the various census sections (educational qualifications, working conditions, overcrowding);
- *Housing difficulty,* related to the housing conditions of dwellings occupied by residents in the various census sections (housing status, lack of functional services such as landline telephone, heating systems and designated parking space).

The indexes have been calculated both at the scale of the census section and at the scale of the neighbourhood. Table 2.1 shows average indexes of household poverty for each neighbourhood.
The analysis of urban poverty, referring to the census sections of the City of Bari, shows that the minimum level of school attendance is reached by 55% of citizens aged over 19. This index shows the highest average percentage in a popular neighbourhood (San Paolo) and the old town (San Nicola), and the lowest percentage in the city centre (Murat).
As regards unemployment the average rate is about 20%; the worst result is 28% in the San Paolo area, and the best result is in the Murat area.

The dwelling space index is 3.42 on average (more or less three inhabitants per 100 square meters). Overcrowding is relevant in the San Nicola area, with five inhabitants per 100 meters (20 square meters per capita) and four inhabitants per 100 meters (25 square meters per capita) in the Stanic and San Paolo areas).
As regards typology of occupation, the index measures the percentage of the total of dwelling occupied by renters. In the City of Bari the average percentage is more or less 29% (while 6% are free used, and 65% are owned by occupants). More specifically, there are 31,558 rented homes, while there are 72,587 owner-occupied homes. Again, San Nicola (47%), San Paolo (46%) and the Madonnella (43%) show the worst conditions, while the peripheral Loseto area shows the best conditions, even if the property values are not highly rated, and its newest properties are often rented to young couples.
When it comes to facilities, on average 17% of households do not have a contract for a landline telephone and 10% of households do not have heating systems. Car parking is lacking, because on average 53% of households do not have a parking place.
The critical condition of the San Nicola area is confirmed also by the lack of facilities: 33% of households do not have a contract for a telephone line, 38% of households have no heating systems and 92% of households do not have a parking place.
The difficulties caused by the lack of facilities are a common problem of the Madonnella and Libertà areas, the most degraded areas in the centre of the city. The Murat area has a good range of facilities, except for parking places. Despite the scarcity of parking places in the Murat area, the property value is quite high, due to the presence of many of historical residential buildings, that have been restored and represent the best quality supply of offices and residence in the city centre area.

Table 2.1. Average of indexes [1] per single area of the City of Bari - 2001

Neighbour-hoods	Index 1	Index 2	Index 3	Index 4	Index 5	Index 6	Index 7
Carbonara	0.61	0.22	3.55	0.26	0.19	0.07	0.51
Carrassi	0.43	0.15	2.92	0.21	0.10	0.07	0.64
Ceglie	0.71	0.25	3.61	0.32	0.22	0.08	0.45
Japigia	0.50	0.20	3.23	0.29	0.12	0.07	0.31
Libertà	0.66	0.22	3.93	0.36	0.18	0.21	0.86
Roseto	0.66	0.25	3.49	0.19	0.20	0.03	0.26
Madonnella	0.53	0.21	3.30	0.43	0.18	0.20	0.84
S.Girolamo F.	0.63	0.21	3.88	0.30	0.21	0.07	0.27
Murat	0.28	0.11	2.28	0.24	0.09	0.06	0.84
Palese	0.52	0.19	3.25	0.23	0.17	0.04	0.26
Picone	0.44	0.16	3.13	0.22	0.11	0.06	0.44
S.Nicola	0.74	0.25	5.00	0.47	0.33	0.38	0.92
S.Paolo	0.84	0.28	4.03	0.46	0.15	0.07	0.36
S.Pasquale	0.43	0.16	3.29	0.23	0.13	0.08	0.53
S.Spirito	0.58	0.20	3.44	0.30	0.20	0.07	0.24
Stanic	0.70	0.27	4.00	0.24	0.27	0.11	0.20
Torre a Mare	0.51	0.19	2.85	0.27	0.26	0.08	0.19
Bari	0.55	0.20	3.42	0.29	0.17	0.10	0.53

Source: Our elaboration of the data from the Population and Housing Census, 2001.

[1] **Index 1-** Index of lack of progress to high school diploma: ratio between the total number of residents aged 19 or over who have not obtained a high school diploma and the total number of residents of the same age group.

Index 2 - Rate of unemployment: the ratio between the total number of residents aged 15 or over who are in search of employment and the workforce of the same age group.

Index 3 – Index of overcrowding: the ratio between the total number of residents and size of dwellings occupied by residents.

Index 4 - Incidence of the number of dwellings occupied by rent-payers: ratio between the number of dwellings occupied by rent-paying residents and the total number of residents.

Index 5 - Incidence of the number of dwellings lacking a landline telephone: ratio between the number of dwellings occupied by residents without a landline telephone and the total number of dwellings occupied by residents.

Index 6 - Incidence of the number of dwellings lacking heating: ratio between the number of dwellings occupied by residents without a heating system and the total number of dwellings occupied by residents.

Index 7 - Incidence of the number of dwellings lacking parking space: ratio between the number of dwellings occupied by residents without a parking space and the total number of dwellings occupied by residents.

3. The geographical zoning of data

3.1 Introduction to spatial clustering

If it is possible to reduce the analysed problem to a few aspects of the observed phenomenon (such as incidence of diseases and/or crimes in a territory), in this case the cartography becomes a thematic map where areas of interests join each other with spatial contiguity. The result is a zoning based on a spatial (or spatial-temporal) clustering, the conceptual aspects of which deserve a better definition.
Knox (1989) in his studies on spatial relationship of epidemic phenomena gave a seminal definition of spatial clustering: a spatial cluster is a non-usual collection/aggregation of real or perceived (social, economic) events; it is a collection of spatial, or spatial/temporally delimited events, an ensemble of objects located in contiguous areas.
From a statistical point of view, in this case, the clustering can be based on the identification of areas where a group of points shows the maximum incidence inside, and at the same time leaves the minimum incidence outside, referring to a given phenomenon. Such operation is obtained by locating a circular window of arbitrary radius, by calculating the probability (risk) p_1, inside the circle, or the probability (risk) p_2, outside the circle, and finally by rejecting the pointless hypothesis:

