# *AN INDEX OF MATERIAL DEPRIVATION FOR GEOGRAPHICAL AREAS*

*A. Testi, E. Ivaldi, A. Busi*

*n°23, Giugno 2004*

***Key words**: health, deprivation, data analysis*

***J.E.L. Classification**: I 10, C40, C81*

*Summary*

Deprivation indexes are widely used in public health care both in epidemiological analyses and in allocation of resources [Dawey Smith et al, 2001].

In the first part this paper examines, in chronological order, some well-known and commonly used deprivation indexes, i.e. the ones proposed by Jarman, Carstairs, Townsend, MATDEP and SOCDEP the Index of Multiple Deprivation (IMD 2000) developed by the U.K. Department of the Environment, Transport and the Regions and, finally, some Italian proposals. The second part presents a new index of material deprivation which has been applied to “small areas” of the city of Genoa.

## Part 1: A review of the literature

Studies on deprivation are related to the analysis of the state of disadvantage suffered by an individual, concerning the living condition of the community where the individual lives. They are based on the assumption that, according to equity's considerations, more deprived areas need a larger amount of resources [Judge and Mays 1994a 1994b*;* Buckingam and Freeman 1997].

Deprivation underlines the characteristics of a group (that is, the population living in a defined geographical area) by measuring similarities and dissimilarities of the individuals forming that group. Similarities may affect not only material conditions but also social and cultural ones.

Deprivation indexes are defined referring to a small, well specified geographical area, where they "measure the proportion of households…with a combination of circumstances indicating low living standards, or a high need for services, or both" [Bartley and Blane, 1994]. Consequently, the deprivation indexes point out the close relationship existing between deprivation and territory. This relationship creates two problems, the first one concerning the definition of territory, the second one linked with indexes interpretation.

Firstly, with regard to the definition of territory, in literature the term "small area" is used [Bartley and Blane 1994; Carr Hill et. al 2002]. However, this definition is not completely clear and unambiguous; in UK indexes are commonly calculated using electoral wards (in Scotland "pseudo-postcode sector"), consisting of about 2000 households (in 1991)[1] but there are also proposals for using enumeration district (smallest census' geographical unit), consisting of about 200 households [Crayford et al, 1995].

Secondly, the use of measures based on geographic areas rather than individual conditions causes the implicit assumption of equality between people living in the same area. Attention must be paid on interpreting the results because "not all deprived people live in deprived wards, just as not everybody in a ward ranked as deprived are deprived themselves" [Townsend et al, 1988]. Though this is a relevant issue, it is impossibile to obtain individual measures of deprivation. Consequently, indexes for small area should identify an

[1] http://census.ac.uk/cdu/Datasets/Census_glossary.htm

"environmental" component underlying social differences like a proxy of unknown individual characteristics. This hypothesis requires that the geographical area is quite small [Cadum et al, 1999].

### ***1.1.*** *Jarman Underprivileged Area Score (UPA 8)*

The Underprivileged Area Score 8 (UPA 8), proposed by Jarman in 1983, was not originally constructed to measure deprivation but to take account of the social factors, derived from census data, that affect the workload of General Practitioners in England and Wales [Jarman *a*]. Since 1988, it was used by the Department of Health and Social Security in the review of RAWP (resource allocation working party) formula, in order to make additional payments to General Practitioners for each patient living in a deprived area. At 1995 the Department of Health set three bands of deprivation payment corresponding to Jarman scores of 30, 40 and 50. Each patient residing in an electoral ward with a Jarman score between 30 and 40 attracted an additional annual payment of £5.85, which increased to £10.20 for each patient residing in an electoral ward with a Jarman score of 50 or more [Crayford et al, 1995]. UPA 8 was also employed in the analysis of admission's rates to mental illness hospital [Carstairs and Morris 1991].

The eight variables comprised in the index are:

1. $X_1$ = pensioners living alone as a percentage of all residents in households
2. $X_2$ = children aged under five years as a percentage of all residents in households
3. $X_3$ = people in households of one person over 16 and one or more children as a percentage of all residents in households

4. $X_4$ = people in households headed by a person in socioeconomic group 5 (unskilled manual workers)[2] as a percentage of all residents in households
5. $X_5$ = people aged 16 or more unemployed as a percentage of economically active adults
6. $X_6$ = people in households living at more than 1 person for room as a percentage of all residents in households
7. $X_7$ = people aged 1 or over with a usual address one year before the census different from the present usual address as a percentage of all residents in households
8. $X_8$ = people in households headed by a person born in the New Commonwealth as a percentage of all residents in households

Note that selected variables, originally derived from 1981 census data, could be divided into two different categories, because they reflect both socio-economic conditions and demographic subgroups.

The index could be expressed as follows:

Let

$$T_1 = \arcsin\sqrt{\frac{X_1}{100}},\ T_2 = \arcsin\sqrt{\frac{X_2}{100}},\ \ldots,\ T_8 = \arcsin\sqrt{\frac{X_8}{100}}$$

and

$$z_1 = \frac{t_1 - \mu_{T_1}}{s_{T_1}},\ z_2 = \frac{t_2 - \mu_{T_2}}{s_{T_2}},\ \ldots,\ z_8 = \frac{t_8 - \mu_{T_8}}{s_{T_8}}$$

the Jarman index is given by the weighted sum of $z_i$

$$\textbf{\textit{UPA 8}} = \sum_{i=1}^{8} z_i w_i$$

The weights $w_i$ applied (table 1), just like selected variables, emerged from a survey of general practitioners. Sample was made up of one in ten general

[2] In U.K. exists a "Registrar General's Social Scale" (renamed in1990 "Social Class based on Occupations", actually to be replaced by the "National Statistics Socio-economic Classifications") which divides people up into seven different groups according to occupations, listed below:
I Professional occupations II Managerial and technical occupations IIIN Skilled non-manual occupations IIIM Skilled manual occupations IV Partly skilled occupations V Unskilled occupations VI Armed forces. [www.hewett.norfolk.sch.uk/curric/soc/class/reg.htm]

practitioners in the U.K, which were asked to weight (on a range from zero to 9) a range of census factors according to the degree which they increased their workload or pressure on their service. The average weights were used [Jarman b].

The index uses an angular transformation (arcsine transformation) and a standardization. The arcsine transformation is given by

$$T = \arcsin\sqrt{\frac{X_i}{100}}$$

where the square root is applied to a decimal value, ranging from zero to one[3]. Firstly, this transformation achieves the normalization of the data, so they become approximately normally distributed; otherwise, you couldn't correctly proceed to standardization. In addition, this type of transformation serves the purpose of bringing back variance to a situation of homogeneity, causing variations more relevant to the extremes of the scale than to the centre [Freeman and Tukey 1950].

Standardization, that expresses each variable in terms of its mean and standard deviation in the specific area, it is used because "if this were not done then items with longer scales would have more weight than those with shorter scales in the overall score" [Bartley and Blane 1994].

The bigger is the value of overall score, the bigger the deprivation suffered by population of interest will be.

With reference to its use in analysis of admission's rates to psychiatric hospitals, this measure has been criticised by Glover. In his study commissioned by the Department of Health, which aimed to develop an index *ad hoc (the Mental Illness Needs Index)* he noted that UPA 8 explained only the 23% of the variance between districts and that two components (elderly living alone and children under five) had very poor predictive capacity [Glover 1998].

