Essays in Health Economics:

Applied and Theoretical Approaches

University of Genoa

Department of Economics

PhD Program in Economics

Candidate

Carolina Bruzzi

Supervisor

Marcello Montefiori

University of Genoa

2021

Contents

Preface

In the present report, I illustrate my research path during the PhD program in Economics at the University of Genova and discuss its results.

The report consists of three independent chapters which reflect the chronological order of the activities carried out during the program.

In the first 2 chapters, two articles are discussed. They belong to the research field inaugurated by Townsend (1987), that is, the analysis of deprivation in urban contexts and the relation between socio economic conditions and health status.

The first article[i] consists in a detailed analysis of the distribution of deprivation in the metropolitan city of Genoa and of its relationship with the health status of the resident population, represented by a generic measure of health such as premature mortality. Information on 14 variables used to assess deprivation was available at the level of the Unità Urbanistiche. These are small administrative units with territorial homogeneity and historical and cultural traditions which appear to be suitable for statistical analyses. An exploratory factorial analysis identified two groups of variables which, according to definitions in the literature, identify the two dimensions of deprivation, material deprivation and social deprivation. For each dimension, two indices were calculated on the basis of two non-compensatory methodologies, the Mazziotta Pareto Index and the Pena Distance Index. Health status was measured through a measure of premature mortality, measured through the calculation of age-standardized SMR. The calculation of the attributable risk was used to evaluate the proportion of the excess mortality observed in various areas which can be attributable to deprivation, and a geographical analysis is presented. The results of the work, in line with the available evidence, confirm the association between material deprivation and premature mortality, but fail to show a relationship with social deprivation.

The second article[ii] presents the results of an analysis of deprivation in Argentine urban areas, with the aim of assessing the presence of a relationship with the average health status of the residents Based on data extracted from the 2017 Survey on Permanent Families in Argentina (EPH), a continuous survey carried out on 32 urban areas in Argentina, five variables were identified, four relating to material and the fifth to social deprivation. The Mazziotta-Pareto Adjusted Index was calculated to provide a synthetic and quantitative measure of the level of deprivation. An analysis of hierarchical clusters was carried out to group different urban areas into deprivation classes and analyse the state of health. The primary measure of the health status used in this study was Life Expectancy at birth. The results, in line with the literature, allow to conclude that the proposed mixed deprivation index accounts for a significant proportion of the variability in Life expectancy at birth across urban areas.

[i] This article has been published in The European Journal of Health Economics in 2020 (https://doi.org/10.1007/s10198-019-01139-x), joint with Enrico Ivaldi and Stefano Landi.

[ii] This article is the result of joint work with Enrico Ivaldi, Paolo Parra Saiani, and Juan José Primosich and it has been published on Social Indicators Research in 2020 (https://doi.org/10.1007/s11205-020-02369-w)

In the third chapter[iii], a study is presented that was developed in in the research field of contract theory. In particular, it refers to contracts in the presence of information asymmetry and the consequent moral hazards and it is aimed at providing policy makers with a tool for an informed use of resources while identifying the sources of inefficiency and waste.

The study is focussed on the problem of defining contracts for the provision of transport services for people with disabilities in a Principal-Agent context. The main issues derive from the hidden actions of the agent and the uncertainty which is due to the type and level of disability of the user which give rise to problems of incomplete information.

The work consists of two parts. In the first, the conceptual framework, the theoretical tools and the main assumptions are presented, including: the description of a principal-agent relation under asymmetric information; the optimization problem of the principal; the description of the causal variables, which are the *effort* applied by the agent in the supply of the service and the disability condition of the user, which, in general terms, is referred to as "*complexity*"; the statistical tool used to model the probability related to the uncertainty which characterizes the service analysed, which is the logistic function; the relation between independent variables and outcome variable; and the agent's risk aversion.

The second part of the study includes the description of two empirical models which, through various simulations, analyse how causal variables affect the probability of the positive result that is the regular performance of the transport service, the expected principal expenditure and the expected agent utility.

The results suggest that, as expected, the more complex the disability condition of the user is, the greater expected expenditure of resources is, but the analysis of the dynamics of the contract, which has actually highlighted the opportunity for unfair behaviours of the agent, suggests that moral hazard is stronger in situations where the user is less complex. As the user's level of disability increases, the agent is less inclined to implement opportunistic behaviours as the uncertainty of the final result increases exogenously.

[iii] This article is the result of joint work with Marcello Montefiori

Non-compensatory aggregation method to measure social and material Deprivation in an urban area: relationship with premature mortality

1. Introduction

Health inequalities can be attributed to the different circumstances under which people grow, live, work, and age. In many countries, higher morbidity and mortality rates have been reported among the lower socioeconomic groups [2, 5, 10, 15, 22, 36, 37, 52, 59, 73, 74].

Studies have shown that areas with lower deprivation levels report better health statuses, such as improvements in infant mortality and cancer survival rates [5, 52], while those facing persistent conditions of deprivation show the worst self-reported health and mortality outcomes [10, 19, 33, 51]. In the United Kingdom (UK), studies on deprivation have been conducted by Castairs and Morris [14], Forrest and Gordon [18], Jarman [28], and Townsend [79]. Deprivation indexes were adopted in the UK starting as measures of the socio-economic conditions of residential areas, based on census or administrative data, and combining several dimensions using statistical procedures to identify priority areas and define targeted public programms [50]. Subsequently, various deprivation indexes were developed in the UK [14, 18, 79] that measured general practitioner workload and capitation, and connected health conditions and health ine- qualities. Today, the indexes are applied in numerous other countries such as Spain [57], Ireland [30], the Netherlands [68], Italy [12, 33], New Zealand [63], North America [77], the USA [3] and Canada [54]. They are used to measure the correlation between deprivation and mortality [8, 23, 64, 74] inequalities in health care services access [16, 34] and the incidence of specific diseases [2, 27, 58, 68, 72, 75].

In the literature, there is no unified definition of deprivation. Deprivation is a wide concept that, on one hand, refers to a disadvantage resulting from the impossibility of achieving functions that are essential for human life [66]. It should be emphasized that health outcomes do not depend solely on the availability of goods and resources, but also largely on the capacity to exploit such goods and resources [28, 66]. This point differentiates the concept of deprivation from that of poverty. Generally, poverty is referred to as the lack of economic sources; in particular, relative poverty, which is closer to deprivation, indicates the condition of people who are economically above a settled threshold within a certain context.

As stated by Townsend [78], the lack of goods, services, resources, and conveniences relates to diet, health, clothing, housing, house-hold facilities, environment, and work, which are normally enjoyed or at least largely accepted as primary goods, and they can be identified with the concept of material deprivation. Material deprivation is, therefore, intended to measure objective conditions as direct variables, since they measure in a direct way how many individuals living in a given geographical area share a certain characteristic [76]. Social deprivation, on the other hand, is defined as "*the condition of those who do not or cannot enter into ordinary forms of family or other relationships*" [79], "*and a non-participation in the roles, relationships, uses, functions, rights, and responsibilities involved in being a member of a society or group*" [78].

These elements make more difficult to identify and quantify deprivation. Given the subjective nature and the difficulty to detect and measure the characteristics described above [35, 78]. Townsend suggests that analysing only material dimension, excluding social deprivation may affect health conditions. Other current analysis, instead, conclude that worse socioeconomic status, considering both deprivation dimensions, is associated worse health conditions [54, 61, 62] and with bad behaviours for health [6]. In particular, Pampalon propose two distinct measures for social and material deprivation and compare them with premature mortality.

In line with these premises, the present work is focused on both deprivation dimensions. The contributions of our study to the literature are: first, it provides a quantitative validation of the subdivision of the variables, chosen based on the previous literature, of social and material deprivation using a quantitative method. Second, it is a complete comparison of material and social deprivation individually for the same geographical area for investigating the relationship between these two domains and health conditions reinforcing the fact that this relationship is not direct and the victims of these conditions may be different subjects. Finally, the proposal of using a non-compensatory aggregation method is to evaluate deprivation in small areas. This approach is novel and has not been used in the previous literature on deprivation. It allows us to a better evaluation of the stability and robustness of deprivation indexes, and, in particular avoidance of compensatory phenomena among partial indicators. A compensatory aggregation could offset the surplus or deficit in one component, thus reducing the relevant information.

2. Materials and methods

2.1. Data

We collected 2011 Italian Census data from the Statistical Office of the Municipality of Genoa.

2.2. Model: The deprivation index

The analysis was developed as follows: (1) choice of the small reference area, (2) choice of socioeconomic variables for the exploratory factor analysis, (3) implementation of two non-compensatory methods for aggregating variables into two different indexes of material deprivation and social deprivation, (4) elaboration of a general health indicator comparing the two deprivation measures, and (5) application of an epidemiological instrument to quantify the effect of deprivation on health status.

Small area

In the literature, the reference statistical units for deprivation analysis are the small areas that are recognizable by administrative borders that preferably present territorial, historical, and cultural characteristics that differentiate them from others. Their size allows them to present groups of individuals that are more homogeneous over certain environmental characteristics and circumstances than the general data can [14]. However, the definition of the geographic area to be taken as a unit of analysis is not clearly formulated or universally accepted. In the examples found in the literature, it is a rather heterogeneous dimension and can also reach a considerable population size [76].

The use of small areas have many positives, but in the interpretation of the results we need to be aware that an area classified as deprived contains people not deprived and, similarly, not all deprived people live in deprived areas [13]. This is the so-called, ecological fallacy, present endemic in the aggregation analysis [32]. Furthermore, the choice is often forced by the nature of available data and non-homogeneous databases, which can only be overcome by linking social and health information to demographic data.

In our work, the choice of the small area was made to priorities characteristics that consider territorial homogeneity and historical, geographic, and cultural traditions and which, at the same time, could be used for analysing the phenomena under study. Specifically, the choice of urban units was determined based on the old independent municipalities included in the City of Genoa starting from 1926, with an average population of about 8500 inhabitants (Genoa's Office of Statistics, 2016). Urban units represent sufficiently disaggregated subdivisions to capture possible territorial differences, for which homogeneous and reliable data exist.

Variables

It is common awareness that a number of socioeconomic phenomena cannot be measured by a single descriptive indicator and that, instead, they should be represented with a multiplicity of aspects or dimensions [43].

From the analysis of the literature, we identified a group of 14 variables, or partial indicators, already used as components of the various previous social or material deprivation indexes although their respective operational definitions may be partially different.

Table 1 presents the variables used, their definition, and the previous studies that employed them. The collection of data for the variables identified is the result of an integration of census data and data derived from the annual statistics compiled by the Municipal Genoa's Office of Statistics. As reference year, we considered the year 2011, that of the latest census available.

Grouping the variables

Factorial analysis assuming a latent structure among the variables estimates unobservable dimensions (latent factors) that somehow influence the observed variables [40]. We applied an exploratory factorial analysis to our 14 variables to analyse the correlations and the latent shared factors among variables. This reflective approach is used merely to identify the common structure between the indicators and not to aggregate or to select them because this method ignores the polarities, namely the meaning of the individual indicators. We try to describe a latent concept, such as the construct of deprivation, using observed variables with formative model. Therefore, factor analysis can be a useful tool for understanding the phenomenon, but a composite index of deprivation must be created following a formative approach [41]. Four factors emerged, namely, the variables 'Low Education', 'Income', 'Overcrowding', and 'Born on the Islands or in South Italy' are positioned on the first component. The second component consists of 'Old Age Index', 'Single Elders', 'Resident Foreigners', and 'Foreigners in Schools'. The third factor includes 'Families Renting', 'Divorced,' 'Unemployment,', and 'Buildings in Mediocre or Bad State'. Finally, 'Single- Parent Families' and 'Single-Person Families' are identified under the fourth factor (Appendix 1).

Table 1 Variables

Variable	Definition	Studies that used an analogous variable
Low education	Population of residents who completed junior school or elementary school are illiterate, out of all residents aged 15 years and above	Julkunen [29] and Pampalon and Hamel [54]
Income	Reciprocal of the average income of the resident population	Julkunen [29]
Overcrowding	Average number of residents per room	Jarman [28], Townsend [79], and Forrest and Gordon [18]
Born on the Islands or in South Italy	Residents born on the Islands or in South Italy	Baumann [6] uses the analogous variable at the European level (nationality not included in Western Europe)
Old age index	Ratio of people aged 65 years and above to people aged 15 years and below	Julkunen [29], Pampalon and Hamel [54], and Baumann [6]
Single elders	Number of people aged 65 years and above and those who live alone	Jarman [28], Forrest and Gordon [18], and Pampalon and Hamel [54]
Resident foreigners	Number of people among the total population who were born abroad	Jarman [28]
Divorced	Proportion of divorced people, compared with the total population	Pampalon and Hamel [54]
Unemployment	Number of unemployed people looking for work, compared with the total active population (aged 15 to 64 years)	Jarman [28], Townsend [79], Forrest and Gordon [18], Julkunen [29], Pampalon and Hamel [54],and Baumann [6]
Buildings in mediocre or bad state	Number of buildings in mediocre or bad state, compared with the total number of buildings	Perez-Mayo [57] uses housing conditions as a variable
Single-parent families	Number of families composed of a single parent and dependent children, compared with all families	Jarman [28], Forrest and Gordon [18], and Julkunen [29]
Single-person families	Number of people living alone, compared with all the families	Julkunen [29] and Pampalon and Hamel [54]
Foreigners in schools	Number of children aged 5 to 14 years and born abroad, compared with the total child population aged 5 to 14 years	Testi and Ivaldi [76]
Families renting	Proportion of families renting a house, compared with the total population	Townsend [79] and Caranci [12]

• These results confirm that there is a common latent structure among two groups of variables and they coherent with Townsend et al. [78] proposal of a bi-dimensional structure for deprivation and is also coherent with the theoretical framework used in literature referring to an economic-material dimension and to socio-demographic characteristics: First dimension, material deprivation: Low Education, Income, Overcrowding, Born on the Islands or in South Italy, Families Renting, Divorced, Unemployment, Buildings in Mediocre or Bad State (first and third factor);

• Second dimension, social deprivation: old Age Index, Single Elders, Resident Foreigners, Foreigners in Schools, Single-Parent Families, Single-Person Families (second and fourth factor).

The first dimension collects objective indicators that have a direct effect on deprivation, while the second one is represented by subjective variables that can affect the uses, functions, rights, and responsibilities involved in being a member of a society or group.

2.3. Construction of deprivation indices with the Mazziotta– Pareto Index (MPI) and Pena Distance Index (DP2) methods

In the construction of composite indicators, the choice of aggregative methodology has an important role on the final results [44]. A key role is played by Compensability among the variables, defined as the possibility of offset any deficit in one dimension with a suitable surplus in another (OECD, 2008). This fundamental consideration determines two connected reasons suggesting the use of non-compensatory methodologies. First, to keep the symmetry or weights of the variables relevant and interpretable, it is necessary to use non-compensatory methods [47]. The second reason concerns the fact that, in some cases, some components are not replaceable. A compensatory aggregation could offset the surplus or deficit in one component, thus reducing the relevant information. The factorial scores in the second and fourth factors (Appendix 1) show positive and negative values, suggesting that social deprivation has different features and social conditions that affect different types of people. These people represent different victims of social deprivation; therefore, they cannot compensate each other, making it better to use non-compensatory methods. To have control over the stability and sensitivity of the measurement of the object of study, both a non-parametric and parametric type of aggregation were used. The MPI is a non-parametric non-compensatory index, which means that it considers variables as non-substitutable and assumes no probability law on the distribution of variables. The DP2 is a parametric index, because its formulation is based on the application of a linear regression model.

Finally, an influence analysis (IA) was performed to assess the robustness of the composite indices [41]. IA aims to empirically quantify the importance of each individual indicator in the calculation of the composite index, simulating the index calculation removing one indicator at a time. The average shift in urban unit rankings is measured. This statistic tells us the relative shift in the position of the entire system of UUs in a single number. It can be calculated as the average of the absolute diferences in UU' ranks with respect to a reference ranking over the 71 UUs: Let GDI be the index value for urban unit c, c=1,…, m

$$\bar{R}_s = \frac{1}{m}\sum_{c=1}^{m} Rank_{reference}(GDI_c) - Rank((GDI_c) \qquad (1)$$

Mazziotta–Pareto Index (MPI)

The MPI [41, 42] is based on the assumption of "non-substitutability" of the dimensions, to which equal importance is attributed, and no compensation between them is allowed. This is a non-compensatory methodology proposed in 2007 to measure healthcare infrastructures under the assumption that each component of the infrastructure is not substitutable by others and, therefore, the values cannot be offset at the aggregate level. Subsequently, the authors applied the same methodology for creating a quality of life measurement capable of maintaining the characteristics of each variable included in the calculation of a synthetic index by penalizing units of analysis that present an unbalanced distribution of some variables compared to the average [44]. The MPI is constructed using the following steps: (1) normalization of the individual indicators by "standardization" and (2) aggregation of the standardized indicators by an arithmetic mean with penalty function based on "horizontal variability", that is, variability of standardized values for each unit. This variability, measured by the coefficient of variation, allows penalization of the score of units that have higher imbalance between the values of the indicators. Finally, use of the standardized deviation in reckoning the synthetic index sets up a measure that is robust and little sensitive to the elimination of a single elementary indicator [42]. We calculate that z_{ij} is the standardized value of each *j*-th variable of each *i*-th urban unit

$z_{ij} = 100 + \frac{x_{ij}-\mu_j}{\sigma_j} 10$ *if the j-th indicator is "positive" (2)*

$z_{ij} = 100 - \frac{x_{ij}-\mu_j}{\sigma_j} 10$ *if the j-th indicator is "negative" (3)*

Where

- $x_{ij}$is the original value of each j-th variable of each i-th urban unit,
- $\mu_j$is the mean of each j-th indicator and
- σ_j is the standard deviation of each jth indicator.

The characteristic of "positive" or "negative" is interpreted with respect to polarity: the polarity is the sign of the relation between the indicator and the phenomenon to be measured, i.e. [39]. In the example of Mazziotta and Pareto 's work, political participation polarity is "positive" (negative) if increasing values of an indicator correspond to positive (negative) variations of political participation [44].

Then, we calculate MPI for each i-statistical unit for both deprivation dimensions as follows:

$MPI_i = \mu_{\bar{z}_{i,k}} - \sigma_{\bar{z}_i} cv_{\bar{z}_i}$ *(4)*

Where

- $\mu_{\bar{z}_{i,k}}$ is the mean value of the standardized values (z_{ij}) of the variables in the k dimension (k=1, 2)
- $\sigma_{\bar{z}_i}$ is the standard deviation of z_{ij} in each dimension and
- $cv_{\bar{z}_i} = \frac{\sigma_{\bar{z}_i}}{\mu_{\bar{z}_{i,k}}}$ i is the coefcient of variation of z_{ij}

This approach is characterized by the use of function ($\sigma_{\bar{z}_i} cv_{\bar{z}_i}$) to penalize the units with "unbalanced" values of the partial composite indices. The penalty is based on the coefficient of variation and it is zero if all values are equal. The purpose is to favour regions that, mean being equal, have greater balance among the different dimensions of deprivation indices according to the MPI method

Pena distance index (DP2)

The DP2 (so-called P2 distance or Peña method) was proposed by Peña [56]. The DP2 method is an iterative procedure that weighs partial indicators depending on their correlation with the global index. The DP2 was used in different arguments: economic and social cohesion [24]; environmental quality [46]; quality of life [69–71], welfare systems [38], political participation [25], and deprivation [33]. This construction solves a large number of problems, for instance, for aggregating variables expressed in different units of measurement, arbitrary weights, treatment of missing values, and duplicate information [46, 56, 69]. Moreover, this method is considered more robust than traditional methods such as principal component analysis and data envelopment analysis as demonstrated by Somarriba and Pena [69] in a research comparing this method with others. The

DP2 distance synthetic indicator has many properties: non-negativity, commutativity, triangular inequality, existence, determination, monotony, uniqueness, transitivity, invariance to change of origin and/or scale of the units in which the variables are defined, invariance to a change in the general conditions, and exhaustiveness and reference base [49, 60]. The basis of the synthetic indicator is a mathematical function expressed as $I = F(x_1, x_2, \ldots, x_n)$,, where I is the synthetic indicator and n is the number of variables, *x*, that contribute information to our index. Starting from a matrix X of order (*i, j*), where *i*—the rows—is the number of the areas considered in this paper and j—the columns—is the number of variables considered. Each element x_{ij} of this matrix represents the state of the variable j in i statistical unit. The DP2 indicator, providing the distance of each area (statistical unit) from a reference base, which corresponds to the theoretical area achieving the lowest value of the variables being studied, is defined for area *i* as follows:

$$DP2_i = \sum_{j=1}^{m} \left\{ \left(\frac{d_{ij}}{\sigma_j} \right) \left(1 - R^2_{j,j-1,j-2,\ldots,1} \right) \right\} \qquad (5)$$

• *i*=1, …, n; (areas) and *j*=1, 2, …, m (variables),

• $d_{ij} = \left| x_{ij} - \widetilde{x_{ij}} \right|$ is the difference between the value taken by the *i*-th variable in area *j*-th (x_{ij}) and $\widetilde{x_{ij}}$ which is the minimum of the variable in the m units considered and

• σ_j is the standard deviation of the variable *j*.

Note that the quantity $\frac{d_{ij}}{\sigma_j}$ is merely a change in the origin and scale, and one may also use zero as the reference *ρ* point and [max (x_{ij}) − mini (x_{ij})] instead σ_j a scaling factor without any adverse efect on the formula [45, 46]. $(1 - R^2_{j,j-1,j-2,\ldots,1})$ is the coefcient of multiple linear correlation squared in the linear regression of X_j over $X_{j-1}, X_{j-2}, \ldots, X_1$, and it indicates the part of the variance of X_j explained linearly by variables $X_{j-1}, X_{j-2}, \ldots, X_1$. This coefficient is an abstract number and is unrelated to the measurement units of the different variables. $(1 - R^2_{j,j-1,j-2,\ldots,1})$ is the correction factor that ensures that the composite synthetic indicator includes only the new information from each variable, avoiding the duplication of information already contained in the preceding variables. Therefore, $R^2_{j,j-1,j-2,\ldots,1}$ with *j>1* is the coefficient of multiple linear correlation squared in the linear regression of the first chosen *j* over the other successive indicators, included one by one. The first variable obtains an absolute weight of unity $(1 - R^2_1)$, the subsequent variable *j* = 2 obtains a weight $(1 - R^2_{2,1})$, the third *j*=3, $(1 - R^2_{3,2,1})$, and in general, the *j*-th variable obtains a weight $(1 - R^2_{j,j-1,j-2,\ldots,1})$. The chosen j is regressed over other indicators included one by one. In this way, the weight assigned to each indicator follows a precise rule that has the goal to reduce the duplicity of information that often afects aggregation methods [46, 69]. Thus, the weights assigned to a variable depend on its ranking, that is, positions in the order, making the DP2- based composite synthetic indexes indeterminate and arbitrary [46, 49].

Montero et al. [46] suggest the following procedure to solve the indeterminacy problem:

1. Initialize the weight vector, $w_j = 1$ *∀j = 1, 2, …, m* and define *ε=0.00001*,

2. Define $\partial_j = \frac{d_{ij}}{\sigma_j}$, *∀j = 1, 2, …, m* and *i=1, 2,…, n*,

3. Obtain $DF_j = \sum_{j=1}^{m} \left(\frac{d_{ij}}{\sigma_j}\right) w_j,\ j = 1, 2, \ldots, m,$

4. Compute Pearson's coefficient of correlation $r(DF_j, \partial_j)$ between DF_j and ∂_j $\forall j = 1, 2, \ldots, m$. Arrange $|r(DF_j, \partial_j)|$ in descending order and re-index the associated variables ∂_j accordingly,

5. Compute $Z_i = \sum_{j=i}^{m} \left[\left(\frac{d_{ij}}{\sigma_j}\right) w_j\right]$; $i = 1,2,\ldots,n$; $w_j = \left(1 - R^2_{j,j-1,\ldots,1}\right)$ for $j = 2,3,\ldots,m$ and $w_j = 1$.

6. If $\sum_{i=1}^{n} (DF_i > Z_i)^2 \geq \varepsilon$ replace DF by Z go to d), Else: stop.

As with the previous method, higher scores represent situations of greater deprivation than lower scores.

Thus defined, the DP_2 is the sum of the distances between the values of variable i in the territory j and the minimum values for the variables in all territories, weighted by the unexplained variance of X_i and the variance X_i. A greater distance from the worst theoretical condition shows a higher DP_2 value, indicating an high index value, whereas, a lower distance from the worst theoretical condition describes a scarce level of DP_2.

Class aggregation

We classified our urban units using hierarchical clustering analysis to facilitate comparisons between the two measures of deprivation. We chose Ward's method because we wanted to minimize within-cluster variance. Specifically, we identified clusters of units based on the distance between units' measures of social and material deprivation [48, 80].

2.4. Mortality ratio standardized by age

We used premature mortality for the comparison of deprivation and health conditions. Premature mortality refers to the death of people under a specific age threshold, which is usually less than the average age of death in a certain population. The cut-off point to define a premature death was set at 65 years of age. This is a commonly used threshold in the literature and in institutional settings [8, 14, 19]. Specifically, we employed the age-standardized mortality ratio (SMR), which is a relative index of mortality, expressing the mortality experience of the study population relative to that of a comparison ('standard') population. SMR is the ratio between the observed deaths in a given territory and the expected deaths calculated with an indirect standardization based on age distribution in that territory. The ratio was calculated using the observed and expected deaths in each *i*-th urban unit for each *t*-th age class. In each urban unit, the SMR is equal to the sum of these ratios:

$$SMR_i = \frac{\sum_t Obs_{it}}{\sum_t Exp_{it}} \qquad (6)$$

Where Obs_{it} are the observed death and Exp_{it} are the expected deaths.

In view of the small size of the units of analysis, SMR was calculated over the 3-year period from 2011 to 2013 to give stability to the indicator, thus avoiding, for example, the possible overestimation of mortality in very small areas due to particular events.

Attributable risk

Attributable risk is an epidemiological tool that identifies the fraction of total disease experience in the population that would not have occurred if the effect associated with the risk factor of interest was absent [11] and allows estimating either the proportion of cases attributed to (or caused by) the exposure factor or the proportion of cases that could be prevented if the exposure factor was eliminated [20]. This approach was used in our work to identify to what extent mortality in each deprivation class can be statistically explained by the level of deprivation of the class, that is, how many deaths can be attributed to a deprived living condition. The procedure involves taking as reference the minimum value of the SMR in the first class and, based on this value, estimating the expected deaths in the other deprivation classes as a product of the residents included in the class and the SMR of the first class. The deaths attributable to a deprivation level greater than the reference level represent the difference between the deaths observed in the class and the expected deaths. The risk attributable to deprivation in each class is, therefore, the percentage ratio between the deaths attributable to deprivation and the observed deaths.

3. Results and discussion

Comparison of methodologies Appendixes 3 and 4 show the social and material deprivation indexes (SDI and MDI) resulting from the aggregation of variables under the two different methodologies. Using Spearman's correlation coefficient—which links the ranks of the two distributions—the aggregation methodologies were compared. Between SDI (MPI) and SDI (DP2), the correlation coefficient is 0.856, and between MDI (MPI) and MDI (DP2) it is 0.983. The two indicators, which use different methodologies, show strongly correlated results. For clarity, the results showed thereafter refer to the indexes calculated using the MPI method, which showed more robust results in influence analysis. As we can see in Table 2, DP2 is the more sensitive index. MPI is less sensitive to inclusion or exclusion of individual indicators. Regarding material and social indices, the first has low ranks shift while social has higher shifts in both aggregation methods.

Table 2 Influence analysis

Material Dimension		
Individual indicator	MPI	DP2
Income	1.86	2.62
Low education	1.83	2.17
Overcrowding	2.39	3.63
Born on the Islands or South Italy	2.39	2.56
Families renting	1.97	3.30
Divorced	3.72	5.04
Unemployment	1.83	1.27
Buildings in mediocre or bad state	3.44	4.65
Mean	2.43	3.15
Standard deviation	0.65	0.95
Social Dimension		
Individual indicator	MPI	DP2
Old age index	10.62	11.30
Single elders	9.46	11.80
Resident foreigners	7.58	13.27
Foreigners in school	8.54	12.00
Single-parent families	11.77	8.45
Single-person families	3.97	9.94
Mean	8.66	11.13
Standard deviation	2.50	1.55

Geographic representation of results

A first comparison of the two dimensions of deprivation (social and material) through Spearman's coefficient shows a low degree of correlation (− 0.110) and suggests that material and social deprivation should be kept separate. Through cluster analysis, five classes of increasing deprivation were identified for each dimension (Appendix 5), from the wealthy to the highly deprived class, as shown in Figs. 1 and Appendix 3. No relationship emerges between the two different dimensions of deprivation at the geographic level. In fact, it is not possible to record a correspondence between high levels for one dimension and the other. Indeed, in some areas, there are discordant levels of material and social deprivation. Materially deprived area are not necessarily also socially deprived and vice versa (Figure 1, Figure 2).

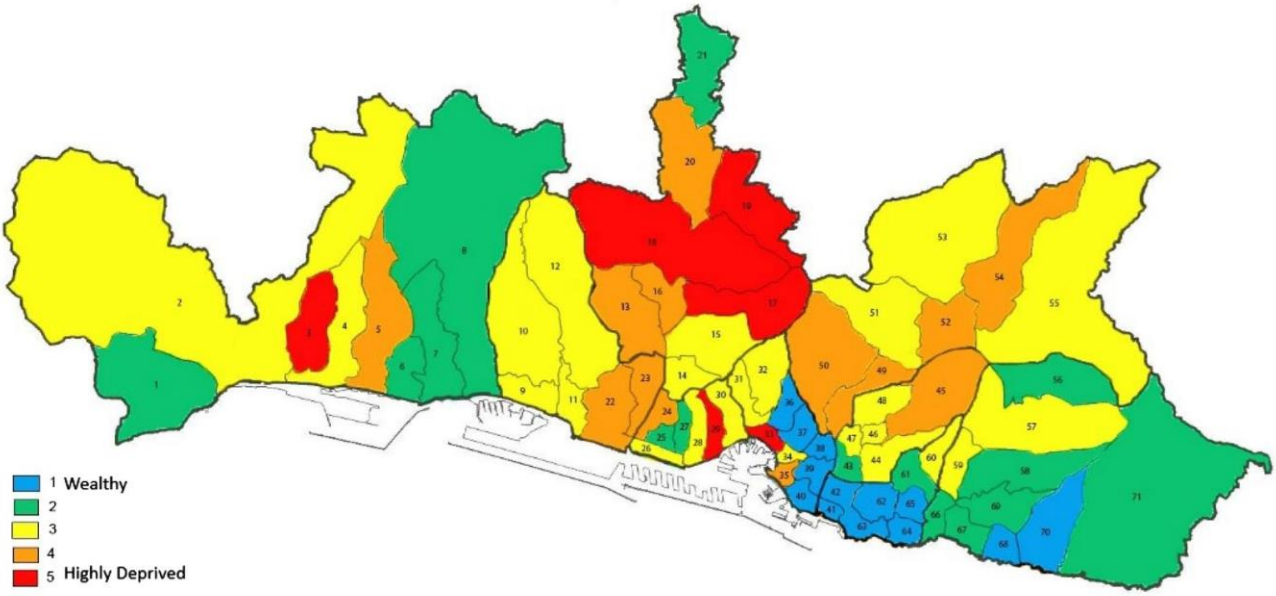


Figure 1 Geographical distribution of Material Deprivation

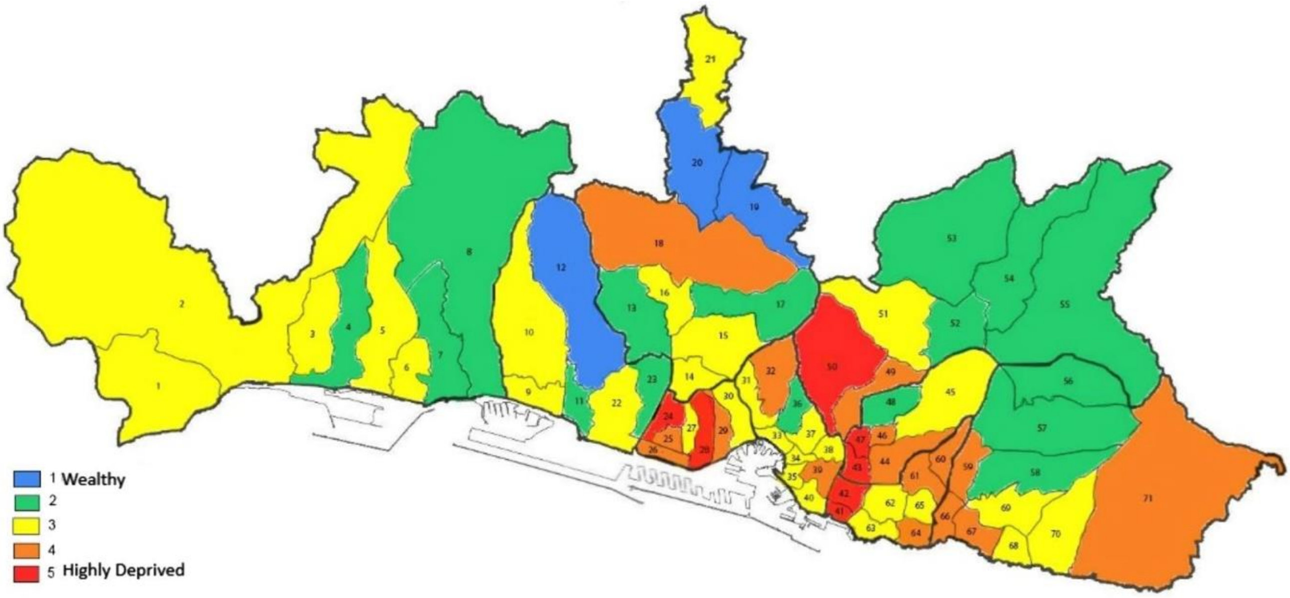


Figure 2 Geographical Distribution of Social Deprivation

Comparison between the deprivation index and SMR

Comparison between the measure of material deprivation and the health indicator was initially explored for all urban units considered individually. Overall, MD and SMR were positively correlated (Spearman's correlation=0.583) in urban unit analysis. When we calculated SMR within deprivation classes, SMR was higher for the classes with higher deprivation. This class analysis shows early mortality in areas with unfavorable material conditions. However, very different results were observed with respect to social deprivation and SMR. There is no correlation between the two measures (Spearman's correlation=−0.098) at the urban unit level, or in the class-level analysis (Table 4). Very different results are found by comparing the social indicator with the health measure. At the urban unit level, there is no linear correlation between the two measures (Spearman's coefficient equal to −0.098), and Table 3 shows that, even at the level of classes, no relationship between social deprivation and health conditions exists (Fig. 3).