$$\begin{cases} H_0: & p_1 \le p_2 \\ H_1: & p_1 > p_2 \end{cases}$$

or rejecting the pointless hypothesis:

$$\begin{cases} H_0: & p_1 \ge p_2 \\ H_1: & p_1 < p_2 \end{cases}$$

if the aim is to identify minimum risk areas. Of all windows, the minimum *p*-value (probability of critical region referring to the test) corresponds to the most important cluster. The identification of a special area can be based on the intensity of a statistical attribute, instead of the number of attribute-characterised elements.
In the field of epidemiological studies many research groups have developed different typologies of software; these are all based on the same approach, but usually differ from each other in the shape of the window.
Among the various methods of zoning, there are SaTScan (Kulldorff 1997) that uses a circular window, FlexScan (Tacahashi et al. 2004), that uses contiguity to build the window, the Upper Level Scan Statistics (ULS: Patil and Taillie 2004), that underpasses the question of geometric shape of the window including aggre-

gate points and finally AMOEBA (A Multidirectional Optimal Ecotope Based Algorithm: Aldstadt and Getis 2006), that uses a similar approach to SaTScan, without the constraint of a circular window.
In this work, the zoning method used is the SaTScan. This method in some cases can produce an imprecise zoning, due to the circular shape of the window, especially in peripheral urban areas, or other areas that are wider than the given size of the radius inside of which is defined the homogeneity of the considered attribute. The research we report in this paper on social-residential problems, based on socio-economic indicators, has led to the identification of small areas with high indexes of poverty although quite heterogeneous in the observed urban context: namely *hot spots*. Regarding this specific aim, SaTScan seems to be quite efficient (release 7.03, freely available at http://www.satscan.org).

4. Determination of the variable for classification

4.1 Methods for spatial clustering

The different scientific research pathways are consequently directed towards the creation of *multidimensional indicators*, sometimes going beyond dichotomized logic in order to move towards a classification which is "*fuzzy*" in nature, in which every unit belongs to the category of poor with a range from 1 to 0, where the value 1 means definitely poor, 0 means not poor at all, and the other values in the interval reflect levels of poverty. Classifying populations simply as either *poor* or *non-poor* constitutes an excessive simplification of reality, negating all shades of difference existing between the two extremes of high level well-being and marked material impoverishment. Poverty is certainly not an attribute which can characterize an individual in terms of presence or absence, but rather is manifested in a range of differing degrees and shades (Cheli and Lemmi 1995).
The development of *fuzzy theory* stems from the initial work of Zadeh (1965), and successively of Dubois and Prade (1980) who defined its methodological basis. Fuzzy theory assumes that every unit is associated contemporarily to all identified categories and not univocally to only one, on the basis of ties of differing intensity expressed by the concept of degrees of association. The use of fuzzy methodology in the field of "poverty studies" in Italy dates back only a few years, thanks to the work of Cheli and Lemmi (1995) who define their method "*Total Fuzzy and Relative*" (TFR) on the basis of the previous contribution from Cerioli and Zani (1990). The TFR approach consists in the definition of the measurement of a *degree of membership* of an individual to the fuzzy totality of the poor, included in the interval between 0 (with an individual not demonstrating clear membership to the totality of the poor) and 1 (with an individual demonstrating clear membership to the totality of the poor). Mathematically such a method consists of the construction of a function of membership to "the fuzzy totality of the poor" continuous in

nature, and "able to provide a measurement of the degree of poverty present within each unit" (Cheli and Lemmi 1995; Lemmi and Pannuzi 1995). Supposing the observation of k indicators of poverty for every family, the function of membership of i_{th} family to the fuzzy subset of the poor may be defined thus (Cerioli and Zani 1990):

$$f(x_{i.}) = \frac{\sum_{j=1}^{k} g(x_{ij}).w_j}{\sum_{j=1}^{k} w_j} \qquad i = 1,.....,n \tag{1}$$

The w_j function in the function of membership are only a *weighting system* (Cheli and Lemmi 1995), as for the generalization of Cerioli and Zani (1990), whose specification is given:

$$w_j = \ln(1/\overline{g(x_j)}) \tag{2}$$

The weighting operation is fundamental for creating synthetic indexes, by the aggregation of belonging (to the ensemble of the poor) functions of each single indicator of poverty. An alternative, by Betti, Cheli and Lemmi (2002) starts from the conjoint use of the coefficient of variation as the first component of the set of weights, with the correlation coefficient, as the second component. The new set of weights, that is proposed for continuous variables, takes into account two factors, described in the following multiplicative form:

$$w_j = w_j^{(a)} * w_j^{(b)} \tag{3}$$

where:

$w_j^{(a)} = \frac{\sigma_j}{\mu_j}$ is given from the coefficient of variation of X_j

$w_j^{(b)} = 1 - \frac{\sum_{l \neq j} \rho(X_j, X_l)}{\sum_{l=1}^{k} \rho(X_j, X_l)}$ is given from the complement to one of the ratio between the sum of all correlation coefficients, left out the j array, and the whole sum of correlation coefficients referring to X_j.

4.2 The result of the Fuzzy Approach

The TFR method has been applied to the values of the described sets of indicators in paragraph 2. From this application we derive average values of ownership functions $g(x_{ij})$ and the corresponding weights w_j .

Table 4.1. Results from the TFR method as regards repartition function and corresponding indexes

Measure of poverty	$\overline{g(x_j)}$	Weighting system w_j (2)	New weighting system w_j (3)
Social difficulty:			
Index 1	0.578	0.548	0.208
Index 2	0.354	1.040	0.351
Index 3	0.348	1.055	0.209
Housing difficulty:			
Index 4	0.431	0.842	0.486
Index 5	0.258	1.354	0.508
Index 6	0.129	2.045	0.821
Index 7	0.478	0.739	0. 903

Source: Our elaboration of the data from the Population and Housing Census, 2001.