***Table 1: Variables and weights of Jarman Index***

[3] This type of transformation has been historically applied with percentages or proportions, that generally follow a binomial (and not a normal) distribution, in which the variance value (pq) depends on the average value (p) [Osborne 2002].

| Variables | Weights |
|---|---|
| $X_1$ | $w_1 = 6,62$ |
| $X_2$ | $w_2 = 4,64$ |
| $X_3$ | $w_3 = 3.01$ |
| $X_4$ | $w_4 = 3,74$ |
| $X_5$ | $w_5 = 3,34$ |
| $X_6$ | $w_6 = 2,88$ |
| $X_7$ | $w_7 = 2,68$ |
| $X_8$ | $w_8 = 2,50$ |

***1.2.*** *Townsend index of deprivation*

This measure was developed by Townsend [Townsend *et al*, 1988], under commission of the *Northern Regional Health Authority,* in order to analyze health measures within the Northern region (counties of Cleveland, Cumbria, Durham, Northumberland and Tyne and Wear [Phillimore et al, 1994]), with particular reference to inequalities in health. In accordance with author's wishes, this index is a measure of material deprivation. Material deprivation is distinguished from social form of deprivation, as Townsend himself stated: "Material deprivation entails the lack of goods, services, resources, amenities and physical environment which are customary, or at least widely approved in the society under consideration. Social deprivation, on the other hand, is non-participation in the roles, relationship, customs, functions, rights and responsibilities implied by a member of a society and its sub-groups. Such deprivation may be attributed to the affects of racism, sexism and ageism…" [Townsend *et al*, 1988].

Townsend index is commonly used also in epidemiological analysis [Carstairs and Morris 1991] and, like the UPA 8, comprises between his four indicators unemployment, suitable to summarize the lack of material resources and economic confidence, and overcrowding, variable reflecting material life's

conditions. Innovative seems to be the insertion of car and household ownership, intended as a proxy respectively of wealth and current income. In this case, the four indicators are unweighted

1. $X_1$ = percentage of economically active people unemployed
2. $X_2$ = percentage of households with more than one person per room
3. $X_3$ = percentage of households with no car
4. $X_4$ = percentage of households not owner-occupied

Townsend points out that the four selected variables, originally obtained from 1981 census and reflecting only socioeconomic circumstances, represent the state or condition of deprivation and for this reason they are called “direct indicator” of deprivation. On the contrary, “indirect indicators” of deprivation represent the victims of those conditions or states, for example, the elderly, ethnic minorities and single parents [Townsend 1987]. Townsend underlines that the belonging to these categories is not, in itself, indicator of deprivation, even if many people among these categories are really deprived. In selecting indicators you have to consider that is the form of deprivation, which has to be measure, and not status of individuals who suffers for it [Townsend *et al* 1988].
Index is calculated as follows:
Let

$$T_1 = \log(X_1 + 1);\ T_2 = \log(X_2 + 1);\ T_3 = X_3;\ T_4 = X_4$$

and

$$z_1 = \frac{t_1 - \mu_{T_1}}{s_{T_1}};\ z_2 = \frac{t_2 - \mu_{T_2}}{s_{T_2}};\ z_3 = \frac{t_3 - \mu_{T_3}}{s_{T_3}} \text{ and } z_4 = \frac{t_4 - \mu_{T_4}}{s_{T_4}}$$

being $\mu_{T_i}\ \ i = 1,\ldots,4$ and $s_{T_i}\ \ i = 1,\ldots,4$ the means and the standard deviations for the whole area of interest.
Townsend index is given by the un-weighted sum of the $z_i$:

$$\textbf{\textit{Townsend Index}} = \sum_{i=1}^{4} z_i$$

Note that the unemployment and overcrowding variables are transformed using a log-transformation, which achieves several effects. Log-transformations are in fact commonly used to stabilize variances, linearize relationship and reduce skewness. In addiction, this type of transformations allows to obtain more normal distributions, especially when a large positive asymmetry exists [Osborne 2002].
Negative values of the overall Townsend scores reflect less deprived areas, positive values reflect more deprived areas.

### *1.3.Carstairs deprivation index*

This index was constructed by Carstairs and Morris in 1991 for evaluating inequalities in health which exist within the population of Scotland. Well known also like SCOTDEP, it is very similar to Townsend index and differs from it only because contains variables more representative of Scotland situation. SCOTDEP, used also in epidemiological analysis, is based on the combination of four socio-economic variables, originally derived from 1981 census. The variables were chosen on the basis of previous works, which examined health and deprivation in the wards of Glasgow and Edinburgh. Accordingly with author's wishes, each of the indicators selected are representative or determinant of material disadvantage [Carstairs and Morris 1991].
Two variables are the same used by Townsend, unemployment is restricted to male (considering the low rate of female employment in Scotland) while being in low social class substitutes housing tenure. This choice is justified by authors observing that "being in low social class…indicates earnings at the lower end of the income scale…Housing tenure does not feature in our list of indicators since this is considered to be of lesser value in Scotland which has a higher proportion of its housing stock in the public sector and lesser variation between areas than occurs in England and Wales" [Carstairs and Morris 1991]. So they concluded that " the variable would not have acted as a discriminator between large sections of the population" [Morris and Carstairs 1991].

The variables selected are listed below:

1. $X_1$ = persons in private households living at a density of >1 person per room as a proportion of all persons in private households
2. $X_2$ = proportion of economically active males who are seeking work
3. $X_3$ = proportion of all persons in private households with head of household in social class four or five
4. $X_4$ = proportion of all persons in private households with no car

Note that "unlike a number of other measures considered all four variables are calculated on the basis of individuals not households; this is considered preferable for the purpose of the analysis of events which relate to individuals but in practice the differences from using the two approaches are likely to be small" [Carstairs and Morris 1991].

The SCOTDEP is an unweighted combination of four standardised variables:

let

$$z_1 = \frac{x_1 - \mu_{X_1}}{s_{X_1}},\ z_2 = \frac{x_2 - \mu_{X_2}}{s_{X_2}},\ z_3 = \frac{x_3 - \mu_{X_3}}{s_{X_3}} \text{ and } z_4 = \frac{x_4 - \mu_{X_4}}{s_{X_4}}$$

being $\mu_{X_i} \quad i = 1,\ldots,4$ and $s_{X_i} \quad i = 1,\ldots,4$ respectively the means and the standard deviations for the whole area of Scotland, the index is given by:

$$\textbf{\textit{SCOTDEP}} = \sum_{i=1}^{4} z_i$$

Bigger the score, greater the deprivation suffered by the area of interest.

The extended distribution of the index was subdivided into seven categories, determined on the basis of distribution standard deviation, that origin a "new" variable called DEPCAT (deprivation category). DEPCAT 7 identifies the greatest deprivation, DEPCAT 1 very affluence.

### ***1.4.*** *MATDEP and SOCDEP*

In 1993, Forrest and Gordon [Forrest and Gordon 1993] developed two different measures of deprivation. Following the distinction between material and social deprivation explicitly stated by Townsend, MATDEP is designed to measure material deprivation, SOCDEP quantifies social deprivation. On this basis all the variables included in MATDEP are "direct indicators" of deprivation, representatives of deprivation state, while those selected for SOCDEP are "indirect indicators", determining the victims of that condition. In fact the variable unemployment, although classified by Townsend as a direct indicator, represents an individual conditions so can be considered similar to indirect indicators. Both indexes use 1991 census data.