Table 3 Material and social deprivation index and standardized mortality ratio classes

Material Deprivation class	SMR (age < 65)	Social Deprivation class	SMR (age < 65)
1	0.71	1	1.40
2	0.85	2	0.91
3	1.01	3	0.99
4	1.15	4	1.05
5	1.58	5	0.97

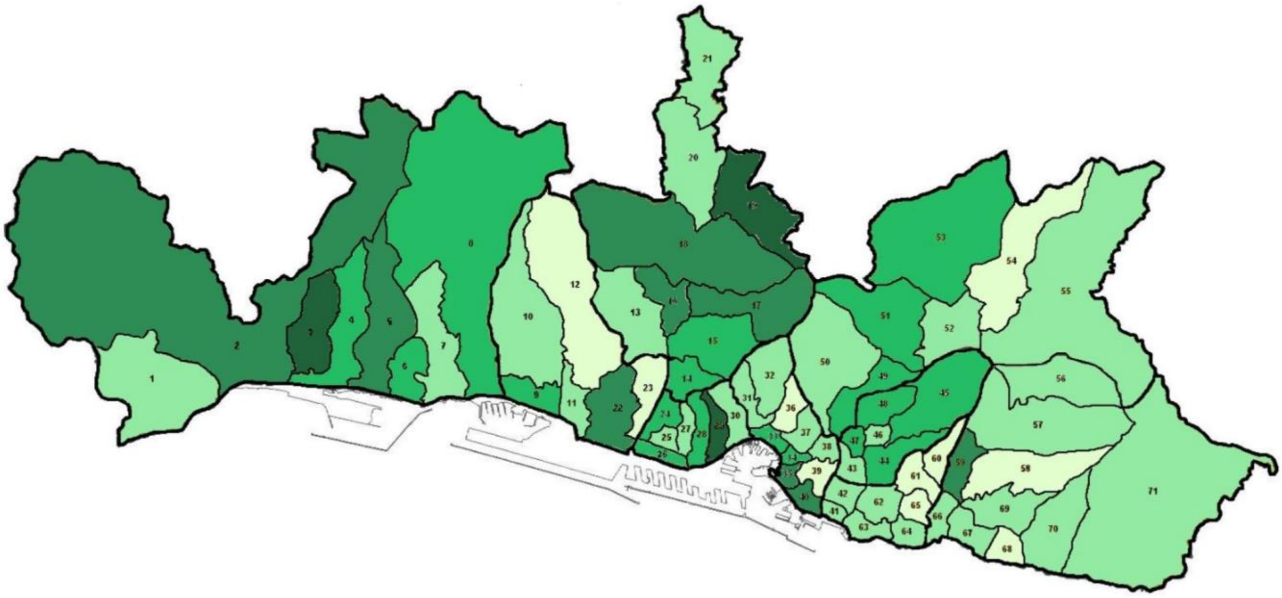


Figure 3 Representation of Standardized Mortality Ratio classes

Material deprivation and attributable risk

In view of the results, an instrument was used to identify the risk attributable to a given factor in the context of reference [11]. This allows us to identify how much mortality in each class of material deprivation can be explained by the material deprivation of the class, that is, how many deaths can be attributed to a materially deprived life (Table 4). The last column in Table 4 shows, for each class, the percentage of premature deaths that can be explained by living in worse deprivation than the first class (i.e. the most favourable condition). In the fifth class, about 154 deaths can be attributed to deprivation: that is, 55% of the observed deaths may be motivated by living in worse conditions. This is a risk related to the reference context, and while statistical processing cannot provide incontrovertible evidence, it can be a useful tool to analyses and quantify the effects.

Table 4 Risk attributable to material deprivation

Material deprivation class	Observed deaths (< 65 years old)	Standardized mortality ratio (SMR)	Relative risk (SMR class/SMR class 1)	Expected deaths if everyone had the same deprivation as class 1 (residents*relative risk)	Deaths caused by deprivation (observed deaths — expected deaths)	Attributable risk (deaths caused by deprivation/observed deaths) (%)
1	296	0.71	1	296	0	0
2	398	0.85	1.19	334.27	63.72	16
3	981	1.01	1.41	694.87	286.13	29
4	498	1.15	1.61	308.07	189.92	38
5	280	1.58	2.22	125.78	154.21	55
Genoa	2453	–	–	1,759.01	693.99	28

The results obtained in this study appear to be in line with the economic, sociological, and epidemiological literature. Material deprivation is found to be correlated with health outcomes in many studies. This relationship is present both when health is measured in a generic way through mortality by any cause [8, 10, 23, 33, 64, 74,9, 31] and when specific pathologies are considered [2, 4, 27, 58, 68, 72, 75]. Our work provides an instant refection of the metropolitan reality of Genoa with an informative purpose that captures the distribution of deprivation. From a material viewpoint, the most deprived areas are located in the industrial or formerly industrial areas and, in particular, public housing neighbourhoods linked to industrial development in the western part of the city. Socially deprived areas are located in some areas of the historic centre and the immediate periphery, which are characterized by high concentration of immigrants and residential neighbourhoods and strong presence of elderly residents. A contribution of our study, through the simultaneous comparison of material and social deprivation in the same geographical area, revealing that in the most materially deprived areas there is no equivalent social deprivation, reinforces the fact that the relationship between the two conditions is not direct and the victims of these conditions are different subjects. As previously found, there are different forms of marginalization [79], which can manifest in different situations of social exclusion, such as lack of social networks, lack of family support, isolation, and racism. The case of social deprivation is more emblematic; Pampalon found an association between social deprivation and premature mortality [54] but, in our case, is not possible to attribute any direct influence on mortality to an unfavourable social condition. Since no trend has been observed, the results suggest that the social index refers to dimensions that are not relevant to mortality. Similar considerations are made by Townsend et al. [78], who excludes the social dimension from the deprivation index and opts for an exclusively material index to explain mortality when building a model of deprivation directly linked to health. It remains to be verified whether living under social deprivation may have long-term health effects or it results in an increased risk to health, as suggested by Baumann et al. [6], or in increased costs for the community [76]. However, it is widely demonstrated that living under material and social deprivation can enhance the behaviours strongly related to the onset of serious pathologies [6, 61, 62]. This study, in addition, underlines the methodological importance of non-compensatory methods applied to represent deprivation at the area level. Each partial indicator should not compensate the others, because each represents a different "source" of deprivation. The use of a non-compensatory aggregation method may be well suited for such types of analyses. This study is not free from limitations. From a methodological viewpoint, it would be important to evaluate how the choice of a small area can impact on the results of the analysis [65] and pay attention to the so-called ecological fallacy [13]. On the other hand, studies have shown that area-level deprivation not only serves as an individual-level proxy when data are unavailable but also has its own impact [21, 72]. Moreover, policy decisions regarding, for example, resource allocation and healthcare organisations, are made on a larger scale than the individual or census section levels. Second, deprivation indexes are a synthetic measure that simply encapsulates the multidimensionality of reality and, therefore, can also present interpretation problems.

Furthermore, the proposed methodology applies solely to the metropolitan urban case, particularly with regard to the choice of variables used. In fact, some variables refer to the social or economic conditions that result in unfavourable situations in an urban context but not a rural one [7]. For example, housing tenure has been indicated as a biased measure of deprivation in the rural context [53], and low level of education in the context of small isolated urban unit towns is often not associated with poor economic conditions. This is because informal training exists in such contexts, which results in individual/family businesses or alternative work solutions that do not work well in urban realities but constitute the basis of societies in nonmetropolitan contexts [67]. However, some studies have shown that it is possible to test the ability of a small area deprivation index to describe deprivation even in rural areas [7]. Even though this study has limitations, it has potential implications for public health and health care policy in particular for the use of appropriate indexes to account for inequalities and resource allocation.

4. Conclusions

The results are in line with the literature, which show that the higher mortality rates are present among lower socioeconomic groups. A part of the health inequalities can be attributed to the different socioeconomic conditions under which people live. Considering a non-compensatory methodology and both deprivation domains (material and social) is important because it provides the ability to better capture the role of each deprivation component and of the two domains of disadvantage that affect health outcomes and create inequalities in life conditions. Material deprivation, as prior literature points out, is associated with lower health outcomes. The index of social deprivation measured in this work shows a weaker association with health outcomes. These results are in line with the considerations of Townsend et al. [78], who excluded the social dimension from the deprivation index to explain mortality when building a model of deprivation directly linked to health. However, the possibilities that living under social deprivation may result in long-term health effects; increased risk to health, as suggested by Baumann [6], and increased costs for the community [76] remain to be verified. Further developments in the analysis of the phenomena of material and social deprivation may concern the methodological and dynamic aspects of the indicators to turn them into usable tools, both politically and technically. Thus, we recommend that future works analyse the two domains separately and examine the relationship between material and social deprivation over time. It could be useful to evaluate deprivation domains over time following the stacking deprivation method [33, 51]. In conclusion, health inequalities are inevitable, but health status differences arising from socioeconomic conditions should be considered unfair and avoided or reduced by employing active social policies [1, 36, 76]. Adequate social and economic policies can be just as important as health care, because income, job position, education, and the environment are all keys to good health.

References

1. Adler, E.A., Newman, K.: Socioeconomic disparities in health: pathways and policies. Health Af. 21(2), 60–76 (2002)
2. Andersen, K. K., Steding-Jessen, M., Dalton, S. O., & Olsen, T. S.: Socioeconomic position and incidence of ischemic stroke in Denmark 2003–2012. A nationwide hospitalbased study. J. Am. Heart Assoc. 3(4), 1–6 (2014).
3. Arcaya, M.C., Arcaya, A.: Inequalities in health: defnitions, concepts and theories. Global Health Actions 8 (2015).
4. Barnett, R., Moon, G., Kearns, R.: Social inequality and ethnic diferences in smoking in New Zealand. Soc. Sci. Med. 59, 129– 143 (2004)
5. Basta, N.O., James, P.W., Gomez-Pozo, B., Craft, A.W., Norman, P.D., McNally, R.J.Q.: Survival from teenage and young adult cancer in northern England, 1968–2008. Pediatr. Blood Cancer 61(5), 901–906 (2014)
6. Baumann, M., Spitz, E., Guillemin, F.: Associations of social and material deprivation with tobacco, alcohol and psychotropic drugs use, and gender: a population study. Int. J. Health Geogr. (2007)
7. Bertin, M., Chevrier, C., Pele, F., Serrano-Chavez, T., Cordier, S., Viel, J.F.: Can a deprivation index be used legitimately over both urban and rural areas? Int. J. Health Geogr. 13(1), 22 (2014)
8. Benach, J., Yasui, Y.: Geographical patterns of excess mortality in Spain explained by two indices of deprivation. J. Epidemiol. Community Health 53(7), 423–431 (1999)
9. Bond Huie, S.A., Krueger, P.M., Rogers, R.G., Hummer, R.A.: Wealth, race and mortality. Soc. Sci. Q. 84(3), 667–684 (2003)
10. Brown, D., Rees, P.: Trends in local and small area mortality and morbidity in Yorkshire and the Humber: monitoring health inequalities. Reg. Stud. 40(5), 437–458 (2006). https://doi. org/10.1080/00343400600757395
11. Bruzzi, P., Green, S., Bylar, D.: Estimating the population attributable risk for multiple risk factors using case–control data. Am. J. Epidemiol. 122(5), 904–914 (1985)
12. Caranci, N., Biggeri, A., Grisotto, L.: L'indice di deprivazione italiano a livello di sezione di censimento: defnizione, descrizione e associazione con la mortalità. Epidemiol. Prev. 34(4), 167–176 (2010)
13. Carr-Hill, R.A., Jamison, J., O'Reilly, D.: Risk adjustment for hospital use using social security data: cross sectional small area analysis. BMJ 324, 390 (2002)
14. Castairs, V., Morris, R.: Deprivation and Health in Scotland. Aberdeen University Press, Aberdeen (1991)
15. Deaton, A., Lubotsky, D.: Mortality, inequality and race in American cities and states. Soc. Sci. Med. 56(6), 1139–1153 (2003)
16. Devaux, M.: Income-related inequalities and inequities in health care services utilisation in 18 selected OECD countries. Eur. J. Health Econ. 16(1), 21–33 (2015). https://doi.org/10.1007/s1019 8-013-0546-4. (Epub 2013 Dec 14)
17. Dillon, W.R., Goldsteinn, M.: Multivariate analysis-methods and applications. Wiley, New York (1984)
18. Forrest, R., Gordon, D.: People and Placet: a 1991 census atlas of England. SAUS, Universty of Bristol (1993)
19. Exeter, D.J., Boyle, P.J., Norman, P.: Deprivation (im)mobility and cause specifc premature mortality in Scotland. Soc. Sci. Med. 72(3), 389–397 (2011)
20. Fuchs, C., Berger, V.W.: Quantifying the proportion of cases attributable to an exposure. J. Modern Appl. Stat. Methods JMASM 3(1), 54–64 (2004)
21. Hagedoorn, P., Vandenheede, H., Vanthomme, K., Willaert, D., Gadeyne, S.: A cohort study into head and neck cancer mortality in Belgium (2001–11): are individual socioeconomic diferences conditional on

area deprivation? Oral Oncol. 61, 76–82 (2016)

22. Hajizadeh, M., Campbell, M.K., Sarma, S.: Socioeconomic inequalities in adult obesity risk in Canada: trends and decomposition analyses. Eur. J Health Econ. 15(2), 203–221 (2014). https://doi. org/10.1007/s10198-013-0469-0. (Epub 2013 Mar 31)
23. Hofmann, R., et al.: Social diferences in avoidable mortality between small areas of 15 European cities: an ecological study. Int. J. Health Geogr. 13(1), 8 (2014)
24. Holgado, M., Molina, M., Salinas Fernández, J. A., Rodríguez Martín, J. A.: A synthetic indicator to measure the economic and social cohesion of the regions of Spain and Portugal'. In: Revista de Economía Mundial, 39, 223–240. Istat (2011). 15° Censimento della popolazione e delle abitazioni 2011 (2015)
25. Ivaldi, E., Bonatti, G., Soliani, R.: An indicator for the measurement of political participation: the case of Italy. Soc. Indic. Res. 132(2), 605–620 (2017)
26. Ivaldi, E., Testi, A.: Genoa Index of Deprivation (GDI): an index of material deprivation for geographical areas. In: Baird, C.M. (ed.) Social indicators: statistics, trends and policy development, pp. 75–98. Nova publisher, New York (2011)
27. Jackson, C., Jones, M., Mishra, G.D.: Educational and homeownership inequalities in stroke incidence: a population-based longitudinal study of mid-aged women. Eur. J. Public Health 24(2), 231–236 (2014)
28. Jarman, B.: Identifcation of underprivileged areas. BMJ 286, 1705–1709 (1983)
29. Julkunen, I.: Social end material deprivation among the unemployed youth in Northen Europe. Soc. Policy Adm. 36(3), 235– 253 (2002)
30. Kelleher, C., Timoney, A., Friel, S.: Indicators of deprivation, voting patterns and health status at area level in the Republic of Ireland. J. Epidemiol. Community Health 56, 36–44 (2002)
31. Lalloué, B., Monnez, J.M., Padilla, C., Kihal, W., Le Meur, N., Zmirou-Navier, D., et al.: A statistical procedure to create a neighborhood socioeconomic index for health inequalities analysis. Int. J. Equity Health 12(1), 21 (2013)
32. Lancaster, G., Green, M.: Deprivation, ill-health and the ecological fallacy. J. R. Statist. Soc. A 165(2), 263–278 (2002)
33. Landi, S., Ivaldi, E., Testi, A.: Measuring change over time in socio-economic deprivation and health in an urban context: the case study of Genoa A. Soc. Indic. Res. (2017). https://doi. org/10.1007/s11205-017-1720-3
34. Landi, S., Ivaldi, E., Testi, A.: Socioeconomic status and waiting times for health services: an international literature review and evidence from the Italian National Health System. Health Policy (2018). https://doi.org/10.1016/j.healthpol.2018.01.003
35. Macintyre, S., Ellaway, A., Cummins, S.: Place efects on health: how we can conceptualise, operationalise and measure them? Soc. Sci. Med. 55, 125139 (2002)
36. Mackenbach, J.P., Bos, V., Andersen, O., Cardano, M., Costa, G., Harding, S., Reid, A., Hemström, O., Valkonen, T., Kunst, A.E.: Widening socioeconomic inequalities in mortality in six Western European countries. Int. J. Epidemiol. 32(5), 830–837 (2003)
37. Mackenbach, J.P.: The persistence of health inequalities in modern welfare states: the explanation of a paradox. Soc. Sci. Med. 75(4), 761–769 (2012)
38. Martinez-Martinez, O.A., Lombe, M., Vazquez-Rodriguez, A., Coronado-Garcia, M.: Rethinking the construction of welfare in Mexico: going beyond the economic measures. Int. J. Soc. Welfare 25, 259–272 (2016)
39. Maggino, F.: Complexity in society: from indicators construction to their synthesis. Springer, Cham, Switzerland (2017)
40. Mazziotta, M., Pareto, A.: Use and misuse of PCA for measuring well-being. Soc. Indic. Res. 142(2), 451–476 (2019)

41. Mazziotta, M., Pareto, A.: Synthesis of indicators: The composite indicators approach. In: Maggino, F. (ed.) Complexity in society: from indicators construction to their synthesis. Social indicators research series, pp. 159–191. Springer, Cham (2017)

42. Mazziotta, M., Pareto, A.: On a generalized non-compensatory composite index for measuring socio-economic phenomena. Soc. Indi. Res. 127(3), 983–1003 (2016)

43. Mazziotta, M., Pareto, A.: Methods for constructing composite indices: one for all or all for one? Riv. Ital. Econ. Demogr. Stat. LXVI I(2), 67–80 (2013)

44. Mazziotta, M., Pareto, A.: A non-compensatory approach for the measurement of the quality of life. In: Maggino, F., Nuvolati, G. (eds.) Quality of life in Italy. Social Indicator Research Series, pp. 27–40. Springer, Netherlands (2012)

45. Mishra, S.K.: A note on the indeterminacy and arbitrariness of Pena's method of construction of synthetic indicator. SSRN. (2012). http://ssrn.com/abstract=2026293

46. Montero, J.M., Chasco, C., Larraz, B.: Building an environmental quality index for a big city: a spatial interpolation approach combined with a distance indicator. J. Geogr. Syst. 12, 435–459 (2010)

47. Munda, G., Nardo, M.: Non-compensatory/non-linear composite indicators for ranking countries: a defensible Setting. Appl. Econ. 41(12), 1513–1523 (2009). https://doi.org/10.1080/0003684060 1019364

48. Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., Hofman, A., Giovannini, E.: Handbook on constructing composite indicators: methodology and user guide. OECD Publishing, Paris (2005)

49. Nayak, P., Mishra, S.K. (2012). Efciency of Pena's P2 distance in construction of human development indices. In: MPRA—Munich Personal RePEc Archive, paper 39022.3

50. Noble, M., Wright, G., Smith, G., Dibben, C.: Measuring multiple deprivation at small area level. Env. Plann. A 38, 169–185 (2006)

51. Norman, P.: Identifying change over time in small area socioeconomic deprivation. Appl. Spat. Anal. Policy 3(2–3), 107–138 (2010)

52. Norman, P., Gregory, I., Dorling, D., Baker, A.: Geographical trends in infant mortality: England and Wales, 1970–2006. Health Stat. Q. 40, 18–29 (2008)

53. O'Reilly, G., O'Reilly, D., Rosato, M., Connolly, S.: Urban and rural variations in morbidity and mortality in Northern Ireland. BMC Public Health 7(1), 123 (2007)

54. Pampalon, R., Hamel, D.: A deprivation index for health planning in Canada. Crhronic Dis. Canada 9(24), 178–191 (2009)

55. Pasetto, R., Sampaolo, L., Pirastu, R.: Measures of material and social circumstances to adjust for deprivation in small area studies of environment and health: review and perspective. Ann. Ist. Super. Sanità 2010, 185–197 (2010)

56. Peña, J.: Problemas de la mediciòn del bienestar y conceptos afnes (Una aplicacion al caso espanol) (1977)

57. Pérez-Mayo, J.: On measuring deprivation in Spain: comparing some techniques. Badajoz, Spain: Department of Applied Economy, University of Extremadura (2002)

58. Petrelli, A., Gnavi, R., Marinacci, C., Costa, G.: Socioeconomic inequalities in coronary heart disease in Italy: a multilevel population-based study. Soc. Sci. Med. 63(2), 446–456 (2006)

59. Regidor, E., Lostao, L., Pascual, C., Martínez, D., Calle, M.E., Domínguez, V.: Income in large residential areas and premature mortality in six countries of the European Union. Health Policy 75(1), 99–108 (2005)

60. Rodríguez Martín, J.A.: An index of child health in the least developed countries (LDCs) of Africa. Soc. Indic. Res. 105, 309–322 (2012)

61. Ross, N.A., Oliver, L.N., Villeneuve, P.J.: The contribution of neighbourhood material and social deprivation to survival: a 22-year follow-up of more than 500,000 Canadians. Int. J. Env. Res. Public Health 10(4), 1378–1391 (2013). https://doi. org/10.3390/ijerph10041378

62. Saint-Jacques, N., Dewar, R., Cui, Y., Parker, L., Dummer, T.J.B.: Premature mortality due to social and

material deprivation in Nova Scotia, Canada. Int. J. Equity Health 13(1), 94 (2014). https ://doi.org/10.1186/s12939-014-0094-2

63. Salmond, E., Crampton, P.: Élaboration de l'indice de défavorisation de le NouvelleZélande (NZDep) et son rôle d'instrument de la politique nationale. Revue Canadienne de Saté Plublique, p. S7–S11 (2012)

64. Santana, P., Costa, C., Marí-Dell'Olmo, M., Gotsens, M., Borrell, C.: Mortality, material deprivation and urbanization: exploring the social patterns of a metropolitan area. Int. J. Equity Health 14(1), 1–13 (2015)

65. Schuurman, N., Bell, N., Dunn, J.R., Oliver, L.: Deprivation indices, population health and geography: an evaluation of the spatial efectiveness of indices at multiple scales. J. Urban Health Bull. N. Y. Acad. Med. 84(4), 591–603 (2007)

66. Sen, A.: On ethics and economics. Basil Blackwell, Oxford (1987)

67. Senior, M., Williams, H., Higgs, G.: Urban–rural mortality diferentials: controlling for material deprivation. Soc. Sci. Med. 51(2), 289–305 (2002)

68. Smits, J., Westert, G., Van den Bos, G.: Socioeconomic status of very small areas and stroke incidence in the Netherlands. J. Epidemiol. Community Health 56, 637–640 (2002)

69. Somarriba, N., Pena, B.: Synthetic indicators of quality of life in Europe. Soc. Indic. Res. 94, 115–133 (2009)

70. Somarriba, N., Zarzosa, P.: Quality of life in Latin America: a proposal for a synthetic indicator. In: Tonon G. (ed.) Indicators of quality of Life in Latin America, social indicators research series, vol. 62, pp. 19–56 (2016)

71. Somarriba, N., Zarzosa, P., Pena, B.: The economic crisis and its efects on the quality of life in the European Union. Soc. Indic. Res. 120, 323–343 (2015)

72. Spadea, T., Zengarini, N., Kunst, A., Zanetti, R., Rosso, S., Costa, G.: Cancer risk in relationship to diferent indicators of adult socioeconomic position in Turin, Italy. Cancer Causes Control 21(7), 1117–1130 (2010)

73. Strand, B.H., Groholt, E.K., Steingrimmsdottir, O.A., Blakely, T., Graf-Iversen, S., Naess, O.: Educational inequalities in mortality over four decades in Norway: prospective study of middle-aged men and women followed for cause specifc mortality, 1960–2000. BMJ 340, c654 (2010)

74. Stringhini, S., Sabia, S., Shipley, M., Brunner, E., Nabi, H., Kivimaki, M., et al.: Association of socioeconomic position with health behaviors and mortality. JAMA 303(12), 1159–1166 (2010)

75. Su, S., Gong, Y., Tan, B., Pi, J., Weng, M., Cai, Z.: Area social deprivation and public health: analyzing the spatial non-stationary associations using geographically weighed regression. Soc. Indic. Res. (2016). https://doi.org/10.1007/s11205-016-1390-6

76. Testi, A., Ivaldi, E.: Material versus social deprivation and health: a case study of an urban area. Eur. J. Health Econ. 10(3), 323 (2009)

77. Torsheim, T., Currie, C., Boyce, W.: Material deprivation and selfrated health: a multilevel study of adolescents from 22 European and North American countries. In: Social Science and Medicine, vol. 59, pp. 1–12 (2004)

78. Townsend, P., Phillimore, P., Battie, A.: Health and deprivation: inequality and the North. Croom Helm, London (1988)

79. Townsend, P. Deprivation. Journal of Social Policy, p. 16/2 pp. 125-146 (1987).

80. Ward, J.: Hierarchical grouping to optimize an objective function. J. Am. Stat. Assoc. 58(301), 236–244 (1963)

Appendix 1

Table A1.1 Analysis of main components on the 14 variables, varimax rotation, Kaiser Normalization, factor structure matrix

Variables	Factor I	Factor II	Factor III	Factor IV
Low education	**0.973**	0.120	0.005	0.023
Income	**0.883**	0.285	0.175	0.293
Overcrowding	**0.873**	0.156	- 0.063	- 0.017
Born on the Islands or	**0.751**	0.337	0.236	- 0.134
Old Age Index	- 0.201	- **0.884**	- 0.115	- 0.128
Single elders	- 0.360	- **0.830**	- 0.049	0.085
Resident foreigners	0.025	**0.618**	0.338	0.614
Foreigners in school	0.273	**0.598**	0.285	0.505
Families renting	0.309	0.268	**0.739**	- 0.147
Divorced	- 0.400	- 0.193	**0.762**	0.088
Unemployment	0.603	0.306	**0.654**	0.010
Buildings in Mediocre or bad state	0.148	0.201	**0.648**	0.107
Single-parent families	- 0.176	- 0.035	0.348	- **0.789**
Single-person families	- 0.390	0.019	0.459	**0.753**

Appendix 2

Table A2.1 Explained total variance: extraction method analysis of main components

	Total	% of variance	Cumulative %
1	4.22	29.444	29.444
2	2.687	19.191	48.635
3	2.539	18.132	66.768
4	1.963	14.020	80.787

Appendix 3

Table A3.1 MPI Index: Material and Social Deprivation

UU	UU Name	MPI Material Deprivation	UU	UU Name	MPI Social Deprivation
39	S.VINCENZO	87,07	19	MOREGO	83,47
40	CARIGNANO	87,16	12	BORZOLI OVEST	89,06
65	PUGGIA	87,43	20	S.QUIRICO	89,64
38	MANIN	88,64	52	S.EUSEBIO	92,48
37	CASTELLETTO	88,76	56	BAVARI	92,65
64	LIDO	88,92	54	DORIA	94,05
62	ALBARO	88,95	23	CAMPI	94,11
41	FOCE	89,33	57	S.DESIDERIO	94,21
63	S.GIULIANO	89,46	13	BORZOLI EST	94,36
36	S.NICOLA	90,15	8	MULTEDO	94,77
68	QUARTARA	90,70	53	MOLASSANA	95,44
42	BRIGNOLE	91,51	55	PRATO	95,50
70	QUINTO	91,98	17	BEGATO	95,63
1	CREVARI	93,42	11	CALCINARA	96,14
7	PEGLI	94,26	58	APPARIZIONE	96,36
67	QUARTO	95,01	4	PALMARO	96,49
56	BAVARI	95,23	36	S.NICOLA	97,20
71	NERVI	95,33	7	PEGLI	97,22
27	BELVEDERE	95,96	48	FORTE QUEZZI	97,44
69	CASTAGNA	96,40	51	MONTESIGNANO	97,85
66	STURLA	96,41	37	CASTELLETTO	98,07
58	APPARIZIONE	97,19	6	CASTELLUCCIO	98,22
25	S.GAETANO	97,33	38	MANIN	98,44
61	S.MARTINO	97,66	3	CA NUOVA	98,45
6	CASTELLUCCIO	97,75	1	CREVARI	98,46
21	PONTEDECIMO	98,13	68	QUARTARA	98,59
43	S.AGATA	98,33	30	S.TEODORO	98,61
8	MULTEDO	98,43	45	QUEZZI	98,61
4	PALMARO	99,00	16	TEGLIA	98,85
44	S.FRUTTUOSO	99,18	2	VOLTRI	98,86
32	OREGINA	99,52	65	PUGGIA	98,89
10	S.GIOVANNI BATTISTA	99,58	21	PONTEDECIMO	98,95
2	VOLTRI	99,87	10	S.GIOVANNI BATTISTA	99,37
55	PRATO	100,06	34	MADDALENA	99,43
30	S.TEODORO	100,12	5	PRA'	99,45
11	CALCINARA	100,19	22	CORNIGLIANO	99,46
12	BOZOLI OVEST	100,49	69	CASTAGNA	99,51
60	CHIAPPETO	100,52	33	PRE'	99,68
15	RIVAROLO	100,77	70	QUINTO	99,87
28	S.BARTOLOMEO	101,07	63	S.GIULIANO	99,95
59	BORGORATTI	101,34	35	MOLO	100,21
9	SESTRI	101,37	31	LAGACCIO	100,35
53	MOLASSANA	101,41	14	CERTOSA	100,39

48	FORTE QUEZZI	101,42	40	CARIGNANO	100,49
31	LAGACCIO	101,61	15	RIVAROLO	100,51
47	MARASSI	101,88	62	ALBARO	100,53
57	S.DESIDERIO	101,93	27	BELVEDERE	100,57
34	MADDALENA	102,14	9	SESTRI	100,80
51	MONTESIGNANO	102,47	60	CHIAPPETO	101,09
46	FEREGGIANO	102,60	64	LIDO	101,17
26	SAMPIERDARENA	102,82	32	OREGINA	101,42
23	CAMPI	103,43	67	QUARTO	101,49
50	PANTALEO	103,50	46	FEREGGIANO	101,81
52	S.EUSEBIO	103,56	71	NERVI	101,95
5	PRA'	103,76	29	ANGELI	101,97
13	BORZOLI EST	103,90	66	STURLA	102,31
14	CERTOSA	103,99	25	S.GAETANO	102,52
54	DORIA	104,08	44	S.FRUTTUOSO	102,63
16	TEGLIA	104,24	49	PARENZANO	102,70
45	QUEZZI	104,96	26	SAMPIERDARENA	102,92
20	S.QUIRICO	105,00	61	S.MARTINO	103,06
24	CAMPASSO	105,10	39	S.VINCENZO	103,26
35	MOLO	105,20	59	BORGORATTI	103,55
22	CORNIGLIANO	105,68	18	BOLZANETO	103,56
49	PARENZANO	105,76	28	S.BARTOLOMEO	104,32
17	BEGATO	108,37	42	BRIGNOLE	104,67
33	PRE'	108,43	41	FOCE	105,12
29	ANGELI	109,28	24	CAMPASSO	105,24
18	BOLZANETO	109,45	50	PANTALEO	105,32
19	MOREGO	109,90	43	S.AGATA	105,36
3	CA NUOVA	126,24	47	MARASSI	107,36

Appendix 4

Table A4 DP2 Index: Material and Social Deprivation

UU	UU Name	DP2 Material Deprivation	UU	UU	DP2 Social Deprivation
40	CARIGNANO	3,60	19	MOREGO	0,41
39	S.VINCENZO	3,68	20	S.QUIRICO	2,33
37	CASTELLETTO	3,83	56	BAVARI	2,99
65	PUGGIA	4,06	12	BOZOLI OVEST	3,68
36	S.NICOLA	4,12	13	BORZOLI EST	4,05
41	FOCE	4,66	35	MOLO	4,16
38	MANIN	4,91	34	MADDALENA	4,17
63	S.GIULIANO	4,96	54	DORIA	4,23
62	ALBARO	5,05	55	PRATO	4,45
64	LIDO	5,19	52	S.EUSEBIO	4,56
42	BRIGNOLE	5,24	58	APPARIZIONE	4,68
68	QUARTARA	5,53	11	CALCINARA	4,70
70	QUINTO	6,10	33	PRE'	4,72
7	PEGLI	6,78	4	PALMARO	4,77
56	BAVARI	7,00	53	MOLASSANA	4,87
1	CREVARI	7,21	8	MULTEDO	4,99
67	QUARTO	7,26	36	S.NICOLA	5,03
71	NERVI	7,48	7	PEGLI	5,10
25	S.GAETANO	7,49	23	CAMPI	5,13
27	BELVEDERE	7,68	38	MANIN	5,26
21	PONTEDECIMO	7,91	57	S.DESIDERIO	5,32
58	APPARIZIONE	7,92	21	PONTEDECIMO	5,40
43	S.AGATA	8,22	37	CASTELLETTO	5,40
6	CASTELLUCCIO	8,26	30	S.TEODORO	5,46
66	STURLA	8,34	48	FORTE QUEZZI	5,47
44	S.FRUTTUOSO	8,56	2	VOLTRI	5,51
69	CASTAGNA	8,59	51	MONTESIGNANO	5,68
8	MULTEDO	8,60	1	CREVARI	5,71
4	PALMARO	8,62	5	PRA'	5,73
31	LAGACCIO	9,00	17	BEGATO	5,75
32	OREGINA	9,00	45	QUEZZI	5,75
61	S.MARTINO	9,03	68	QUARTARA	5,84
2	VOLTRI	9,14	29	ANGELI	5,85
11	CALCINARA	9,16	16	TEGLIA	5,96
30	S.TEODORO	9,22	14	CERTOSA	5,98
10	S.GIOVANNI BATTISTA	9,31	27	BELVEDERE	5,99
55	PRATO	9,42	9	SESTRI	6,00
15	RIVAROLO	9,51	22	CORNIGLIANO	6,03
12	BOZOLI OVEST	9,63	31	LAGACCIO	6,05
9	SESTRI	9,65	63	S.GIULIANO	6,08
53	MOLASSANA	9,78	6	CASTELLUCCIO	6,11
60	CHIAPPETO	9,99	15	RIVAROLO	6,14
51	MONTESIGNANO	10,01	10	S.GIOVANNI BATTISTA	6,14

48	FORTE QUEZZI	10,09	71	NERVI	6,17
28	S.BARTOLOMEO	10,13	60	CHIAPPETO	6,21
26	SAMPIERDARENA	10,33	39	S.VINCENZO	6,29
57	S.DESIDERIO	10,36	70	QUINTO	6,31
47	MARASSI	10,49	64	LIDO	6,37
59	BORGORATTI	10,62	69	CASTAGNA	6,47
5	PRA'	10,63	32	OREGINA	6,47
46	FEREGGIANO	10,71	40	CARIGNANO	6,47
52	S.EUSEBIO	10,77	49	PARENZANO	6,50
23	CAMPI	11,11	46	FEREGGIANO	6,56
14	CERTOSA	11,20	66	STURLA	6,57
22	CORNIGLIANO	11,34	65	PUGGIA	6,62
34	MADDALENA	11,37	62	ALBARO	6,66
16	TEGLIA	11,38	44	S.FRUTTUOSO	6,74
50	PANTALEO	11,40	67	QUARTO	6,77
24	CAMPASSO	11,47	61	S.MARTINO	6,92
54	DORIA	11,53	25	S.GAETANO	7,01
13	BORZOLI EST	11,77	59	BORGORATTI	7,01
45	QUEZZI	11,82	26	SAMPIERDARENA	7,01
20	S.QUIRICO	12,00	28	S.BARTOLOMEO	7,06
49	PARENZANO	12,03	18	BOLZANETO	7,20
35	MOLO	12,73	42	BRIGNOLE	7,26
18	BOLZANETO	13,39	43	S.AGATA	7,27
17	BEGATO	13,41	50	PANTALEO	7,33
29	ANGELI	14,56	41	FOCE	7,59
33	PRE'	14,77	24	CAMPASSO	7,95
19	MOREGO	15,16	47	MARASSI	8,12
3	CA NUOVA	22,57	3	CA NUOVA	8,44