The value of weights w_j according to the basic method, varies according to the degree of importance in determining the level of poverty. Taking for example the set of indexes of housing difficulty, since there is more property with heating than property with parking place, it is appropriate to give a more important weight for the first indicator (w_i = 2.045), than for the second one (w_i = 0.739), because it is supposed that the more widespread positive condition identifies a more discriminating factor when the same condition is not verified in determining estate poverty. Notice that in the new system of weighting the definition of the level of poverty is less discriminating than the former approach both for social and housing difficulty.

As regards the distribution of results deriving from the analysis of 1,312 census sections, the more fuzziness is closer to 1, this identifies the condition of social and housing difficulty; on the contrary, the more fuzziness is closer to 0 the more the value identifies the condition of social and residential welfare.

Table 4.2. Composition of absolute values and percentage values of the census sections for conditions of poverty in 2001

	Absolute values		*Percentage values*	
Fuzzy Value	**Social difficulty**	**Housing difficulty**	**Social difficulty**	**Housing difficulty**

0,0 ┤0,2	583	643	44%	49%
0,2 ┤0,4	49	361	4%	27%
0,4 ┤0,6	370	145	28%	11%
0,6 ┤0,8	6	73	1%	6%
0,8 ┤1,0	304	90	23%	7%
Totale	*1.312*	*1.312*	*100*	*100*

Source: Our elaboration of the data from the Population and Housing Census, 2001.

Regarding to the set of *social difficulty indexes*, 23% of observed census sections show fuzzy values in the range (0.8-1.0), representing a clear condition of poverty, counterbalanced to 44% of census sections belonging to fuzzy cluster of non-poor, identified by the range (0.0-0.2).
The extreme conditions of *housing difficulty* are less widespread: more specifically, 7% of observed sections belong to the range (0.8-1.0), representing the condition of unquestionable poverty, and 49% of the observed sections belong to the range (0.0-0.2), representing the non-poor.

5. The identification of hot spots of social and housing difficulty

5.1 The clustering of social difficulty

As reported above, the core question of this paper derives from the need to identify hot spots presenting conditions of social and housing difficulty in the urban area of Bari, on the basis of data emerging from a set of indicators. With the use of the SaTScan method a possible identification of hot spots of discomfort has been obtained from the data generated by a fuzzy analysis starting from two sets of indicators (the first for social character, the second for residential character of difficulty).
The use of SaTScan on variables referring to *social difficulty*, from the TFR method, leads to the zoning of hot spots representing extreme poverty areas. In detail we have identified four clusters, composed of a different number of census sections, for a total number of 491, where the identification of difficulty is given by the mean inside; the higher the mean value, the higher is the level of poverty. A further aspect of interest is given by the *p*-value, that is the probability of the critical region of the test, where the lower the values shown, the better defined is the cluster.

Table 5.1. Composition and description of cluster referring to social difficulty sets

Cluster	**Number of**	**Mean inside**	**Mean outside**	**Standard de-**	***p*-value**

	cases			viation	
1	8	0.92	0.39	0.33	0.0620
2	64	0.76	0.38	0.33	0.0010
3	268	0.62	0.34	0.32	0.0010
4	151	0.59	0.37	0.33	0.0010
Total	*491*				

The interpretation of data analysis shows that the value of mean inside (included in the interval 0 - 1, that is to say "no poverty - definite poverty") are very high. Optimal values are shown also by *p*-values of the four clusters.

The four clusters are shown on a map by a shades of green ranging from maximum social difficulty (the darkest grey) to minimum social difficulty (the lightest grey) so that the represented reality is immediately understandable.

The dark shades denote a more depressed condition, located in such peripheral areas that can be defined "central peripheries" (Pace 2006) identified by the central popular quarters of the past. The first risk area for poverty (mean inside 0.92) is identifiable in a part of the Madonnella Quarter, where there is a public housing estate (The "Duca degli Abruzzi" Estate). The second area (mean inside 0.76) is represented by the historic mediaeval town centre called San Nicola, that is still characterised by phenomena of social marginalisation. The third cluster is composed of the largest number of sections (268, with a mean inside corresponding to 0.62), represented by central peripheries of the past, like Libertà, and popular peripheries of the 1960s (San Girolamo, San Paolo) or outlying neighbourhoods, like Palese, that have been less interested by urban renewal (Rotondo and Selicato 2006).

The situation of the fourth group (mean inside 0.59) is represented again by three neighbourhoods that have been neglected by urban policies of renewal, in spite of the presence of a large number of public housing estates.

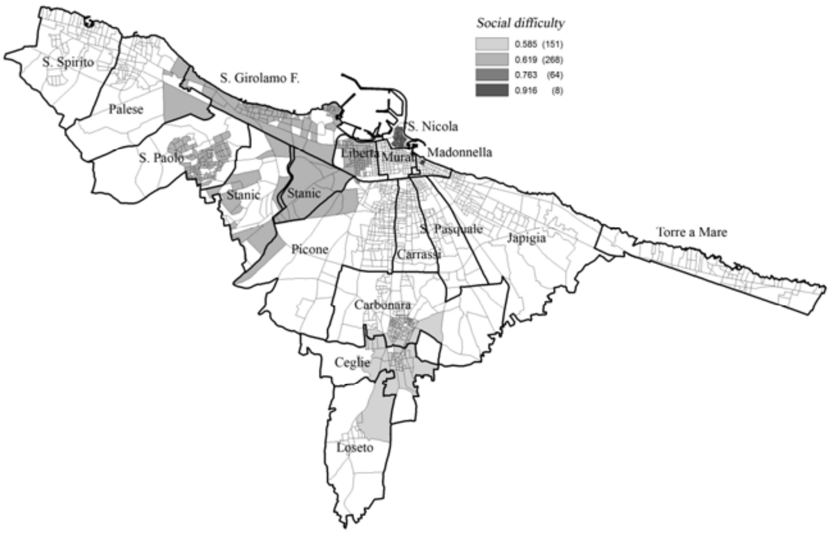

Fig. 5.1. Geographical representation of hot spots of social difficulty in the City of Bari

5.2 The clustering of housing difficulty

The individuation of hot spots of *housing difficulty*, shows a higher level of selectivity than social discomfort. The four identified clusters are composed by a total of 286 census sections. Among the four clusters, the first and the third are mainly discriminating, since the second and the fourth cluster, although they show high values of means inside, have a high *p*-value (about 0.5).