Variables included in MATDEP are:

1. $X_1$ = percentage of household with more than one person per room
2. $X_2$ = percentage of households lacking or sharing use of a bath/shower and/or inside Wc
3. $X_3$ = percentage of household with no central heating
4. $X_4$ = percentage of household with no car

Index is formulated as follows:

let $\max(X_i)$, $i = 1,\ldots,4$ the maximum values for each indicator in the whole area of interest, the overall score is given by the unweighted summation of each $x_i$ divided by his maximum, that is

$$\textbf{\textit{MATDEP}} = \sum_{i=1}^{4} \frac{x_i}{\max(X_i)}.$$

Variables included in SOCDEP follow:

1. $X_1$ = percentage of economically active population unemployed.
2. $X_2$ = percentage of economically active 16-24 year olds unemployed.
3. $X_3$ = lone parent households as a proportion of all households.
4. $X_4$ = percentage of households containing a single pensioner.

5. $X_5$ = percentage of households containing a person with limiting long-term illness.
6. $X_6$ = percentage of households containing dependants only (e.g. single pensioners with long term illness)

SOCDEP score is, like MATDEP, the summation of the unweighted standardised scores for each variable, so being $\max(X_i)$, $i = 1,\ldots,6$ the maximum values for each indicator in the whole area of interest,

$$\textbf{\textit{SOCDEP}} = \sum_{i=1}^{6} \frac{x_i}{\max(X_i)}$$

Both indexes are therefore sums of values between zero and one. Consequently, the maximum theoretical scores are four for MATDEP and six for SOCDEP, both correspondent to the greatest deprivation.

### ***1.5.*** *Index of Multiple Deprivation 2000 (IMD 2000)*

The Index of Multiple Deprivation 2000, commissioned in 1998 by the Department of the Environment, Transport and the Regions (DETR), was developed by the Department of Social Policy and Social Work at the University of Oxford in order to give information useful to local authorities [DETR 2000]. This is a ward level, innovative and detailed index, which reviews and update 1991 Index of Local Conditions (1991 ILC) and 1998 Index of Local Deprivation (1998 ILD). IMD 2000 is made up of six separate “domains” of deprivation, listed below:

1. Income
2. Employment
3. Health Deprivation and Disability
4. Education, Skills and Training
5. Housing
6. Geographical Access to Services.

The index is based on the premise that deprivation is made up of separate dimensions, each one reflecting different aspects of deprivation. For this reason an appropriate index is calculated for each dimension: indexes are made up of a number of indicators statistically robust, up to date and available at a ward level for the whole of England. Above all, each index should directly measure a major aspect of the dimension of deprivation under consideration [DETR 2000].

Methodology to calculate the index is complex, but we can summarize the main aspects as follows:

1. Income and Employment index are presented as a simple rate, so "if a ward scores 38,6 in the Income Domain, this means that 38,6% of the ward's population are income deprived,…and is it possible to say that Ward X with a score of 40% is twice deprived as Ward Y with a score of 20%" [DETR 2000]
2. the other four indexes are obtained by combining indicators using factor analysis
3. then for each index a rank of one is assigned to the most deprived ward, and a rank of 8414 is assigned to the least deprived ward.

The IMD 2000 uses 32 indicators derived from a great number of data sources[4], and between them stand out some previously untapped, such as Department of Social Security benefits data and University and Colleges Admission Service (UCAS) data. Note that most of the indicators can be updated regularly, just because they are not derived only from census data, quickly out of date [DETR 2000].

Finally, the six separated indexes are combined into an overall index, following those steps:

1. Indexes scores are transformed to a standard distribution, using exponential transformation

[4]Listed in DETR 2000, Appendix B.

2. Then indexes are summed after the weights listed below have been applied[5]
    1. Income [25%];
    2. Employment [25%];
    3. Health Deprivation and Disability [15%];
    4. Education, Skills and Training [15%];
    5. Housing [10%]
    6. Geographical Access to Services [10%];

Again, the bigger the IMD 2000 score, the more deprived the ward, and the overall index is then ranked in the same way of the Domain Indices.
Besides six summarizing measures of the IMD 2000 have been produced at district level; they focus on different aspects of multiple deprivation and give information to Local Authorities about the form of deprivation suffered by area of their interest.

### ***1.6.*** *Italian proposals*

Deprivation indexes were developed relatively late in Italy: the first attempt to describe a national deprivation index was made by Cadum and colleagues [ Cadum et al 1999] in 1999. They realized an index calculated at municipal level by using 1991 census data base. Their study was designed to analyse general mortality according to deprivation category and to measure the size of ecological bias using the Turin Longitudinal Study. They realized an index which comprises five variables, listed below:

5. $X_1$ = low education
6. $X_2$ = unemployment
7. $X_3$ = renter occupier housing
8. $X_4$ = no indoor bathroom
9. $X_5$ = lone parent with childhood

[5] Weights are applied following two criteria: firstly, Income and Employments domains should carry more weights, a position supported by academic literature; secondly it was also proposed that the most robust domains should carry the most weight [DETR 2000].

Variables were selected on the basis of the results of a factorial analysis, performed in order to identify the most valuable from a list of census indicators of inequalities.

The index is an unweighted combination of five standardised variables:

let

$$z_1 = \frac{x_1 - \mu_{X_1}}{s_{X_1}},\ z_2 = \frac{x_2 - \mu_{X_2}}{s_{X_2}},\ z_3 = \frac{x_3 - \mu_{X_3}}{s_{X_3}},\ z_4 = \frac{x_4 - \mu_{X_4}}{s_{X_4}} \text{ and}$$

$$z_5 = \frac{x_5 - \mu_{X_5}}{s_{X_5}}$$

being $\mu_{X_i} \quad i = 1,\ldots,5$ and $s_{X_i} \quad i = 1,\ldots,5$ respectively the means and the standard deviations for the whole area of Italy, the index is given by:

$$\textbf{\textit{Index}} = \sum_{i=1}^{5} z_i$$

This measure has been calculated for all the 8100 Italian municipalities present at 1991 census and has proved that increasing deprivation is significantly associated with mortality from all causes.

Other measures of deprivation were although developed at local level, for examples the IDS/IDM [Biggeri et al 1998] and the IAS [Valerio Vitullo 2000]. The IDS/IDM was developed in order to evaluate differentials in mortality which exist between north-west areas and the rest of Tuscany considering socio-economical differences.

The IDM is a material deprivation index which uses variables derived from 1961,1971,1981 and 1991 census. The variables aren't defined in an unambiguous manner because census were changed by passing time; although they refer to education level, income, lacking of private facilities and household ownership. The IDS uses the same variables included in IMD with the addition of lone parent households. IDM/IDS are calculated at municipal level, and they are given by the sum of 4/5 z-scores.

The IAS was proposed in 2000 [Valerio e Vitullo 2000] with the objective of evaluate the relationship between health state and social economical circumstances and guide allocation of resources.