Appendix 5

Table A5 Cluster analysis results: classes of deprivation

Class of Material Deprivation	UU	UU Name	MPI Material deprivation	Class of Social Deprivation	UU	UU Name	MPI Social deprivation
1	39	S. Vincenzo	87.07	1	19	Morego	83.47
1	40	Carignano	87.16	1	12	Borzoli ovest	89.06
1	65	Puggia	87.43	1	20	S. Quirico	89.64
1	38	Manin	88.64	1	52	S. Eusebio	92.48
1	37	Castelletto	88.76	1	56	Bavari	92.65
1	64	Lido	88.92	1	54	Doria	94.05
1	62	Albaro	88.95	1	23	Campi	94.11
1	41	Foce	89.33	1	57	S. Desiderio	94.21
1	63	S. Giuliano	89.46	1	13	Borzoli Est	94.36
1	36	S. Nicola	90.15	1	8	Multedo	94.77
1	68	Quartara	90.7	1	53	Molassana	95.44
1	42	Brignole	91.51	1	55	Prato	95.5
1	70	Quinto	91.98	1	17	Begato	95.63
2	1	Crevari	93.42	2	11	Calcinara	96.14
2	7	Pegli	94.26	2	58	Apparizione	96.36
2	67	Quarto	95.01	2	4	Palmaro	96.49
2	56	Bavari	95.23	2	36	S. Nicola	97.2
2	71	Nervi	95.33	2	7	Pegli	97.22
2	27	Belvedere	95.96	2	48	Forte quezzi	97.44
2	69	Castagna	96.4	2	51	Montesignano	97.85
2	66	Sturla	96.41	2	37	Castelletto	98.07
2	58	Apparizione	97.19	2	6	Castelluccio	98.22
2	25	S. Gaetano	97.33	2	38	Manin	98.44
2	61	S. Martino	97.66	2	3	Ca Nuova	98.45
2	6	Castelluccio	97.75	2	1	Crevari	98.46
2	21	Pontedecimo	98.13	2	68	Quartara	98.59
2	43	S. Agata	98.33	2	30	S. Teodoro	98.61
2	8	Multedo	98.43	2	45	Quezzi	98.61
3	4	Palmaro	99	3	16	Teglia	98.85
3	44	S. Fruttuoso	99.18	3	2	Voltri	98.86
3	32	Oregina	99.52	3	65	Puggia	98.89
3	10	S. Giovanni battista	99.58	3	21	Pontedecimo	98.95
3	2	Voltri	99.87	3	10	S. Giovanni battista	99.37
3	55	Prato	100.06	3	34	Maddalena	99.43
3	30	S. Teodoro	100.12	3	5	Pra’	99.45
3	11	Calcinara	100.19	3	22	Cornigliano	99.46
3	12	Borzoli ovest	100.49	3	69	Castagna	99.51
3	60	Chiappeto	100.52	3	33	Pre’	99.68
3	15	Rivarolo	100.77	3	70	Quinto	99.87
3	28	S. Bartolomeo	101.07	3	63	S. Giuliano	99.95
3	59	Borgoratti	101.34	3	35	Molo	100.21
3	9	Sestri	101.37	3	31	Lagaccio	100.35
3	53	Molassana	101.41	3	14	Certosa	100.39
3	48	Forte quezzi	101.42	3	40	Carignano	100.49
3	31	Lagaccio	101.61	3	15	Rivarolo	100.51
3	47	Marassi	101.88	3	62	Albaro	100.53
3	57	S. Desiderio	101.93	3	27	Belvedere	100.57
3	34	Maddalena	102.14	3	9	Sestri	100.8
3	51	Montesignano	102.47	3	60	Chiappeto	101.09
3	46	Fereggiano	102.6	3	64	Lido	101.17
3	26	Sampierdarena	102.82	3	32	Oregina	101.42
4	23	Campi	103.43	4	67	Quarto	101.49
4	50	Pantaleo	103.5	4	46	Fereggiano	101.81
4	52	S. Eusebio	103.56	4	71	Nervi	101.95

4	5	Pra’	103.76	4	29	Angeli	101.97
4	13	Borzoli EST	103.9	4	66	Sturla	102.31
4	14	Certosa	103.99	4	25	S. Gaetano	102.52
4	54	Doria	104.08	4	44	S. Fruttuoso	102.63
4	16	Teglia	104.24	4	49	Parenzano	102.7
4	45	Quezzi	104.96	4	26	Sampierdarena	102.92
4	20	S. Quirico	105	4	61	S. Martino	103.06
4	24	Campasso	105.1	4	39	S. Vincenzo	103.26
4	35	Molo	105.2	4	59	Borgoratti	103.55
4	22	Cornigliano	105.68	4	18	Bolzaneto	103.56
4	49	Parenzano	105.76	4	28	S. Bartolomeo	104.32
5	17	Begato	108.37	5	42	Brignole	104.67
5	33	Pre’	108.43	5	41	Foce	105.12
5	29	Angeli	109.28	5	24	Campasso	105.24
5	18	Bolzaneto	109.45	5	50	Pantaleo	105.32
5	19	Morego	109.9	5	43	S. Agata	105.36
5	3	Ca Nuova	126.24	5	47	Marassi	107.36

Health and Deprivation: A New Approach Applied to 32 Argentinian Urban Areas

1. Introduction

That poverty and deprivation are relative in both time and space is widely agreed since the pioneering work of Townsend (1962) and Runciman (1966). What they consist of varies over time and is dependent on the social situation in which it is experienced. Cognitive, economical, and relational, all are possible dimensions of the concept of 'deprivation' or 'poverty'. From these first lines, it is clear that to measure poverty or deprivation accurately, surveys or censuses must be used that establish both the normal or average standard of living of the majority in a society/culture and any "enforced" reductions in this standard due to lack of resources (Gordon 1995). The objective is to classify high-risk populations, or the areas in which they live, to inform policy makers to supply a combination of social and material resources to enhance their quality of life.

The paper is divided into several parts. In the first, we provide an overview of the concepts of deprivation and poverty; the second describes the variables we used and the technics of aggregation. The results, derived from the application of the proposed methodological approach, are then presented and discussed in the last part of the paper.

2. Poverty and Deprivation

Poverty has been one of the phenomena that attracted the most attention of social research- ers—suffice it to think to the work of Booth (1889), Du Bois (1899), Rowntree (1901, 1941), Rowntree and Kendall (1917), and Bowley (1915)—and one of the first for which efforts have been made to define it, mainly through income levels and the amount of the expense. Many authors (Cullen 1979; Rose 1972; Simey and Simey 1960)—erroneously— consider Charles Booth the father of the concept of "line of poverty"[1], even if he never explicates how he arrived to determine that threshold:

> By the word 'poor' I mean to describe those who have a fairly regular though bare income, such as 18 s. [shillings] to 21s. per week for a moderate family, and by 'very poor' those who fall below this standard, whether from chronic irregularity of work, sickness, or a large number of young children. […] The proportion of the population shown to be above the line of poverty, I make to be 65 per cent, that on the line 22 per cent, while those falling chronically below it into the region of distress are 13 per cent (Booth 1887, pp. 328, 375).

Rowntree exposed in a more complex way his proposal; while recognizing the importance of the multiple facets of the concept, he kept a focus on physical health:

> Expenditure needful for the development of the mental, moral, and social sides of human nature will

[1] As noted by Gillie (1996, p. 728), the criterion was already mentioned in the 1870 Elementary Education Act and used by McDougall (1885, pp. 3–4). Curiously, in the works of Booth there is no mention of the expression 'poverty line', "although it is used synonymously with 'line of poverty' in one of his notebooks compiled in 1887 (Gillie 1996, p. 715). If Gillie merely deny the use by Booth of the expression 'poverty line', Vaughan (2007, p. 242) goes as far to state that in his works there is so sign of the concept, neither

not be taken into account at this stage of the inquiry. Nor in thus estimating the poverty line will any account be taken of expenditure for sick clubs or insurance. We confine our attention at present simply to an estimate of minimum necessary expenditure for the maintenance of merely physical health. This may be discussed under three heads: food, house rent (including rates), household sundries (such as clothing, light, fuel, etc.) (1901, pp. 87–88).

However, poverty, as stated by Townsend (1979), "can be defined objectively and applied consistently only in terms of the concept of relative deprivation". "Deprivation may be defined as a state of observable and demonstrable disadvantage relative to the local community or the wider society or nation to which an individual, family or group belongs. The idea has come to be applied to conditions (that is, physical, environmental and social states or circumstances) rather than resources and to specific and not only general circumstances, and therefore can be distinguished from the concept of poverty" (Townsend 1987, p. 5). Following these arguments, the line of poverty or the disadvantage relative to the local com- munity depends from the characteristics of each country and from the needs considered as fundamental in each society. In this sense, deprivation is not context free. The point was already proposed by Adam Smith:

By necessaries I understand not only the commodities which are indispensably necessary for the support of life, but whatever the custom of the country renders it indecent for creditable people, even of the lowest order, to be without. A linen shirt, for example, is, strictly speaking, not a necessary of life. The Greeks and Romans lived, I suppose, very comfortably though they had no linen. But in the present times, through the greater part of Europe, a creditable day-labourer would be ashamed to appear in public without a linen shirt, the want of which would be supposed to denote that disgraceful degree of poverty which, it is presumed, nobody can well fall into without extreme bad conduct. (1776, pp. 399–400).

Pursuing these arguments, deprivation can be associated to lack of education, to bad conditions of dwellings, and even the lack of social relations. The well-known approach by Sen shifts its attention from the deprivation of income to the deprivation of capabilities: "Poverty is not a matter of low well-being, but the inability to pursue well-being precisely because of the lack of economic means" (1992, p. 110).

Many scientists dispute the use of composite indicators that lead to the determination of a single value for each geographic area, preferring the so-called dashboard. In this case, it is possible to identify various dimensions of the phenomenon, all relevant, without that they are further aggregated. From the statistical point of view, it is an incontrovertible choice but from the standpoint of political and media is a heavy limitation. The easy-disclosure in the media and the immediate understanding by the user are certainly the strengths of a unique index (Mazziotta and Pareto 2013).

Deprivation indexes are a quite simple, cheap instrument to measure socioeconomic differentials: they are generally made up of census indicators, easily available, combined using different types of statistical procedures (Bartley and Blane 1994). The history of census-based area deprivation indices dates to at least until 1971, when the Department of the Environment (DOE) in the United Kingdom used data taken from the census to identify localities where a high proportion of households were exposed to adverse social and economic conditions (Department of Health and Social Security 1988). In the last three decades, deprivation indexes based upon the characteristics of areas of residence were widely used in epidemiology and public health (Davey Smith et al. 2001; Salmond and Crampton 2002). Developed at first in UK (Carstairs and Morris 1991; Forrest and Gordon 1993; Jarman 1983; Townsend et al. 1988), nowadays they are used in the whole world such as Spain (Pérez-Mayo 2003), Ireland (Kelleher et al. 2002), Netherlands (Smits et al. 2002), France (Havard et al. 2008), Italy (Caranci et al. 2010; Landi et al. 2018a), New Zealand (Salmond and Crampton 2002), North America (Arcaya et al. 2015; Pampalon et al. 2009a, 2012; b; Torsheim et al. 2004) and Argentina (Durán and Condorí 2019). They are used to measure the correlation between deprivation and mortality (Benach and Yasui 1999; Bond Huie et al. 2003; Santana et al. 2015), waiting-time inequalities for health services (Landi et al. 2018b) and the incidence of specific diseases (Andersen et al. 2014; Jackson et al. 2014; Spadea et al. 2010; Su et al. 2017). These indexes have emphasized, among other things, the multidimensional aspect of urban deprivation as related to its absolute versus relative nature and its material versus social content.

The use of deprivation's indices on a geographical base is possible on the assumptions that (1) we can infer individual's characteristics from those of the area to which he belongs, and that (2) socio-economic conditions of that geographical area can systematically deter- mine different risks of morbidity and mortality under the same conditions[2]—context effect: see Macintyre et al. (2002). An adequate index is possible only if we explicit the meaning of a phenomena determined by the social and temporal context (Jesuit et al. 2003). Having a certain amount of resources in a context economically and socially active has a different meaning than disposing the same amount in a stagnating reality, for the individual (or the family) may have a different degree of capabilities in developing his life.

Although health inequalities depend on socioeconomic circumstances (Carstairs and Morris 1991; Sloggett and Joshi 1994) there is still a lack of understanding of possible reasons underlying the deprivation indexes' good performances (Carstairs 2001). Conse- quently, like Gordon has suggested (1995), answering the key question "which index is the best?" is often not easy. Indexes differ in items included and in the kind of deprivation they consider: following Townsend (1987), it is possible to distinguish between *material* and *social* form of deprivation. The former "entails the lack of goods, services, resources, amenities and physical environment which are customary, or at least widely approved in the society under consideration" while the latter "is non-participation in the roles, relation- ships, customs, functions, rights and responsibilities implied by a member of a society and its sub-groups." The components of the indexes are direct indicators, when representing conditions or states, and indirect indicators, expressing the victims of those conditions or states (Townsend 1987), the possession of commodities being valuable only to the extent that it enables the person to do or be a range of things (Sen 1987).

3. The Argentinian Context

[2] Both assumptions are called into question: the first by those who warn against ecological fallacy (Lancaster and Green 2002), the second by those who think that the effect is irrelevant and not quantifiable with only census information (Sloggett and Joshi 1994): people inhabiting in a geographical area with given characteristics may have different needs, under the same conditions

Argentina is politically divided into twenty-three *provincias* (provinces) and one autonomous city, Ciudad Autónoma de Buenos Aires, the Federal Capital of the nation. Moreover, five different geographical regions are distinguished in its territory. The "NOA" region includes the northwestern provinces: Jujuy, Salta, Tucumán, Santiago del Estero, Catamarca and La Rioja. The "NEA" region includes the northeastern provinces of the country: Formosa, Chaco, Misiones, Corrientes and Entre Ríos. The "Cuyo" region includes the provinces located in the mountainous area in the center-west of the country: Mendoza, San Juan and San Luis. The "Pampeana" region includes the City of Buenos Aires, the province of Buenos Aires, Córdoba, Santa Fe and La Pampa. The "Patagonia" region includes the southern provinces of Argentina: Neuquén, Río Negro, Chubut, Santa Cruz and Tierra del Fuego. Besides, each province is divided into departments, a smaller unit, with a total of 5127 of them throughout its territory.

According to the information from the Encuesta Permanente de Hogares[3] (EPH, Permanent Household Survey), in the second half of 2017, the percentage of households below the poverty line was 17.9%, the 25.7% of the population. Within this percentage, 3.5% are indigent households, which represents 4.8% of the population. According to these figures, 1,611,001 households, which include 7,079,764 people, were below the poverty line (LP); and of that group, 316,350 households were, in turn, under the indigence line (LI), which included 1,323,747 indigent people. Here we encounter a big problem when studying poverty in Argentina, because the survey used to measure poverty does not have a national coverage (Eguía 2017). The Encuesta Permanente de Hogares (EPH) is carried out in only 32 main urban areas (Aglomerados) and is representative of only 6.3 out of 10 residents.[4]

The available data allow us to confirm the existence of a demographic dimension of inequalities and poverty that most affect poorest people. This demographic vulnerability is expressed through a rate of fertility, adolescent and unwanted fecundity higher than in the rest of the population, with morbidity and maternal-infant mortality rates much higher. It can also be found many young or old women as head of households as well as single parent households. This fact affects not only the accumulation of economic capital, but also the human capital. Thus, poor people are disallowed from developing their capacities, deteriorating their chances of getting out of poverty.

"The demographic dynamics of poverty" (Pantelides and Moreno 2009) affects in a special way groups that are particularly neglected, as indigenous peoples. Their lack of access to basic services such as education and health (including sexual and reproductive health) is at the base of their lower life expectancy. Thus, the

[3] The Permanent Household Survey (EPH), undertaken monthly by the National Institute of Statistics and Censuses (INDEC), is a continuous survey that takes place in 32 urban areas of Argentina (provincial capitals and cities above 100,000 inhabitants). The Survey produces quarterly data for agglomerates of the population and is the basic source for these kinds of studies and statistical simulations. INDEC publishes the data quarterly.

[4] Furthermore, "since 2010, other urban areas with population over 2000 inhabitants (where a third of people live), are also sampled in the third quarter of the year. However, poverty estimates are not extended to this sample and this dataset is not yet available since 2014, so it is impossible to calculate the poverty rate for this group. In addition, although the rural population represents a relatively small fraction of the total population (9%, according to the 2010 Population, Household, and Housing Census), an urban-only household survey, even if the smaller urban areas were included, may underestimate national poverty: rural households are twice as likely to have at least one unsatisfied basic need (18.2% compared to 8.3% nationally). The exclusion of rural areas also means that certain vulnerable groups are systematically excluded: for example, there are proportionally twice as many indigenous people in rural areas. In addition, available datasets for urban areas (such as the EPH) do not include questions that would allow identifying indigenous peoples or afro descendants. The Population Census is the only source of information to characterize both rural population and ethnic groups in terms of living standards (dwelling characteristics, education attainment, among others). Yet this source does not allow an assessment of whether they are more deprived in terms of incomes or employment than others in the population" (World Bank 2018, p. 90).

Patagonia area is at the minimum rate, a 6.8% lower than the national average. In contrast, the Northeast region shows the highest proportion of poverty: 29% higher than the national average. This region is mainly of aboriginal or peasant population (Bolsi et al. 2005).

In reference to some of the dimensions of deprivation, in 2017 the incidence of food insecurity reached 4.1% of urban households in Argentina, affecting about 2 out of every 10 households residing in shantytowns and little more than 10% of households with very low socioeconomic status. In this same year, 20% of households did not have health cover- age and had to cut expenses on medical care or medication (Salvia 2017).

The deprivation in access to basic services was 30% in 2017, showing a regular decline since 2010 when it was 36%. Probably due to the public investment made between 2010 and 2017, it was the households of low and low-middle strata that saw their socio-urban integration improved. In any case, almost 7 of every 10 homes in shantytowns did not have access to basic services. Almost half of the households of marginal working class and very low socioeconomic status were in the same situation. There is an association between deprivation and lower social class. The working classes tend to suffer from deprivation, and all ethnic and minority groups have higher levels of deprivation than the white majority (Fieldhouse and Tye 1996). In the dimension of access to a decent house, the deficit reached 17.6% in 2017. It should be noted that 6 of every 10 homes in shantytowns are in poor housing conditions and 4 out of every 10 households belong to very low socioeconomic stratum (Salvia 2017).

4. Methods

The analysis of the literature offers several solutions to deduce a priori what should be the most appropriate variables to be included within an index even if the choice obviously is influenced by the availability of data, the purpose of the indicator and the significance criteria (Carstairs and Morris 1991; Fu et al. 2015; Ivaldi et al. 2016; Jarman 1983, 1984; Landi et al. 2018a; Noble et al. 2003; Pampalon et al. 2009a, b; Parra Saiani 2009; Townsend et al. 1988; Townsend 1987; Whelan et al. 2010); furthermore, some of the variables have changed over time. One objective of this paper is to propose a new approach to construct an estimate Argentinian's deprivation, based on the use of indicators from which the main dimensions of deprivation are derived from the literature (Table 1). This approach is applied to the 32 Aglomerados to examine the spatial structure of deprivation dimensions as well as its intensity in 2017—corresponding with the latest EPH.

Our analysis is based on data that allows delimiting small geographical units, with homogenous social and economic characteristics (Pampalon et al. 2009a, b; Schuurman et al. 2007). The data are extracted from 2017 EPH, a sample survey about the characteristics of families. It covers 32 urban agglomerations with more than 100,000 inhabitants, where 70% of the urban population of Argentina lives (Pampalon et al. 2009a, b).

From the analysis of the literature, we identified a group of 14 indicators, already used as components of the various deprivation indices, even though their respective operational definitions may be partially different. Table 1 lists and defines the variables most commonly used in other indices of deprivation.

Deprivation has a significant influence on life expectancy in many studies (Benach and Yasui 1999; Bond

Huie et al. 2003; Ivaldi and Testi 2011; Lalloué et al. 2013; Landi et al. 2018a, b; Morris and Carstairs 1991; Santana et al. 2015; Stringhini et al. 2010; Townsend et al. 1988). Therefore, we have selected six potential indicators on the basis of their correlation (Pearson Correlation Coefficient ρ) to life expectancy (See Appendix Fig. 2, Tables 4, 5). Variables presenting an inverse linear relation with life expectancy at birth are low education (− .285), income (− .348), overcrowding (−.365), buildings in mediocre or bad state (−.302), single-parent families (−.452).

To justify the exclusion of the rest of the variables we may advance several rea- sons. As for the Born on the North-East, although those who were born in this region (along with those who were born in the northwest region) belong to areas with higher internal migration rates (mainly to Buenos Aires) and have subsistence economies or monocultures, their life expectancy does not differ from the rest of the population.

When considering foreigners in general, it should be noted that according to data from the 2010 national census the number of immigrants is low (3.7%), and almost 70% of them were born in neighbouring countries or in non-bordering countries (Peru, 8.5%) but culturally close to the Argentine community. Unlike immigration in Europe (9.8%) or in North America (14.9%), the most immigrants in Argentina have a mestizo aspect with physical and cultural features that make them more similar to the populations of the provinces of the northwest and northeast of Argentina, which relativizes the xenophobic phenomena. The divorced, they do not exceed 6% of the total population. For the rest of the variables (Old age index, Single elders, Families renting, Unemployment), it could be considered the existence of different family arrangements, such as large households—a family strategy aimed at increasing the labour force available in households with scarce resources—or the creation of family/social networks that allow the flow of material or emotional goods and services, and whose functioning is especially important in situations of family crisis (loss of employment, serious economic problems, serious illnesses, disabilities, deaths, etc.).

Hence, our index is composed of four variables aimed at represent material deprivation (Low education, Income, Overcrowding, Buildings in mediocre or bad state), and of one variable of social deprivation (Single parent families), highlighting a context defined more by differences in the “objective” living conditions than by social exclusion (Baumann et al. 2007; Julkunen 2002; Testi and Ivaldi 2009).

Table 5 Indicators and variables used in other indices of deprivation

Variable	Definition	Also, used in deprivation index
Low education	Population of residents who completed junior school, elementary school or are illiterate, out of all residents aged 15 years and above	Pampalon and Hamel (2009)
Income	Reciprocal of the average income of the resident population	Julkunen (2002)
Overcrowding	Average number of residents per room	Jarman (1983), Townsend (1987), Forrest and Gordon (1993), Duran e Condori (2019)
Born on the Nord-East	Residents born on the Northest Argentina	Baumann (2007) uses the analogous variable at the European level (nationality not included in Western Europe)
Old age index	Ratio of people aged 65 years and above, to people aged 15 years and below	Julkunen (2002), Pampalon and Hamel (2009), Baumann (2007)
Single elders	Number of people aged 65 years and above, and who live alone	Jarman (1983), Forrest and Gordon (1993), Pampalon and Hamel (2009)
Resident foreigners	Number of people from the total population who were born abroad	Jarman (1983)
Foreigners in schools	Number of children aged 5 to 14 years born abroad, compared to total child population aged 5 to 14 years	Testi and Ivaldi (2009);
Families renting	Proportion of families renting a house, compared to the total population	Townsend (1987); Testi et Ivaldi (2009); Duran e Condori (2019)
Divorced	Proportion of divorced people, compared to the total population	Pampalon and Hamel (2009)
Unemployment	Number of unemployed people looking for work, compared to the total active population (aged between 15 and 64 years)	Jarman (1983), Townsend (1987), Forrest and Gordon (1993), Julkunen (2002), Pampalon and Hamel (2009); Duran e Condori (2019)
Buildings in mediocre or bad state	Number of buildings in mediocre or bad state, compared to the total number of buildings	Perez-Mayo (2002); Fusco et al (2013); Duran e Condori (2019)
Single-parent families	Number of families composed of a single parent and dependent children, compared to all families	Jarman (1983), Forrest and Gordon (1993), Julkunen (2002), Duran e Condori (2019)
Single-person families	Number of people living alone, compared to all families	Julkunen (2002), Pampalon and Hamel (2009) Duran e Condori (2019)

Source Our elaboration

4.1. The Calculation of the Index

A fundamental issue, concerning composite index construction, is the degree of compensability or substitutability of the individual indicators. The components of a composite index are called ‘substitutable’ if a deficit in one component may be compensated by a surplus in another (e.g., a low value of “Proportion of people who have participated in religious or spiritual activities” can be offset by a high value of “Proportion of people who have participated in meetings of cultural or recreational associations” and vice versa). Similarly, the components of a composite index are called ‘non-substitutable’ if a compensation among them is not allowed (e.g., a low value of “Life expectancy at birth” cannot be offset by a high value of “GDP per capita” and vice versa). Thus we can define an aggregation approach as ‘compensatory’ or ‘non-compensatory’ depending on whether it permits compensability or not (Casadio Tarabusi and Guarini 2013).

Compensability is closely related with the concept of unbalance, i.e., a disequilibrium among the indicators that are used to build the composite index. In any composite index each dimension is introduced to rep- resent a relevant aspect of the phenomenon considered, therefore a measure of unbalance among dimensions may help the overall understanding of the phenomenon. In a non-compensatory or partially compensatory approach, all the dimensions of the phenomenon must be balanced and an aggregation function that takes unbalance into account, in terms of penalization, is often used (Mazziotta and Pareto 2017, 2019).

Due the nature of the variables used we opted for partially-compensatory approach: the Adjusted Mazziotta–Pareto Index (AMPI). The AMPI is a composite index that allows comparability of the data across units and over time (Mazziotta and Pareto 2018). It is a variant of the Mazziotta–Pareto Index (MPI), based on a re-scaling of the individual indicators by a Min–Max transformation, in contrast with the classic MPI where all the indicators are normalized by a linear combination of z-scores (De Muro et al. 2011).

AMPI make the indicators independent from the unit of measure, therefore, all the individual indicators are assigned equal weights and absolute time comparisons are allowed (Mazziotta and Pareto 2016). The steps for computing the Adjusted MPI (AMPI) are given below (Mazziotta and Pareto 2016, 2017).

Normalization

Given the matrix X = $\{x_{ij}\}$ with n rows (units) and m columns (individual indicators), it's possible to calculate the normalized matrix R = $\{r_{ij}\}$ as follow:

$$r_{ij} = \frac{x_{ij} - Min_{x_j}}{Max_{x_j} - Min_{x_j}} 60 + 70$$

where xij is the value of the indicator *j* for the unit *i*; $Min_{x_j}$and $Max_{x_j}$are the 'goalposts' for the indicator *j*.

If the indicator j has negative 'polarity', the complement of r_{ij} with respect to 200 is calculated.

To facilitate the interpretation of results, the 'goalposts' can be fixed so that 100 represents a reference value. A simple procedure for setting the 'goalposts' is the following.

Let $Inf_{x_j}$and Sup_{x_j}be the overall minimum and maximum of the indicator *j* across all units and all periods considered. Denoting with Ref_{x_j} the reference value for the indicator *j*, the 'goalposts' are defined as:

$$\begin{cases} Min_{xj} = Ref_{x_j} - \\ Max_{xj} = Ref_{x_j} + \end{cases}$$

where $Ref_{x_j} = \frac{Sup_{x_j} - Inf_{x_j}}{2}$.

The normalized values will fall approximately in the range (70; 130), where 100 represents the reference value.

Aggregation

Denoting with M_{r_i} and S_{r_i} , respectively, the mean and the standard deviation of the normalized values of the unit i, the generalized form of the AMPI is given by:

$$AMPI_i = M_{r_i} - S_{r_i} \cdot cv_i$$

Where

$$cv_i = \frac{S_{r_i}}{M_{r_i}}$$

is the coefficient of variation for the unit *i*.

The AMPI has the same properties than the MPI. Nevertheless, the AMPI allows to compute the score of each unit independently of the others, in contrast to the MPI where the mean and standard deviation of the individual indicators are requested (Mazziotta and Pareto 2019). The 'price' to pay for having scores comparable over time is that indicators with different variability are aggregated. However, normalized indicators in an identical range have much more similar variability than original ones (Mazziotta and Pareto 2016).

The final scores are comparable over and normalized indicators in an identical range have much more similar variability than original ones. [1]

AMPI is useful to evaluate deprivation changes during different years (Landi et al. 2018a; Mazziotta and Pareto 2016; Norman 2010). The 'price' to pay for having final scores comparable over time is that individual indicators with different variability are aggregated. However, normalized indicators in an identical range have much more similar variability than original ones.

Using AMPI, different types of deprivation are identified and a general measure of dep- rivation, called Argentina Deprivation Index (ADI), is proposed to take into account the multidimensional nature of the concept.

Class Aggregation

In order to offer a geographical representation of the results in our work we applied to both the social and material indexes a hierarchical cluster analysis—the Ward method—which is directed to minimizing variance within the groups, and defines the distance between two classes C1 and C2 as the increase of the sum of squares when the classes are joined (Ward, 1963). In this case, the partition is better considered the more homogeneous each class and the more different the classes are from each other. In other words, the higher the variance among the classes, and the lower internal variance (in the classes).

Finally, an influence analysis was performed to assess the robustness of the composite indices (Mazziotta and Pareto 2017). Influence analysis aims to quantify the importance of each individual indicator in the calculation of the composite index, simulating the index calculation removing one indicator at a time. The average shift in Aglomerados rankings is measured. This statistic tells us the relative shift in the position of the entire system of Aglomerados in a single number. It can be calculated as the average of the absolute differences in Aglomerados' ranks with respect to a reference ranking over the 32 Aglomerados:

Let ADI be the index value for Aglomerados *c, c = 1,…,M*

$$R_c = \frac{1}{M}\sum_{c=1}^{M}\left|Rank_{reference}(ADI_c) - Rank(ADI_c)\right|$$

5. Results and Discussion

The results show large variability on the Argentinian territory both in terms of deprivation both in a health perspective. Deprivation has a range variation of 44.85 points between the most deprived Aglomerado S. del

[1] For the mathematical properties of AMPI, see Mazziotta and Pareto (2016, 2018).

Estero–La Banda and the less deprived Ciudad de Bs As. The mean of the deprivation index is 98.47 (with a standard deviation of 8.02) which is quite near to the median score of 97.25. It's notable that the major variation is in the tails of the index distribution (Table 2).

The health indicator goes from a value of 74.25 of expected life years to 77.17 sug- gesting relevant differences in health conditions among various areas: around in the ter- ritory we see areas where population has a life expectation almost three (2.92) years lower than others, even if it refers only to the Ciudad de Bs As Aglomerado, that can be considered a particular case being isolated in the cluster results (Fig. 1). However, in the remaining units the variation in life expectancy (1.56 years) is not irrelevant between the best and the worst deprivation score (Table 2).

The comparison between the measure of material deprivation and the health indicator was initially explored for all urban units considered individually. By calculating Spearman's coefficient, a correlation of − .598 highlights the existence of a negative statistical co-graduation between the condition of deprivation and life expectancy at birth.

Table 6 ADI, class of deprivation and Life expectancy at birth.

Aglomerado	ADI	Class of deprivation	Life expectancy per class
S. del Estero–La Banda	118.09	1	74.25
Gran San Juan	113.27	2	75.06
Gran Resistencia	109.35		
Gran Tucumán–T. Viejo	108.06		
Partidos del GBA	106.06		
Concordia	105.19		
Formosa	104.65		
Corrientes	103.15		
Salta	101.91		
Gran Mendoza	99.77	3	75.52
Jujuy–Palpalá	99.56		
La Rioja	99.56		
Gran Catamarca	99.05		
Mar del Plata–Batán	98.94		
San Luis–El Chorrillo	97.70		
San Nicolás–Villa Constitución	97.33		
Viedma–Carmen de Patagones	97.25		
Cdro. Rivadavia–R. Tilly	97.04		
Posadas	97.02		
Santa Rosa–Toay	96.45		
Gran La Plata	96.43		
Neuquén–Plottier	96.14		
Bahía Blanca–Cerri	95.82		
Rawson–Trelew	95.60		
Gran Córdoba	95.44		
Gran Rosario	94.63		
Río Cuarto	94.26		
Gran Santa Fé	94.14		
Río Gallegos	93.96		
Gran Paraná	86.54	4	75.81
Ushuaia–Río Grande	85.51		
Ciudad de Bs. As.	73.24	5	77.17

Source: Our elaboration

The decrease in the level of deprivation clearly appears to correspond to a growing trend of life expectancy: in areas with more unfavourable material conditions, the opposite hap- pens. Life expectancy at birth increases as the class is less deprived (Table 2); in particular, it increases by 1.36 years moving from units of Class 4 to the Ciudad de Buenos Aires, the only one in Class 5; and an increase of almost 3 years moving from the more deprived to the less deprived Class.

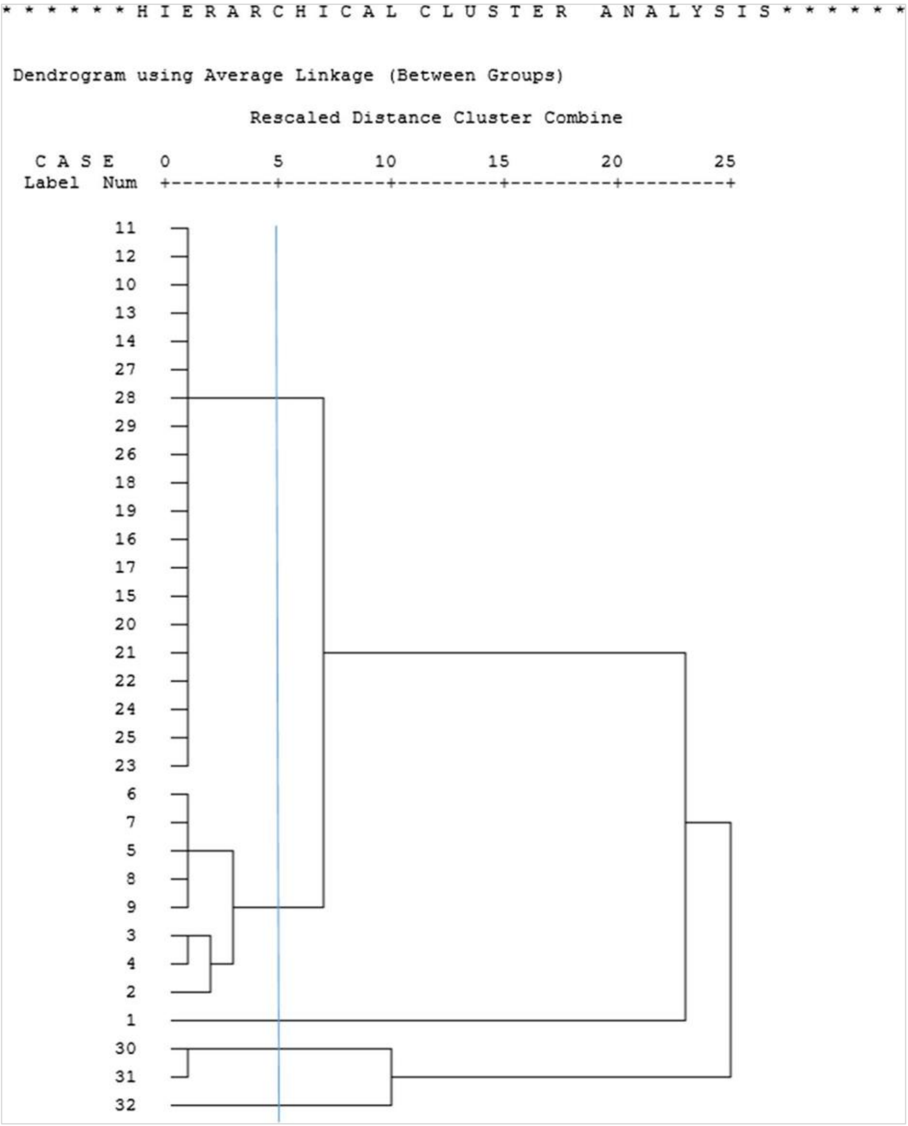


Figure 4 Hierarchical cluster analysis

Validation step aims to assess the robustness of the composite index, in terms of capacity to produce correct and stable measures, and its discriminant capacity (Saisana et al. 2005). Statistical analyses should be conducted to explore the robustness of rankings to the inclusion and exclusion of individual indicators and setting different decision rules to construct the composite index (Freudenberg 2003). ADI is not very sensitive to inclusion or exclusion of individual indicators as shown in Table 3.