Table 5.2. Composition and description of cluster referring to the set of housing difficulty

Cluster	Number of cases	Mean inside	Mean outside	Standard deviation	*p*-value
1	65	0.71	0.24	0.23	0.0010
2	8	0.64	0.26	0.25	0.5150
3	129	0.46	0.24	0.24	0.0001
4	84	0.38	0.26	0.25	0.5500
Total	*286*				

As regards the spatial mapping, the darkest brown areas represent the highest level of housing difficulty, and include the old town centre of San Nicola (mean inside 0.71), and a piece of Ceglie (mean inside 0.64).

The secondary cluster (characterised by high mean inside and high *p*-value), identify as areas with housing problems Libertà, Stanic and Picone areas (129 census sections, mean equal to 0.46) and Carbonara (84 census sections, mean equal to 0.38).

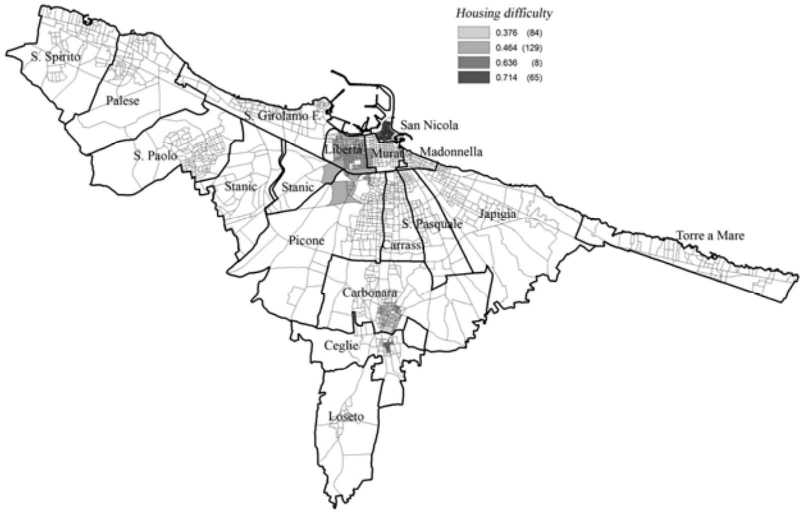

Fig. 5.2. Geographical representation of hot spots of housing difficulty in the City of Bari

6. Housing difficulty and property values

6.1 Housing market segmentation and multidimensional ranking

The fuzzy analysis carried out on the Census Sections utilizes a wide number of spatial elements in the light of select clusters according to an appropriate statistical identification. At the urban level, at least two further level of spatial classification exist, that can be helpful to describe a relationship between the urban poverty and spatial-physical attributes of the urban estates. The first one is corresponding to the division of the city in quarter and neighbours, that are the result of the urban development during the time. The second one is the classification of the real estates, in the light of their owning to a cluster of the housing market, according to those socioeconomic (Bayliss, 1968), architectural, spatial and situational (Rosen, 1977) attributes that generate their market value. Urban poverty in this case could represent an indirect way to identify urban areas characterized by low-price housing estate. This relationship consider that the availability of high-price estate is obviously limited for poor.

Property value is added according to the supposition that the value of a dwelling represents the greatest indicator of limitation in terms of market availability to residential services. Indeed, this addition is the basis for the shift in the geographical dimension of the analysis from census sections to localities.
The quality of residential services, and consequently the price, is further associated with physical aspects and external quality, in addition to recognisable aspects of the social context which may be aligned with the significance of intrinsic relative and variable marginal prices which appreciate according to the social and environmental context.
A main difficulty is to make the spatial window referring to urban poverty corresponding with a well identified spatial contest, such as a quarter. Cities are historically subdivided in quarters that are not well bounded according to their physical characters, and the social status of hits inhabitants. In this case the interesting question is about how to connect the urban and architectural character of housing estate with the socio economical data deriving from the analysis of social and housing difficulty.
In our experiment, the attempt is to test the correspondence between a possible ranking of the urban quarters, in the light of their attributes of social and housing difficulty, with a market segmentation implicitly represented by estimated real estate values. The experiment has been carried out by the application of a fuzzy multicriterial ranking, with the consciousness that the spread of values in a quarter, as above reminded, is to be considered imperfectly bounded.

6.2 The method for a fuzzy ranking of the urban quarters

The fuzzy multimendsional evaluation proposed by Munda (1995), employs the construction of a discrete evaluation with multiple criteria to which relative value judgements can be expressed through levels of verbal and quantitative grading.
The components of a discrete multi-criteria valuation can be described in the following way: E is a finite set of n elements; m is the number of different criteria considered relevant in a evaluation problem, where:
the element X is assessed as preferable than Y (both belonging to the set E) according to the m-th point of view if $\mathrm{Rank}_m(X) > \mathrm{Rank}_m(Y)$, or, alternatively, the element X is assessed as indifferent respect to Y according to the m-th point of view if $\mathrm{Rank}_m(X) = \mathrm{Rank}_m(Y)$.
The rank of the m-th criterion Rank_m is expressed in our case by a quantitative intensity of preference g_m; therefore if $\mathrm{Rank}_m(X) > \mathrm{Rank}_m(Y)$, this means that:

$$g_m(X) - g_m(Y) > s$$

and, if $\mathrm{Rank}_m(X) = \mathrm{Rank}_m(Y)$, this means that:

$$g_m(X) - g_m(Y) \leq s$$

s is a positive number, the so-called "indifference threshold". This implies that a grey area exists in the interval $(0,s)$, where, in spite of a preference $g_m(X) > g_m(Y)$, we obtain as result of a pairwise comparison, the collocation in the same rank of two generic elements X and Y. This is the representation of non perfect transitivity of such kind of rankings, that was historically evidenced by Luce (1956).
Note that it is possible as well to face with a problem of incomparability between X and Y, but we assume that incomparability does not exists in this case of study.
Therefore, the elements composing a multi-criteria evaluation (finalised to have a ranking) are:

- the intensity of preference (when preference is expressed by quantitative criterion scores);
- the number of criteria in favour of a given alternative;
- the eventual weight associated to each single criterion;
- the kind of relationship of each single alternative with all the other alternatives (that is to say preference, indifference or incomparability).