IAS comprised five variables (ageing, unemployment, low education level, lone parent households and lacking of private facilities) and has been calculated at municipal level within Basilicata, using the z-score method.
Innovative seems to be the validation of the scores, realised by measuring IAS correlation with taxable income

### *1.7. A comparison of deprivation indexes*

Table 2 summarizes the indexes reviewed, with the exception of IMD 2000 (excluded for his particular typology):

***Table 2: Indexes of deprivation***

| | **Jarman (UPA 8)** | **Townsend** | **Carstairs** | **MATDEP** | **SOCDEP** |
|---|---|---|---|---|---|
| **Primary application** | Allocating resources to GP's | Analysis of inequalities in health | Analysis of inequalities in health | | |
| **Further application** | Analysis admission rates to psychiatric hospitals | Epidemiological analysis | Epidemiological analysis | | |
| **Purpose(Y)** | GP's workload | Material deprivation | Material deprivation | Material deprivation | Social deprivation |
| **Independent variables (X)** | | | | | |
| $X_1$ Elderly alone | √ | | | | √ |
| $x_2$ Children under 5 | √ | | | | |
| $x_3$ Lone parent households | √ | | | | √ |
| $X_4$ Households social class 5 | √ | | | | |
| $x_5$ Unemployment | √ | √ | √ (Male) | √ | √<br>√ (Youth) |
| $X_6$ Overcrowding | √ | √ | √ | √ | |
| $x_7$ Moved | √ | | | | |
| $x_8$ Ethnic minorities | √ | | | | |
| $x_9$ Households no car | | √ | √ | √) | |
| $x_{10}$ Housing tenure | | √ | | | |

| | | | | | |
|---|---|---|---|---|---|
| $x_{11}$ Households social class 4 o 5 | | | √ | | |
| $x_{12}$ Households no amenity | | | | √ | |
| $x_{13}$ Households no central heating | | | | √ | |
| $x_{14}$ Households limiting long-term illness | | | | | √ |
| $x_{15}$ Households dependants only | | | | | √ |
| **Index formulation** | Additive | Additive | Additive | Additive | Additive |
| **Weights** | Yes | No | No | No | No |
| **Transformation used** | $\arcsin\sqrt{\frac{X}{100}}$ | Log of 2 x | None | None | None |
| **Standardization** | Yes | Yes | Yes | No | No |
| **Normalization** | No | No | No | Yes | Yes |
| **Data sources** | Census | Census | Census | Census | Census |

These indexes present two common characteristics:

1. they are made up of variables which identify a condition of disadvantage
2. their formulation is additive

All the selected variables appear to be correct in describing a disadvantage state. The question is to establish the methodological reasons underlying their selection.

Given that the problem of data availability limits the choice to indicators available from census, we have to remember that all indexes are constructed leaving aside a universally accepted definition of the dependent variable which has to be measured.

For this reason while Jarman (that however was not primary interested in measuring deprivation) solves this problem adopting a consensual approach, Townsend links the needs for primary cares to a concept of material deprivation, that himself defines theoretically and distinguishes from social deprivation, making a distinction also between direct and indirect indicators of deprivation.

Carstairs, Forrest and Gordon follow the line traced by Townsend, but differences which they introduce point out on one hand the need of take “face validity” into account (as we have already seen this is Carstairs’s reason for considering only male unemployment and for excluding housing tenure), on the other hand mark non-existence of a well defined theoretical reference’s basis (for example remember the double meaning of unemployment, inserted in SOCDEP).

In the choice of indicators, further limitation is due to statistical considerations, so highly inter-correlated variables and indicators with poor predictive capacity have to be excluded.

With reference to variables concretely used, note that although indexes could be re-calculated by use more recent census data, exogenous changes occurred with the passing of time could diminish some variables significance (e.g. nowadays car ownership seems to be less relevant as a proxy for income).

Weights are applied only to Jarman index (again on the basis of a consensual approach), while in other indices the same weight is given to each indicators, probably because doesn’t seem possible to evaluate accurately influences that different determinants have on deprivation.

The use of transformations meets statistical needs, although has been criticized (in the case of Jarman) because “obscures the original policy intent” [Carr Hill and Sheldon 1991].

Variables are standardized or normalized: first solution seems to be preferable, just because it is possible to give an unambiguous meaning to zero value, which indicates an average situation of deprivation.

# Part 2: An index of deprivation for Genoa: a proposal

## *2.1. Materials and methods*

The aim of this paper is to propose an index of deprivation following the British experiences. It was decided to develop a material deprivation index, which represents a neutral objective measure, independent from the consequences of the individual standard of life. Thus among the indexes presented above, the ones of Towsend and Carstairs were taken into consideration. They do not appear, however, completely appropriate in describing deprivation in today's society. Some variables are obsolete, for instance "car ownership", some are not meaningful in Italian reality, such as "social class". Consequently, a new index is proposed.

First of all, the definition of a small geographic unit in which the deprivation index will be calculated was required. This unit is commonly called "small area".

In this work the "Unità Urbanistica" (UU) has been chosen: UU represents the smallest subdivision suitable for the purpose, which is identifiable within the city. Larger areas do not ensure sufficient homogeneity, whereas for the smaller ones only some data is available. UUs mainly reflect historical origin when many independent towns existed around the old city (and in 1926 they were included in the same metropolitan area).

Nowadays the area of Genoa is subdivided into 71 UUs. The average population of UU is 9557 individuals, with a minimum of 1143 and a maximum of 30981 individuals (1991 Census).

Secondly, the variables included in the index must be chosen. The choice is strictly related to the concept of material deprivation and the availability of reliable and updated data.

The variables should be, as said above, "direct" measures of deprivation [Townsend et. al.1988]. Variables reflecting demographic sub-groups, such as ethnic minorities and elderly people living alone, were, therefore, excluded.

The chosen variables are the following: i) unemployment, ii) housing ownership, iii) overcrowding, iv) low education level.

Unemployment represents a state of economical insecurity and lack of resources; housing ownership could be intended as a proxy for wealth, while overcrowding has been inserted for its potential capacity to synthesize living conditions.

The fourth variable, "low education level", is used as a proxy to represent social position and indirectly it gives information about current income, as national studies on deprivation [Cadum et al 1999, Parodi et al 2003] suggest.

The choice of the variables has been limited to the ones derivable from census data, updated at a municipal level when possible. Consequently, if some variables, though relevant, were not available at UU level, they were excluded.

The four variables selected were quantified as indicated below (1991 census)[6]:

1. $X_1$ = percentage of economically active people unemployed (Unemployment)
2. $X_2$ = percentage of households not owner occupied (Housing ownership)
3. $X_3$ = average number of occupants per house (Overcrowding)
4. $X_4$ = percentage of people with secondary or lower level study certificate (Low education level)

An additional index was computed, summing the partial indicators referring to each variable selected.

Differences in variables scales indicate the use of standardization. Table 3, which lists in the second column variables value for UU Crevari, in the third Genoa mean and in the fourth Genoa standard deviation, shows that variables are not homogeneous. For this reason they could not be combined together as they are. Standardization, accordingly to several experiences [Jarman, Townsend et al, Carstairs and Morris, Cadum et al], seems to be a better solution.

[6] Overcrowding is derived from a municipal updating (2000).

Because of non-normality of original distributions[7], the use of a transformation capable to achieve an approximate normal distribution for each variable is required. To find an appropriate transformation the Box Cox method [BOX and COX 1964] was used.