Excluding a variable each time the rank varies by only three units with a standard deviation of .85. The excluded variables that show a greater variation on the average rank are the recipient of the income and the one-parent families with a variation slightly higher than 4 positions while the exclusion of the variable Buildings in mediocre or bad state shows an average variation of the index of 2.25 positions.

Table 7 Robustness of ADI

Individual Indicator excluded	Rank shift*
Buildings in mediocre or bad state	2.25
Overcrowding	2.81
Low education	3.19
Income	4.38
Single-parent families	4.38
Mean	3.40
SD	0.85

*Average shift in Aglomerados when excluding an indicator

6. Conclusions

The aim of this paper was to examine and measure the dimensions of deprivation in Argentina in relation to health in the urban Argentinian areas, and to propose a new approach to the study of the deprivation. Although the relationship between variations in health status and well-being on the one hand and social inequalities on the other is well known, there remains considerable uncertainty on their reciprocal interdependence and their implication on health policy (Fuchs 2004). With our work we are suggesting that it would be more suit- able to use composite indexes of deprivation and that the relation between the latter and health must be studied considering the geographical level, and not only the individual one. The results are encouraging, and they allow us to state that the deprivation index proposed explain in a significant way the variability of life expectancy at birth (an increase of almost three years moving from the more deprived to the less deprived Class).

This study is not free from limitations. First, Aglomerados do not have the same number of inhabitants. Second, the ADI is built to cover deprivation in an urban area and thus it is sensitive to urban–rural differences (Barnett et al. 2001; Bertin et al. 2014; Gilthorpe and Wilson 2003), although it is quite common, especially for Argentina, to focus attention only on urban areas, as Battiston et al. did (2013). While the proposed approach can be replicated in an urban context, the indicators used may be not able to describe deprivation in rural areas. For example, housing tenure has been indicated as a biased measure of deprivation in the rural context (O'Reilly et al. 2007). Either critiques of indices tend to focus primarily on the subjectivity of the indicator election or the weighting algorithms associated with their aggregation (Carr-Hill and Sheldon 1991; Frohlich and Mustard 1996; Talbot 1991). Auxiliary problems also include lack of or difficulty accessing sufficient individual longitudinal health data, which necessitates relying on area-based statistics to quantify individual health patterns at a larger scale (Frohlich and Mustard 1996; Schuurman et al. 2007).

In turn, this study presents typical limitations from those based on an ecological design, which means that the observations and results are based on sample sections and cannot be extrapolated to individuals. Hence, it does not determine the effects of con- text between individuals and the deprivation levels defined for the area. Moreover, it is important to interpret results while considering the risk of ecological fallacy because "not all deprived individuals live in deprived wards", just as not everybody in a ward ranked as deprived are themselves deprived" (Townsend et al. 1988). Although a relevant issue, it is difficult to obtain individual deprivation measures and both individual- and area-level deprivation are important determinants of health status (Hagedoorn et al. 2016; Spadea et al. 2010).

Territorial differences in Argentinians' health are intertwined with a "dual epidemiological situation", since typical diseases of developing countries persist and re-emerge along with other, such cardiovascular diseases or cancer, typical of societies with higher life expectancy at birth (Kessler 2014; UNDP 2011). Furthermore, we have to widen our sight, including—other than diseases in themselves– the impact that they have on household budget in the absence of mechanisms to reduce and to facilitate treatment (Kessler 2014). The proportion of households that claim to have cut expenses in medical care or purchase of medicines (for economic reasons) exhibits a decrease between 2010 and 2011 with increases and subsequent adjustments; households declaring having cut health expenses for economic reasons reaches 29.3% in 2018 (Bonfiglio et al. 2019,). Even though this study has limitations, it has potential implications for public health and health care policy in particular for the use of appropriate indexes to account for inequalities and resource allocation. Health inequalities are maybe inevitable, but health status differences arising from socioeconomic conditions should be considered unfair and avoided or reduced by employing active social policies (Adler and Newman 2002; Mackenbach 2012; Mackenbach et al. 2003; Testi and Ivaldi 2009), since poverty and deprivation cease to be a residual or cyclical problem, remediable through market expansion. In contrast, it is increasingly transformed into a persistent and long-term problem, "disconnected from macroeconomic trends and fixated upon disreputable neighbourhoods of relegation in which social isolation and alienation feed upon each other as the chasm between those consigned there and the rest of society deepens" (Wacquant 1999).

References

Adler, N. E., & Newman, K. (2002). Socioeconomic disparities in health: Pathways and policies. *Health Affairs, 21*(2), 60–76. https://doi.org/10.1377/hlthaff.21.2.60.

Andersen, K. K., Steding-Jessen, M., Oksbjerg Dalton, S., & Skyhøj Olsen, T. (2014). Socioeconomic position and incidence of ischemic stroke in Denmark 2003–2012: A nationwide hospital-based study. *Journal of the American Heart Association, 3*(4), e000762. https://doi.org/10.1161/JAHA.113.000762.

Arcaya, M. C., Arcaya, A. L., & Subramanian, S. V. (2015). Inequalities in health: Definitions, concepts, and theories. *Global Health Action, 8*(1), 27106. https://doi.org/10.3402/gha.v8.27106.

Barnett, S., Roderick, P., Martin, D., & Diamond, I. (2001). A multilevel analysis of the effects of rurality and social deprivation on premature limiting long term illness. *Journal of Epidemiology and Community Health, 55*(1), 44LP–51LP. https://doi.org/10.1136/jech.55.1.44.

Bartley, M., & Blane, D. (1994). Commentary: Appropriateness of deprivation indices must be ensured. *BMJ, 309*(6967), 1479.

Battiston, D., Cruces, G., Lopez-Calva, L. F., Lugo, M. A., & Santos, M. E. (2013). Income and beyond: Multidimensional poverty in Six Latin American countries. *Social Indicators Research, 112*(2), 291–314. https://doi.org/10.1007/s11205-013-0249-3.

Baumann, M., Spitz, E., Guillemin, F., Ravaud, J.-F., Choquet, M., Falissard, B., et al. (2007). Associa- tions of social and material deprivation with tobacco, alcohol, and psychotropic drug use, and gen- der: A population-based study. *International Journal of Health Geographics, 6*(1), 50. https://doi.org/10.1186/1476-072X-6-50.

Benach, J., & Yasui, Y. (1999). Geographical patterns of excess mortality in Spain explained by two indices of deprivation. *Journal of Epidemiology and Community Health, 53*(7), 423LP–431LP. https://doi.org/10.1136/jech.53.7.423.

Bertin, M., Chevrier, C., Pelé, F., Serrano-Chavez, T., Cordier, S., & Viel, J.-F. (2014). Can a deprivation index be used legitimately over both urban and rural areas? *International Journal of Health Geographics, 13,* 22. https://doi.org/10.1186/1476-072X-13-22.

Bolsi, A., Paolasso, P., & Longhi, F. (2005). El Norte Grande Argentino entre el progreso y la pobreza. *Población & Sociedad, 12–13,* 227–283.

Bond Huie, S. A., Krueger, P. M., Rogers, R. G., & Hummer, R. A. (2003). Wealth, race, and mortality. *Social Science Quarterly, 84*(3), 667–684. https://doi.org/10.1111/1540-6237.8403011.

Bonfiglio, J. I., Vera, J., & Salvia, A. (2019). *Pobreza monetaria y vulnerabilidad de derechos. Inequidades de las condiciones materiales de vida en los hogares de la Argentina urbana (2010–2018)*. Ciudad Autónoma de Buenos Aires. http://wadmin.uca.edu.ar/public/ckeditor/ObservatorioDeudaSocial/Documentos/2019/2019-OBSERVATORIO-DOC-ESTADISTICO-POBREZA-MONETARIA.pdf.

Booth, C. (1887). The inhabitants of Tower Hamlets (School Board Division), their condition and occupations. *Journal of the Royal Statistical Society, 50*(2), 326–401.

Booth, C. (1889). *Labour and life of the people*. London: Williams & Norgate.

Bowley, A. L. (1915). *The nature and purpose of the measurement of social phenomena*. London: P.S. King & Son. Caranci, N., Biggeri, A., Grisotto, L., Pacelli, B., Spadea, T., & Costa, G. (2010). L'indice di deprivazione italiano a livello di sezione di censimento: definizione, descrizione e associazione con la mortalità.

Epidemiologia e Prevenzione, 34(4), 167–176.

Carr-Hill, R. A., & Sheldon, T. (1991). Designing a deprivation payment for general practitioners: The UPA(8) wonderland. *BMJ: British Medical Journal, 302*(6773), 393–396.

Carstairs, V. (2001). Socio-economic factors at areal level and their relationship with health. In P. Elliott, J. Wakefield, N. Best, & D. Briggs (Eds.), *Spatial epidemiology: Methods and applications* (pp. 51–67). Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198515326.003.0004.

Carstairs, V., & Morris, R. (1991). *Deprivation and health in Scotland*. Aberdeen: Aberdeen University Press.

Casadio Tarabusi, E., & Guarini, G. (2013). An unbalance adjustment method for development indicators.

Social Indicators Research, 112(1), 19–45. https://doi.org/10.1007/s11205-012-0070-4.

Cullen, M. J. (1979). Charles Booth's poverty survey: Some new approaches. In T. C. Smout (Ed.), *Search for wealth and stability: Essays in economic and social history presented to M. W. Flinn* (pp. 155–174). London: Macmillan.

Davey Smith, G., Whitley, E., Dorling, D., & Gunnell, D. (2001). Area based measures of social and economic circumstances: Cause specific mortality patterns depend on the choice of index. *Journal of Epidemiology and Community Health, 55*(2), 149–150.

De Muro, P., Mazziotta, M., & Pareto, A. (2011). Composite indices of development and poverty: An application to MDGs. *Social Indicators Research, 104*(1), 1–18.

Department of Health and Social Security. (1988). *Review of the resource allocation working party formula.*

Final report by the NHS Management Board. London: Dept. of Health and Social Security.

Du Bois, W. E. B. (1899). *The Philadelphia Negro: A social study* (1996th ed.). Philadelphia: University of Pennsylvania Press.

Durán, R. J., & Condorí, M. Á. (2019). Deprivation index for small areas based on census data in Argentina.

Social Indicators Research, 141(1), 331–363.

Eguía, A. (2017). Miradas sobre la pobreza en Argentina. *Cuadernos FHyCS-UNJu, 51,* 71–83.

Fieldhouse, E. A., & Tye, R. (1996). Deprived people or deprived places? Exploring the ecological fallacy in studies of deprivation with the Samples of Anonymised Records. *Environment and Planning A, 28*(2), 237–259.

Forrest, R., & Gordon, D. (1993). *People and places: A 1991 census atlas of England*. Bristol: SAUS. Freudenberg, M. (2003). *Composite indicators of country performance: A critical assessment* (No.

2003/16). Paris. https://doi.org/10.1787/405566708255.

Frohlich, N., & Mustard, C. (1996). A regional comparison of socioeconomic and health indices in a

Canadian province. *Social Science and Medicine, 42*(9), 1273–1281. https://doi.org/10.1016/0277-9536(95)00220-0.

Fu, M., Exeter, D. J., & Anderson, A. (2015). "So, is that your 'relative' or mine?" A political-ecological critique of census-based area deprivation indices. *Social Science and Medicine, 142,* 27–36. https://doi.org/10.1016/j.socscimed.2015.07.036.

Fuchs, V. (2004). Reflections on the socio-economic correlates of health. *Journal of Health Economics, 23*(4), 653–661.

Fusco, A., Guio, A.-C., & Marlier, E. (2013). Building a material deprivation index in a multinational context: Lessons from the EU Experience BT—Poverty and Social Exclusion around the Mediterranean Sea. In V. Berenger & F. Bresson (Eds.), *Poverty and social exclusion around the Mediterranean Sea: Economic studies in inequality, social exclusion and well-being* (pp. 43–71). Boston, MA: Springer, US. https://doi.org/10.1007/978-1-4614-5263-8_2.

Gillie, A. (1996). The origin of the poverty line. *The Economic Historic Review, 49*(4), 715–730.

Gilthorpe, M. S., & Wilson, R. C. (2003). Rural/urban differences in the association between deprivation and healthcare utilisation. *Social Science and Medicine, 57*(11), 2055–2063. https://doi.org/10.1016/S0277-9536(03)00071-6.

Gordon, D. (1995). Census based deprivation indices: Their weighting and validation. *Journal of Epidemiology and Community Health, 49*(Suppl 2), S39–S44. https://doi.org/10.1136/jech.49.Suppl_2.S39.

Hagedoorn, P., Vandenheede, H., Vanthomme, K., Willaert, D., & Gadeyne, S. (2016). A cohort study into head and neck cancer mortality in Belgium (2001–11): Are individual socioeconomic differences conditional on area deprivation? *Oral Oncology, 61*, 76–82. https://doi.org/10.1016/j.oraloncology.2016.08.014.

Havard, S., Deguen, S., Bodin, J., Louis, K., Laurent, O., & Bard, D. (2008). A small-area index of socioeconomic deprivation to capture health inequalities in France. *Social Science and Medicine, 67*(12), 2007–2016. https://doi.org/10.1016/j.socscimed.2008.09.031.

Ivaldi, E., Bonatti, G., & Soliani, R. (2016). *The measurement of well-being in the current era*. New York: Nova Science Publishers.

Ivaldi, E., & Testi, A. (2011). Genoa Index of Deprivation (GDI): An index of material deprivation for geographical areas. In C. M. Baird (Ed.), *Social indicators: statistics, trends and policy development* (Chap. 4, pp. 75–98). Hauppauge, NY: Nova Science Publishers. ISBN: 978-1-61122-841-0.

Jackson, C. A., Jones, M., & Mishra, G. D. (2014). Educational and homeownership inequalities in stroke incidence: A population-based longitudinal study of mid-aged women. *European Journal of Public Health, 24*(2), 231–236. https://doi.org/10.1093/eurpub/ckt073.

Jarman, B. (1983). Identification of underprivileged areas. *British Medical Journal, 286*(6379), 1705– 1709. https://doi.org/10.1136/bmj.286.6379.1705.

Jarman, B. (1984). Underprivileged areas: Validation and distribution of scores. *British Medical Journal (Clinical Research Edition), 289*(6458), 1587–1592. https://doi.org/10.1136/bmj.289.6458.1587.

Jesuit, D., Rainwater, L., & Smeeding, T. (2003). Regional poverty within the rich countries. In J. A. Bishop & Y. Amiel (Eds.), *Research on economic inequality* (Vol. 9, pp. 345–377). Emerald: Bingley. https://doi.org/10.1016/s1049-2585(03)09016-1.

Julkunen, I. (2002). Social and material deprivation among unemployed youth in northern Europe. *Social Policy & Administration, 36*(3), 235–253. https://doi.org/10.1111/1467-9515.00249.

Kelleher, C., Timoney, A., Friel, S., & McKeown, D. (2002). Indicators of deprivation, voting patterns, and health status at area level in the Republic of Ireland. *Journal of Epidemiology and Community Health, 56*(1), 36–44. https://doi.org/10.1136/jech.56.1.36.

Kessler, G. (2014). *Controversias sobre la desigualdad. Argentina, 2003–2013*. Buenos Aires: Fondo de Cultura Económica.

Lalloué, B., Monnez, J.-M., Padilla, C., Kihal, W., Le Meur, N., Zmirou-Navier, D., et al. (2013). A statistical procedure to create a neighborhood socioeconomic index for health inequalities analysis. *International Journal for Equity in Health, 12*(1), 21. https://doi.org/10.1186/1475-9276-12-21.

Lancaster, G., & Green, M. (2002). Deprivation, ill-health and the ecological fallacy. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 165*(2), 263–278. https://doi.org/10.1111/1467-985X.00586.

Landi, S., Ivaldi, E., & Testi, A. (2018a). Measuring change over time in socio-economic deprivation and health in an urban context: The case study of Genoa. *Social Indicators Research, 139*(2), 745– 785. https://doi.org/10.1007/s11205-017-1720-3.

Landi, S., Ivaldi, E., & Testi, A. (2018b). Socioeconomic status and waiting times for health services: An international literature review and evidence from the Italian National Health System. *Health Policy, 122*(4), 334–351. https://doi.org/10.1016/j.healthpol.2018.01.003.

Macintyre, S., Ellaway, A., & Cummins, S. (2002). Place effects on health: how can we conceptual- ise, operationalise and measure them? *Social Science and Medicine, 55*(1), 125–139. https://doi.org/10.1016/S0277-9536(01)00214-3.

Mackenbach, J. P. (2012). The persistence of health inequalities in modern welfare states: The explana- tion of a paradox. *Social Science and Medicine, 75*(4), 761–769. https://doi.org/10.1016/j.socscimed.2012.02.031.

Mackenbach, J. P., Bos, V., Andersen, O., Cardano, M., Costa, G., Harding, S., et al. (2003). Widening socioeconomic inequalities in mortality in six Western European countries. *International Journal of Epidemiology, 32*(5), 830–837. https://doi.org/10.1093/ije/dyg209.

Mazziotta, M., & Pareto, A. (2013). Methods for constructing composite indices: One for all or all for one? *Rivista italiana Demografia e Statistica, 13*(2), 67–80.

Mazziotta, M., & Pareto, A. (2016). On a generalized non-compensatory composite index for meas- uring socio-economic phenomena. *Social Indicators Research, 127*(3), 983–1003. https://doi.org/10.1007/s11205-015-0998-2.

Mazziotta, M., & Pareto, A. (2017). Synthesis of indicators: The composite indicators approach. In F.

Maggino (Ed.), *Complexity in society: From indicators construction to their synthesis* (pp. 159– 191). Cham: Springer.

Mazziotta, M., & Pareto, A. (2018). Measuring well-being over time: The adjusted Mazziotta–Pareto index versus other non-compensatory indices. *Social Indicators Research, 136,* 967–976. https://doi.org/10.1007/s11205-017-1577-5.

Mazziotta, M., & Pareto, A. (2019). Use and misuse of PCA for measuring well-being. *Social Indicators Research, 142*(2), 451–476. https://doi.org/10.1007/s11205-018-1933-0.

McDougall, A. (1885). *Drink and Poverty*. Manchester.

Morris, R., & Carstairs, V. (1991). Which deprivation? A comparison of selected deprivation indexes.
Journal of Public Health Medicine, 13(4), 318–326.

Noble, M., Wright, G., Lloyd, M., Dibben, C., Smith, G. A. N., Ratcliffe, A., et al. (2003). *Scottish index of deprivation 2003: Summary report*. Edinburgh. http://www.scotland.gov.uk/library5/socia l/siod-00.asp.

Norman, P. (2010). Identifying change over time in small area socio-economic deprivation. *Applied Spa- tial Analysis and Policy, 3*(2), 107–138. https://doi.org/10.1007/s12061-009-9036-6.

O'Reilly, G., O'Reilly, D., Rosato, M., & Connolly, S. (2007). Urban and rural variations in morbidity and mortality in northern Ireland. *BMC Public Health, 7*(1), 123. https://doi. org/10.1186/1471-2458-7-123.

Pampalon, R., Hamel, D., & Gamache, P. (2009a). A comparison of individual and area-based socio-economic data for monitoring social inequalities in health. *Health Reports, 20*(3), 85–94.

Pampalon, R., Hamel, D., Gamache, P., Philibert, M. D., Raymond, G., & Simpson, A. (2012). Un indice régional de défavorisation matérielle et sociale pour la santé publique au Québec et au Canada. *Health Canadian Journal of Public Health, 103*(S2), S17–S22.

Pampalon, R., Hamel, D., Gamache, P., & Raymond, G. (2009b). A deprivation index for health plan- ning in Canada. *Chronic Diseases in Canada, 29*(4), 178–191.

Pantelides, E., & Moreno, M. J. (2009). *Situación de la población en Argentina*. Buenos Aires: Pro- grama Naciones Unidas para el Desarrollo – PNUD - UNFPA.

Parra Saiani, P. (2009). *Gli indicatori sociali*. Milano: Franco Angeli.

Pérez-Mayo, J. (2003). On measuring deprivation in Spain: Comparing some techniques. In *3rd confer- ence on the capability approach, September 7–9.* (p. 33). Pavia, Italy.

Rose, M. E. (1972). *The relief of poverty, 1834–1914*. London: Macmillan

Rowntree, B. S. (1901). *Poverty. A study of town life* (2000th ed.). London: MacMillan and Co. Rowntree, B. S. (1941). *Poverty and progress. A second social survey of York*. London: Longmans,
Green and Co.

Rowntree, B. S., & Kendall, M. (1917). *How the labourer lives. A study of the rural labour problem*.
London: Thomas Nelson and Sons.

Runciman, W. G. (1966). *Relative deprivation and social justice: A study of attitudes to social inequality in* *twentieth-century England*. London: Routledge and Kegan Paul.

Saisana, M., Saltelli, A., & Tarantola, S. (2005). Uncertainty and sensitivity analysis techniques as tools for the quality assessment of composite indicators. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 168*(2), 307–323. https://doi.org/10.1111/j.1467-985X.2005.00350.x.

Salmond, C., & Crampton, P. (2002). Heterogeneity of deprivation within very small areas. *Journal of Epidemiology and Community Health, 56*(9), 669–670.

Salvia, A. (2017). *Análisis multidimensional de la situación de pobreza en hogares desde un enfoque de derechos. Argentina urbana: 2010–2017*. http://bibliotecadigital.uca.edu.ar/repositorio/investigacion/analisis-situacion-pobreza-hogares-2017.pdf.

Santana, P., Costa, C., Marí-Dell'Olmo, M., Gotsens, M., & Borrell, C. (2015). Mortality, material deprivation and urbanization: Exploring the social patterns of a metropolitan area. *International Jour- nal for Equity in Health, 14*(1), 55. https://doi.org/10.1186/s12939-015-0182-y.

Schuurman, N., Bell, N., Dunn, J. R., & Oliver, L. (2007). Deprivation indices, population health and geography: An evaluation of the spatial effectiveness of indices at multiple scales. *Journal of Urban Health, 84*(4), 591–603. https://doi.org/10.1007/s11524-007-9193-3.

Sen, A. (1987). *On ethics and economics*. Oxford and New York: Basil Blackwell. Sen, A. (1992). *Inequality reexamined*. Oxford: Oxford University Press.

Simey, M. B., & Simey, T. S. (1960). *Charles Booth, social scientist*. London: Oxford University Press.

Sloggett, A., & Joshi, H. (1994). Higher mortality in deprived areas: Community or personal disadvantage? *BMJ, 309*(6967LP), 1470LP–1474LP. https://doi.org/10.1136/bmj.309.6967.1470.

Smith, A. (1776). *An inquiry into the nature and causes of the wealth of nations*, Vol. 2 (1976th ed.).Chicago: Chicago University Press.

Smits, J., Westert, G. P., & van den Bos, G. A. M. (2002). Socioeconomic status of very small areas and stroke incidence in the Netherlands. *Journal of Epidemiology and Community Health, 56*(8), 637–640. https://doi.org/10.1136/jech.56.8.637.

Spadea, T., Zengarini, N., Kunst, A., Zanetti, R., Rosso, S., & Costa, G. (2010). Cancer risk in relation- ship to different indicators of adult socioeconomic position in Turin. *Italy. Cancer Causes & Con- trol, 21*(7), 1117–1130. https://doi.org/10.1007/s10552-010-9539-0.

Stringhini, S., Sabia, S., Shipley, M., Brunner, E., Nabi, H., Kivimaki, M., et al. (2010). Association of socioeconomic position with health behaviors and mortality. *JAMA, 303*(12), 1159–1166. https://doi.org/10.1001/jama.2010.297.

Su, S., Gong, Y., Tan, B., Pi, J., Weng, M., & Cai, Z. (2017). Area social deprivation and public health: Analyzing the spatial non-stationary associations using geographically weighed regression. *Social Indicators Research, 133*(3), 819–832. https://doi.org/10.1007/s11205-016-1390-6.

Talbot, R. J. (1991). Underprivileged areas and health care planning: implications of use of Jarman indicators of urban deprivation. *BMJ (Clinical research ed.), 302*(6773), 383–386. https://doi.org/10.1136/bmj.302.6773.383.

Testi, A., & Ivaldi, E. (2009). Material versus social deprivation and health: A case study of an urban area. *The European Journal of Health Economics*, *10*(3), 323. https://doi.org/10.1007/s10198-008-0136-z.

Torsheim, T., Currie, C., Boyce, W., Kalnins, I., Overpeck, M., & Haugland, S. (2004). Material deprivation and self-rated health: A multilevel study of adolescents from 22 European and North American countries. *Social Science and Medicine*, *59*(1), 1–12. https://doi.org/10.1016/j.socscimed.2003.09.032.

Townsend, P. (1962). The meaning of poverty. *British Journal of Sociology*, *5*(2), 130–137.

Townsend, P. (1979). *Poverty in the United Kingdom. A survey of household resources and standards of living*. Harmondsworth: Penguin Books.

Townsend, P. (1987). Deprivation. *Journal of Social Policy*, *16*(2), 125–146. https://doi.org/10.1017/S0047279400020341.

Townsend, P., Phillimore, P., & Beattie, A. (1988). *Health and deprivation: Inequality and the North*. London: Croom Helm.

UNDP. (2011). *El sistema de salud argentino y su trayectoria de largo plazo. Logros alcanzados y desafíos futuros*. Buenos Aires. http://www.paho.org/arg/images/Gallery/publicaciones/Elsistemadesaludargentino-pnudopscepal.pdf?ua=1.

Vaughan, L. (2007). The spatial syntax of urban segregation. *Progress in Planning*, *67*(3), 205–294. https://doi.org/10.1016/j.progress.2007.03.001

Wacquant, L. (1999). Urban marginality in the coming millennium. *Urban Studies*, *36*(10), 1639–1647. https://doi.org/10.1080/0042098992746.

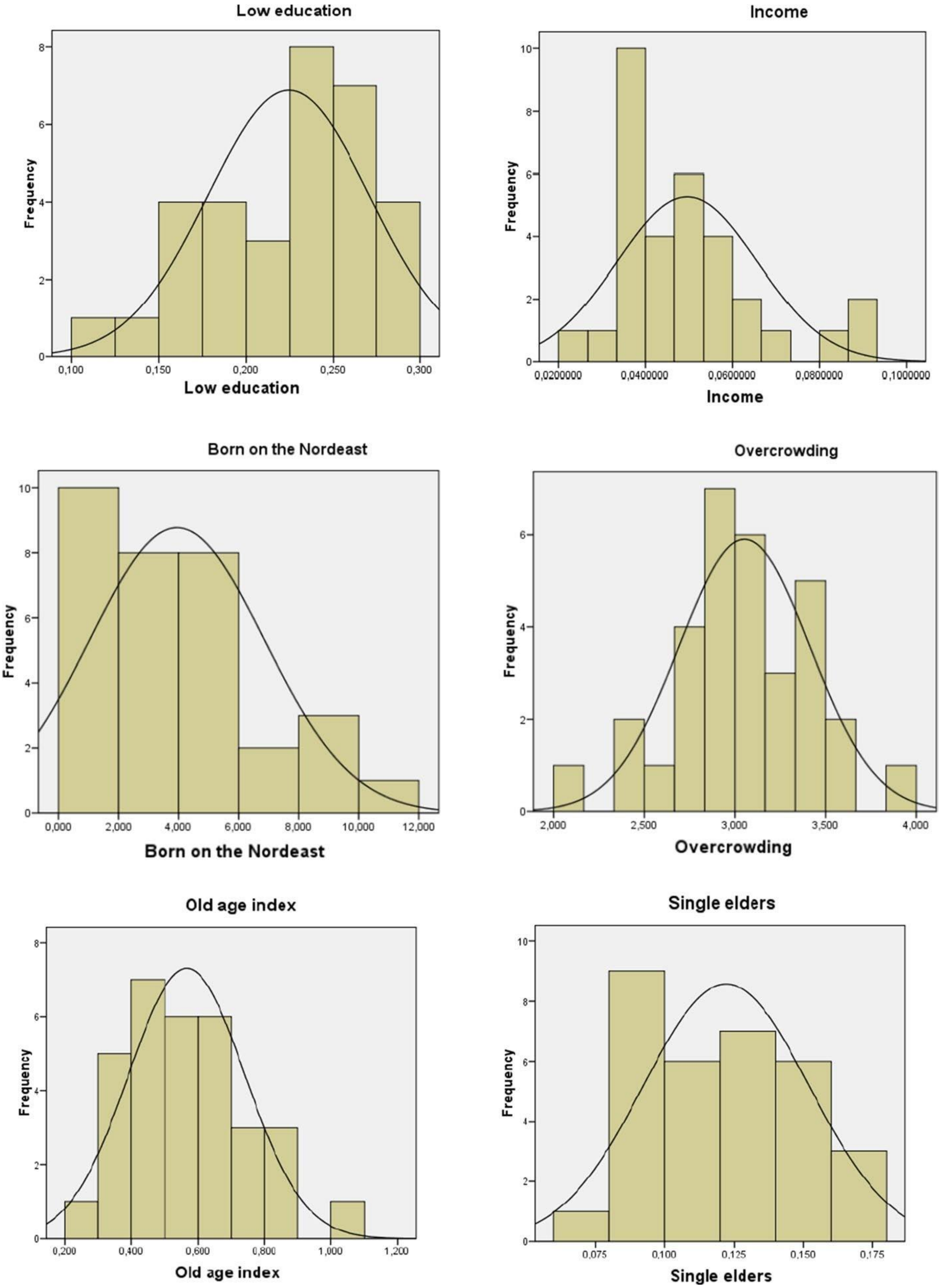
Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, *58*(301), 236–244. https://doi.org/10.1080/01621459.1963.10500845.

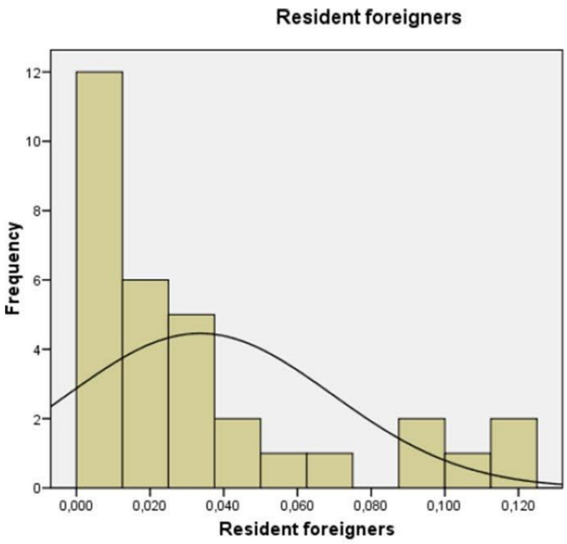
Whelan, C. T., Lucchini, M., Pisati, M., & Maître, B. (2010). Understanding the socio-economic distribution of multiple deprivation: An application of self-organising maps. *Research in Social Stratification and Mobility*, *28*(3), 325–342. https://doi.org/10.1016/j.rssm.2010.03.003.

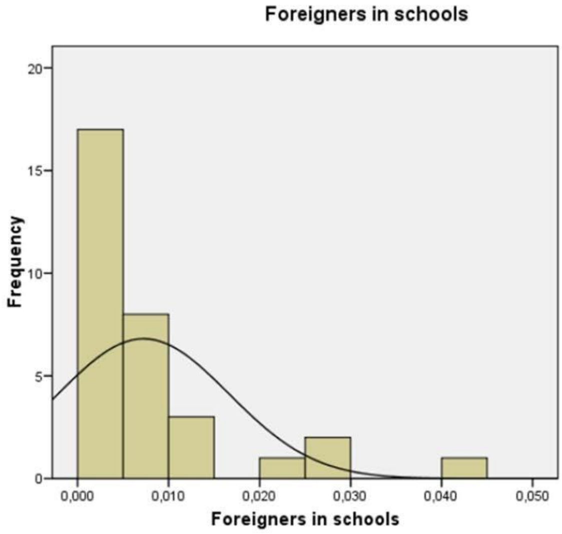
World Bank. (2018). *Argentina: Escaping crises, sustaining growth, sharing prosperity (English)*. Washington, D.C. http://documents.worldbank.org/curated/en/696121537806645724/Argentina-Escaping-Crises-Sustaining-Growth-Sharing-Prosperity.