As the regards the last point of above, in order to give a better definition of such area of indifference, some authors introduce the dual concept of "strong preference" and "weak preference" (Roy, 1985).
The "strong preference" and "weak preference", are represented by a couple of thresholds of indifference, instead of one: in this case, if $\text{Rank}_m(X) > \text{Rank}_m(Y)$, this means that:

$$g_m(X) - g_m(Y) > s_1 + s_2$$

or this can mean as well that:

$$g_m(X) - g_m(Y) > s_1 .$$

In the first case we speak of "strong preference", represented by the overcoming of the sum of two thresholds (s_1 and s_2, representing the weak and strong preference thresholds); in the second, we speak of "weak preference" (s_1, representing only the weak preference threshold).
The final result of the application is that in the two-levels preference the intensity of preference g is associated to a pseudo-ranking of a set of element ordered by pseudo-criteria.
We speak of pseudo-criteria because the ranking is affected by a special kind of uncertainty.
In a second step, other authors (Munda, 1995) identify the possibility that the preference of an alternative with respect to another can be formulated through a fuzzy measure of the difference between the value judgements expressed for the alternative in question; leading to a quantitative transposition for the evaluation of credibility, or rather, the value of the function of the fuzzy membership.

The credibility of the ranking relations between two generic alternatives, *X* and *Y*, according to a generic criterion j, can be expressed by judgements (and relationships) as follows:
$\mu_{>>}(X,Y)_j$ defines the credibility of absolute preference for X with respect to Y (μ (X,Y)=1);
$\mu_{>}(X,Y)_j$ defines the credibility of moderate preference for X with respect to Y (μ $(X,Y)_j$ is between 0,5 and 1);
$\mu_{\approx}(X,Y)_j$ defines the credibility of moderate indifference for X with respect to Y (μ $(X,Y)_j$ s near by 0,5);
$\mu_{=}(X,Y)_j$ defines the credibility of absolute indifference for X with respect to Y (μ $(X,Y)_j$ =0,5);
$\mu_{<}(X,Y)_j$ defines the credibility of moderate preference for Y with respect to X (μ $(X,Y)_j$ is between 0 and 0,5);
$\mu_{<<}(X,Y)_j$ defines the credibility of absolute preference for Y with respect to X (μ $(X,Y)_j$ =0).
In this way, we set a fuzzy problem, where the different expressions of $\mu(X,Y)_j$ are the elements of the "Universe of Discourse".
In absence of fuzzyness, only one of the above listed expressions of the universe of $\mu(X,Y)_j$ should differ from 0. In our case obviously it will possible to have more than one relationship of preference/indifference different to 0.
In the final evaluation of the alternatives with respect to all criteria, the pairwise comparison, criteria by criteria, is aggregated. Such aggregation is performed by defining the minimum threshold of the value of the relationships $\mu_j(X,Y)$ on the base of a credibility test α (Bezdek, 1983). According to such credibility test, when the value of the fuzzy preference according to a criterion j of the quarter *X* compared to the quarter *Y* exceeds the threshold value, it can be deduced that the judgement has a credibility equal to 1; in the opposite case such judgement is considered to have no credibility:
$0 \leq \mu\ (X,Y) \leq 1$ if $\mu_j\ (X,Y) > \alpha$ for the majority of criteria j
$\mu\ (X,Y) = 0$ if $\mu_j\ (X,Y) \leq \alpha$ for all the criteria j
$\mu\ (X,Y) = 1$ if $\mu_j\ (X,Y) \geq \alpha$ for all the criteria j and μ (X,Y)j > α for at least one of criteria j.
For every quarter, two function are defined. The function $\Phi^+(X)$ indicates the prevalence of *X* over all Y_k with an index of values included in the interval (0, 1), while the function $\Phi^-(X)$ indicates the non-prevalence of the quarter *X* with respect to others, yet again with an index of values included in the interval (0,1). The expressions of $\Phi^+(X)$ and $\Phi^-(X)$ represent two non symmetric expression of Rank(*X*); they derive from the below formulas.

$$\Phi^+(X) = \frac{\sum_{k=1}^{m-1} \left[\mu_{>>}(X,Y_k) \wedge C_{>>}(X,Y_k) + \mu_{>}(X,Y_k) \wedge C_{>}(X,Y_k) \right]}{\sum_{k=1}^{m-1} C_{>>}(X,Y_k) + \sum_{k=1}^{m-1} C_{>}(X,Y_k)} \qquad \textbf{(4)}$$

$$\Phi^{-}(X)=\frac{\sum_{k=1}^{m-1}\left[\mu_{<<}(X,Y_k)\wedge C_{<<}(X,Y_k)+\mu_{<}(X,Y_k)\wedge C_{<}(X,Y_k)\right]}{\sum_{k=1}^{m-1}C_{<<}(X,Y_k)+\sum_{k=1}^{m-1}C_{<}(X,Y_k)} \qquad \textbf{(5)}$$

$C_{>>}$ and $C_{>}$ represent, respectively, the number of criteria for which Rank*(X)*>Rank*(*Y_k*)*, according to a strong or a weak preference ($\mu_{>>}$, and $\mu_{>}$) for the function Φ^{+}(X).
Instead, $C_{<<}$ and $C_{<}$ represent, respectively, the number of criteria for which Rank*(X)*<Rank*(*Y_k*)*, according to a strong or a weak preference ($\mu_{<<}$, and $\mu_{<}$) for Φ^{-}(X).