This method turns to a family of power transformations given by

$$x(\lambda) = \frac{(x^{\lambda} - 1)}{\lambda} \qquad \lambda \neq 0$$

$$x(\lambda) = \ln(x) \qquad \lambda = 0$$

and it plans, in order to select the value of the parameter $\lambda$, to use that value which, given an observations vector $\mathrm{x} = x_1, x_2, \ldots, x_n$ , maximizes the logarithm of the likelihood function

$$f(x, \lambda) = -\frac{n}{2}\ln(\sigma^2_{X(\lambda)}) + (\lambda - 1)\sum_{i=1}^{n}\ln(x_i)\,.$$

***Table 3: Example of deprivation score calculation for the UU Crevari***

| Variable | Crevari's values | Mean of Genoa | S.D. of Genoa | Z-score |
|---|---|---|---|---|
| Overcrowding | 2,20 | 2,29 | 0,11 | 0,32 |
| Unemployment | 7,11 | 5,29 | 1,69 | 1,00 |
| Low education level | 78,34 | 67,56 | 11,78 | 1,07 |
| Housing ownership | 37,26 | 37,48 | 7,77 | -0,71 |
| **Index score** | | | | 1,69 |

After variables have been transformed, z-scores were calculated for each observation. They are obtained by subtracting the mean of Genoa from the observed transformed value and dividing the result for the standard deviation of Genoa.

[7] Non normality is substantially due to two factor, that is the nature of data and the number of observations. In fact, percentages, that generally follow not a normal but a binomial distribution, were used. In addition only 71 observations were considered.

The index score for the UU Crevari is given by the unweighted combination of four z-scores, as showed in the fifth column of table 3..

### ***2.2*** *Results and discussion*

Index has been calculated for all UUs within the city of Genoa as explained in the previous paragraph.

The index scores range from -4,41 to 8,49. The distribution does not exhibit absolute normality, because a positive asymmetry exists (skewness is 0,55). For this reason, the population is not distributed equally on both sides of the mean: 60% is found below and 40% above the mean.

Index values have been ranked: a rank of 1 was assigned to the most affluent UU and a rank of 71 to the most deprived.

In order to better distinguish between different levels of deprivation, the index distribution has been divided into six classes determined on the basis of standard deviation. Class 1 (very affluent) identifies less deprived UUs; on the contrary, Class 6 (great deprivation) contains UUs characterized by strong deprivation.

Table 4 shows definition of classes and boundaries. Note that the classes are unequal; in fact they were not designed to ensure equality of numbers within classes but to retain the discriminatory features of the distribution. Other methods of determining the boundaries of classes (e.g. septiles or quintiles) would effect no substantial changes in general patterns observed.

***Table 4: Class boundaries and definition***

| Class number | Definition | Boundaries |
|---|---|---|
| 1 | Very affluent | (-4,41;-σ) |
| 2 | Affluent | (-σ;-σ/2) |
| 3 | Not deprived | (-σ/2; 0) |
| 4 | Low deprivation | (0;+σ/2) |
| 5 | Middle deprivation | (+σ/2; σ) |
| 6 | High deprivation | (σ; +8,49) |

According to the above mentioned criteria, in Table 5 results for each UU are presented: index score, ranking and class number. The second column lists the UU Census Identification number.

***Table 5: Deprivation scores and deprivation classes***

| UU | ID | Index score | Rank | Class |
|---|---|---|---|---|
| **CASTAGNA** | 69 | -4,41 | ***1*** | **1** |
| **FOCE** | 41 | -4,26 | ***2*** | **1** |
| **QUINTO** | 70 | -4,19 | ***3*** | **1** |
| **LIDO** | 64 | -4,01 | ***4*** | **1** |
| **PUGGIA** | 65 | -4,01 | ***5*** | **1** |
| **S.VINCENZO** | 39 | -4,00 | ***6*** | **1** |
| **ALBARO** | 62 | -3,52 | ***7*** | **1** |
| **QUARTARA** | 68 | -3,44 | ***8*** | **1** |
| **S.BARTOLOMEO** | 28 | -3,31 | ***9*** | **1** |
| **QUARTO** | 67 | -3,14 | ***10*** | **1** |
| **STURLA** | 66 | -3,08 | ***11*** | **1** |
| **S.MARTINO** | 61 | -3,02 | ***12*** | **1** |
| **S.GIULIANO** | 63 | -2,99 | ***13*** | **1** |
| **PEGLI** | 7 | -2,80 | ***14*** | **1** |
| **BELVEDERE** | 27 | -2,79 | ***15*** | **1** |
| **MANIN** | 38 | -2,48 | ***16*** | **2** |
| **BRIGNOLE** | 42 | -2,19 | ***17*** | **2** |
| **NERVI** | 71 | -2,16 | ***18*** | **2** |
| **CASTELLETTO** | 37 | -2,14 | ***19*** | **2** |
| **CARIGNANO** | 40 | -2,13 | ***20*** | **2** |
| **S.NICOLA** | 36 | -2,01 | ***21*** | **2** |
| **MULTEDO** | 8 | -1,46 | ***22*** | **2** |
| **CALCINARA** | 11 | -1,45 | ***23*** | **2** |
| **SESTRI** | 9 | -1,36 | ***24*** | **3** |
| **CHIAPPETO** | 60 | -1,32 | ***25*** | **3** |
| **S.AGATA** | 43 | -1,32 | ***26*** | **3** |
| **MARASSI** | 47 | -1,27 | ***27*** | **3** |
| **FORTE QUEZZI** | 48 | -1,20 | ***28*** | **3** |
| **PALMARO** | 4 | -0,78 | ***29*** | **3** |
| **S.TEODORO** | 30 | -0,71 | ***30*** | **3** |
| **S.FRUTTUOSO** | 44 | -0,70 | ***31*** | **3** |
| **OREGINA** | 32 | -0,68 | ***32*** | **3** |
| **APPARIZIONE** | 58 | -0,68 | ***33*** | **3** |
| **RIVAROLO** | 15 | -0,49 | ***34*** | **3** |
| **BORGORATTI** | 59 | -0,44 | ***35*** | **3** |
| **S.GAETANO** | 25 | -0,40 | ***36*** | **3** |
| **S.G.BATTISTA** | 10 | -0,19 | ***37*** | **3** |
| **FEREGGIANO** | 46 | -0,12 | ***38*** | **3** |
| **CERTOSA** | 14 | 0,20 | ***39*** | **4** |
| **S.PANTALEO** | 50 | 0,49 | ***40*** | **4** |