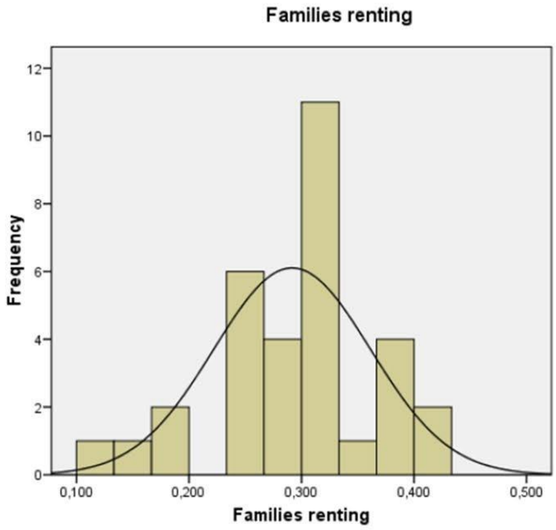
Appendix 1

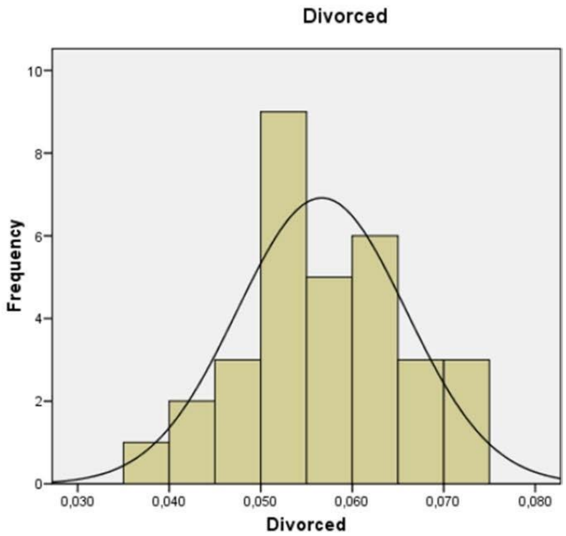
Figure 2 Variables Distribution

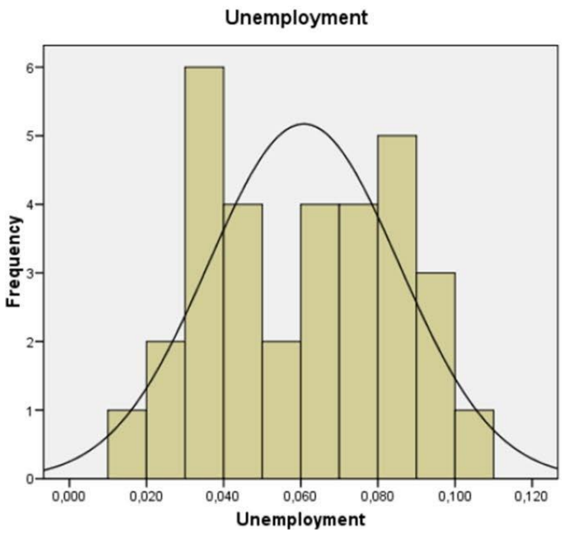


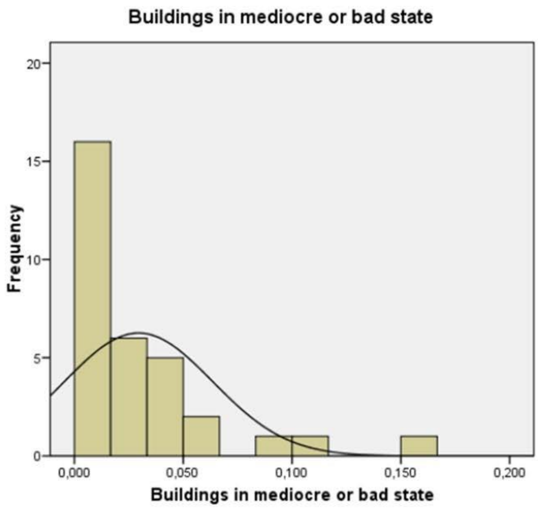

Resident foreigners
Frequency
Resident foreigners

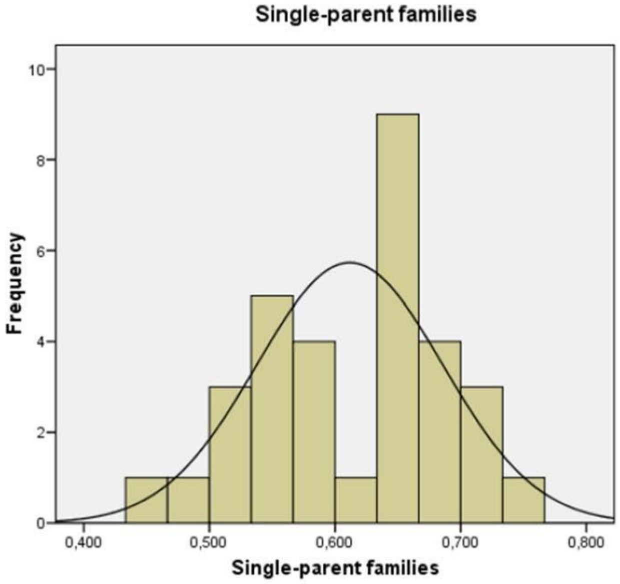

Foreigners in schools
Frequency
Foreigners in schools

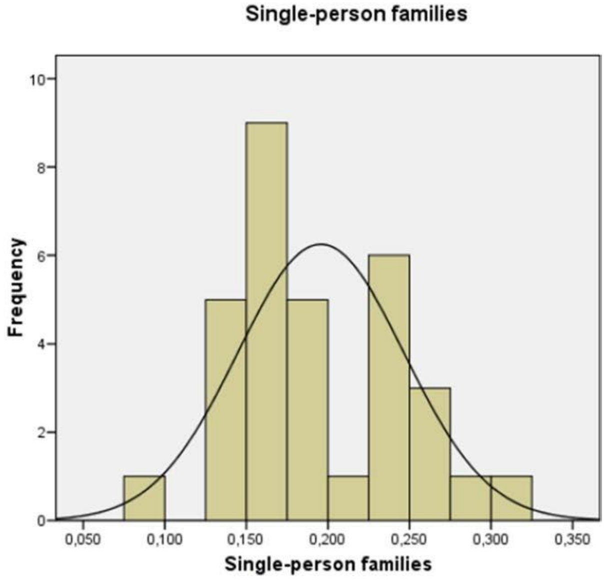

Families renting
Frequency
Families renting


Divorced
Frequency
Divorced


Unemployment
Frequency
Unemployment


Buildings in mediocre or bad state
Frequency
Buildings in mediocre or bad state


Single-parent families
Frequency
10
8
6
4
2
0
0,400
0,500
0,600
0,700
0,800
Single-parent families


Single-person families
Frequency
10
8
6
4
2
0
0,050
0,100
0,150
0,200
0,250
0,300
0,350
Single-person families

Appendix 2

Table 8 One-Sample Kolmogrov-Smirnov test

	Low education	Income	Born on the Nord-east	Over-crowding	Old age index	Single elders	Resident foreigners	Foreigners in schools	Families renting	Divorced	Unemployment	Buildings in mediocre state	Single-parent families	Single-person families
N	32	32	32	32	32	32	32	32	32	32	32	32	32	32
Normal Paramenters (a,b)														
Mean	.22450	.049572668	3.94785	3.05219	.56614	.12212	.03355	.00725	.29162	.05666	.06076	.02946	.61172	.19561
SD	.046356	.0161374380	2.909729	.360341	.174980	.029779	.035792	.009379	.069692	.009228	.024683	.033993	.074248	.051073
Most extreme differences														
Absolute	.098	.151	.123	.106	.111	.107	.221	.226	.117	.099	.123	.231	.162	.147
Positive	.066	.151	.123	.100	.111	.107	.221	.226	.102	.099	.123	.231	.093	.147
Negative	−.098	− .123	− .092	− .106	− .077	− .054	− .188	− .220	− .117	− .076	− .106	− .199	− .162	− .139
Kolmogorov– Smirnov Z	.555	.852	.693	.599	.629	.603	1.250	1.277	.663	.560	.697	1.308	.918	.829
Asymp. Sig (2-tailed)	.918	.463	.723	.866	.824	.861	.088	.077	.771	.912	.716	.065	.368	.498

Appendix 3

Table 9 Correlations

	Low education	Income	Born on the Nordeast	Overcrowding	Old age index	Single elders	Resident foreigners	Foreigners in schools	Families renting	Divorced	Unemployment	Buildings in mediocre or bad state	Single-parent families	Single-person families	Life expectancy at birth
Low education	1	,238	,172	,149	,075	,176	-,337	-,304	-,391	,053	,059	,134	-,036	,006	-,285
Income	,238	1	-,199	-,092	,367	,419	-,303	-,176	-,165	-,108	-,002	-,127	-,219	,226	-,348
Born on the Nordeast	,172	-,199	1	-,010	-,284	-,261	,032	-,216	-,033	-,049	-,361	-,276	-,075	,102	,210
Overcrowding	,149	-,092	-,010	1	-,546	-,483	-,354	-,543	-,511	-,584	-,127	,564	,784	-,772	-,365
Old age index	,075	,367	-,284	-,546	1	,973	-,090	,291	,355	,450	,271	-,203	-,748	,731	,159
Single elders	,176	,419	-,261	-,483	,973	1	-,144	,235	,308	,433	,277	-,172	-,720	,693	,152
Resident foreigners	-,337	-,303	,032	-,354	-,090	-,144	1	,658	,379	,228	,098	-,300	-,177	,136	,447
Foreigners in schools	-,304	-,176	-,216	-,543	,291	,235	,658	1	,422	,439	,155	-,200	-,401	,353	,339
Families renting	-,391	-,165	-,033	-,511	,355	,308	,379	,422	1	,523	,328	-,211	-,566	,550	,594
Divorced	,053	-,108	-,049	-,584	,450	,433	,228	,439	,523	1	,133	-,345	-,580	,610	,277
Unemployment	,059	-,002	-,361	-,127	,271	,277	,098	,155	,328	,133	1	,192	-,244	,232	,197
Buildings in mediocre or bad state	,134	-,127	-,276	,564	-,203	-,172	-,300	-,200	-,211	-,345	,192	1	,515	-,528	-,302
Single-parent families	-,036	-,219	-,075	,784	-,748	-,720	-,177	-,401	-,566	-,580	-,244	,515	1	-,954	-,452
Single-person families	,006	,226	,102	-,772	,731	,693	,136	,353	,550	,610	,232	-,528	-,954	1	,367
Life expectancy at birth	-,285	-,348	,210	-,365	,159	,152	,447	,339	,594	,277	,197	-,302	-,452	,367	1

Uncertainty, asymmetric information and contract design for transport service intended for disabled people.

1. Reference Literature framework

In the present work a principal-agent problem is presented. In particular, it refers to the provision of a service for a specific class of users who have it for free, because it is financed by general taxation.

We describe a scenario of delegation of tasks and functions (Laffont, 2009) by a public principal to a private agent, where the latter is selected on for his specific skills and knowledge of the service to be performed, even if these may not be verifiable by the principal (Arrow, 1968). Then we consider an agency context with asymmetric information: as Arrow said, the agency problem is interesting "*when there is uncertainty at some point"* (Arrow, 1984) and when there is a conflict between principal and agent objective. Uncertainty refers to the fact that one party's or both have incomplete information because the contract is not able to define and control all the relevant aspects for the realization of the service. The conflict refers to the fact that the agent act on behalf of the principal but their objectives are not easily alignable (Hart, 1987).

Asymmetric information indicates that information is not equally distributed among the participants and individuals may not have complete knowledge, or effective control, about the causal relationships between actions and results. Literature on asymmetric information agrees to assess the existence of two types of asymmetry. The first refers to the case where, the outcome is affected by an unobservable action of the agent who, then, is able to implement strategic behaviours (hidden action) creating situations defined as moral hazard (Arrow, 1984). The second refers to the fact that the agent has private (hidden) information about his attributes (Spence, 1973), which can produce averse selection with respect to contractual specifications. We describe a classical moral hazard situation in which there is a conflict of objectives, hidden action and uncertain results, then a risk must be shared and a bilateral relationship must be defined through a contract.

The conflict between agent and principal aims is the prerequisite because in the absence of which the actors would find a shared strategy towards optimality for both. The agent effort, being not directly observable by the principal, creates the opportunity for strategic behaviours. The result is affected by the agent's action but also by the state of the nature, i.e. some random circumstances which are not under the control of the parties. Therefore, the need for incentives to implement a correct behaviour is necessary to make the positive result more likely, knowing that the final outcome realization, is not completely determined by principal or agent actions.

We shortly recall the literature inherent moral hazard, risk aversion, and incentive contract theory.

The concept of moral hazard, known very early in the insurance sector, refers to situations in which there is a problem of risk sharing between two contractors and, thanks to an insurance policy, one of the actors is able to implement a strategic behaviour i.e. he is able to obtain a benefit for which he does not bear the cost, because the cost are paid by the insurance. Arrow (1963)analysing the medical care market identifies as source of failure of the market, "*special structural characteristics of the medical-care market are largely attempts to overcome the lack of optimality due to the non-marketability of the bearing of suitable risks and the imperfect* *marketability of information*". Institutional support and the availability of insurance policies explain non-competitive behaviours, such as moral hazard, which ultimately hinds the path to optimality.

With a different view and a more orthodox approach, Pauly (1968), in addition to not recognizing in moral hazard a source of market failure, with a position less influenced by moral judgments than Arrow, suggests that moral hazard is part of natural human behaviour.

Many are the definitions of moral hazard in the insurance field. "*Moral hazard problems arise when there is imperfect information concerning the actions of these who purchase insurance, because those actions cannot be perfectly monitored and the insurance contract cannot specify all of the actions which the insured is to undertake"* (Stiglitz, 1983)**,** and finally Moral Hazard is "*the ability of insured agents to affect the probabilities of insured events***"** (Laffont, 2009)**.**

The theory of incentives has developed to answer to these problems and it aims to identify methods to encourage correct behaviour, as proposed by (Stiglitz, 1983), this fields is characterized by information gap, risk, and insurance.

If there were no asymmetric information the contract would specify every aspects of the action to be implemented. In the absence of risk against which to insure, the problem would not arise. If both actors were risk neutral and without seeking guarantees against losses (no insurance), there would be no problem of incentives because each would bear the consequences of their actions.

Considering the subject of our work, we are interested in definitions of moral hazard that allow us to expand the subject of insurance to different situations, in which the insurance dynamics is realized by an employment contract or an agency relationship. Therefore in a more general perspective *moral hazard refers to the likely misconduct of an individual who can obtain benefits that are partially or fully paid for by others* (Cutler, 2000).

In agency relationships, when behaviours of the participants are not completely observable and consequently the contract is incomplete, a problem of moral hazard arises, in addition to the uncertainty deriving from the fact that they are not able to completely determine the outcome. Arnott and Stiglitz (1991) note that moral hazard problems are not limited to situations of explicit insurance contracts, but affect also implicit contracts (associated with labour markets, land markets, capital markets and product markets), and that forms of potential inefficiency occur in all these situations. Since the beginning of the debate, the central question regards the impossibility, in the presence of Moral Hazard, to reach Pareto-optimal equilibria (first best outcome). For Holmstrom (1979) an optimal risk sharing solution is precluded. It is possible to reach a Second-Best solution, *"which trades off some of the risk-sharing benefits for provision of incentives".* A possible solution to the information asymmetry is monitoring the actions, but this could be not possible or excessively expensive. What Holmstrom demonstrates is that any information concerning the actions of the agent or state of nature, even if imperfect, improve contracts allowing a better evaluation of the actions.

Stiglitz (1983) describes the non-quasi-concave indifference curves and the feasible locus with an irregular shape, typical of this analytical context, and taking up what Pauly (1968) has already said, it demonstrates that if individuals have the possibility to buy additional insurance at the same relative price of insurance in terms of premium-benefit, they will always to do it. These conditions and the non-complete observability of agent's actions, determine that an equilibrium may not exist or that if it exist could be different from the standard competitive equilibrium.

As broadly discussed and described, the main candidates to represent moral hazard actions are effort variables which positively affect the level of agent's production but at the same time create disutility for the agent (Laffont, 2009).

A strand of literature aims to analyse the problem of incentives, not strictly in explicit insurance contracts, but in service provision contracts in order to efficiently reach adequate service standards, which could be obtained not only with technical variables, but also through effort variables.

In the analysis of the health care provision, Ma (1994)**,** considers that the agent, the health care provider, can exert effort in various way in order to maximize his own utility, i.e. increasing production or quality, reducing costs, dumping patients extremely expensive. Ma, comparing cost reimbursement and prospective payment demonstrates that the payment system induce the agent to mix efforts between strategies of cost reduction and quality provision. Payment systems are analysed with respect to the problem of choosing the best type of contract with respect to the criterion of efficiency, which is not the same between different healthcare systems. Chalkley and Malcomson (1996) suggest that the choice of the best contract depends on the context of knowledge of the actors involved (such as cost structure, quality of service, and number of users) and substantially state that one method may not always be better than the others. Theory is not able to define which contracts is better than others but can help to collect information useful to achieve e a contract in line with the purposes.

The financing of health expenditure is a problem whose relevance is gaining space on the political agenda both with respect to pressures on the demand for services, and on the supply side (Brouwer, 2019), much literature has focused the analysis on payment systems of health care service and on healthcare insurance characteristics and consequences.

There is a wide debate about the health care financing system's ability to act on the moral hazard mechanism, because it involves a trade-off between risk-taking and adequate incentives. A broader insurance coverage induces opportunistic behaviours increasing the risk and, at the same time, increasing the expenditure. It is because people in generally increase their consumption if subsidized. It has been debated since the 1970s whether the demand for healthcare services was influenced by the price to be paid. Some study find a certain elasticity to the price of the demand for health services is recognized, therefore it is accepted that health insurance determines a moderate moral hazard (Cutler, 2000). However, positions are still conflicting. Since, it is not easy to isolate the effects of moral hazard and selection on decisions regarding the consumption of health services and the purchase of health insurance, recent studies have observed that the largest buyers of supplementary health insurance have better health and lower consumption of health care (Bolhaar, 2012) and even that there is not direct relation between high risk individuals and healthcare insurance choices (Di Novi, 2008).

The problem of moral hazard is gaining increasing attention because today a large part of public spending derives from delegation contracts to produce goods and services offered to citizens (Brown, 2016), consequently, the theory of contracts is applied not only as regards healthcare but in other areas.

As mentioned above, a fundamental element of the contract theory considered here is the risk aversion of the agents resulting from the uncertainty of the results. Applying the theory of expected utility, which states that individuals choose between risky options or uncertain prospects by comparing their expected utility values, i.e. the weighted sums obtained from the sum of the utility value of the possible outcomes multiplied by the respective probabilities, we define the agent's preference system and therefore his choices with respect to the benefits and costs involved in the contract. Not entering into the merits of the debates concerning this theory (Mongin, 1997)and the assumption that all individual are almost risk neutral if the risk is quite negligible (Rabin, 2013), we take, in an instrumental way for our purposes, a von Neumann-Morgenstern expected utility function and we will assume a broadly applied form of the agent's utility function characterized by risk aversion.

In our work, as we will show beyond, the agent's utility function depends on the contract reimbursements and the effort variable. Since the agent takes the reimbursement as given, the only choice variable is the effort. We assume effort, as described by Koopman (1953), with some properties, three in particular, that allow it to be included in quantitative problems. Effort can be attributed a quantitative measure and therefore effort can be expressed by numerical values. The effects of effort are assumed to be additive and the product of the effort is formally represented by a function of the effort trough a particular "return function**"** (Koopman, 1953)**.** Although, the problem of obtaining an appropriate functional form for the effort returns function appears obvious. Since it is an exercise in abstraction it seems useful in the choice of the analytic form to apply the three criteria suggested by Koopman (1953): the mathematical simplicity of the expression, the generalization of formulations applied to concrete cases and finally simply represent the way in which the effort produces effect.

The introduction of effort variables is relevant because it creates a moral hazard problem caused by their not observability and their ambiguous impact on the outcome variables. In particular, first effort produce asymmetric information between agent and principal. Secondly, on one side effort represents a cost for the agent because directly produce disutility for him; but on the other side it affects positively the probability of the final outcome and then it is a benefit for the principal and for the agent in terms of the final revenue.

How the probability of the outcome is defined and from which variables and how it is affected remain problematic and can be considered unsolved. We, based on some assumptions about our variables, will refer to functional models widely used to describe limited growth trends.

There are numerous models describe the dynamics of growth. Verhulst (1838), deeming that a stable population would consequently have had a characteristic of the level of saturation, introduced a new growth curve model: the equation of logistic growth ((Tsoularis, 2002). Verhulst's logistic function model formed the basis for several models in various research fields (Berkson, 1944). It has been used to model many biological systems (Carlson, 1913; Pearl R. , 1930; Annadurai, 2000; Wacheneheim, 2003); to describe animal and herd behaviour (Morgan, 1976; Krebs, 1985). Logistic equation has found even economic applications, Fisher and Fry (1971)have used it to describe the penetration of new products and new technologies into the market. Finally, it should be noted that the logistic function is at the basis of inferential logistic regression models so widespread and robust tool to analyse dichotomous outcome variables, in the epidemiological, economic and many other fields (Cramer, 2002; Peng, 2002).

In this work we present a problem of incomplete information, in line with the literature just described, in which moral hazard can create the opportunity for strategic behaviour that can lead to non-optimal solutions. Our analysis describes an agency relationship for the provision of a public transport service, in which there are three participant: the principal who designs the contract and pay an agent to perform a service on his behalf, the user of the service is a specific class of individuals who has the service for free. The asymmetric information arises from two elements. First, the subject of the contract depends on the agent effort, which is not observable by the principal. Second, the service user, is the bearer of specific characteristics that are not controllable which have a direct influence on the probability of the outcome, then the agent is not able to establish a priori the level of effort that guarantees the service carried out.

The service we will considered is the transport service intended for people with disabilities identified in advance by the principal on the basis of institutional classification. We identify the principal as a municipality. The service must be scheduled in advance on a regular basis.[i]

The novelty of this work is that, to the best of our knowledge, it is the very first time that the service of transportation intended to disabled people is investigated on the light of asymmetry of information and, very crucial to our analysis, uncertainty.

The object of this work is the analysis of the problem of defining an efficient contract for the provision of a public service and a critical discussion on how causal variables affect the probability of a particular public service (and consequently the expected expenditure of the principal) and provide technical and informative tools to support decision of the policy maker.

The paper is organized as follow: In Section 2 material and methods are presented. First, we describe the theoretical agency-model in its general formulation, secondly we go into a more particular dimension and describe the two empirical models applied to the transport service intended to disabled people. In Model 1 the problem is presented in his simplest form considering only the effort as causal variable. In Model 2 we introduce the type of the user as further causal variable and the logistic function as formalization of the probability of the service. The results will be appropriately presented in Section 3. Section 4 is devoted to discussion of results.

[i] We consider the scheduled service so we exclude the uncertainty associated to the on call service (such as characterize the taxi service) and every day the number of service to provide is known.
Users are identified on the basis of the national system of disability certification

2. Material and Methods

The Theoretical Model

The aim of this work consists in handling the problem related to moral hazard in the context of public transport service provision by a private agent for disabled people. So we apply the incentive contract theory in the new framework. To do that, we recall the agency theory with asymmetric information and we present the general solutions of the optimization problem of the principal. After that, we design two empirical models to analyse how the variables of interest affect the realization of the service.

The Principal Agent Problem

In a basic agency context, we describe the entrustment of a public service to a private agent-carrier who will be remunerated by the principal. The service is aimed at a specific user who benefits from it free of charge. We describe the agency relation, widely discussed in the literature, in which the main issues derive from the hidden action of the agent and from the uncertainty affecting results. The time schedule of the Moral Hazard we are interested in is presented in Figure 1: principal have a service to be provided and he wants it made by an agent, then he establishes the rules of the contract, before the agent's action, specifying the reimbursement as a function of observable results (Arrow, 1984). The agent can accept (or reject) the contract. Once, the contract is signed, the agent chooses his action (in particular we focus on the effort the agent puts in his job) between a set of alternatives. The asymmetric information arises from the fact the effort is not verifiable by the principal, because he is not able to observe the agent's action carrying out his task, and that this action is not able to determine completely the outcome, because in that case the principal could deduce the agent's action by the outcome obtained. This uncertainty derives from the characteristics of the users which, despite the classification of users on the basis of the complexity of which they are carriers, remain exogenous.

Then, neither the principal nor the agent are able to have a complete information on the object of the negotiation which finally depends on the state of the world realized. Defining a complete contract is not possible and it would still be insufficient to guarantee the desired result on both sides (Brown, 2016).

Therefore, on the first hand there is a separation or contrast of interests between the principal and the agent, the principal wishes a high effort in order to guarantee a high probability of success, while the agent is negatively affected by the effort in terms of utility. On the other, the asymmetric information and the uncertainty of the outcome leave unregulated space for opportunistic behaviours, since the variable effort is the agent's private information.

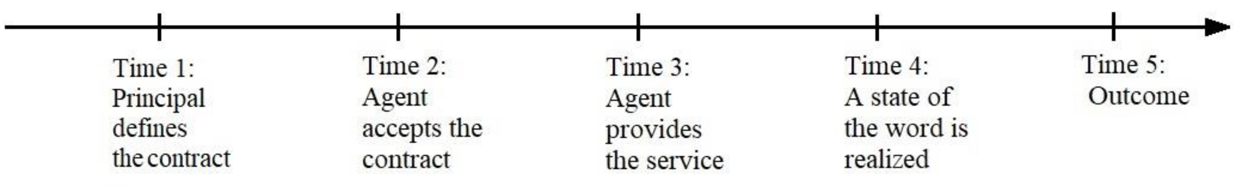


Figure 1 Timing of the contract

We consider a situation in which there are only two possible outcomes w of the service provision, and no intermediate solutions:

$$w = \begin{cases} 1 \ if\ the\ service\ is\ performed \\ 0\ otherwise \end{cases}$$

Each event takes place with probability p and *(1-p)* respectively.

In the main general formulation, we assume the probability that the service is provided depends on only two variables: first, a decision variable, the agent's effort $ef \in [0,1]$, and second an exogenous variable, $g \in [0,1]$[i] that summarize the characteristics of the user that can influence the probability of the successful completion. For both variables we normalize to 0 the minimum level and to 1 the maximum level the two variables can assume. So we define the probability *p=p(ef,g)* which is negatively influenced by the type of patient (g) and positively by the effort (*ef*) of the agent $\frac{\partial p}{\partial g} < 0, \frac{\partial^2 p}{\partial g^2} > 0, \frac{\partial p}{\partial ef} > 0, \frac{\partial^2 p}{\partial ef^2} < 0$

The Principal objective is to guarantee as many users as possible to benefit from the service, minimizing the expenditure. This result is indirectly and positively affected by the variable *ef*, according to our assumptions, through the conditioning of the effort on the probability *p*. The agent aims to maximize its utility that, as we will define shortly, depends positively on the reimbursement and negatively on the effort.

Considering the principal risk neutral (Rogerson, 1985), we define his objective expected cost function:

$$ObjP(\alpha, \beta) = p(ef, g)\alpha + (1 - p(ef, g))\beta \qquad (7)$$

Where α and β are the reimbursements (denoted by *y*) that the agent receives according to the outcome:

$$y_i = \begin{cases} \alpha \quad if\ w = 1 \\ \beta\ if\ \ w = 0 \end{cases} \qquad (8)$$

We assume the agent is risk averse, with a utility function additive with respect to *y* and *ef:*

$$U(y, e) = u(y) - v(ef) \qquad (9)$$

The utility of the agent depends positively on the reimbursement and negatively on the effort. We assume that this disutility, deriving from the effort, is event-independent (Stiglitz, 1983).

Then, the agent's expected value of the contract is:

$$EU(y, ef) = p(ef, g)u(\alpha) + \big(1 - p(ef, g)\big)u(\beta) - v(ef) \qquad (10)$$

[i] Let g a random variable such that cannot be controlled by either the principal or the agent.

Therefore, we face a problem of "incomplete contract", as named in the literature (Tirole, 1999; Schmitz, 2003), deriving from the asymmetric information and from the consequent Moral Hazard. Principal aims to define contractual terms suitable to satisfy the participation constraint and the incentive-compatibility constraint guaranteeing, with the minimum cost, the service to the largest possible number of users. The first constraint requires that the contract was defined in such a way to guarantee the agent an expected utility at least equal to its reservation utility (U_0). The incentive-compatibility constraint, which represents the principal's limited information on the agent's behaviour (Holmstrom, 1979), requires the definition of contractual terms able to induce, even if in a context of asymmetric information, the agent to adopt the behaviour desired by the principal. In our model we are interested in controlling the variable effort (*ef*) which

$$\begin{gathered}\min_{\{\alpha,\beta\}} ObjP(\alpha,\beta) \\ subjec\ to \\ participation\ contraint: EU(y,ef) \geq U_0 \\ incentive-compatibility\ constraint:\ ef^* \in arg\max_{ef}\{EU(y,ef)\}\end{gathered} \tag{11}$$

is neither observable, as we said, nor "objectively" verifiable, not even ex post, by the Principal and therefore cannot be found a "direct" implementation through the contract.

So the Principal, who is able to observe only the outcome, faces a constrained minimization problem, as defined by Holmstrom (1979) in the following general formulation:

Where:

U_0 is the agent's reservation utility

ef * is the optimal level of effort

Given our assumptions on utility and cost functions, we face a concave programming problem. Therefore, the first order conditions are necessary and sufficient for the identification of a point of absolute optimum.

Disabled people transportation service

The agency problem introduced is referred to a particular public service. We analyse the transport service for disabled people which represent a wide category of users for their characteristics and needs, because disabled people are not a homogeneous group, they present differences in many aspects, such as the type and severity of the disability (Gant, 1992). We do not ignore the problems and the open discussions surrounding the definition and measurement of disability (Slater, 1974; Fujiura, 2001) as a complex and multidimensional concept (Altman, 2001) and without entering into the debate concerning the classification and the measurement of disability, we assume the provision of the transport service is affected by uncertainty linked to specific individual characteristics represented by g, these characteristics may create obstacles to the regular provision of the service. [i]

Since the existence of *The International Classification of Functioning, Disability and Health* - ICF, implemented also in Italy, in a completely instrumental way, we use a broad and generic conceptualization of disability which allow us to summarize in a single variable g the complexity condition of the user, and consequently following the ICF classification g is determined for each user. Assuming a limited scale $g \in [0,1]$. The minimum value of g is set equal to zero and it indicates a low degree of complexity which has a low negative impact on the probability of the service, while the value 1 (maximum value of g) indicates a high complexity which has a greater negative impact on the probability of the service,

Empirical Model 1: The probability depends only on the agent effort

Assuming all the users homogeneous in terms of complexity, we consider that the probability p of the service provision depends exclusively on *ef*, that is the variable that describes the agent's effort: *p=p(ef).*

The agent chooses the effort level, which maximizes his expected utility function. Deriving by *ef* we get:

$$p'(ef)[u(\alpha) - u(\beta)] = v'(ef) \qquad (12)$$

That can be written as:

$$p'(ef) = \frac{v\prime(ef)}{u(\alpha)-u(\beta)} \qquad (13)$$

From equation 7 we can deduce that:

- if $u(\alpha) = u(\beta)$, the incremental variations of probability $p'(ef)$ with respect to the effort tends to infinity and then *ef→0,* if $u(\alpha) > u(\beta)$, then *ef>0* and if $u(\alpha) - u(\beta) \to \infty$ *,* then *ef→1.*

To every level of effort is associated a map of indifference curves describing the agent preferences in relation to α and β with respect to different utility values (Figure 1).

[i] For example different types of disability that could affect the probability of the realization of the service in many ways. Limits in the movements (limitations in walking, wheelchair) require specific technical skills in the provider and equipment (such as the elevator) in the vehicles. Behavioural problems, in a very wide meaning, may determine difficulties in communication, in the relation and stressing situations for both providers and users. Then, it may need more soft skills than technical ones. Furthermore behavioural problems could be related to some situation in which the user in sometimes is not collaborative and is not willing to use the service. The combination of these elements are summarized in the variable g.

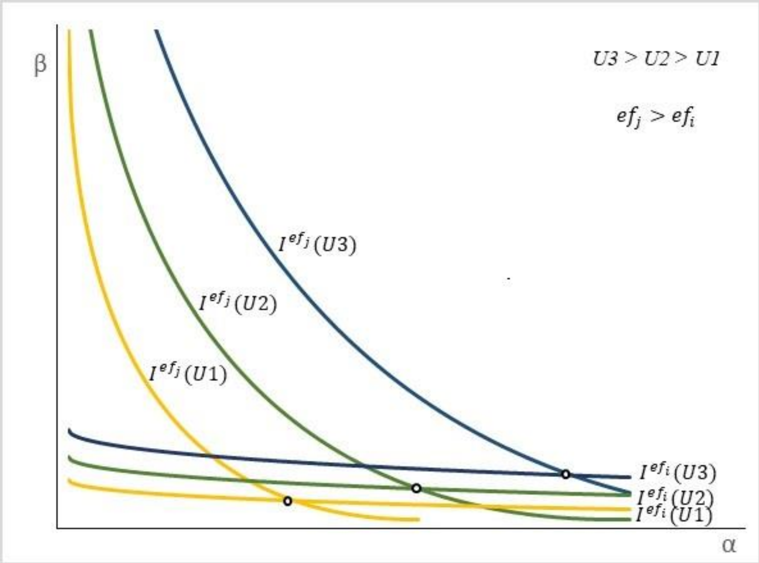


Figure 2 Indifference curves with two different levels of effort ($ef_j > ef_i$) and three levels of utility

We can observe (Figure 2) that for each effort level, the slope of the indifference curves represents the agent's marginal rate of substitution between α and β, and it is given by:

$$\frac{d\beta}{d\alpha}|I^*(u) = s^*(\alpha,\beta) = -\left(\frac{\frac{\delta EU(y,ef^*)}{\delta\alpha}}{\frac{\delta EU(y,ef^*)}{\delta\beta}}\right) = -\frac{u'_\alpha\, p^*}{u'_\beta(1-p^*)} < 0 \qquad (14)$$

We can deduce that, being decreasing and having different slopes for every effort level the agent's indifference curves relative to the same utility level only cross once, and the single-crossing condition is satisfied. In the intersection point the agent obtain the same utility with different effort levels which are consequently considered equivalent. Without this property the preference system, the agent will always choose the less expensive in terms of effort indifference curve and then there will be no moral hazard (Stiglitz, 1983).Discrete effort: with only two possible values

In this scenario we consider only two possible levels of effort: ef^L low effort and ef^H high effort, such that $ef^L < ef^H$.

Denoting $p(ef^H)$ and $p(ef^L)$, the probabilities associated with each effort level, cetris paribus, we will have $p(ef^H) > p(ef^L)$.

From equation (7), given the utility level, the marginal rate of substitution with respect to the j-th effort, is:

$$s^j(\alpha,\beta) = -\frac{u'_\alpha p^j}{u'_\beta(1-p^j)} \quad with\ j = L, H \qquad (15)$$

Then,

$$s^L(\alpha,\beta) = -\frac{u'_\alpha\, p(ef^L)}{u'_\beta\left(1-p(ef^L)\right)} < -\frac{u'_\alpha\, p(ef^H)}{u'_\beta(1-p(ef^H))} = s^H(\alpha,\beta) \qquad (16)$$

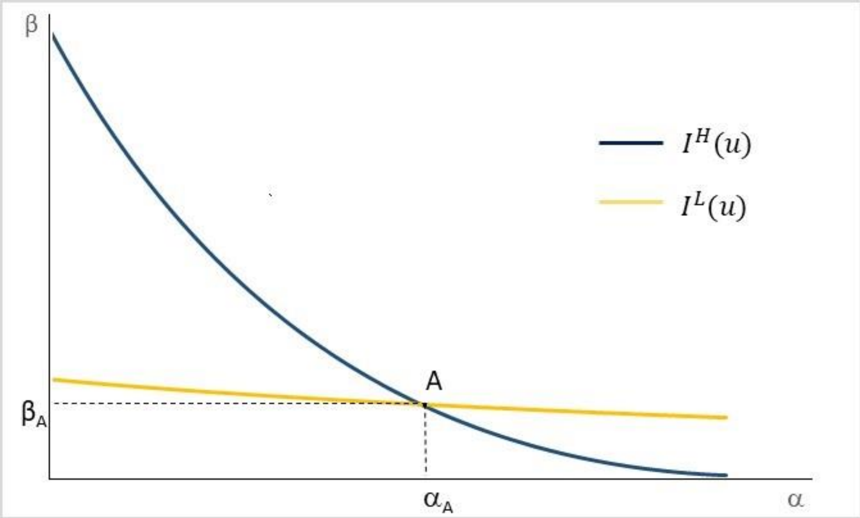


Figure 3 Indifference curves ef^L and ef^H for the same level of utility

Figure 3 shows that the slope of the indifference curve $I^H(u)$ is steeper than $I^L(u)$. This implies that, for the same level of utility , the indifference curve will have a point of intersection (point A) in which the agent is indifferent between ef^H and ef^L :

$$EU^H(\alpha_A, \beta_A) = EU^L(\alpha_A, \beta_A) \quad (17)$$

Then the point A is the switching point: in A we observe the transition from one curve to the other in the agent's preferences, but also, as it will be discussed shortly, in the contract that the principal proposes. Figure 4 shows how to identify the set of contracts that would solve the problem of minimizing spending under the constraints of participation and incentive compatibility. On the left of point A, along the curve I_{H0}, in the point B, that represent the contract (α_B, β_B), the agent will find it profitable to choose ef^L, obtaining the greater expected utility associated with the I_{L1} curve (greater than the expected utility associated with I_{L0}).