Neighbourhoods	**Reale Estate value**	**Φ^{+}(X)**	**Φ^{-}(X)**
Carbonara	0.61	0.22	3.55
Carrassi	0.43	0.15	2.92
Ceglie	0.71	0.25	3.61
Japigia	0.50	0.20	3.23
Libertà	0.66	0.22	3.93
Roseto	0.66	0.25	3.49
Madonnella	0.53	0.21	3.30
S.Girolamo F.	0.63	0.21	3.88
Murat	0.28	0.11	2.28
Palese	0.52	0.19	3.25
Picone	0.44	0.16	3.13
S.Nicola	0.74	0.25	5.00
S.Paolo	0.84	0.28	4.03
S.Pasquale	0.43	0.16	3.29
S.Spirito	0.58	0.20	3.44
Stanic	0.70	0.27	4.00
Torre a Mare	0.51	0.19	2.85

In order to provide a further control of the stability of the ranking, it will be possible to assess the "semantic distance" (Munda, 1995). The semantic distance has a general expression of which in our case we use a appropriate reduction.
Let's give a j-th quantitative criterion of a set of m criteria; let's suppose that $f_j(X)$ and $g_j(Y)$ represent the value functions of the criterion to express $\text{Rank}_j(X)$ and $\text{Rank}_j(Y)$.

In the most general case f_j and g_j can be crisp numbers (this means that the function give a certain result), probabilistic values (this means that f_j and g_j represent expected values), or fuzzy numbers.
In this last case (that is our case) we deal with fuzzy numbers represented by the area bordered by the function f_j and g_j (e.g. a left-right number could be related by a pseudo-Gaussian integral of value equal to 1), as the number "about 1800" and "about 2400", represented in figure 6.1, that have a non empty intersection .

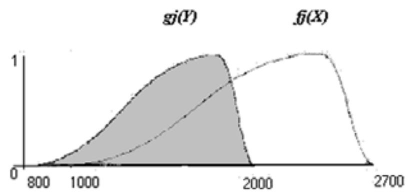

Fig. 6.1. The fuzzy numbers "about 1800" and "about 2400", has a non empty intersection

The fuzzy number and $g_j(1800)$="about 1800" has a domain of the ownership function corresponding to the interval (800, 2000), while the fuzzy number $f_j(2400)$="about 2400" has a domain of the ownership function corresponding to the interval (1000, 2700); the maximum value of "about 1800", that is 2000, is bigger than the minimum value than about 2400, that is 1000.
The "Semantic Distance", in this case, is represented by the sum of two double integral:

$$S_d(f_j(x), g_j(y)) = \int_{-\infty}^{+\infty}\int_{X}^{+\infty} |Y - X| g_j(Y), f_j(X) dYdX + \int_{-\infty}^{+\infty}\int_{-\infty}^{X} |X - Y| f_j(X), g_j(Y) dYdX$$

In case of empty intersection the hypothesis $f_j(X) < g_j(Y)$ is not proved, and the Semantic Distance is represented by $|\max(f_j(X)) - \max(g_j(Y))|$, that coincides with the expected value of $|X-Y|$ (Munda, 2005).
Our multidimensional ranking is based on the calculation of the values of the fuzzy variable $\mu(X,Y)$ and of the semantic distance Sd for each couple of quarters, referring to each criterion of social/housing difficulty, and to the property value of housing estates.
In the following tables the expected value of $f_j(X)$- $g_j(Y)$ (in Tab. 6.1) the semantic distance $Sd(X,Y)$ (in Tab. 6.2), and the ratio between the absolute value of the expected value of $f_j(X)$- $g_j(Y)$ and the semantic distance $Sd(X,Y)$ (in Tab. 6.3) are represented as an example, as regards the criterion "real estate value".
In our assessment the real estate value is a Left-Right fuzzy number, expressing the possible market price per square meter of the housing property.
Data are referring to the year of the last housing census (2001) and measured in the old monetary unit (thousand of liras per square meter). As above explained, a partial correspondence is supposed to exists between quarters and market segmen-

tation, due to the partial homogeneity of urban and architectural character, together with

Table 6.1. Expected value of the difference $f_j(X)$ - $g_j(Y)$ between neighborhoods according: $f_j(X)$ is the fuzzy number "real estate value" of X and $g_j(Y)$ is the fuzzy number "real estate value" of Y .

Neigh-borhoods	Torre a Mare	Stanic	S.Spirito	S.Pasquale	S.Paolo	S.Nicola	Picone	Palese	Murat	S.Girolamo	Madonnella	Loseto	Libertà	Japigia	Ceglie	Carrassi
Carbonara	20	479	-458	-491	154	466	-825	-541	-1554	21	-362	154	-379	-341	179	-512
Carrassi	533	992	54	21	666	979	-312	-29	-1041	533	150	667	133	170	691	
Ceglie	-158	300	-637	-670	-24	287	-1004	-720	-1733	-158	-541	-24	-558	-520		
Japigia	362	820	-116	-150	495	-808	-483	-200	-1212	362	-20	495	-37			
Libertà	400	858	-79	-112	533	845	-445	-162	-1175	400	16	533				
Loseto	-133	324	-612	-645	0	312	-979	-695	-1708	-133	-516					
Madonnella	383	841	-95	-129	516	829	-462	-179	-1191	383						
S.Girolamo	0	458	-479	-512	133	445	-845	-562	-1575							
Murat	1575	2033	1095	1062	1708	2020	729	1012								
Palese	562	1020	83	50	695	1008	-283									
Picone	845	1304	366	333	979	1291										
S.Nicola	-445	12	-924	-958	-312											
S.Paolo	-133	324	-612	-645												
S.Pasquale	512	970	33													
S.Spirito	479	937														
Stanic	-458															

Table 6.2. Semantic distance revealed between neighborhoods according to the criterion "real estate value" (expressed in euro/mq)