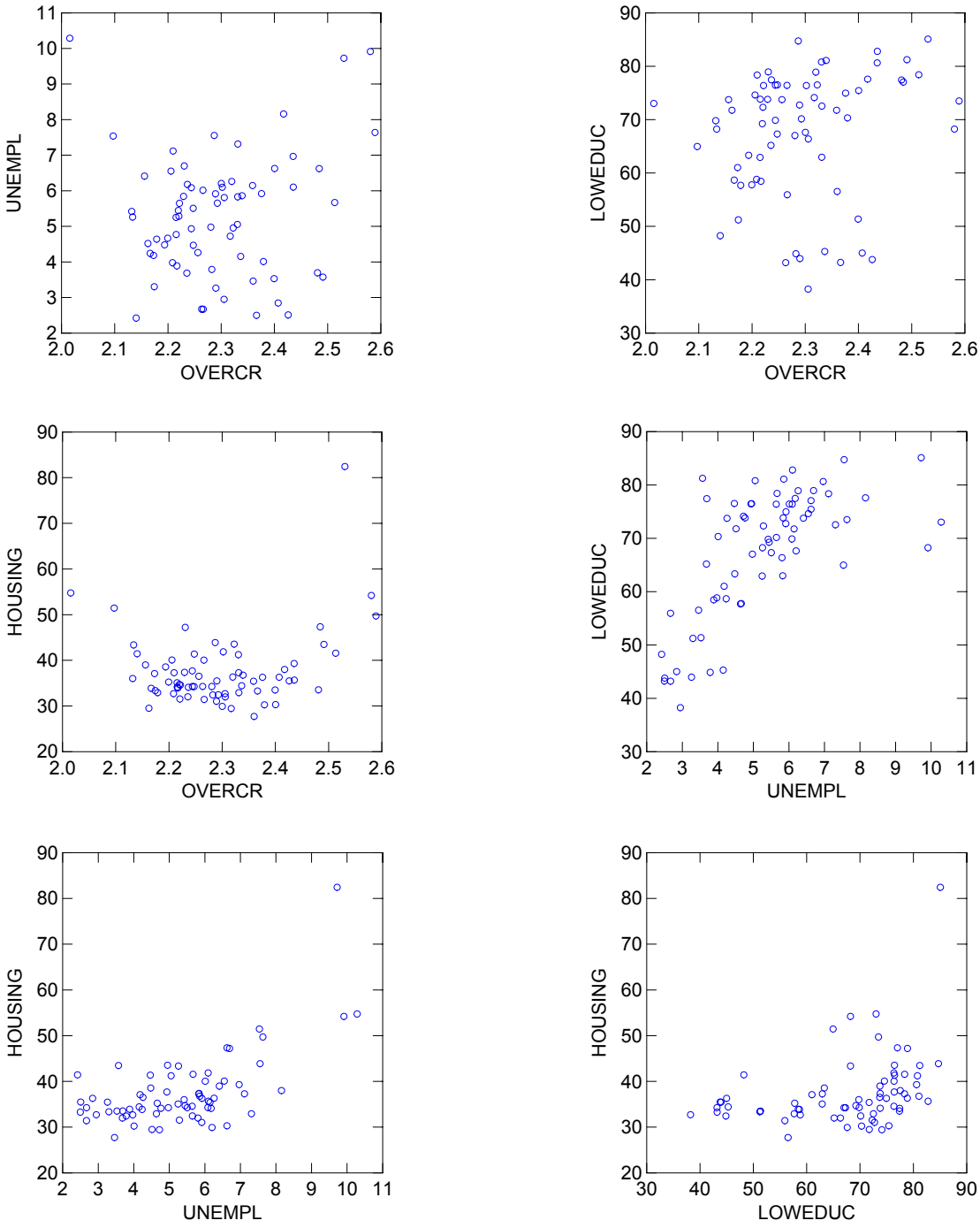
| | | | | |
|---|---|---|---|---|
| **CASTELLUCCIO** | 6 | 0,52 | ***41*** | 4 |
| **MADDALENA** | 34 | 0,63 | ***42*** | 4 |
| **PONTEDECIMO** | 21 | 0,64 | ***43*** | 4 |
| **CAMPASSO** | 24 | 0,66 | ***44*** | 4 |
| **ANGELI** | 29 | 0,67 | ***45*** | 4 |
| **BAVARI** | 56 | 0,84 | ***46*** | 4 |
| **VOLTRI** | 2 | 0,92 | ***47*** | 4 |
| **BORZOLI OVEST** | 12 | 0,99 | ***48*** | 4 |
| **LAGACCIO** | 31 | 1,21 | ***49*** | 4 |
| **PARENZO** | 49 | 1,33 | ***50*** | 4 |
| **SAMPIERDARENA** | 26 | 1,38 | ***51*** | 4 |
| **CREVARI** | 1 | 1,69 | ***52*** | 5 |
| **MOLO** | 35 | 1,82 | ***53*** | 5 |
| **PRA'** | 5 | 1,83 | ***54*** | 5 |
| **QUEZZI** | 45 | 1,94 | ***55*** | 5 |
| **MOLASSANA** | 53 | 2,13 | ***56*** | 5 |
| **PRATO** | 55 | 2,15 | ***57*** | 5 |
| **MONTESIGNANO** | 51 | 2,42 | ***58*** | 5 |
| **BORZOLI EST** | 13 | 2,47 | ***59*** | 5 |
| **TEGLIA** | 16 | 2,56 | ***60*** | 5 |
| **BOLZANETO** | 18 | 2,88 | ***61*** | 6 |
| **MOREGO** | 19 | 3,12 | ***62*** | 6 |
| **CAMPI** | 23 | 3,38 | ***63*** | 6 |
| **DORIA** | 54 | 4,03 | ***64*** | 6 |
| **S.DESIDERIO** | 57 | 4,04 | ***65*** | 6 |
| **CORNIGLIANO** | 22 | 4,19 | ***66*** | 6 |
| **S.QUIRICO** | 20 | 4,38 | ***67*** | 6 |
| **BEGATO** | 17 | 4,71 | ***68*** | 6 |
| **S.EUSEBIO** | 52 | 5,65 | ***69*** | 6 |
| **PRE'** | 33 | 6,31 | ***70*** | 6 |
| **CA' NUOVA** | 3 | 8,49 | ***71*** | 6 |

From the methodological point of view, the index distribution for 71 UUs has a mean of zero and a standard deviation of 2,79. This value of the standard deviation is due to the existence of correlation between selected variables, as table 6 shows. In particular, a quite strong linear relationship exists between low education level and housing ownership.

***Table 6: Correlation matrix***

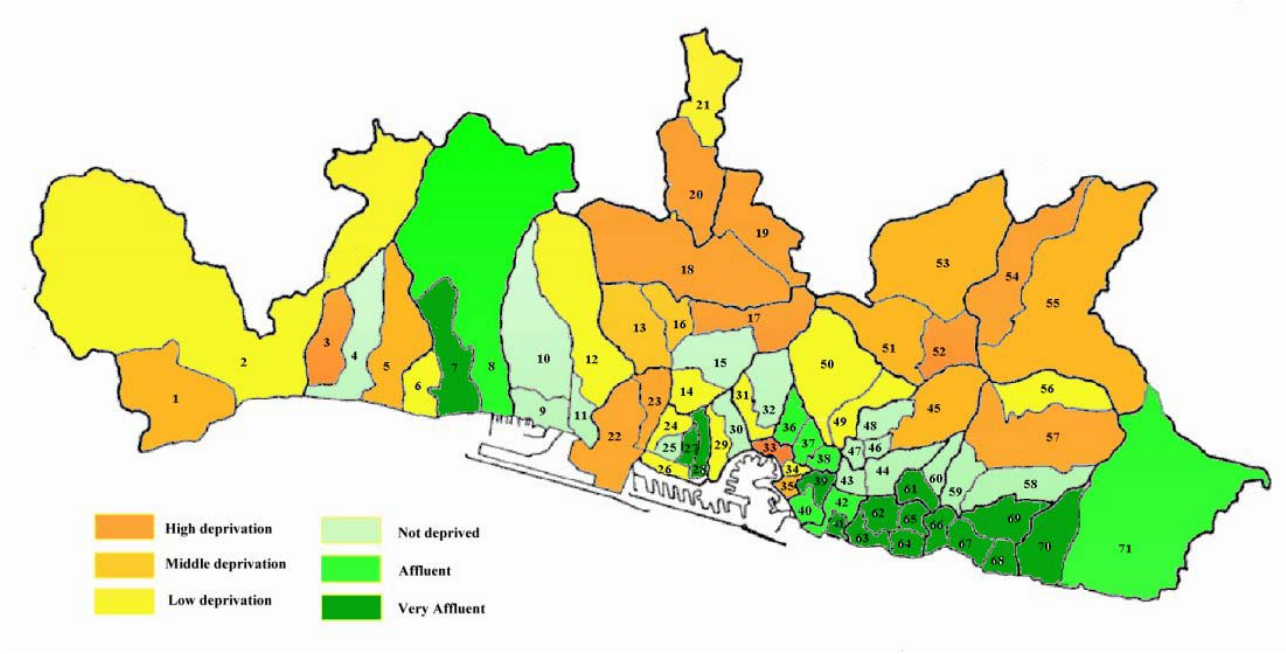
| | Overcrowding | Unemployment | Low education level | Housing ownership |
|---|---|---|---|---|
| Overcrowding | 1 | | | |
| Unemployment | 0,13 | 1 | | |
| Low education level | 0,16 | 0,66 | 1 | |
| Housing ownership | 0,22 | 0,60 | 0,34 | 1 |

***Figure 1: Correlation plots***




Finally, a deprivation map of the city of Genoa, representing UUs (identified by their ID number), is obtained (Figure 2).