Notice that, in the point B, the incentive compatibility constraint, aiming at obtaining a higher level of effort, would not be satisfied.

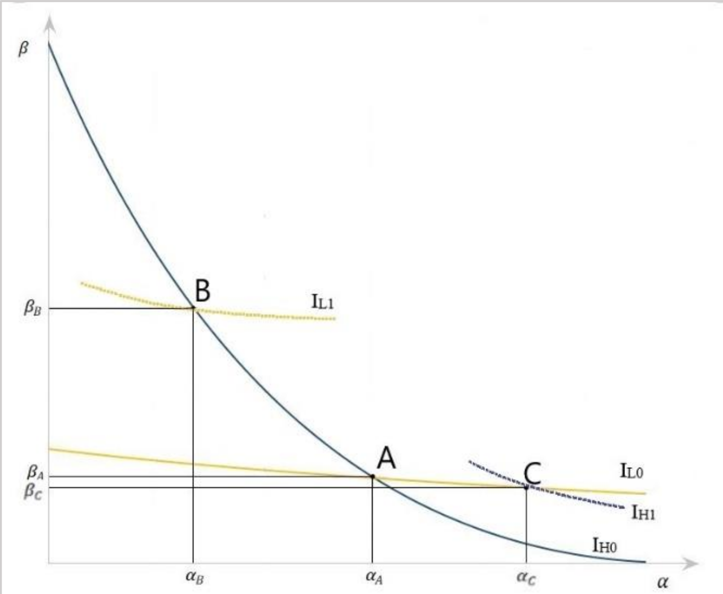


Figure 4 Indifference curves and equilibrium points

The following Figure 5 shows the line (dots line) of the feasible contracts Θ_{LH}, given a level of utility

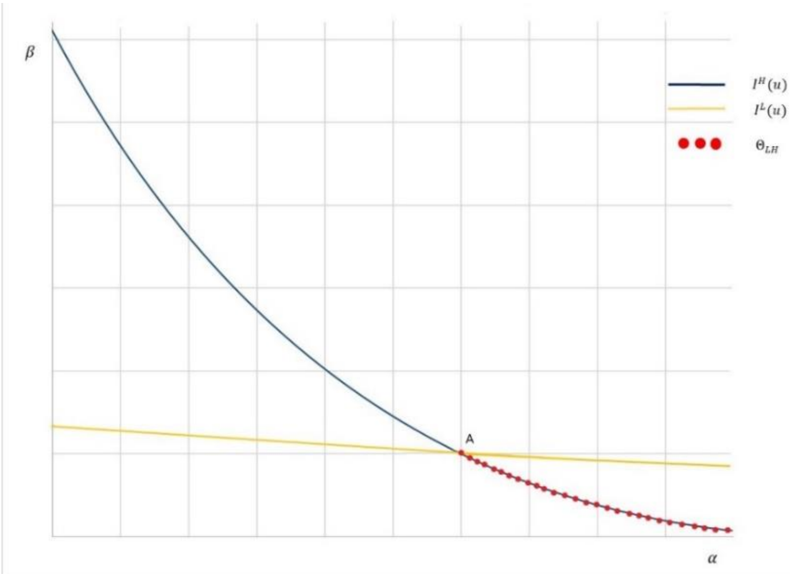


Figure 5 Eligible contracts line

In this scenario, in which we consider only two possible levels of effort, we rewrite the problem of constrained maximization referring to the principal (5) as follows:

$$\max_{\{\alpha,\beta\}}[-ObjP(\alpha,\beta)] = \max_{\{\alpha,\beta\}} \; -[(p(ef^H)\alpha) + (1-p(ef^H))\beta]$$
Subject to
$$u(\beta) + p(ef^H)\big(u(\alpha) - u(\beta)\big) - v(ef^H) \geq U_0$$
$$(p(ef^H) - p(ef^L))\big(u(\alpha) - u(\beta)\big) \geq v(ef^H) - v(ef^L) \qquad (18)$$

We look for a contract which solve the problem (12) by setting Kuhn-Tucker (1951) [i] system of equations.

[i] Appendix A

Following Wallace's method for a simple solution of optimization problems with inequality constraints (Wallace, 2004), we differentiate the Langrangean[4] of the problem with respect to the α, β, and finding the solutions of the system of equality and inequality known as the necessary Kuhn-Tucker condition we obtain the candidate as solution of the problem. Since the sufficient Kuhn-Tucker conditions are satisfied, i.e., the objective function is differentiable and concave and the constraints are differentiable and convex, we are able to find a solution of the problem. We can write the conditions of the optimal contract:

$$u_\alpha'(\alpha) = -\frac{p^H}{\lambda p^H + \mu(p^H - p^L)} \quad (19)$$

$$u_\beta'(\beta) = -\frac{1-p^H}{\lambda(1-p^H) + \mu(p^H - p^L)} \quad (20)$$

Which can be written in the following way in order to allow for a better understanding of the properties of the optimal contract:

$$\frac{1}{u_\alpha'(\alpha)} = \lambda + \mu\left(1 - \frac{p(e^L)}{p(e^H)}\right) \quad (21)$$

$$\frac{1}{u_\beta'(\beta))} = \lambda + \mu\left(1 - \frac{1-p(e^L)}{1-p(e^H)}\right) \quad (22)$$

From these result we note that $\mu \neq 0$, because we are in a situation of asymmetric information. If it would be 0 then the incentive compatibility constraint cannot be satisfied. This result shows how the existence of the moral hazard problem implies a cost for the principal.

We also note that, since $\mu > 0$[i], exactly as we assumed, the reimbursement of the agent varies with respect to the result obtained and in indirectly determined by the quotient $\frac{p(e^L)}{p(e^H)}$. Greater is the ratio $\frac{p(e^L)}{p(e^H)}$ and smaller will be the agent reimbursement.

Constant Relative Risk Aversion utility function

To better understand the theoretical model, we provide hereafter simulations by the use of an "explicit" functional form for the agent's utility function. The function that has to be chosen must satisfy the conditions previously defined.

We assume a risk averse agent characterized by a Constant Relative Risk Aversion (CARA) utility function:

$$u(y) = \frac{y^{1-r}}{1-r}, with\ r \neq 1; \quad (23)$$ [ii]

Where:

r is the relative coefficient of risk aversion;

$k \geq 1$ is the parameter representing the marginal effect of ef in terms of utility

[i] The Kuhn-Tucker condition impose that the multiplier μ of the incentive compatibility constraint must be non-negative.

[ii] Where $u'(y) = y^{-r}$ and $u''(y) = -ry^{-r-1}$ and the relative coefficient of risk aversion is $r = -y\frac{u''(y)}{u'(y)}$

We face a concave programming problem that can be written as:

$$\max_{\{\alpha,\beta\}} \{-[(p(e^H))\alpha + (1-p(e^H))\beta]\}$$

Subject to

$$\frac{\beta^{1-r}}{1-r} + p(ef^H)\left(-\frac{\beta^{1-r}}{1-r}\right) - (ef^H)^k \geq U_0$$

$$p(ef^H)\left(\frac{\alpha^{1-r}}{1-r}\right) + (1-p(ef^H))\left(\frac{\beta^{1-r}}{1-r}\right) - (ef^H)^k > p(ef^L)\left(\frac{\alpha^{1-r}}{1-r}\right) + (1-p(ef^L))\left(\frac{\beta^{1-r}}{1-r}\right) - (ef^L)^k \qquad (24)$$

From which we derive the expressions of the two tariffs (α and β) that solve the problem of the principal:

$$\alpha = \left[(1-r)\left(\frac{1-p(ef^L)}{p(ef^H)-p(ef^L)}\right)v(ef^H) - \left(\frac{1-p(ef^H)}{p(ef^H)-p(ef^L)}\right)v(ef^L) + U_0\right]^{\frac{1}{1-r}} = \left[(1-r)\left(\frac{1-p(ef^L)}{p(ef^H)-p(ef^L)}\right)(ef^H)^k - \left(\frac{1-p(ef^H)}{p(ef^H)-p(ef^L)}\right)(ef^L)^k + U_0\right]^{\frac{1}{1-r}} \qquad (25)$$

$$\beta = \left[(1-r)\left(\frac{p(ef^H)}{p(ef^H)-p(ef^L)}\right)v(ef^L) - \left(\frac{p(ef^L)}{p(ef^H)-p(ef^L)}\right)v(ef^H) + U_0\right]^{\frac{1}{1-r}} = \left[(1-r)\left(\frac{p(ef^H)}{p(ef^H)-p(ef^L)}\right)(ef^L)^k - \left(\frac{p(ef^L)}{p(ef^H)-p(ef^L)}\right)(ef^H)^k + U_0\right]^{\frac{1}{1-r}} \qquad (26)$$

These expression are the solution of the minimization problem of the principal that, together with the partial derivatives (provided in Appendix B), allow us to understand how each variable impacts on the expected expenditure of the principal after the contract is set.

Empirical Model 2: The probability of the service depends on the effort and the type of the user

In this model we consider the probability depending both on the $ef \in [0,1]$ and the type/severity of the user, such that by the complexity of his psycho-physical situation represented by the variable$g \in [0,1]$. Then, as in equation (1) we have *p=p(ef,g).*

Let us recall some assumption we made. A higher level of effort determines a higher probability of success, while the type *g* of the individual negatively affects the probability. The agent's effort is endogenous, that is, a choice variable that the agent can arbitrarily set to maximize its utility, while patient's severity is assumed to be exogenous and randomly distributed across patients. Since, due to the wide spectrum of implications and facets that a condition of disability can assume, the endogeneity of the variable that represents complexity contains the uncertainty deriving from the indispensable collaboration and acceptance of the service by the user.

Since the variable describing the outcome of the service is binary (1 if the service is provided, 0 otherwise), it will have a binomial probability distribution. Each service has its own specific probability: every time the agent chooses a certain effort *ef*, the probability of outcome 1 ultimately will depend on the user type *g*.

Figure 6.1 and Figure 6.2 show how the outcomes of the service (*w*) would be distributed with respect to the variables *ef* and g. The black dots indicate the "binary" outcome *w* of the service, according to the level of *ef*: as effort increases, more likely outcome equal to 1 occurs.

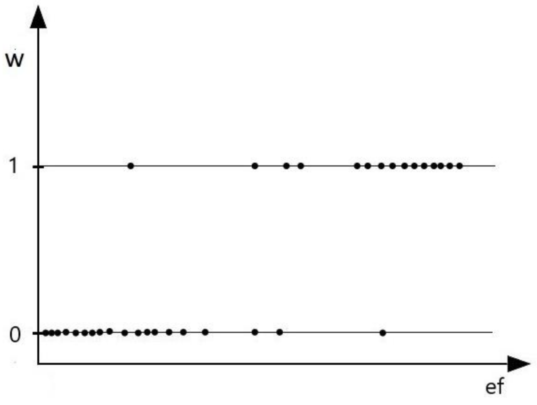


Figure 6.1 Scatter plots representing the outcome variable and the effort variable ef

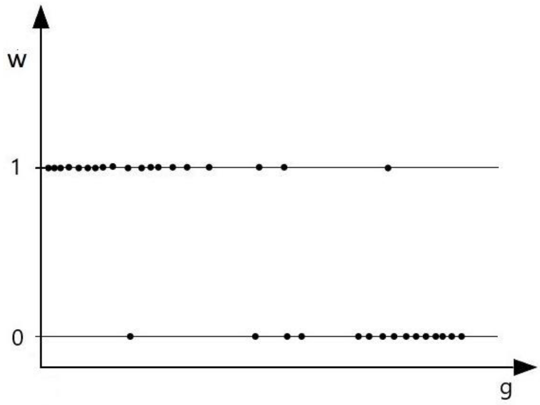


Figure 6.2 Scatter plots representing the outcome variable and the user characteristics variable g.

What it is shown in figure 6.1 is that for low values of *ef* many outcomes are 0 and for high effort values many outcomes are equal to 1.This figure explain tha assumption that effort may have an influence on the probability that different outcome may be achieved. Similarly, but in the opposite direction, is the influence of variable representing the the user complexity (g). In Figure 6.2 is shown that for low values of g many successes are obtained (w=1) and for high complexity values many outcomes equal to 0 are obtained. .

Given W the dichotomous variable, and p probability of positive outcome / success $p = p(W = 1)$, the expected value of W equal p in fact : $\mathbb{E}[W|X_1 \dots X_2] = \mathbb{P}[W = 1|X_1 \dots X_2] = p$.[i]

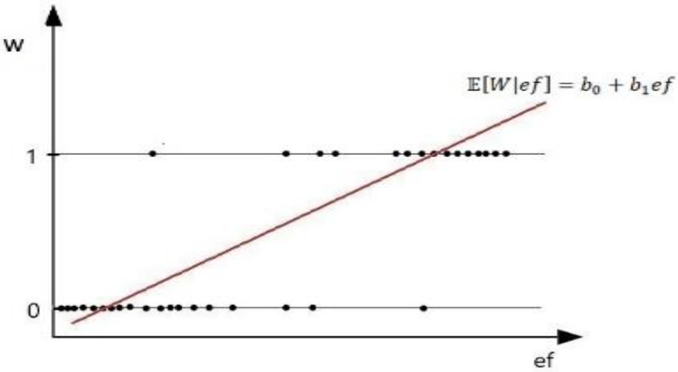


Figure 7 Linear probability Model

So the regression tool, estimating the expected value of the dependent variable, estimates a model of probability. In our study we consider only two explanatory variables then we model $\mathbb{E}[W|ef, g] = \mathbb{P}(w = 1) = \mathbb{P}(w = 1|ef, g)$[ii].

The application of a multiple linear regression model to a binary dependent variable is not adequate because it results in two types of problems. First, it establishes a linear relationship between the result variable and the explanatory variables: the model $p(w = 1|ef, g) = b_0 + b_1 ef + b_2 g$, as shown in Figure 7, represents a linear relation between the variables that implies a constant effect of the regressor on the outcome variable.

[i] $\mathbb{E}[W|X_1 \dots X_2] = p * 1 + (1 - p) * 0 = p = \mathbb{P}[W = 1|X_1 \dots X_2]$

[ii] The regression estimates the expected values of the dependent variable $\mathbb{E}[W]$

The second issue, directly deriving from the first, is that the model could return values greater than 1 or smaller than 0, which is not admissible for a probability.

Our response variable being a probability is limited both below and above, but it shows a peculiarity: with respect to the causal variables that we take into consideration, it does not necessary have a linear trend, therefore it grows as the effort increases in a non-necessarily-linear way (and its shape is also affected by the value assumed by *g*). Based on these considerations, we decided to use the logistic function to represent the relationship described above.

Logistic function

The logistic function originally was proposed to represent the evolution of a population over time by Verhulst (1838)as an alternative to the Malthus model. It found many applications in various research fields (Berkson, 1944; Cramer, 2002).

Considering x the independent variable and *z(x)* a dependent variable outcome (or performance), we write a generic non-linear growth model of z(x)

$$\dot{z(x)} = \frac{dw(x)}{dx} = bz(x) - \varphi\big(z(x)\big) \qquad (27)$$

Where $\varphi\big(z(x)\big)$ is the non linear component of the evolution of z(x).

The following quadratic differential equation defines the logistic equation:

$$\dot{z(x)} = Az(x) - Bz(x)^2 \qquad (28)$$

Assuming that the growth of w depends on a constant γ>0 and on the distance to its extreme values z_{inf} e z_{sup}, we rewrite the equation (14):

$$\dot{z(x)} = \gamma[z(x) - z_{inf}][z_{sup} - z(x)] \qquad (29)$$

Solving the differential equation we obtain:

$$z(x) = \frac{z_{sup}e^{\gamma(z_{sup}-z_{inf})(x-c)} + w_{inf}}{1+e^{\gamma(z_{sup}-z_{inf})(x-c)}} = \frac{z_{sup}}{1+e^{-\gamma(z_{sup}-z_{inf})(x-c)}} + \frac{z_{inf}}{1+e^{\gamma(z_{sup}-z_{inf})(x-c)}} \qquad (30)$$

Setting $z_{inf} = 0$ e $\gamma^* = \gamma\big(z_{sup}\big)$:

$$z(x) = \frac{z_{sup}}{1+e^{-\gamma^*(x-c)}} \qquad (31)$$

The generic formulation of the logistic function

$$z(x) = \frac{k}{1+e^{a-bx}} \qquad (32)$$

Where $a = c\gamma^*, b = \gamma^*, k = z_{sup}$.

The logistic function has the following properties (Winsor, 1932): the inflection point, where the curve changes its concavity direction and with respect to which the function presents a symmetric behaviour, is at the point $x = \frac{a}{b}$ and $w = \frac{k}{2}$. In the interval $(0;\frac{a}{b})$ the increase of x determines the increasing marginal variations in z. For$x > \frac{a}{b}$,, increasing x causes decreasing variations of z.

The point of intersection with the vertical axis is given by: $x = 0$ e $z = \frac{k}{1+e^{a}}$. The growth rate of z is $\frac{dz}{dx} = \frac{b}{k}\, z(x)(k - z(x))$.

Generally the logistic function is used in its standard formula, uniquely identified by two parameters, μ and s, and with the upper asymptote equal to 1 (k = 1):

$$z(x) = \frac{1}{1+e^{-\left(\frac{x-\mu}{s}\right)}} \qquad (33)$$

Set μ, s> 0 the position and scale parameters respectively: μ defines the center of symmetry of the curve and *s* acts on the shape of the curve.

The position parameter determines the inflection point A (Figure 8.1) defined by $x = \mu$ and $z(\mu) = \frac{1}{2}$. Increasing μ, the inflection point moves to the right and therefore the value of x increases so that the yields

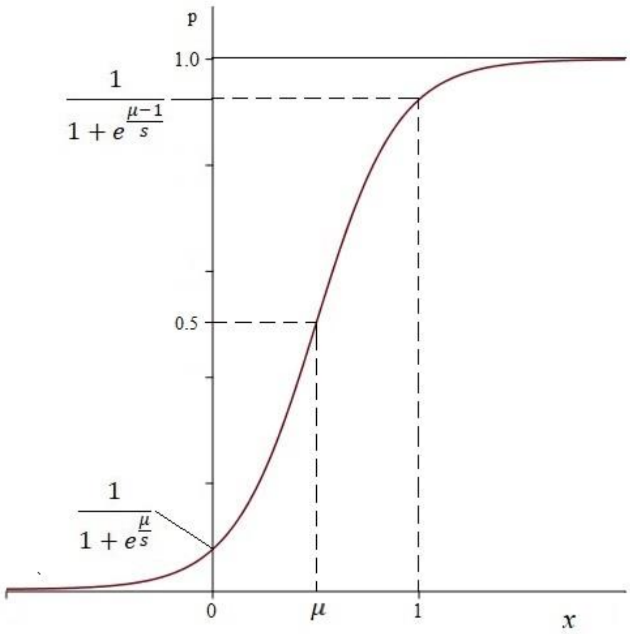


Figure 8.1Standard Logistic Function

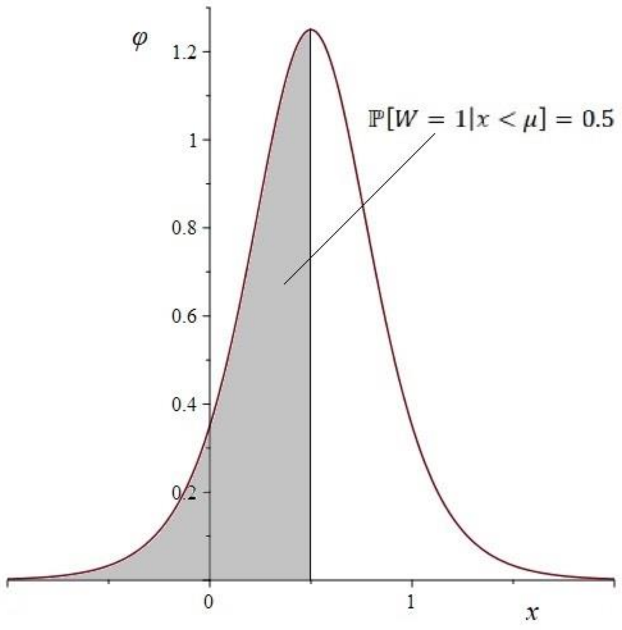


Figure 8.2 Density Function of the logistic distribution

begin to decrease. Moreover, considering $x \in [0,1]$[i], with the same number of s, increasing μ the points $z(0) = \frac{1}{1+e^{\frac{\mu}{s}}}$ and $z(1) = \frac{1}{1+e^{\frac{\mu-1}{s}}}$ go down.

The scale parameter influences the shape of the curve and we note that increasing s and keeping μ fixed, the point $z(0) = \frac{1}{1+e^{\frac{\mu}{s}}}$ rises and the point $z(1) = \frac{1}{1+e^{\frac{\mu-1}{s}}}$ goes down.

When the joint distribution of two variables tends to have a linear shape, equal variations of *x* correspond to constant variations of the variable w. We interpret the shape of the curve as the degree of sensitivity of the outcome to changes in the explanatory variable. So, in the case of the logistic function, for large values of *s*, in the interval $x \in [0.1]$ the curve relationship tends to be straight and therefore the relationship between the

[i] We use $ef \in [0,1]$

variables tends to be linear. On the other hand, when s is small and the curve has a sigmoidal shape, unitary variations of x correspond to non-constant variations of z.

Application to the transport service

The scenario described in the present study presents characteristics which satisfy the conditions proposed by

Pearl and Reed (1920) that make the logistic function suitable to represent the probability we are interesting to investigate in our model.

Assuming a dichotomous outcome variable, the model estimate a probability. Obviously, the results are limited between zero (lower bound) and one (upper bound), the probability's values vary continuously as the explanatory variables vary; and, as we assume a non-linear relationship with the explanatory variables probability varies in a non costant way.

In particular, for effort levels below a given threshold (i.e., below the inflection point), the marginal effect of a increase of *ef* on *p* is bigger with respect to points located above the afore-mentioned threshold.

Since we have assumed the type of patient g to be exogenous we analyse the probability, according to the effort, given different levels of g.

Then, according to our assumptions, the probability of the positive outcome is inversly correlated to the complexity of the patient. This assumptions can be represented by a higher value of μ. To each complexity level g_j corresponds a given μ_j. In this way, for each j-th type of user, we can represent the distribution of probability as a logistic function, i.e., for each value g_j we define *p* as a function of the effort:

$$(W = 1|g_j, ef) = p_j(ef) = \frac{1}{1+e^{-\left(\frac{ef-\mu_j}{s}\right)}} \qquad (34)$$

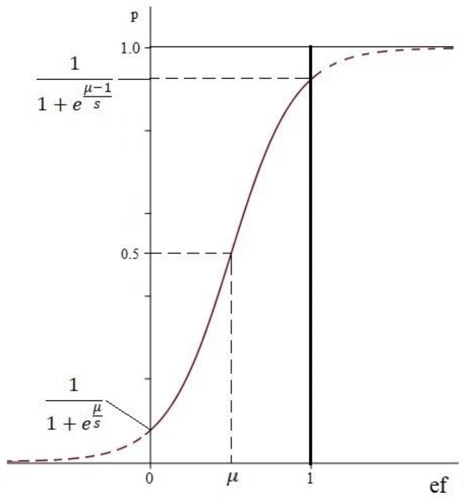


Figure 9 The logistic model of the probability as function of the effort

Provided that we are considering a market characterized by asymmetry of information and uncertainty any outcome can be granted for sure. Infact for any value assumed by the parameters μ and *s*, and considering $ef \in [0,1]$ the probability *p* never reaches levels 0 and 1, (in which there is no uncertainty) (Figure 9 and Figure 10)

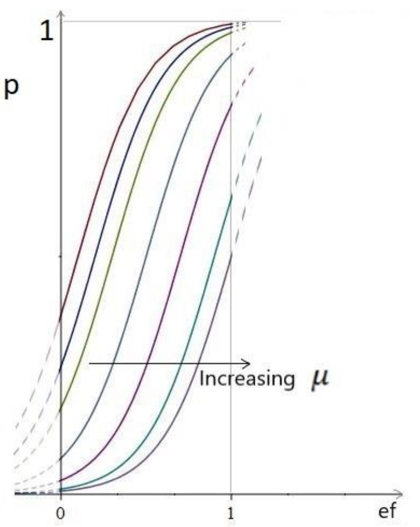


Figure 50 Shape of the logistic curve according to μ

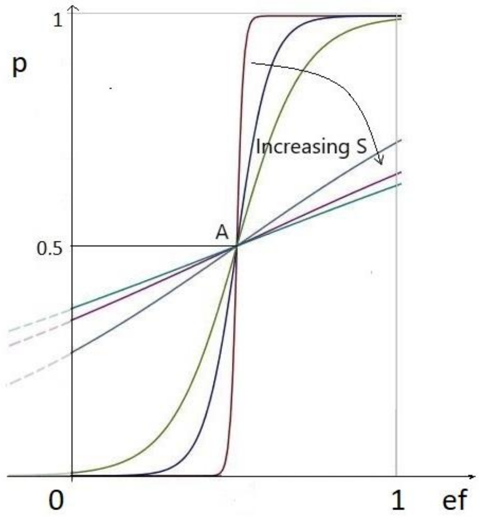


Figure 6 Shape of the logistic curve according to s

Looking at figures 10 and 11 we can make some consideration on the effect of μ and s on the probability.

The marginal effect of effort in terms of probability increase when *ef* <μ and decrease when *ef*> μ. The parameter μ allows to identifiy the effort threshold under which the increase in *ef* is more effective and beyond which is less effective. Therefore, provided a given value of *s*, an increase of μ determines a shift to the right of the curve and, as a consequence, for each level of *ef*, an absolute reduction of expected probability (Figure 10). Then, we assume that μ_j indicates the complexity of the j-th patient because having assumed that the complexity of the patient affects, indipendently from other variables, the probability of the outcome, the ability of the effort of be effective in obtaining positive results that is what parameter μ represent as just described. . In other words, the more complex the user is, the less is the influence of the agent effort on the probability of a positive outcome.

Leaving μ unchanged, increments in the parameter *s* causes a reduction in the variability of the marginal effet of effort on the probability[i]. In the interval *ef*$\in[0,1]$ the curve tends to flatten, we interpret this factc as the variations in probability become less sensitive to variations in effort, because increasing *s*, the effort returns tend to become constant (Figure 11) and consequently the probability is less affected by variations in the effort. We use this observation to describe several types of services: a low specialization service will have a probability that is little affected by the agent's effort and then curve will be straight or even with downward concavity (*s* high); vice versa, a highly specilization service could be more influenced with a very marked sigmoidal curvature (*s* low).

Based on these considerations, we assume the paramenter s describes the type of service.

For a given level of user complexity (μ), we describe three cases which, depending on the level assumed by s, represent highly specialized transport (typically disabled transport), medium specialized transport and little or no specialized transport (ordinary buses), in which two types of patients are compared (patient A with

[i] In figure 11 is show that the curve tends to become a straight line increasing s. This implies that constant marginal variations in probability.

a lower degree of complexity and patient B with a greater degree of complexity). User A with a low level of complexity type $\mu_A = 0{,}5$ and user B is an high severity type $\mu_B = 0{,}8$. In the case of high specialization service we consider for $s^* = 0{,}2$, for medium specialization service $s^* = 0{,}5$, and for low specialization service: public transport (bus or train) $s^* = 0{,}9$.

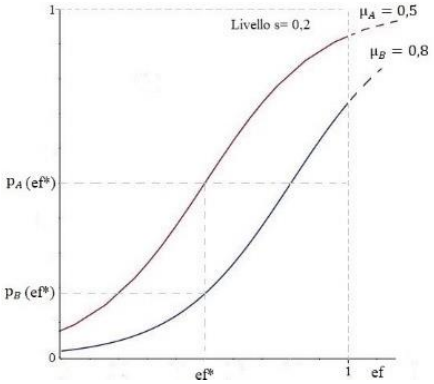


Figure 12 High specialization service

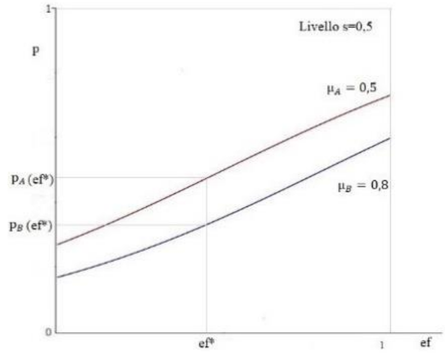


Figure 13 Medium specialization service

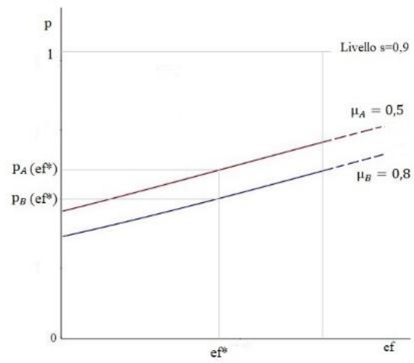


Figure 14 Low specialization service

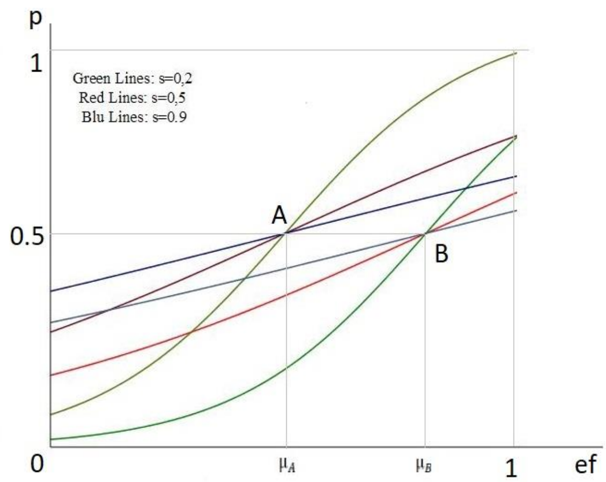


Figure 15 Comparison of different scenarios varyng the parameters μ and s

3. Results: Simulation Analysis

Although the attempt to simplify the reality under investigation, the service at stake contains several intrinsic characteristics that determine a considerable complexity, for this reason we have implemented various simulation analyses to understand and describe the model.

Let us briefly recall the model and our assumptions. The dependent variable is the outcome of the transport service that we consider can be performed or not, that is, it can assume only two values zero and one. We consider two causal variables, that influence the probability of the outcome in different ways: the agent effort, $ef \in [0,1]$ which positively affects the probability and the variable describing the complexity of the user of the transport service, $g \in [0,1]$ which instead has a negative impact on probability. The principal is risk neutral and aims at guaranteeing the service; in other words he looks to get high probability, minimizing the expected costs. We assume that the principal, for the definition of the optimal contract, can consider different values of effort.

The risk averse agent acts with reference to his own expected utility function (17), of CARA type, which is increasing in the reimbursement obtained for the work done and negatively affected by the effort applied because it creates disutility.

In Table 1 is presented the simulation referring to the Empirical Model 1 in which the probability depends only on the agent's effort. Risk aversion, probabilities value and reservation utility are varied separately to assess their impact on the values at stake.

The subsequent simulations refer to the Empirical model 2 that considers the probability as function of both the agent effort and the user health complexity. The logistic functional form is applied to the probabilities, and in the simulations presented in Tables 2,3,4 parameters are varied "independently" from each other.

The last simulations present the dynamic of the contract with reference to two different scenarios: the first is referred to the scenario in which the user presents a medium health complexity level, while the second considers a high complexity user.

3.1. Simulation Empirical Model 1: Probability depends only on the agent effort

Two effort levels

The first simulation refers to the scenario in which the probability only depends on the variable *effort* and the users show the same complexity level. We consider the case, presented above, in which *effort* can assume only two values $ef^L = 0\,, ef^H = 1$ (see Section 2 "Empirical Model 1"). The agent by choosing the effort level to be applied ef^L or ef^H, determines respectively p^L and p^H, where $p^L < p^H$, i.e., the probability that the service is performed. We analyse what happens to the reimbursement values and to the principal expected costs, varying separately the probabilities values p^H *and* p^L, the reservation utility (*U₀*) and the agent risk aversion coefficient (*r*) .The results reported in Table 1 are derived solving the minimization problem of the principal. (Software code shown in Appendix C).

Table 10 Results of the simulation with two levels of effort $ef^H = 1, ef^L = 0$

	Varying p^H				Varying p^L				Varying U_0			Varying r			
r	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	**0,2**	**0,3**	**0,6**	**0,8**
k	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
p^H	**0,5**	**0,8**	**0,9**	**0,999**	0,8	0,8	0,8	0,8	0,8	0,8	0,8	0,8	0,8	0,8	0,8
p^L	0,2	0,2	0,2	0,2	**0,3**	**0,4**	**0,5**	**0,7**	0,2	0,2	0,2	0,2	0,2	0,2	0,2
U_0	10	10	10	10	10	10	10	10	**5**	**1**	**0,5**	10	10	10	**10**
α	40,11	32,11	31,04	30,26	32,49	33,06	34,03	42,25	10,03	1,36	0,84	15,73	19,27	43,76	59,83
β	21,77	23,36	23,59	23,76	22,09	20,25	17,36	2,25	5,44	0,11	0,01	12,90	15,35	29,40	27,01
S[§]	***30,94***	***30,36***	***30,30***	***30,25***	***30,41***	***30,50***	***30,69***	***34,25***	***9,11***	***1,11***	***0,67***	***15,17***	***18,49***	***40,89***	***53,27***
$\alpha - \beta$	18,34	8,75	7,45	6,49	10,40	12,81	16,67	40,00	4,58	1,25	0,83	2,84	3,92	14,36	32,82
α/β	1,842	1,37	1,32	1,27	1,47	1,63	1,96	18,78	1,84	12,25	121,78	1,22	1,26	1,49	2,22
λ^*	5,5	5,5	5,5	5,5	5,5	5,5	5,5	5,5	3	1	0,75	1,72	2,39	9,24	23,91
η^{**}	1,38	0,22	0,09	8E-04	0,32	0,5	0,89	8	0,22	0,22	0,22	0,02	0,04	0,55	3,31

[§] Principal Expected Cost; [*] Participation Constraint Lagrange multiplier (shadow prize); ** Incentive Compatibility Constraint Lagrange multiplier (shadow prize)

Looking at Table 1 we note that as p^H increases, principal expected costs decrease, in fact, ceteris paribus, if the probability of the transport with high effort increases, there are two element to be considered: first, when agent puts high effort, the risk of failure in absolute decreases, then the risk borne by the agent is reduced; second, the high effort level become relatively more effective in producing positive outcomes. As a consequence, the reduced risk determines a reduction in the risk premium, so the difference between the two reimbursements tends to decrease significantly. In particular α decreases and β increases; furthermore, the strength of the incentive-compatibility constraint is reduced by the fact that the agent is already incentivised by the resulting higher probability, p^H and it is shown by the reduction of the weight of the incentive compatibility constraint, i.e. the parameter η.