Neighborhoods	**Torre a Mare**	**Stanic**	**S.Spirito**	**S.Pasquale**	**S.Paolo**	**S.Nicola**	**Picone**	**Palese**	**Murat**	**S.Girolamo**	**Madonnella**	**Loseto**	**Libertà**	**Japigia**	**Ceglie**	**Carrassi**
Carbonara	328	523	579	540	320	493	809	677	1582	328	464	320	395	429	316	557
Carrassi	655	1050	476	448	694	1004	531	501	1052	655	458	694	383	396	744	
Ceglie	348	326	742	695	270	329	1063	826	1762	348	595	270	558	583		
Japigia	500	860	450	420	543	833	572	489	1237	500	396	543	313			
Libertà	482	848	443	432	524	862	618	472	1224	482	404	524				
Loseto	339	401	722	648	310	383	979	711	1705	339	550					
Madonnella	533	866	534	486	569	871	660	558	1219	533						
S.Girolamo	384	471	644	611	333	488	943	724	1620							
Murat	1620	2080	1105	1132	1706	2035	785	977								
Palese	731	1118	520	500	730	1049	514									
Picone	952	1378	624	611	984	1319										
S.Nicola	478	161	974	989	379											
S.Paolo	339	401	722	648												
S.Pasquale	610	986	529													
S.Spirito	659	1023														
Stanic	478															

Table 6.3. Ratio between the expected value of the difference $f_j(X)$ - $g_j(Y)$ and their semantic distance according to the criterion "real estate value"

Neigh-bour-hoods	Torre a Mare	Stanic	S.Spirito	S.Pasquale	S.Paolo	S.Nicola	Picone	Palese	Murat	S.Girolamo	Madonnella	Loseto	Libertà	Japigia	Ceglie	Carrassi
Carbonara	0,06	0,92	0,79	0,91	0,48	0,95	1,02	0,80	0,98	0,06	0,78	0,48	0,96	0,79	0,57	0,92
Carrassi	0,81	0,94	0,11	0,05	0,96	0,98	0,59	0,06	0,99	0,81	0,33	0,96	0,35	0,43	0,93	
Ceglie	0,45	0,92	0,86	0,96	0,09	0,87	0,94	0,87	0,98	0,45	0,91	0,09	1,00	0,89		
Japigia	0,72	0,95	0,26	0,36	0,91	0,97	0,84	0,41	0,98	0,72	0,05	0,91	0,12			
Libertà	0,83	1,01	0,18	0,26	1,02	0,98	0,72	0,34	0,96	0,83	0,04	1,02				
Loseto	0,39	0,81	0,85	1,00	0,00	0,81	1,00	0,98	1,00	0,39	0,94					
Madonnella	0,72	0,97	0,18	0,27	0,91	0,95	0,70	0,32	0,98	0,72						
S.Girolamo	0,00	0,97	0,74	0,84	0,40	0,91	0,90	0,78	0,97							
Murat	0,97	0,98	0,99	0,94	1,00	0,99	0,93	1,04								
Palese	0,77	0,91	0,16	0,10	0,95	0,96	0,55									
Picone	0,89	0,95	0,59	0,55	0,99	0,98										
S.Nicola	0,93	0,07	0,95	0,97	0,82											
S.Paolo	0,39	0,81	0,85	1,00												
S.Pasquale	0,84	0,98	0,06													
S.Spirito	0,73	0,92														
Stanic	0,96															

7. Results for Urban Policies

7.1 Introduction

Using the SaTScan methodology to find the hot spots of social and housing hardship, provides certain considerations for future research in the social field and for urban planning of regeneration areas, now relevant in the European Union policies agenda.
Starting from information obtained by the cluster intersection of social and housing hardship, it could be possible to obtain useful indications for planning urban regeneration policies, making decisional process more transparent and scientifically supported.

7.2 Cluster intersection of social and housing hardship to identify the "target" areas to which urban regeneration policy should be directed

It could be obvious that social difficulty and housing physical degradation present some intersections, places where there are the highest relations between social and cultural characteristics of residents and their housing conditions. In the case study there are 195 Census sections (the darkest areas) in which there is a critical situation in terms of social and housing conditions, as shown in Figure 7.1. In the city of Bari there are three areas where urban poverty could be measured in both of its two principal facets of social difficulty and physical degradation. They are the old centre of the city, called "San Nicola", and the two peripheral areas "Libertà" and "Carbonara".

As represented in Figure 7.1, the emerging areas are very few. In these areas the urban poverty level measured in terms of social and housing difficulty is the highest. The representation is clear and the smooth gradualism shown in Figures 5.1 and 5.2, mixed in Figure 7.1, expresses fairly small differences between different clusters, all characterised by a high level of social and housing difficulty (as shown by the mean inside values of Tables 5.1 and 5.2).

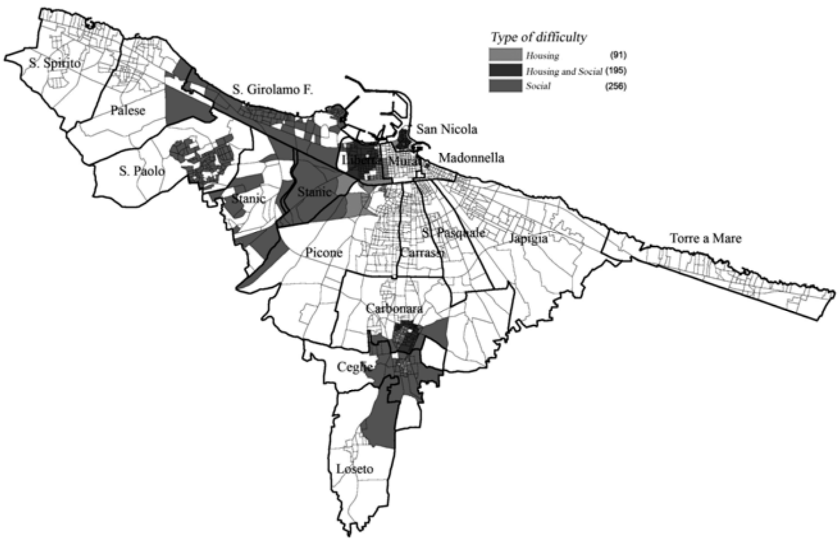

Fig 7.1. Geographical distribution of social and housing difficulty hot spots, in the city of Bari. The darkest Census sections represent the Cluster intersection of social and housing hardship

Comparing results with previous works (Perchinunno et al. 2008), where we used the same data but we have applied a fuzzy model to evaluate the urban poverty indexes distributed on the Census sections, the benefits of using the SaTScan model appear immediately. In fact, comparing the social and housing difficulty distribution obtained in the previous work with the same representation shown in Figures 5.1 and 5.2, we can see a better “resolution” of the urban poverty image achieved in a simpler way, highlighting immediately the more significant areas interested by the phenomenon.