*Figure 2: Deprivation classes in Genoa*




According to what may be considered as “face validity”, the most affluent areas are concentrated in the southeast nearby the sea, while the most deprived UUs are in the industrial or ex-industrial areas where are dislocated many council houses.

To use the index in the resource allocation a rigorous validation was required. Following the past experiences [e.g. Cadum et al., 1999] correlation between index and mortality was studied. Data collected by Dott.ssa Vercelli’s équipe (Istituto Tumori Genoa) for the years 1998-2000 were used. *Standardised Mortality Ratios* (*SMR*) for all causes were calculated for each UU. *SMR* measure relative risk, that is the probability of death for people living in a specified UU related to the same probability calculated in a reference group, which is the city of Genoa.

*SMR* could be interpreted like a measure of increase/decrease of the probability of death with respect to the reference group. They have been calculated for people aged 0-64 and over 65.

The results are substantially encouraging because they confirm index effectiveness in explaining variations of the socioeconomic circumstances within

Genoa. The predictor role of the socioeconomic circumstances with respect to variations in mortality patterns seems verified.

Table 7 shows that this happens particularly for the people aged 0-64. The association between deprivation classes and *SMR* is very strong: the relative risk decreases by 16% if you live in a "very affluent" zone, while increase by 23% if you live in a "very deprived" zone.

***Table 7: Correlation between mortality and deprivation***

| ***Deprivation class*** | ***Observed deaths*** | ***Expected deaths*** | ***SMR*** |
|---|---|---|---|
| *Male-Female 0-64 anni* | | | |
| 1 Zona ricca | 685 | 812,95 | 0,84 |
| 2 Benestante | 423 | 490,20 | 0,86 |
| 3 Non deprivata | 1118 | 1121,74 | 0,99 |
| 4 Bassa deprivazione | 741 | 738,89 | 1,002 |
| 5 Media deprivazione | 485 | 466,64 | 1,04 |
| 6 Alta deprivazione | 544 | 438,90 | 1,23 |
| | | | |
| *Male-Female 65-ω* | | | |
| 1 Zona ricca | 4698 | 4919,67 | 0,95 |
| 2 Benestante | 2928 | 2708,78 | 1,08 |
| 3 Non deprivata | 6231 | 6474,51 | 0,96 |
| 4 Bassa deprivazione | 3715 | 3831,18 | 0,96 |
| 5 Media deprivazione | 1997 | 2115,27 | 0,94 |
| 6 Alta deprivazione | 2227 | 1869,87 | 1,19 |

Figure 3 and 4 show correlation between index values and *SMR* calculated at UU and Circoscrizione level respectively.

***Figure 3 –SMR versus deprivation index, UU level (71 observations)***

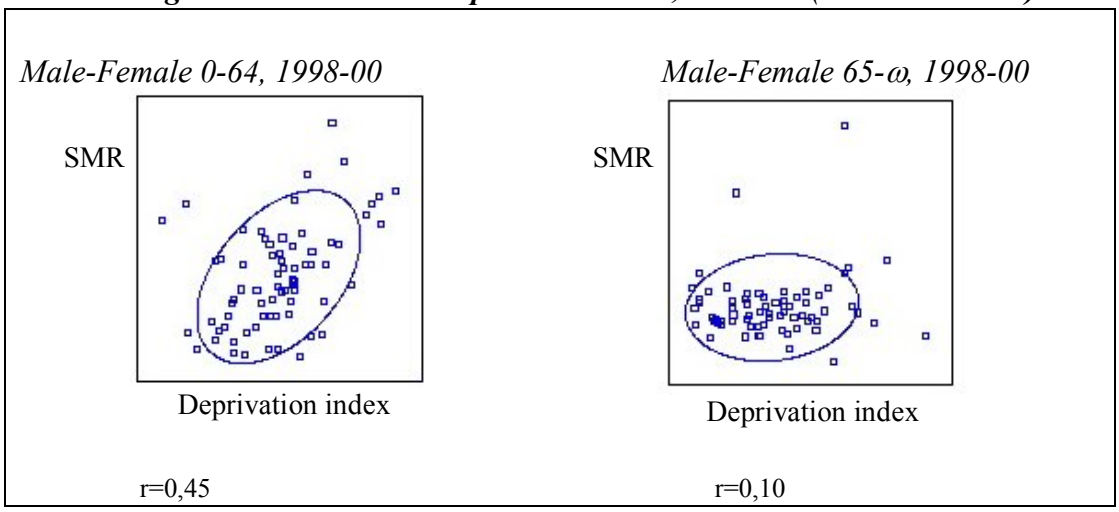
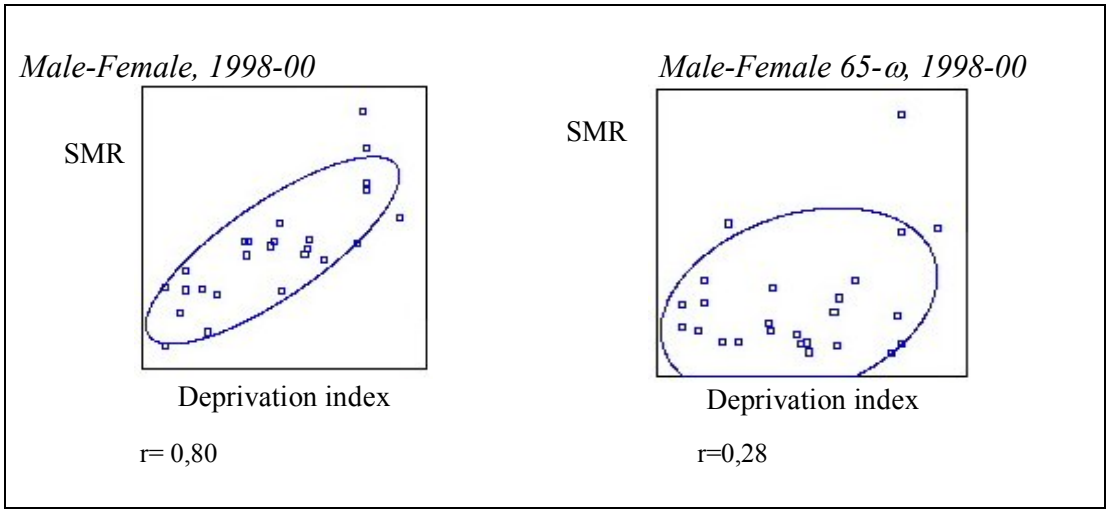


***Figure 4: SMR versus deprivation index,Circoscrizione level (25 observations)***




Correlation is still more strong for the age class 0-64: for people aged 0-64 socioeconomic circumstances seem able to explain a 20% of *SMR* variation at UU level and a 64% at Circoscrizione level.