On the opposite increasing the probability with low effort exploited, p^L, while keeping all the other variables constant, implies that the low effort becomes relatively more able to guarantee the transport and that the risk of negative outcomes is reduced in absolute.[i] In this case grows the probability of a success even with low effort, then, the incentive in exploiting high effort must be stronger than the opportunity deriving from low effort and it is observable in the increment of weight of the incentive constraint represented by η. This implies, that α increases and β decreases. As a consequence *(α-β)* increases and expected costs *S* increase.

Clearly λ remains constant provided that it "depends" on the reservation utility. In fact the reservation utility is assumed equal to 10 in the first two scenarios. In the 4[th] scenario λ varies because the aversion to risk, and as a consequence the utility function, of the agent varies. On the other hand η varies in the same direction of the expected costs variation implying that a lower premium to risk, in order to meet the incentive compatible constraint, is required.

By reducing the reservation utility U_0, the expected costs function is reduced, the weight of the participation constraint λ decreases and μ, the weight of the Incentive-Compatibility constraint, remains constant.

As the risk aversion coefficient r increases, the expected cost function increases. This is in line with the fact that as the agent's risk aversion increases, his risk premium and therefore the principal's expenditure increases. λ increases and η increases as well.

3.2. Simulation Empirical Model 2: Probability depends on effort and complexity of user variables

In the second empirical model, that considers the probability as a function of both the effort, $ef \in [0,1]$, and the user complexity $g \in [0,1]$, we define the base scenario as follows: in the logistic formulation of probability to provide the transportation service, and considering that we are interested in analysing highly specialized services of transportation for disabled people, we assume *s*=0,2[ii].

---

[i] Because, given $p^L \leq p^H$, and $(1-p^L) \geq (1-p^H)$, if p^Lincreases, $(1-p^L)$ decreases and even $(1-p^H)$ decreases

[ii] Remember that, since the scale parameter of the logistic function represent the shape of the curve and then contains the sensibility of the result variable to the causal variable variations, we assumed that the *s* of represent the specialization of the service: the probability of a positive outcome is less influenced by the agent effort for highly specialized service (*s* low) than for service with low specialization (*s* high) (for more details see section *Logistic Function- Application to the transport service*).

Since we analyse a public service created and offered to selected categories of users, we consider as the base case the one in which the user has an average complexity μ=0,5 (See section 2 *"Empirical Model 2"* and Figure 10 , 11, 12, 13 and 14 for details on how the parameters are able to capture these variables). We propose hereafter several simulations. First, we analyse what happens to expected costs and reimbursement values when there are only three relevant effort values: the extremes (*ef*=0, *ef*=1) and the point at which the logistic function changes concavity (the inflection point) that is where *ef*=μ. To be noticed that, on the left hand side of the inflection point probability (modelled as a logistic function) shows increasing marginal returns while on the right hand side marginal returns begin to decrease (Figure 9).

In the second simulation we start by the definition of the level of effort necessary to guarantee a certain probability level. Subsequently, using these values, we compare what happens to principal's expected costs with respect to three different users (with different severity).

In the third simulation we focus on the effect of marginal effort variations on the variables of interest for the principal, i.e., probabilities, reimbursement values, expected costs, parameter *s*.

Finally, the contract dynamics in two scenarios is presented. The base scenario (in which we consider an highly specialized service *s*=0,2 and medium complexity users, μ=0,5) and another scenario where the service is highly specialized (*s*=0,2) and the user present high level of complexity (μ=0,8)[i].

Comparison of different complexity users

We analyse the expected cost minimization problem, with a probability function with the logistic form (equation 28), comparing a low level and a desired level *ef^* of effort. We assume that the relevant effort values are only three: the minimum *ef*=0, the maximum *ef=1* and inflection point *ef*=μ.[ii] Then we analyse what happens to the principal's expected costs when he sets the minimum level of effort *ef=0* and compare it with the desirable *ef=1* and *ef*=μ separately (in Table 2 we refer to Case 1 and Case 2); and when he sets the lower effort level *ef*=μ and compares it with *ef=1* (Case 3). Considering only highly specialized services, represented by the parameter *s* set equal to 0,2, we compare three users who differ each other by complexity levels: user A is assumed low complexity (μ=0,3), user B medium complexity *(*μ=0,5), and user C high complexity *(*μ =0,8*)*.

Assuming, also in this framework, an agent with constant risk aversion (CARA) utility (17) characterized by *r*=0,5, *k*=1 we compare the three users A, B and C previously described, with respect to different ef^ principal's desired effort (Table 2).

[i] See Note 12 and section *Logistic Function*

[ii] The inflection point is considered relevant because, having modelled a sigmoidal growth of probability, for effort values lower than the inflection point μ, there are increasing marginal returns while for higher values marginal returns begin to decrease

Table 11 In the context of highly specialized service (s=0,2) Principal expenditure for different type of users

	Case 1			Case 2			Case 3		
$ef^L=0$	ef^=1			$ef^L=0$	ef^= μ		$ef^L=\mu$	ef^=1	
	User A	User B	User C	User A	User B	User C	User A	User B	User C
μ	0,3	0,5	0,8	0,3	0,5	0,8	0,3	0,5	0,8
$p(ef^L)$	0,18	0,08	0,02	0,18	0,08	0,02	0,50	0,50	0,50
$p(ef$^)	0,97	0,92	0,73	0,50	0,50	0,50	0,97	0,92	0,73
α	9,11	9,27	10,17	8,32	9,27	10,99	9,13	9,27	9,71
β	5,68	6,02	6,19	5,82	6,02	6,17	5,19	6,02	7,02
λ°	*3,00*	*3,00*	*3,00*	*2,56*	*2,75*	*2,9*	*3,00*	*3,00*	*3,00*
$\eta^{\circ\circ}$	*0,02*	*0,05*	*0,19*	*0,37*	*0,35*	*0,43*	*0,04*	*0,10*	*0,37*
α-β	3,43	3,25	3,98	2,5	3,25	4,82	3,94	3,25	2,69
*S**	9,01	9,02	9,10	7,07	7,65	8,58	9,01	9,01	8,98
*EU***	5,03	5,08	5,27	4,80	5,00	5,30	5,03	5,07	5,25

° Participation Constraint Lagrange multiplier (shadow prize); °° Incentive Compatibility Constraint Lagrange multiplier (shadow prize)**Principal expected Cost,* ***Agent Expected Utility*

Of course, it must be pointed out that Case 2 and Case 3 can only provide information on the trend of the variables of interest, because their formulation depends on the complexity of each user and on a priori knowledge of this value.

The table can be summarized as follow. Increasing the complexity, within each case, the reimbursement values, α and β, and the agent expected utility, *EU,* increase. The principal expected costs, S, and the difference between α and β have different trends with respect to the increment of the user complexity, in particular, *S* increases in Case 1 and Case 2 but tends to decreases in Case 3; the difference (α-β) has not a constant course in Case 1in which it decreases from user A to user B and increases from user B to user C; it increases in Case 2 and decreases in Case 3. It is observable that for user B the reimbursement values are the same in each case, and consequently their difference.

Isolating each user we note that principal expected costs decrease from case 1 to case 2 and increase to case 3. The expected Utility has the same trend for user and B but opposite do user C, for which increases from case 1 to case 2 and decreases to Case 3. The trend of the difference between the reimbursement values (α-β) for user A decreases from Case 1 to Case 2 and increases to Case 3, for user B it behaves in the opposite way and for user C it is decreasing from Case 1 to Case 3.

For the principal the most convenient strategy, for every user complexity, is comparing $ef^L=0$ with *ef*^= μ (Case 2) because the expected costs are always lower than in other cases. But this implies reaching the probability *p(ef^)*=0,5. For the agent Case 2 represent the case in which the expected utility has more variability with respect to the increment of user complexity. Furthermore, for the low complexity user (user A) the expected utility is lower than reservation value, this means that participation in contract could be not ensured.

As seen in the previous simulation (Table 1) variations in probabilities, $p(ef^L)$ *and p(ef^)=* $p(ef^H)$, ceteris paribus, have implications on the absolute risk of failure and on the relative ability of one effort level respect to the other one to guarantee the positive outcome.

In Case 1, similarly to the previous simulation (Table 1), are compared $ef^L = 0$ and $p(ef^{\wedge})=1$, and, as expected, increasing the complexity of the user, the respective probabilities decrease. This implies that the risk of failure (*1-p*) increase in absolute both with respect to high effort and low effort. As a consequence, the risk-averse agent, may require an higher reimbursement β (which is received when the service is not performed) as compensation for negative results even if agent applies high effort. Furthermore, if $p(ef^{\wedge})$ is reduced the agent needs higher incentive to apply higher effort, because high effort has become relatively less convenient than low effort, and then α must increase then the strength of the incentive-compatibility constraint is augmented as shown by the fact that the respective Lagrange multiplier η increases.

In Case 2, in which principal incentives $ef^{\wedge}=\mu$ against $ef^L=0$, the desired probability $p(ef^{\wedge})$, is constant among different complexity users and $p(ef^L)$ decreases with the increment of user complexity. If $p(ef^L)$ decreases determines a higher risk of failure and a relative less opportunity in applying lower effort. As consequence of the augmented risk, β increases. The reduction in the relative convenience in applying low effort would imply a slack of the incentive constraint and then a reduction in α. But, the increasing in user complexity determines exogenous uncertainty in the results that with the risk aversion of the agent determine that α increases. Finally the difference *between* the reimbursement values, α and β increases. We justify this result with a consideration: the reduction in the opportunity in applying low effort implies a slack of the incentive constraint that is weaker than the increased risk of not performing the service, therefore the risk premium increases.

In Case 3 (in which principal incentives $ef^{\wedge}=1$ against $ef^L=\mu$), the lower probability value $p(ef^L)$ remains constant and $p(ef^{\wedge})$ decreases. This implies that the risk of failure increases and that the higher effort level becomes relatively less suitable in ensuring a successful outcome. Consequently the weight of the incentive compatibility constraint increases, and α increases. Furthermore, the increased risk of negative outcomes determines that β increases as compensation.

We observe that increasing the complexity of the user, both $p(ef^L)$ and $p(ef^{\wedge})$ decrease. Then there is in general more risk and the If $p(ef^L)$ and $p(ef^{\wedge})$ decrease at different rates and their distance tends to increase when the transport is successfully done, it most likely that it comes from the application of high effort.

Derivation of effort level to guarantee a given probability level

As before, *s* specifies the degree of specialization of the service that is a highly specialized one, so *s*=0,2, and $\mu \geq 0$[i] represents the complexity of patient, the probability function is the logistic form (equation 28). The agent utility function is CARA (equation 17) and we consider U_0 =5 the value of the reservation utility.

In this simulation first we determine the value of the agent effort *ef* able to guarantee a certain level of probability *p** desired by the principal for each type (in terms of complexity) of user. Secondly some considerations are made on the outcome variables such as reimbursement values, expected cost, expected utility.

[i] See Note 12 and section *Logistic Function*

The problem is to determine $ef^{\wedge}$ that satisfies $p(ef^{\wedge}) \geq p*$

That is, for every user complexity *j*, we solve the following inequality:

$$p_j(ef^{\wedge}) = \frac{1}{1+e^{-\left(\frac{ef^{\wedge}-\mu_j}{s}\right)}} \geq p*$$

For example, if we set *p**=0,8 the probability level desired by the principal the problem is:

$$p_j(ef^{\wedge}) = \frac{1}{1+e^{-\left(\frac{ef^{\wedge}-\mu_j}{0,2}\right)}} \geq 0,8$$

$$ef^{\wedge}{}_j \geq \mu_j + 0,277$$

This result[i] suggest that the effort necessary to ensure a certain level of probability depends on the user complexity. Since we assumed *ef* ∈ [0,1], the inequality $\mu_j + 0,277 \leq 1$ must be satisfied, and then $\mu_j \leq 0,723$. So, if $\mu_j > 0,723$, the principal cannot guarantee the level of probability *p** because it would require *ef>1* . Then, fixed the probability level p^*, the principal for each user complexity, can estimate the effort needed to guarantee that probability level.

In the last column is shown that if the principal considers the worst situation of the complexity, i.e. $\mu_j = 0,8 > 0,723$,and he set *ef*=1* he will never reach the probability level desired.

If he takes account of the best case scenario, i.e. $\mu_j = 0$ (no complexity in the user) then he will set *ef^=0,277.*

Here we first impose the level of probability $\bar{p} = 0,8$ as the level aimed by the principal, then we solve the minimization problem for different users with some complexity (μ_j).

[i] Results for different objective level of probability are presented in Appendix in Table A1

Table 12 In the context of highly specialized service (s=0,2), Principal expected expenditure for different users, fixed desired probability

p^*	0,8	0,8	0,8	0,8	0,799
μ	0,3	0,5	0,6	0,7	0,8
ef^L	0	0	0	0	0
$p(ef^L)$	*0,18*	*0,08*	*0,05*	*0,03*	*0,02*
$ef^{\wedge} = ef^H$	0,58	0,777	0,877	0,977	1
α	8,3	8,97	9,33	9,7	10,09
β	5,83	6,04	6,11	6,16	6,19
α-β	2,47	2,93	3,22	3,54	3,9
S *	7,81	8,38	8,69	8,99	9,31
S variation (%)		7%	4%	4%	4%
EU**	5,57	5,77	5,87	5,97	6,07
EU variation (%)		4%	2%	2%	2%

**Expected cost* ***Agent Expected Utility*

We can see that the effort level desirable in order to guarantee a certain level of probability is not necessary always one. This is notable because, to define efficient contracts the principal knowing that could set less expensive reimbursements for less complex users.

In Table 3 we see that, fixed p^*, both the reimbursement values and expected costs increase with the complexity of the user. Variations are decreasing. The difference between the reimbursement values increase.

If the principal set the probability level p^* and based on this defines the contract, what changes with the complexity of the user is the lower probability $p(ef^L)$, which decreases, and the effort required to guarantee p^* Then, on the one hand reimbursement values could compensate the relative increment in the convenience of applying the higher effort respect lower effort (that would mean reduction of α) but the increment of the user complexity determines more uncertainty and finally that α increases. On the other hand the reduction in $p(ef^L)$ implies an absolute increment of the risk of failure, and then the participation constraint make β increasing.

If the agent would be able to apply exactly the necessary effort to guarantee the probability defined by the contract, he will increase his expected utility over the reservation value U_0.

Guaranteeing the probability level p^* implies the definition of the effort level necessary to that, but anyway it could not be sufficient as shown in the last column, in which it can be seen that beyond certain levels of patient complicity it is not possible to reach the predetermined level of probability.

How effort affect probabilities of the transport and expected costs of the principal

This simulation is intended to understand the implications of effort variation in terms of payments, probabilities and expected expenditure. In Table 4 the variations in expected expenditure and in probabilities in the base scenario (highly specialized service, defined with $s=0,2;$ and medium complexity user, defined with $\mu=0,5$[i]) are presented. We maintain the previous assumptions: the logistic functional form of the probability (as in equation 28), the CARA as agent utility function (equation 27) and the agent reservation utility is U_0=5.

[i] See Note 12 and section *Logistic Function*

Each column of Table 4 presents the optimal contract values (i.e. reimbursements, expected costs, and relative variations) which solves the principal problem comparing two near effort levels (the difference between ef^L and ef^H is 0,1). From left to right, it is implemented a growth of the efforts by steps of range 0,1, i.e. from one column to the next are compared effort values which are respectively increased by 0,1, for each of them in parenthesis, is indicated the relative probability.

Table 13 Variations in Expected Cost and Probabilities deriving from effort variation

ef^L ($p(ef^L)$)	0 *(0,076)*	0,1 *(0,119)*	0,2 *(0,182)*	0,3 *(0,269)*	0,4 *(0,378)*	0,5 *(0,500)*	0,6 *(0,622)*	0,7 *(0,731)*	0,8 *(0,818)*	0,9 *(0,881)*
ef^H ($p(ef^H)$)	0,1 *(0,119)*	0,2 *(0,182)*	0,3 *(0,269)*	0,4 *(0,378)*	0,5 *(0,500)*	0,6 *(0,622)*	0,7 *(0,731)*	0,8 *(0,818)*	0,9 *(0,881)*	1 *(0,924)*
α	12,71	10,54	9,44	8,91	8,72	8,72	8,84	9,03	9,26	9,53
β	5,82	6,03	6,22	6,38	6,48	6,48	6,31	5,89	5,07	3,74
S *	6,64	6,85	7,08	7,33	7,6	7,88	8,16	8,45	8,76	9,09
$\Delta\alpha(\%)$	-	-17%	-10%	-6%	-2%	0%	1%	2%	3%	3%
$\Delta\beta$ (%)	-	4%	3%	3%	2%	0%	-3%	-7%	-14%	-26%
ΔS *(%)	-	3%	3%	4%	4%	4%	4%	4%	4%	4%
$\frac{S^*}{ef^H}$	66,40	34,25	23,60	18,33	15,20	13,13	11,66	10,56	9,73	9,09
$\frac{S^*}{p(ef^H)}$	55,70	37,55	26,33	19,42	15,20	12,66	11,16	10,34	9,95	9,84

******Expected cost*

What is interesting is that reimbursements α and β have opposite growth trend: first α present an increasing variation, contrary β increases with a decreasing tendency. The inflection point is when lower probability $p(ef^L) = 0{,}5$ and higher probability is $p(ef^H) = 0{,}6$. The probability increases with effort but with a decreasing trend. The expected expenditure grows in each effort increment with an increasing trend.

Analisys of contracting dynamics

Given the previous assumption on probability[i], expected cost and utility[ii] function, as before, we fix the parameters of the agent utility function, r=0,5 ; k=1, and the reservation utility $U_0 = 5$.

In the context of highly specialized transport service (s=0,2) we analyse how it could be possible reaching an optimal contract in two different scenarios reducing the problems related to moral hazard and uncertainty. The base scenario (referred to a medium complexity user μ=0,5) and the high complexity user scenario (with μ=0,8) with respect two strategies of the principal that are comparing different effort values to determine the reimbursement values. In Strategy A principal comparer ef^L=0 *and* $ef^H = 1$. In Strategy B, are considered effort values with an infinitesimal difference, ef^L=0,999 *and* $ef^H = 1$.

The base scenario (s=0,2; μ =0,5[17])

Strategy A In the first case, the principal compare ef^H=1 with ef^L=0.

[i] Logistic function form

[ii] CARA agent utility function

STEP 1 the principal solves his minimization problem and, with respect to p^H=0,9241, p^L=0,075, determines the contract α = 9,2702, β = 6,0284 (point E^1 in Figure 16), for which the expected cost is *S* = 9,0243 and expected agent utility *EU*=5,0000000012.

STEP 2 The agent, who knows that the effort cannot be observed by the principal and the reimbursements α and β values, implementing a strategic behaviour, determines the effort that maximizes his expected utility with respect to the contract (α= 9,2702 and β = 6,0284), them he could obtain with *ef*=0,75, *EU*=5,07 but it is notable that this effort level determines *p*=0,78 which is lower than that foreseen by the contract

The problem, therefore, is the fact that through a strategic behaviour, with less effort the agent obtains a better result, but the principal, on the other hand, remains "screwed" because a lower probability is determined.

In Figure 16 are represented the indifference curves:

The Red curves refers to the effort level desired by the agent, *ef*=0,75; the Blue curves to *ef*=1 desired by the principal. The continuous line refers to reservation utility value U_0=5 and the asterix curve (**) considers a higher expected utility value, *U*=5,07 the achievable one with the agent strategic behaviour.

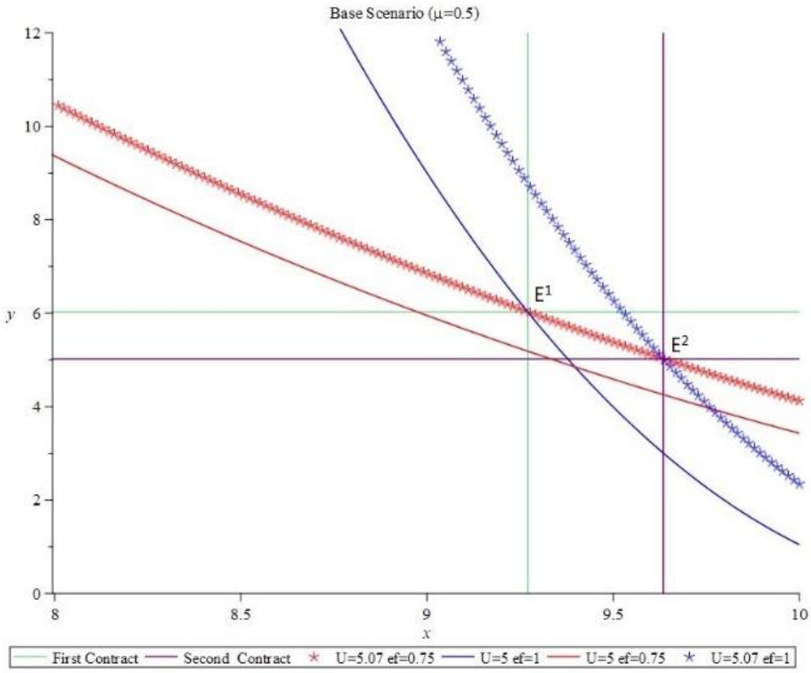


Figure 16 Base scenario: The compromise contract

We might describe the situation as following:

- The principal, who would like to be on the curve Blue Line, is willing, in order to have *p* = 0.92, to spend something more and to give the agent *U* = 5,07 and therefore pass on the Blue Asterix curve;

- the agent , who would like to be on curve Red Asterix, could agree to commit a little more and then exert *ef* = 1, to receive *U* = 5,07, then he accept to go down reaching curve Blue Asterix.

Therefore, the intersection point between the Blue Asterix and Red Asterix curves determines the compromise contract (point E^2 in Figure 16): *x* = 9,634483916 =α*, *y*= 5,022030=β* and the new value of the expected expenditure *S*= 9,28.

However, the solution just presented is not an equilibrium because following the same reasoning above, the agent could always respond with a strategic behaviour (exerting *ef*=0,75) which leads to the smaller probability

p = 0,783. The principal could only notice the rip-off (a posteriori) through the expense in fact in the case by observing the expense incurred. in fact, once the reimbursements are established on the basis of the effort exercised, obviously the expense differs ($S_{ef==1}$=9.284591555, $S_{ef=0,75}$ = 8.635651806).

Table 14 Dynamic of contract in base scenario-Strategy A

ef^H=1	$ef^L = 0$			
	STEP 1 Blu Line	STEP 2 Red Asterix	STEP 3 Red Asterix	STEP 4 Blu Asterix
α	9,2702	9,2702	9,6345	9,6345
β	6,0284	6,0284	5,02203	5,02203
ef	1	0,7500	0,7500	1
p	0,9241	0,7773	0,7773	0,9241
EU**	5,0000	5,0769	5,0735	5,0770
S*	9,0243	8,5483	8,6073	9,2846

*****_Expected cost,_ ******_Expected Utility_

Strategy B: If the principal could take as a comparison a very close lower limit of effort ef^L=, (we assume an infinitesimal effort difference ef^H=1 with ef^L=0,999) the probabilities would be very close: p^H=0,9241, p^L=0, 9237

STEP 1 the contract resulting from the minimization problem of the principal, is: α = 9,6595, β = 2,8377 with the expected expenditure S = 9,1420 and the agent expected utility EU=5.

STEP 2 The agent with respect to the contract α = 9,6595 and β = 2,8377 knows that if he behaves correctly (therefore ef = 1), he would draw the expected utility of 5. And if he implements a strategic behaviour, applying ef = 0,9995 will get: EU=5,0000005 really near to the reservation level. In this case the principal would have an expected expenditure equal to 9,1408 and p=0,924.

Table 15 Dynamic of contract in base scenario with close effort levels- Strategy B

$ef^H = 1$	$ef^L = 0{,}999$			
	STEP 1 Blu Line	STEP 2 Red Asterix	STEP 3 Red Asterix	STEP 4 Blu Asterix
α	9,6595	9,6595	9,66028	9,66028
β	2,8377	2,8377	2,8328	2,8328
ef	1	0,9995	0,9995	1
p	0,9241	0,9240	0,9240	0,9241
EU**	5,0000000	5,0000005	5,0000005	5,000001
S*	9,1420	9,1408	9,1412	9,1424

*****_Expected cost,_ ******_Expected Utility_

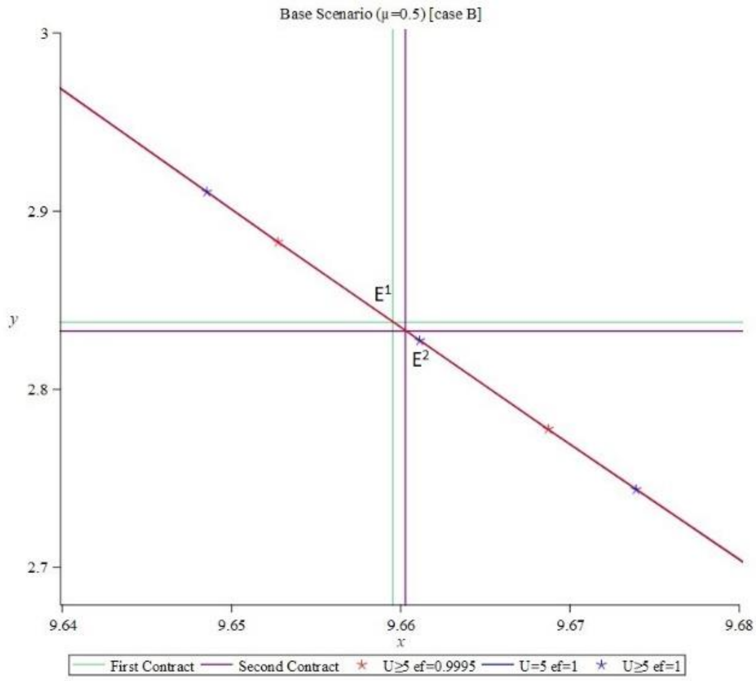


Figure 17 Base Scenario with close effort levels

With this strategy the agent has less reasons to implement strategic a behaviour, and even if he do the principal expected results are not so dramatically worst.

High complexity user scenario (s=0,2; μ =0,8 [i]*)*

In this scenario the only change regards the complexity of the user, then we analyse what happens when the user presents an high level of complexity then μ *=0,8* with respect to the same Strategies A and B.

Strategy A: the principal, in defining the contract uses ef^L= 0 and ef^H= 1.

SETP 1. The principal by solving his constrained minimization problem in which it encourages the high effort level which determines p^H=0,731, defines a contract with α= 10,167 and β = 6,187 and from which the expected expenditure is *S* = 9,0967 and the agent expected utility is 5.

STEP 2 Knowing the reimbursements α and β established by the contract, the agent takes action in order to maximize its expected utility. So, given α = 10,167 and β = 6,187, he determines the effort that would allow him to maximize his expected utility. But the resulting effort level would be an *ef* = 1,114 which is not possible because we assumed $ef \in [0,1]$, so we set a constraint such that the effort satisfies the initial assumptions $ef \leq 1$. Therefore *ef*=1 and then *p*=0,731 *EU*=5,0000000022.

[i] See Note 12 and section *Logistic Function*

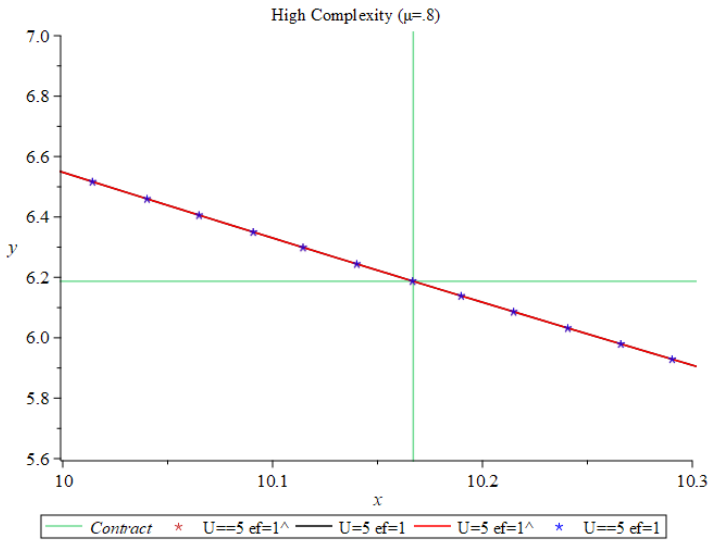


*Figure 18 High complexity User: a possible contract*Table 16 Dynamic of contract in high complexity scenario- Strategy A

ef^H =1 $ef^L = 0$	STEP 1 Blu Line	STEP 2 Red Asterix
α	10,167	10,167
β	6,187	6,1871
ef	1	*1*
p	0,7311	0,7311
*EU***	5,0000000012	*5,0000000022*
*S**	9,0967	9,096668

***** *Expected Cost,* ****** *Expected Utility*

Strategy B: As before, we consider the situation in which the principal in the participation constraint uses as low an effort value which is infinitesimally smaller than the desired one: $ef^H = 1$ and $ef^L = 0{,}999$.

STEP 1 In this situation the resulting contract is described by: p^H=0,731, p^L=0,730, α = 9,838 β = 6,909 and consequently *S* = 9,051 and *U*=5,0000000011.

STEP 2 The agent, maximizing its expected utility, implements a strategic behaviour applying an effort level equal to 0,9995 and therefore reducing the probability to 0,730. In this way, his expected utility is equal to 5,0000002880 which is very little greater than which one deriving from correct behaviour.

Table 17 Dynamic of contract in high complexity scenario with close effort levels –Strategy B

$ef^H = 1$ $ef^L = 0{,}999$	STEP 1 Blu Line	STEP 2 Red Asterix	STEP 3 Red Asterix	STEP 4 Blu Asterix
α	9,838	9,838	9,838	9,838
β	6,909	6,909	6,908	6,908
ef	1	0,99950128	0,99950128	1
p	0,7311	0,7306	0,7306	0,7311
*EU***	5,00000000118	5,00000028801	5,0000002906	5,0000002877
*S**	9,05074466	9,04930786	9,0493658	9,0508033

***** *Expected Cost,* ****** *Expected Utility*

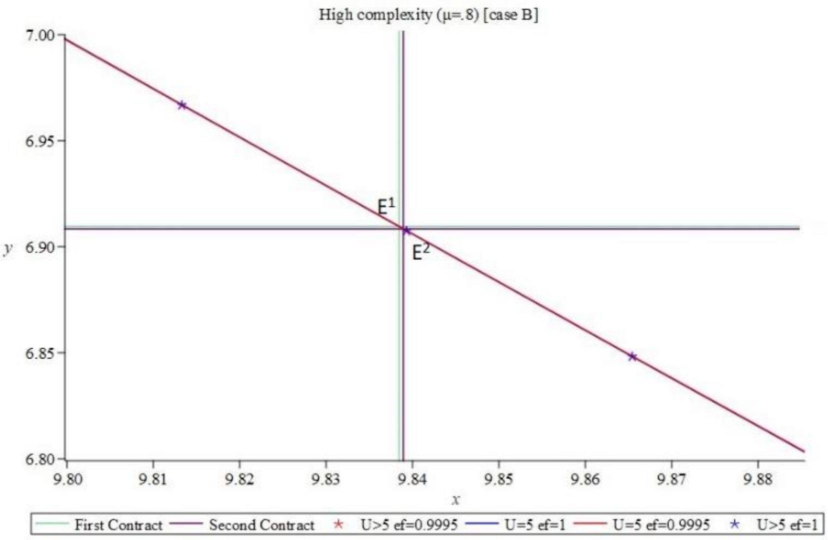


Figure 19 High complexity User with close effort levels: the compromise contract

As in the basic scenario, this strategy reduces the opportunity for misconduct by the agent. And similarly to the *strategy A*, the moral hazard problem does not appear much relevant in terms neither of probability nor of expected expenditure. In step 2 the expected utility provides very small incentives in implementing unfair behaviours and consequently, it could be achieve a stable contract, without great inefficiencies even without looking for the compromise described in steps 3 and 4 which lead to almost the same results (Table 8).

Results of model 1, summarized in Table 1, suggest that absolute and relative risk influence the weight that constraints (participation and incentive-compatibility) have on reimbursement values and principal expected costs and this appears in line with what might have been expected. In particular, the weight of the participation constraint seems to be mainly affected by variations in absolute risk of failure, because, being the agent risk averse must be supported the participation in the contract. The importance of incentive-compatibility constraint is related to the relative convenience of the effort values. If the probability deriving from the high effort (p^H) increases, it becomes relatively more convenient for the agent to apply the high effort then the low effort, therefore the incentive to ef^H has less weight. Conversely, if the probability of success even with low effort increases, to boost ef^H becomes more important.

In model 2 we introduced a variable representing the type of user, *g*, which for assumption affects negatively the probability of the service. Since our outcome variable is dichotomous (the transport in exploited or not), and that consequently its expected value coincides with the probability of the success, we modelled the probability of the outcome of the service with a logistic function which is widely applied for limited outcome variables. Given our focus on a highly specialized transport services, we fixed the parameter of scale *s*, as indicator of the specialization of the job, and let vary the position parameter μ, assumed as indicator of the type of the user. The parameter μ allows to identify the effort threshold under which the increase in *ef* is more effective and beyond which is less effective. Given these assumption and comparing three different users (in Table 2), we can see that growing μ both α and β and expected utility of the agent increase. But the variation in the expected cost differs with the strategy applied by the principal in defining the contract. In fact assuming the principal is able to know and compare not only the extremes but also relevant effort values, such as the inflection point of the probability distribution, he could define different contracts as shown in Table 2 and obtaining different results in terms of reimbursement values, expected cost and utility.

Results in Table 3 show what happens if the principal fixes a level of probability he wants guaranteed. The increment of the expected expenditure decreases with higher complexity of users, even if the difference between the values of the reimbursement increases.

Furthermore, in Table 4 the contribution of effort variation is assessed: infinitesimal effort variations (equal to 0,1 points) determine quite constant percentage variations in the expected costs. Respectively, reimbursement variations have symmetrical trends and tend to compensate each other.

The last set of simulations provides a representation of the dynamics of contract implementation with respect to two scenarios: the base scenario with a medium complexity user and the provision of the service for user with high complexity.

The base scenario allows us to understand the moral hazard and the opportunity for strategic behaviour that the agent has. When the user has medium complexity and the principal establishes the contract considering the minimum (0) and maximum (1) effort levels, what happens is that the agent is able to apply less effort to obtain a greater level of expected utility than that it would result from maintaining correct behaviour. In this case the expected expenditure will decrease but even the probability. Then the principal to maintain the probability level desired has to offer a compromise contract which requires higher expected costs (setting a higher reimbursement in the case of success and a lower otherwise, i.e. increasing the gap between the two payments as in Table 5).