7.3 Directions for urban regeneration policies

The preliminary question leading towards the identification of town planning and architectonic solutions to the problem of urban regeneration, at a time when there is a lack of public resources for investment, focuses on the identification of areas with the highest level of urban poverty in order to direct the choices of political decision-makers in a transparent, well thought-out and objective manner.
The authors believe that the model used in this study is able to provide relevant data for the identification of such areas.
The present study provides certain considerations for the future. The first stems from the importance of in-depth research based on methods which privilege groups of indicators of a limited number, as demonstrated above. The effectiveness of such a method is to some degree demonstrated by the specific case which can only lead to the wish to widen the investigation. In future studies these indicators could be integrated with further relational elements of, for example, the availability of social services provision by the city council to families in need or the number of requests for support in rent payments for those most at need. This type of data is, however, available at a different level of aggregation from that analyzed here, being available by street and house number of requesting families rather than from census sections and, therefore would need to be made homogenous with data prior to use.
The second consideration regards the possibility of using this model as a form of evaluation “ex post” of the effectiveness of urban policy, to verify the consequences of urban regeneration on areas characterized by high levels of poverty, examining the nature of the variations measured according to the same indicators , and monitoring the housing property market.

References

Aldstadt J, Getis A (2006) Using AMOEBA to create spatial weights matrix and identify spatial clusters. Geographical Analysis, 38: 327—343

Betti G, Cheli B, Lemmi A (2002) Studi sulla povertà. Franco Angeli, Milano

Betti G, Cheli B, Lemmi A (2002) Occupazione e condizioni di vita su uno pseudo panel italiano: primi risultati, avanzamenti e proposte metodologiche. Working paper n. 17, Dip. di Scienze Statistiche, Padova.

Cerioli A, Zani S (1990) A Fuzzy Approach to the Measurement of Poverty. In: Dugum C, Zenga M (eds.) Income and Wealth Distribution, inequality and Poverty. Springer Verlag, Berlin.

Cheli B, Lemmi A (1995) A "Totally" Fuzzy and Relative Approach to the Multidimensional Analysis of Poverty. Economic Notes

Desai ME, Shah A (1988) An econometric approach to the measurement of poverty. Oxford Economic Paper vol 40, n° 3: 505-522

Dubois D, Prade H (1980) Fuzzy sets and systems. Academic Press, San Diego CA

Gailly B, Hausman P (1984) Désavantages relatifs à une mesure objective de la pauvreté. In: Sarpellon G (ed.) Understanding poverty. Franco Angeli, Milano

Gerundo R (2000) I Programmi Urbani Complessi. Graffiti, Napoli

Knox EG (1989) Detection of clusters. In: Elliott P (Ed) Methodology of enquiries into disease clustering. London: Small Area Health Statistics Unit: 17-20

Kulldorff M (1997) A spatial scan statistic. Statistics: Theory and Methods, n.26: 1481-1496.

Kulldorff M, Nagarwalla N (1995) Spatial disease clusters: detection and inference. Statistics in Medicine, n.14: 799-810

Lemmi A, Pannuzi N (1995) Continuità e discontinuità nei processi demografici. L'Italia nella transizione demografica, 4. Rubettino, Arcavacata di Rende: 211-228

Lemmi A, Pannuzi N, Mazzolli B, Cheli B, Betti G (1997) Misure di povertà multidimensionali e relative: il caso dell'Italia nella prima metà degli anni '90,. In: Quintano C (ed) Scritti di Statistica Economica, 3, Istituto di Statistica e Matematica, Istituto Universitario Navale di Napoli. Quaderni di Discussione, n. 13, Curto, Napoli: 263-319.

Luce RD (1956) Semiorders and a theory of utility discrimination, Econometrica, 24:178-191.

Mack JE, Lansley S (1985) Poor Britain. Allen and Unwin, London

Montrone S, Perchinunno P, Torre CM (2007) Modelli per la localizzazione della povertà urbana in ambito metropolitano: il caso della città di Bari. In: Annali del Dipartimento di Scienze Statistiche "Carlo Cecchi", Facoltà di Economia, Università degli Studi di Bari, n. 6 Tomo primo: 139-169, Cacucci Editore, Bari

Munda G (1995) Multicriteria evaluaiton in a fuzzy environment. Phisica-Verlag, The Hague.

Munda G (2007) Social multi-criterial evaluation for a sustainable Economy. Springer Verlag, Heidelberg

Patil GP, Taillie C (2004) Upper level set scan statistic for detecting arbitrarily shaped hotspots. Environmental and Ecological Statistics, n. 11: 183-197.

Perchinunno P, Rotondo F, Torre CM (2008) A multivariate fuzzy analysis for the regeneration of urban poverty areas. Lecture Notes in Computer Science, Vol. 5072: pp. 137-152

Poterba J (1984) Tax Subsidies to Owner-occupied Housing: An Asset Market Approach. Quarterly Journal of Economics 99: 729-52

Rotondo F, Selicato F (2006) Periferie meridiane: laboratori del nuovo progetto urbano. In: Poceedings of Conference: Territori e città del Mezzogiorno. "Quante periferie? Quali politiche di governo del territorio", 22/23 marzo 2006, Napoli, San Giovanni a Peduccio. http:\\www.inu.it. Accessed 28 June 2008

Roy B (1985) Méthodologie multicritere d' aide à la decision. Economica, Paris.

Smith N, Williams P (2006) Gentrification of the City. Routledge, New York

Takahashi K, Yokoyama T, Tango T (2004) FleXScan: Software for the flexible spatial scan statistic. National Institute of Public Health, Japan.

Towsend P (1979) Poverty in The United Kingdom. Penguin, Harmondsworth

Zadeh LA (1965) Fuzzy sets. Information and Control., 8(3): 338-353