Table 8 – Circoscrizioni and Unità Urbanistiche (UU)

| Circoscrizioni | Ex-Circoscrizioni | Unità urbanistiche |
|---|---|---|
| I CENTRO EST | 11 OREGINA-LAGACCIO | 31 Lagaccio |
| | | 32 Oregina |
| | 12 PRÈ-MOLO-MADDALENA | 33 Prè |
| | | 34 Maddalena |
| | | 35 Molo |
| | 13 CASTELLETTO | 36 S. Nicola |
| | | 37 Castelletto |
| | | 38 Manin |
| | 14 PORTORIA | 39 S. Vincenzo |
| | | 40 Carignano |
| II CENTRO OVEST | 9 SAMPIERDARENA | 24 Campasso |
| | | 25 S. Gaetano |
| | | 26 Sampierdarena |
| | | 27 Belvedere |
| | | 28 S. Bartolomeo |
| | 10 S. TEODORO | 29 Angeli |
| | | 30 S. Teodoro |
| III BASSA VAL BISAGNO | 16 S. FRUTTUOSO | 43 S. Agata |
| | | 44 S. Fruttuoso |
| | 17 MARASSI | 45 Quezzi |
| | | 46 Fereggiano |
| | | 47 Marassi |
| | | 48 Forte Quezzi |
| IV VAL BISAGNO | 18 STAGLIENO | 49 Parenzo |
| | | 50 S. Pantaleo |
| | 19 MOLASSANA | 51 Montesignano |
| | | 52 S. Eusebio |
| | | 53 Molassana |
| | 20 STRUPPA | 54 Doria |
| | | 55 Prato |
| V VAL POLCEVERA | 5 RIVAROLO | 13 Borzoli Est |
| | | 14 Certosa |
| | | 15 Rivarolo |
| | | 16 Teglia |
| | | 17 Begato |
| | 6 BOLZANETO | 18 Bolzaneto |
| | | 19 Morego |
| | 7 PONTEDECIMO | 20 S. Quirico |
| | | 21 Pontedecimo |

| Circoscrizioni | Ex-Circoscrizioni | Unità urbanistiche |
|---|---|---|
| VI MEDIO PONENTE | 4 SESTRI | 9 Sestri |
| | | 10 S. Giovanni Battista |
| | | 11 Calcinara |
| | | 12 Borzoli Ovest |
| | 8 CORNIGLIANO | 22 Cornigliano |
| | | 23 Campi |
| VII PONENTE | 1 VOLTRI | 1 Crevari |
| | | 2 Voltri |
| | 2 PRA’ | 3 Ca’ Nuova |
| | | 4 Palmaro |
| | | 5 Pra’ |
| | 3 PEGLI | 6 Castelluccio |
| | | 7 Pegli |
| | | 8 Multedo |
| VIII MEDIO LEVANTE | 15 FOCE | 41 Foce |
| | | 42 Brignole |
| | 22 S. MARTINO | 60 Chiappeto |
| | | 61 S. Martino |
| | 23 S. FRANCESCO D’ALBARO | 62 Albaro |
| | | 63 S. Giuliano |
| | | 64 Lido |
| | | 65 Puggia |
| IX LEVANTE | 21 VALLE STURLA | 56 Bavari |
| | | 57 S. Desiderio |
| | | 58 Apparizione |
| | | 59 Borgoratti |
| | 24 STURLA-QUARTO | 66 Sturla |
| | | 67 Quarto |
| | | 68 Quartara |
| | | 69 Castagna |
| | 25 NERVI-QUINTO-S. ILARIO | 70 Quinto |
| | | 71 Nervi |

## References


BARTLEY M, BLANE D. Commentary: Appropriateness of deprivation indices must be ensured. *BMJ* 1994; 309: 1479.

BIGGERI A, BENVENUTI A, MERLER E, NARDULLI M, OLMASTRONI L. La mortalità per condizione socioeconomica e professionale nello studio longitudinale toscano. Comuni di Livorno e Firenze. INFORMAZIONI STATISTICHE settore Studi e Ricerche n° 5 Maggio 1998.

BOX G, COX D. An analysis of transformations (with discussion). *J.R. Statist. Soc. B* 1964 26, 211-252.

BUCKINGHAM K, FREEMAN P. Sociodemographic and morbidity indicators of need in relation to the use of community health service: observational study. *BMJ* 1997; 315: 994-996.

CADUM E, COSTA F, BIGGERI A, MARTUZZI M. Deprivazione e mortalità: un indice di deprivazione per l'analisi delle disuguaglianze su base geografica. *Epid Prev* 1999; 23: 175-187.

CARR HILL R, SHELDON T. Designing a deprivation payment for general practitioners: the UPA(8) wonderland. *BMJ* 1991; 302: 393-396.

CARR HILL R, JAMISON J, O'REILLY D, STEVENSON M, REID J. Risk adjustment for hospital use using social security data: cross sectional small area analysis. *BMJ* 2002; 324: 390.

CARSTAIRS V, MORRIS R. Deprivation and Health in Scotland. 1991 *Aberdeen University Press*.

CRAYFORD T, SHANKS J, BAJEKAL M, LANGFORD S. Analysis from inner London of deprivation payments based on enumeration districts rather than wards. *BMJ* 1995; 311: 787-788

DAWEY SMITH G, WHITLEY E, DORLING D, GUNNEL D. Area based measures of social and economic circumstances: cause specific mortality patterns depend on the choice of index. *Journal Epidemiologic Community Health* 2001;55: 149-150.

DEPARTMENT OF ENVIRONMENT, TRANSPORT AND THE REGIONS. Indices of Deprivation 2000. *Regeneration Research Summary.* N. 31; 2000

FREEMAN M, TUKEY J. *Ann. Math. Statist.* 1950; 21, 607-611.

FORREST R, GORDON D. People and Places: a 1991 Census atlas of England.1993 *SAUS. University of Bristol*

GLOVER G, ROBIN E, EMAMI J, ARABSCHEIBANI G. A needs index for mental health care. *Soc Psychiatry Psichiatr Epidemiol* 1998; 33: 89-96.

JARMAN B. Identification of underprivileged areas. *BMJ* 1983; 286: 1705-09.

JARMAN B. Underprivileged areas: validation and distribution of scores. *BMJ* 1984; 289: 1587-92.

JUDGE K, MAYS N. Equity in the NHS Allocating resources for health and social care in England. *BMJ* 1994; 308: 1363-1366.

JUDGE K, MAYS N. Equity in health care. *BMJ* 1994; 309: 673.

MORRIS R, CARSTAIRS V. Which deprivation? A comparison of selected deprivation indexes. *Journal of Public Health Medicine* 1991; 13: 318-326.

OSBORNE J. Notes on the use of data transformations. *Practical Assessment, Research & Evaluation* 2002; 8(6).

PHILLIMORE P, BEATTIE A, TOWNSEND P. Widening inequality of health in northern England, 1981-91. *BMJ* 1994; 308: 1125-1128.

TOWNSEND P. Deprivation. *Journal of Social Policy* 1987; 16, 2, 125-146.

TOWNSEND, PHILLIMORE, BEATTIE. Health and deprivation: inequality and the North. 1988 London *Croom Helm*

VALERIO M, VITULLO F. Sperimentazione di un indice di svantaggio sociale in Basilicata. Epid Prev 2000; 24: 219-223.