If the principal is able to anticipate the opportunistic behaviour of the agent, considering in the definition of the contract a higher value of low effort (very close to the high desired one as shown in Table 6), of course he will expect higher expenditure but he reduce the space for opportunistic behaviours. Furthermore, although it is clear that there is a margin for the agent to implement strategic agent behaviour, in this situation the differences are so small that a fairly satisfactory balance could be found for both parties.

In the second scenario, referring to high complexity users, the problem of moral hazard is less evident. The simulation with the software suggested that, the only way for the agent to strategically exploit the contract would be to implement an effort level higher than 1 which is not possible. In Table 7 we see how in this case the problem of moral hazard is not relevant even if the agent manages in some way to obtain an expected utility even if slightly higher than that foreseen by the contract.

Also in this case, as for the user of medium complexity, the differences in the final result between a correct behaviour and an opportunist one, are so small that a fairly satisfactory balance could be found for both parties.

In both scenarios (scenario base and highly complex scenario) a contract that anticipates the agent's misbehaviour is more expensive than the one with milder criteria. Furthermore, it should be noted that the more expensive contract in the baseline scenario is more expensive than both contracts calculated for high complexity patients (Table 8).

4. Conclusions

What emerges from our work is that that the public service aimed at people with disabilities that we have presented, addresses two types of problems that are costly both in monetary terms and in respect of the resulting outcome. The uncertainty of the result is attributable to the specific characteristics of the user and some inefficiency related to the incomplete information about the agent behaviour, which is able to worsen the result. The principal, in order to reduce the expected expenditure and guarantee certain levels of probability, has to reduce the possibility for opportunistic behaviours, incentivising the high effort levels but reaching both the objectives (reduced expected costs and high level probability) is not ensured and in some cases it could be very difficult to find a contract able to get over the moral hazard.

To tackle the problem we have posed, first we set up a theoretical model that takes into account the information asymmetry with assumptions that simplify the analysis, subsequently these assumptions were "relaxed" so as to address more problems and criticalities related. Finally we proposed some simulations, which on the basis of the theoretical model and respecting the required conditions, try to provide a comprehension, even if only partial, of a complex reality.

The inclusion of the variable *g* that describes the type of patient makes the description of the problem more realistic and allows to define the contract in a more coherent way with the actual task that the agent has to perform and to justify the expense incurred by the principal. Obviously, it does not allow to reduce the uncertainty of the result, since the final result depend not only on agent’s action, but also on a random variable, (that is the type of users), it is also a random variable, but taking into account of the random causal variable allows to highlight from which component the uncertainty and negative results derive, and consequently to evaluate whether they can be reduced or not

The choice of probability formulation as a function of effort and complexity of the patient in the context of disabled transport, intended as a highly specialized service, seemed useful and usable, even is evident that is one of the possible way to represent the problem.

First, we have, effectively highlighted the possibility of moral hazard behaviours thus confirming the real problem inherent this transport service: possible incorrect behaviours lead to compromise solutions that require a higher expected expenditure.

Secondly, the fact that the agent implements strategic behaviours in situations of less serious patients in which the probability that the service is performed is less affected by the effort. In situations where patients have high levels of complexity, the maximum probability obtainable with the maximum level of effort is already reduced and the agent's effort is less able of influencing it. Therefore, it appears that the agent, when users are very complex, is less likely to engage in incorrect behaviour, which would further reduce the probability of success.

Our analysis confirmed the trade-off between the principal goals, to ensure that the service is performed with high probability and to contain expenditure, basically because, high probability requires high effort and then an increase in the expenditure. And even that the increment in expected costs is attributable both to complexity of the user and in opportunistic misbehaviour of the agent.

The user complexity, on one side reduces the probability of a positive outcome, even in case the agent puts high effort in his job. This can lead to a variation (reduction) in the relative convenience of the efforts and consequently in the need of more incentive to high effort. On the other side, user complexity is related with the growth of exogenous uncertainty and then determine higher absolute risk of failure in performing the service, then the contract could compensate the risk averse agent.

The unfair actions of the agent could create higher expected cost because a compromise contract which try to avoid moral hazard, increases the reimbursement values and leads to a worse result for the principal in term of expenditure.

Of course the complexity of the user in unchangeable so primarily the policy maker should focus on reducing the moral hazard opportunity.

We show that a possible strategy, to reduce the margin of opportunity for incorrect behaviours, could be the definition of contracts that consider very close efforts in the incentive-compatibility constraint.

However, since we have seen that the base scenario lead to a compromise contract which is more expensive (in expected terms), than the high complex user scenario which is worse from the probability point of view. In first scenario, seems prevailing the monetary lost due to the moral hazard. In the case with high complexity user what is more concerning is the lower probability, because the expected cost are lower than in the base scenario.

Given these observations, a possible way to reduce the inefficiency deriving from user complexity is take into account that different complexity levels affect in various way the outcome. And a solution to face the exogeneity of the user complexity and to limit the problems resulting from the incomplete information, could be to define different contracts for different types of users, so as to isolate the respective relative criticisms and therefore reach more appropriate solutions.

The first evident limitation of this discussion concerns is not having considered a principal's benefit function. In fact, as principal objective we used the minimization of expected cost. The main reason, which partially justify this lack, derives from the fact that the work has a more empirical than theoretical value and describes a phenomenon of reduced dimensions. This is because we describe a service that, in each municipal reality has its own consistency in terms of expenditure, but from which few people benefit, therefore inserting a welfare function seemed redundant.

Another important limitation that derives from the lack of administrative data is the absence of an empirical analysis. But as mentioned in the previous point, the small number of users in each municipality has not, so far, motivated the possibility of forming consistent databases.

We also note that contract theory inherent the provision of healthcare service, from which we started, often uses the quality as unobservable outcome variable, which we ignored. This lack is motivated from the fact that the service considered here is less articulated than health care, is limited in the time and it is more similar to the taxi service, for which we assume the transport realized by car from one point to another is most important aspect.

These limitations suggest the further work would first focus on the collection and analysis of data on disabled people transportation service derived from municipalities. Secondly, it seems relevant comparing the various solutions implemented in different municipalities with respect the definition the user of the service, the contract offered. Finally, based on the features derived from the evidences collected from reality, it could be interesting trying to define more appropriate and precise contractual solutions.

References

Altman, B. (2001). Disability definitions, models, classification schemes, and applications. In S. K. Albrecht G.L., *Handbook of Disability Studies* (pp. 97-122). SAGE.

Annadurai, G. B. (2000). Development of mathematical models (logistic, Gompertz and Richards models) describing the growth pattern of Pseudomonas putida. *Bioprocess Eng., 23*, 607-612.

Arnott, R. S. (1988). The Basic Analytics of Moral Hazard. *NBER WORKING PAPER 2484*. doi:10.3386/w2484

Arnott, R. S. (1991). The Welfare Economics of Moral Hazard. In L. H., *Risk, Information and Insurance*. Springer. doi: https://doi.org/10.1007/978-94-009-2183-2_5

Arrow, K. (1963). Uncertainty and the Welfare Economics of Medical Care. *The American Economic Review, 53*(5), 941-973.

Arrow, K. (1968, Jun). The Economics of Moral Hazard: Further Comment. *The American Economic Review, 58*(3), 537-539.

Arrow, K. (1984). The economics of Agency. In J. Pratt, & R. Zeckhauser, *Principals and Agents: The Structure of Business* (pp. 37-51). Boston: Harvard Business School Press.

Berkson, J. (1944). Application of the Logistic Function to Bio-Assay. *Journal of the American Statistical Association, 39*(227), 357-365.

Bolhaar, J. L. (2012). A Dynamic Analysis of the Demand for Health Insurance and Health Care. *European Economic Review, 56*(4), 669-690.

Brouwer, W. v. (2019). When is it too expensive? Cost-effectiveness thresholds and health care decision-making. *The European Journal of Health Economics*, 175–180.

Brown, T. P. (2016, April). Managing Complex Contracts: A Theoretical Approach. *Journal of Public Administration Research and Theory, 26*(2), 294–308.

Carlson, T. (1913). Über geschwindigkeit und grösse der hefevermehrung in würze. *Biochem. Z, 57*, 313.

Chalkley, M. M. (1996). Contracts For The National Health Service. *The Economic Journal, 106*, 1691–1701.

Cramer, J. (2002). The Origins of Logistic Regression. *Tinbergen Institute Discussion Paper, No. 02-119/4*.

Cutler, D. Z. (2000). The Anatomy of Health Insurance. In C. A. JP, *The Handbook of Health Economics.*

Di Novi, C. (2008). Adverse selection in the U.S. health insurance markets: Evidence from the MEPS. *Department of Public Policy and Public Choice – POLIS*(Working Paper n.113).

Fisher, T. F. (1971). Tech. Forecasting Soc. *Changes 3:75*.

Fujiura, G. R.-K. (2001). Counting disability. In S. K. Albrecht G.L., *Handbook of disability studies.* SAGE.

Gant, R. (1992). Transport for the Disabled. *Geography, 77*(1), 88-91.

Hart, O. H. (1987). The theory of contract. *The advances in economic theory: 5th world congress.* New York: Cambridge University Press.

Herman, R. E. (1972). A Manner of Characterizing the Development of Countries. *Proc. Nat. Acad. Sci., 69*(10), 3019-3023.

Holmstrom, B. (1979). Moral Hazard and Observability. *Bell Journal of Economics,, 10*, 74-91.

Koopman, B. (1953, Feb-). The Optimum Distribution of Effort. *Journal of the Operations Research Society of America, 1*(2), 52-63.

Krebs, C. J. (1985). Species diversity. In K. C. J., *Ecology: the experimental analysis of distribution and abundance3th ed* (p. 507 534). New York: Harper and Row New York.

Kuhn, H. W. (1951). Nonlinear programming . *Proceedings of 2nd Berkeley Symposium* (pp. 481-492). Berkeley: University of California Press.

Laffont, J. M. (2009). *The Theory of Incentives: The Principal-Agent Model.* Princeton: Princeton University Press.

Ma, C. (1994). Health care payment systems: Cost and quality incentives. *J. Econ. Manage. Strategy, 3*(1), 93-112.

Marchetti, C. N. (1980). *The Dynamics of Energy Systems and the Logistic Substitution Model.* United States: Final report.

Mongin, P. (1997). Expected Utility Theory. In J. H. Davis. London: Edward Elgar.

Morgan, B. (1976). Stochastic models of groupings changes. *Adv. Appl. Probability, 8*(1), 30-57.

Pauly, M. (1968). The Economics of Moral Hazard. *Quarterly Journal of Economics, 88,* 44-62.

Pearl, R. (1930). *Introduction of Medical Biometry and Statistics.* Philadelphia: Saunders.

Pearl, R. R. (1920). On the rate of growth of the population of the US since 1790 and its mathematical representation . *Proc. Nat. Acad.Sci., 6*, 275-288.

Peng, C. J. (2002). An Introduction to Logistic Regression Analysis and Reporting. *The Journal of Educational Research, 96*(1), 3-14.

Rabin, M. (2013). Risk Aversion and Expected-Utility Theory: A Calibration Theorem. In Z. W. MacLeanL.C:, *Handbook of the Fundamentals of Financial Decision Making* (pp. 241-252).

Rogerson, W. P. (1985). The First-Order Approach to Principal-Agent Problems. *Econometrica, 53*(6), 1357-1367.

Rutkowski-Kmitta, F. G. (2001). Counting disability. In S. K. Albrecht G.L., *Handbook of disability studies.,* (pp. 69-96). SAGE.

Schmitz, P. (2003). The Hold –Up Problem and incomplete contracs: a survey of recent topics in contract theory. *Bulletin of economic research, 53*(1).

Slater, S. B. (1974). The Definition and Measurement of Disability. *Social Science and Medicine, 8,* 305-308.

Spence, M. (1973, Aug.). Job Market Signaling. *The Quarterly Journal of Economics,, 87*(3), 355-374.

Stiglitz, J. (1983). Risk, Incentives and Insurance: The Pure Theory of Moral Hazard. *The Geneva Papers on Risk and Insurance, 8*(26), 4-33.

Tirole, J. (1999). Incomplete Contracts: Where do we stand? *Econometrica, 67*(4), 741-781.

Tsoularis, A. W. (2002). Analysis of logistic growth models. *Mathematical Biosciences, 179,* 21–55.

Verhulst, P. (1838). Notice sur la loi que la population suit dans son accroissement. *Curr. Math. Phys., 10*(1838), 113.

Wacheneheim, D. P. (2003). Analysis of the logistic function model: derivation and applications specific to batch cultured microorganisms. *Bioresource technology, 86*(2), 157-164.

Wallace, B. (2004). Constrained Optimization: Kuhn-Tucker Conditions. *Royal Holloway.Egham*.

WHO. (2001). *The International Classification of Functioning, Disability and Health (ICF).* Geneva, Switzerland: World Health Organization.

Winsor, C. (1932). The Gompertz curve as a growth curve. *Proc. Nat. Acad. Sci*, *18*, 1-8.

Appendix A

The optimization problem**.** The Lagrangian of the problem:

$$L = -\beta - p(ef^H)(\beta - \alpha) + \lambda\big(u(\beta) + p(ef^H)\big(u(\alpha) - u(\beta)\big) - v(ef^H) - U_0\big) + \mu((p(ef^H) - p(e^L))\big(u(\alpha) - u(\beta)\big) - v(ef^H) + v(ef^L))$$

The Kuhn-Tucker conditions (1951) are represented in three conditions for each variable and for each constraint:

$$\frac{dL}{d\alpha} \leq 0\ ;\ \alpha \geq 0;\ \alpha\frac{dL}{d\alpha} = 0$$

$$\frac{dL}{d\beta} \leq 0\ ;\ \beta \geq 0;\ \beta\frac{dL}{d\beta} = 0$$

$$u(\beta) + p(ef^H)\big(u(\alpha) - u(\beta)\big) - v(ef^H) - U_0 \geq 0;\ \lambda \geq 0;\ \lambda\big(u(\beta) + p(ef^H)\big(u(\alpha) - u(\beta)\big) - v(ef^H) - U_0\big) = 0$$

$$(p(ef^H) - p(ef^L))\big(u(\alpha) - u(\beta)\big) - v(ef^H) + v(e^L) \geq 0;\ \mu \geq 0; \mu\ \Big((p(ef^H) - p(ef^L))\big(u(\alpha) - u(\beta)\big) - v(ef^H) + v(ef^L)\Big) = 0$$

The Kuhn-Tucker conditions are necessary for a point to be a solution of the problem: each point that satisfies the K-T conditions, i.e. solves the equation systems, is a candidate to be a maximum.

$$\frac{\frac{\partial u(\alpha)}{\partial \alpha}}{\frac{\partial u(\beta)}{\partial \beta}} = \frac{\lambda - \mu\frac{p(ef^H) - p(ef^L)}{1 - p(ef^H)}}{\lambda + \mu\frac{p(ef^H) - p(ef^L)}{p(ef^H)}}$$

$$u(\alpha) = \frac{1 - p(ef^L)}{p(ef^H) - p(ef^L)}v(e^H) - \frac{1 - p(ef^H)}{p(ef^H) - p(ef^L)}v(ef^L) + U_0$$

$$u(\beta) = \frac{p(ef^H)}{p(ef^H) - p(ef^L)}v(ef^L) - \frac{p(ef^L)}{p(ef^H) - p(ef^L)}v(ef^H) + U_0$$

From which we derive the following conditions $p(ef^H) \neq p(ef^L)$ e $v(ef^H) \geq v(ef^L)$

Appendix B

Code MAPLE Derivatives

$ExpCOST := ph \cdot a + (1 - ph) \cdot b;$

$$Obj := ph\left(\frac{(1-r)\ (1-pl)\ vh}{ph-pl} - \frac{(1-ph)\ vl}{ph-pl} + Uris\right)^{\frac{1}{1-r}} + (1 - ph)\left(\frac{(1-r)\ ph\ vl}{ph-pl} - \frac{pl\ vh}{ph-pl} + Uris\right)^{\frac{1}{1-r}}$$

$a := \left((1-r)\cdot\left(\frac{1-pl}{ph-pl}\right)\cdot vh - \left(\frac{1-ph}{ph-pl}\right)\cdot vl + Uris\right)^{\frac{1}{1-r}}$

$$a := \left(\frac{(1-r)\ (1-pl)\ vh}{ph-pl} - \frac{(1-ph)\ vl}{ph-pl} + Uris\right)^{\frac{1}{1-r}}$$

$b := \left((1-r)\cdot\left(\frac{ph}{ph-pl}\right)\cdot vl - \left(\frac{pl}{ph-pl}\right)\cdot vh + Uris\right)^{\frac{1}{1-r}}$

$$b := \left(\frac{(1-r)\ ph\ vl}{ph-pl} - \frac{pl\ vh}{ph-pl} + Uris\right)^{\frac{1}{1-r}}$$

$vl := 0; vh := 1;$

$diff(a, ph)$

$$-\frac{\left(\frac{(1-r)\ (1-pl)}{ph-pl} + Uris\right)^{\frac{1}{1-r}}\ (1-pl)}{(ph-pl)^2\left(\frac{(1-r)\ (1-pl)}{ph-pl} + Uris\right)}$$

$diff(a, pl)$

$$\left(\left(\frac{(1-r)\ (1-pl)}{ph-pl} + Uris\right)^{\frac{1}{1-r}}\left(-\frac{1-r}{ph-pl} + \frac{(1-r)\ (1-pl)}{(ph-pl)^2}\right)\right) \Big/ \left((1-r)\left(\frac{(1-r)\ (1-pl)}{ph-pl} + Uris\right)\right)$$

$diff(a, Uris)$

$$\frac{\left(\frac{(1-r)\ (1-pl)}{ph-pl} + Uris\right)^{\frac{1}{1-r}}}{(1-r)\left(\frac{(1-r)\ (1-pl)}{ph-pl} + Uris\right)}$$

$diff(ExpCOST, ph)$

$$\left(\frac{(1-r)\,(1-pl)}{ph-pl}+Uris\right)^{\frac{1}{1-r}}$$

$$-\frac{ph\left(\frac{(1-r)\,(1-pl)}{ph-pl}+Uris\right)^{\frac{1}{1-r}}(1-pl)}{(ph-pl)^2\left(\frac{(1-r)\,(1-pl)}{ph-pl}+Uris\right)}-\left(\right.$$

$$\left.-\frac{pl}{ph-pl}+Uris\right)^{\frac{1}{1-r}}$$

$$+\frac{(1-ph)\left(-\frac{pl}{ph-pl}+Uris\right)^{\frac{1}{1-r}}pl}{(1-r)\,(ph-pl)^2\left(-\frac{pl}{ph-pl}+Uris\right)}$$

$diff(ExpCOST, pl)$

$$\left(ph\left(\frac{(1-r)\,(1-pl)}{ph-pl}+Uris\right)^{\frac{1}{1-r}}\left(-\frac{1-r}{ph-pl}\right.\right.$$

$$\left.\left.+\frac{(1-r)\,(1-pl)}{(ph-pl)^2}\right)\right)\Bigg/\left((1-r)\left(\frac{(1-r)\,(1-pl)}{ph-pl}\right.\right.$$

$$\left.\left.+Uris\right)\right)$$

$$+\frac{1}{(1-r)\left(-\frac{pl}{ph-pl}+Uris\right)}\left((1-ph)\left(\right.\right.$$

$$\left.\left.-\frac{pl}{ph-pl}+Uris\right)^{\frac{1}{1-r}}\left(-\frac{1}{ph-pl}-\frac{pl}{(ph-pl)^2}\right)\right)$$

$diff(ExpCOST, Uris)$

$$\frac{ph\left(\frac{(1-r)\,(1-pl)}{ph-pl}+Uris\right)^{\frac{1}{1-r}}}{(1-r)\left(\frac{(1-r)\,(1-pl)}{ph-pl}+Uris\right)}$$

$$+\frac{(1-ph)\left(-\frac{pl}{ph-pl}+Uris\right)^{\frac{1}{1-r}}}{(1-r)\left(-\frac{pl}{ph-pl}+Uris\right)}$$

$diff(ExpCOST, r)$

$$ph\left(\frac{(1-r)\,(1-pl)}{ph-pl}+Uris\right)^{\frac{1}{1-r}}\left(\frac{\ln\left(\frac{(1-r)\,(1-pl)}{ph-pl}+Uris\right)}{(1-r)^{2}}-\frac{1-pl}{(1-r)\,(ph-pl)\left(\frac{(1-r)\,(1-pl)}{ph-pl}+Uris\right)}\right)+\frac{(1-ph)\left(-\frac{pl}{ph-pl}+Uris\right)^{\frac{1}{1-r}}\ln\left(-\frac{pl}{ph-pl}+Uris\right)}{(1-r)^{2}}$$

Appendice C

Code MAPLE Table 1

$r := 0.5; k := 1; ph := 0.8; pl := 0.2; el := 0; eh := 1; Uris := 10;$

$$Lag := -ph\cdot(a)-(1-ph)\cdot b-\lambda\cdot\left(\frac{ph\cdot\left(a^{1-\mathrm{r}}-b^{1-r}\right)}{1-r}+\frac{b^{1-r}}{1-r}-eh^{k}-Uris\right)-\mu\cdot\left(ph\cdot\left(\frac{a^{1-r}}{1-r}\right)+(1-ph)\cdot\left(\frac{b^{1-r}}{1-r}\right)-eh^{k}-pl\cdot\left(\frac{a^{1-r}}{1-r}\right)-(1-pl)\cdot\left(\frac{b^{1-r}}{1-r}\right)+el^{k}\right);$$

$FOC1 := diff(Lag, a);$
$FOC2 := diff(Lag, b);$
$FOC3 := diff(Lag, \lambda);$
$FOC4 := diff(Lag, \mu);$

Sol := solve({FOC1 = 0, FOC2 = 0, FOC3 = 0, FOC4 = 0})=

={a = 32.11111111, b = 23.36111111, lambda = -5.500000000, mu = -0.2222222222}

ExpCOST := ph (a) + (1 - ph) (b)= 30.36111111

Appendix D

Code MAPLE Table 2

restart

$$p(ef) := \frac{1}{1 + \mathrm{e}^{\frac{(m - ef)}{0.2}}};$$

$$L := p(ef1)\cdot(a) + (1 - p(ef1))\cdot b + \lambda\cdot\left(p(ef1)\cdot\left(\frac{a^{1-\mathrm{r}}}{1-r}\right) + (1 - p(ef1))\cdot\left(\frac{b^{1-r}}{1-r}\right) - (ef1)^{k} - Uris\right) + \mu\cdot\left(p(ef1)\cdot\left(\frac{a^{1-r}}{1-r}\right) + (1 - p(ef1))\cdot\left(\frac{b^{1-r}}{1-r}\right) - ef1^{k} - p(ef2)\cdot\left(\frac{a^{1-r}}{1-r}\right) - (1 - p(ef2))\cdot\left(\frac{b^{1-r}}{1-r}\right) + (ef2)^{k}\right);$$

$m := 0.3; r := \frac{1}{2}; k := 1; ef1 := 0.3; ef2 := 0; Uris := 10; p(ef1); p(ef2);$

$FOC1 := diff(L, a); FOC2 := diff(L, b); FOC3 := diff(L, \lambda); FOC4 := diff(L, \mu);$

$sol := solve(\{FOC1 = 0, FOC2 = 0, FOC3 = 0, FOC4 = 0\});$

ExpCOST:= p(ef1)(29.0107742) + (1 - p(ef1))(24.14577369)= 26.57827394

Appendix E

Definition of the effort desired by the principal given the probability desired

Table A1

p=0,5	$p(ef) > 0,5$	$ef > \mu ef$
p=0,7	$p(ef) > 0,7$	$ef > \mu + 0,1695$
p=0,8	$p(ef) > 0,8$	$ef > \mu + 0.2772$
p=0,9	$p(ef) > 0,9$	$ef > \mu + 0.4394$
p=0,999	$p(ef) > 0,99$	$ef > \mu + 1.381$

Appendix F

Definition of the effort desired by the principal given the probability desired (Table 3)

Code MAPLE Table 3

$$p(ef) := \frac{1}{1+\mathrm{e}^{\frac{(m-ef)}{0.2}}};$$

#imposing p(ef1)=0.8 we conside ef1=m+0.277

$r := 0.5; k := 1; m := 0.8; ef1 := m + 0.277; ef2 := 0; Uris := 5;$
$p(ef2);$

$$Lag := (0.8)\cdot(a) + (0.2)\cdot b + \lambda\cdot\left((0.8)\cdot\left(\frac{a^{1-r}}{1-r}\right) + (0.2)\cdot\left(\frac{b^{1-r}}{1-r}\right) - ef1^{k} - Uris\right) + \mu\cdot\left((0.8)\cdot\left(\frac{a^{1-r}}{1-r}\right) + (0.2)\cdot\left(\frac{b^{1-r}}{1-r}\right) - ef1^{k} - p(ef2)\cdot\left(\frac{a^{1-r}}{1-r}\right) - (1-p(ef2))\cdot\left(\frac{b^{1-r}}{1-r}\right) + ef2^{k}\right);$$

$FOC1 := diff(Lag, a); FOC2 := diff(Lag, b); FOC3 := diff(Lag, \lambda); FOC4 := diff(Lag, \mu);$

$sol := solve(\{FOC1 = 0, FOC2 = 0, FOC3 = 0, FOC4 = 0\});$

sol:={a=10.08838311, b=6,188226268, λ=-3.038500000,μ=-0.140889289}

$S := (10.08838211)\cdot 0.8 + (6.188226268)\cdot 0.2$ = 9.308350942

ì

Appendix G

How the effort affect probabilities of the transport and Expected expediture of the principal

Code MAPLE Table 4

#s=0.2; mi=0.5;
#efL=efH − 0.1
#r=0.5; k=1; Uris=5

for *ef1* **from** 0.1 **by** 0.1 **to** 1 **do** $p(ef1) := \frac{1}{1+e^{\frac{(0.5-ef1)}{0.2}}}$; $s :=$ 0.2; $m := 0.5$; $Lag := \frac{1}{1+e^{\frac{(m-ef1)}{s}}}\cdot(a) + \left(1 - \frac{1}{1+e^{\frac{(m-ef1)}{s}}}\right)\cdot b + \lambda\cdot\left(\frac{1}{1+e^{\frac{(m-ef1)}{s}}}\cdot\left(\frac{a^{1-0.5}}{1-0.5} - \frac{b^{1-0.5}}{1-0.5}\right) + \frac{b^{1-0.5}}{1-0.5} - ef1 - 5\right) + \mu\cdot\left(\frac{1}{1+e^{\frac{(m-ef1)}{s}}}\cdot\left(\frac{a^{1-0.5}}{1-0.5}\right) + \left(1 - \frac{1}{1+e^{\frac{(m-ef1)}{s}}}\right)\cdot\left(\frac{b^{1-0.5}}{1-0.5}\right) - ef1 - p(ef1-0.1)\cdot\left(\frac{a^{1-0.5}}{1-0.5}\right) - (1-p(ef1-0.1))\cdot\left(\frac{b^{1-0.5}}{1-0.5}\right) + (ef1-0.1)\right)$; $FOC1 := diff(Lag, a)$; $FOC2 := diff(Lag, b)$; $FOC3 := diff(Lag, \lambda)$; $FOC4 := diff(Lag, \mu)$; $solution :=$ $solve(\{FOC1 = 0, FOC2 = 0, FOC3 = 0, FOC4 = 0\})$; $T(ef1) :=$ $table\ (p(ef1), solution)$; **end do**

for *ef1* **from** 0.1 **by** 0.1 **to** 1 **do** *print T(ef1)* **end do**

Appendix H

Analisys of dynamics of the contracting

restart;

$ph := \dfrac{1}{1+\mathrm{e}^{\frac{-(eh-\mu)}{s}}}$; $pl := \dfrac{1}{1+\mathrm{e}^{\frac{-(el-\mu)}{s}}}$;

$Obj := ph\cdot(a) + (1-ph)\cdot b;$

#*evalf*(*ph*);

$IC := ph\cdot\left(\dfrac{(a^{1-r})}{(1-r)}\right) + (1-ph)\cdot\left(\dfrac{(b^{1-r})}{(1-r)}\right) - eh^k \geq pl\cdot\left(\dfrac{(a^{1-r})}{(1-r)}\right) + (1-pl)\cdot\left(\dfrac{(b^{1-r})}{(1-r)}\right) - el^k$

$IP := \dfrac{ph\, a^{1-r}}{1-r} + (1-ph)\cdot\left(\dfrac{b^{1-r}}{1-r}\right) - eh^k \geq Uris;$

$EF := eh \geq el;$

$\mu := 0.5; s := 0.2;\ r := 0.5; k := 1; Uris := 5; eh := 1; el := 0; ph; pl;$

Obj; *ph*;

$Vinc1 := (IP,\ IC, EF)$

with(*Optimization*) :

$A := [Minimize(Obj,\ \{\mathrm{Vinc1}\},\ assume = nonnegative)]$

$A := [[9.02435560584122243,\ [a = 9.27027570426536,\ b = 6.02843550388639]]]$

assign(*A*[1, 2])

a; *b*;

Obj; *IC*; *IP*;

$$9.02435560584122$$

$$4.99999999987821 \leq 5.00000000121789$$

$$5 \leq 5.00000000121789$$

###

RED ASTERIX CURVE Agent utility maximization problem######

restart;

$UTAG := p\cdot\left(\dfrac{(a^{1-r})}{(1-r)}\right) + (1-p)\cdot\left(\dfrac{(b^{1-r})}{(1-r)}\right) - em^k;$

$p := \dfrac{1}{1+\mathrm{e}^{\frac{-(em-\mu)}{s}}}$; $\mu := 0.5; s := 0.2;\ r := 0.5; k := 1;$

$a := 9.27027570426536; b := 6.02843550388639;$

with(*Optimization*) :

$A := [Maximize(\mathrm{UTAG},\ assume = nonnegative)]$

$$A := [[5.07696815667237722,\ [em = 0.757175844312259]]]$$

assign(*A*[1, 2])

em; *UTAG*; *p*;

$$0.757175844312259$$

$$5.07696815667238$$

$$0.783448888016142$$

$UTAGc := p \cdot \left(\frac{(x^{1-r})}{(1-r)} \right) + (1-p) \cdot \left(\frac{(y^{1-r})}{(1-r)} \right) - em^k = UTAG;$

$UTcost := solve(\{UTAGc\}, \{y\})$; *(*REDASTERIX*)*

$$UTcost := \{y = 13.08882016\, x - 97.46910462 \sqrt{x} + 181.4568891\}$$

###
######################################

BLU` `LINE CURVE Ipotetic Indifference Curve of Agent with correct behaviour U=5 and ef=1

$UTteo := 5.00000000121789;$

$eteo := 1; \mu := 0.5; s := 0.2;\ r := 0.5; k := 1;$

$p := \frac{1}{1 + e^{\frac{-(eteo - \mu)}{s}}};$

$UTAGteo := p \cdot \left(\frac{(x^{1-r})}{(1-r)} \right) + (1-p) \cdot \left(\frac{(y^{1-r})}{(1-r)} \right) - eteo^k$
= UTteo;

$UTAGteo := 1.848283639 \sqrt{x} + 0.1517163606 \sqrt{y} - 1$
$= 5.00000000121789$

$UTcostteo := solve(\{UTAGteo\}, \{y\})$; *(*BLU LINE*)*

$UTcostteo := \{y = 148.4131579\, x - 963.5739110 \sqrt{x}$
$+ 1564.003312\}$

#####################################

RED LINE CURVE Ipotetic Indifference Curve of Agent with scorrect behaviour ef=.75 and lower utility U=5

#####################################

restart;

$UTtest := 5.00000000121789;$

$etest := 0.757175844312259260; \mu := 0.5; s := 0.2;\ r := 0.5; k := 1;$

$p := \frac{1}{1 + e^{\frac{-(etest - \mu)}{s}}};$

$UTAGtest := p \cdot \left(\frac{(x^{1-r})}{(1-r)} \right) + (1-p) \cdot \left(\frac{(y^{1-r})}{(1-r)} \right) - etest^k$
= UTtest;

$UTAGtest := 1.566897776 \sqrt{x} + 0.4331022236 \sqrt{y} - 0.7571758443$
$= 5.00000000121789$

$UTcosttest := solve(\{UTAGtest\}, \{y\})$; *(*RED LINE*)*

$UTcosttest := \{y = 13.08882018\, x - 96.18322339 \sqrt{x} + 176.7006562\}$

####################################

BLU ASTERIX CURVE Ipotetic Indifference Curve of Agent with correct behaviour ef=.75 and higher Utility U=5.07

####################################

restart;

$UTtest2 := 5.076968158;$

$etest2 := 1; \mu := 0.5; s := 0.2; r := 0.5; k := 1;$

$$p := \frac{1}{1 + e^{\frac{-(etest2 - \mu)}{s}}};$$

$$UTAGtest2 := p \cdot \left(\frac{(x^{1-r})}{(1-r)} \right) + (1-p) \cdot \left(\frac{(y^{1-r})}{(1-r)} \right) - etest2^{k} = \text{UTtest2};$$

$$UTAGtest2 := 1.848283639 \sqrt{x} + 0.1517163606 \sqrt{y} - 1 = 5.076968158$$

UTcosttest2 := *solve*({*UTAGtest2*}, {*y*}); ***(*BLU ASTERIX*)***

$UTcosttest2 := \{y = 148.4131579\, x - 975.9346623 \sqrt{x} + 1604.386833\}$

plota := *plot*([[0, 6.02843550388639127], [10, 6.02843550388639127]], *x* = 7 ..10, *y* = 0 ..12, *color* = *aquamarine*, *legend* = ["First step Contract"]) :

plotb := *plot*([[9.27027570426535696, 0], [9.27027570426535696, 12]], *x* = 7 ..10, *y* = 0 ..12, *color* = *aquamarine*) :

plotX := *plot*([[0, 5.0220296], [10, 5.0220296]], *x* = 7 ..10, *y* = 0 ..12, *color* = *purple*, *legend* = ["Second step Contract"]) :

plotY := *plot*([[9.634483916, 0], [9.634483916, 12]], *x* = 7 ..10, *y* = 0 ..12, *color* = *purple*) **:**

(*RED ASTERIX*)

plotm := plot([13.08882016 x − 97.46910462 √x + 181.4568891], x = 7 ..10, y = 0 ..12, color = red, style = point, symbol = asterisk, legend = ["U=5.07 ef=0.75"]) ***:***

(*BLU lINE*)

plotc := plot([148.4131579 x − 963.5739110 √x + 1564.003312], x = 7 ..10, y = 0 ..12, color = blue, legend = ["U=5 ef=1"]) ***:***

(*RED LINE *)

***plotd := plot*([13.08882018 *x* − 96.18322339 √*x* + 176.7006562], *x* = 7 ..10, *y* = 0 ..12, *color* = *red*, *legend* =** ["U=5 ef=0.75"]) **:**

(*BLU ASTERIX*)

$plotf := plot\left(\left[148.4131579\,x - 975.9346623\,\sqrt{x} + 1604.386833\right], x = 7..10, y = 0..12, color = \mathbf{blue}, style = \mathbf{point}, symbol = \mathbf{asterisk}, \mathbf{legend} = \left[\text{"U=5.07 ef=1"}\right]\right):$

$display([plota, plotb, plotX, plotY, plotm, plotc, plotd, plotf]);$