UNIVERSITÀ DEGLI STUDI ROMA TRE

CREI
Centro di **R**icerca
Interdipartimentale
di **E**conomia delle **I**stituzioni

CREI Working Paper no. 3/2008

FAIRNESS OF NATIONAL HEALTH SERVICE IN ITALY: A BIVARIATE CORRELATED RANDOM EFFECTS MODEL

by

Antonello Maruotti
University of Roma Tre (Italy)

available online at http://host.uniroma3.it/centri/crei/pubblicazioni.html
ISSN 1971-6907

Fairness of national health service in Italy: a bivariate correlated random effects model

Antonello Maruotti[*]
DIPES - Università di Roma Tre

Abstract

In this paper we consider a possible way of measuring equity in health as the absence of systematic disparities in health (or in the major social determinants of health) between groups with different levels of underlying social advantage/disadvantage. Starting from the fairness approach developed by the World Health Organization, we propose to extend the analysis of fairness in financing contribution through a generalized linear mixed models framework by introducing a bivariate correlated random effects model. We aim at analyzing the burden of health care payment on Italian households by modeling catastrophic payments and impoverishment due to health care expenditures. For this purpose, we describe a bivariate model for binary data, where association between the outcomes is modeled through outcome-specific latent effects which are assumed to be correlated; we show how model parameters can be estimated in a finite mixture context. By using such model specification, the fairness of the Italian national health service is investigated.

Keywords: fairness, health care, random effects models, binary data, non parametric maximum likelihood.

1 Introduction

This paper focuses on the problem of fairness in financing contribution to the Italian National Health Service (NHS). Our interest is in analyzing and understanding the distribution of health system contributions across households and the role of excessively or catastrophic households' health payments. The analysis of the consequences of household contributions can be divided in two broad approaches: the income approach (van Doorslaer et al., 1999) and the burden approach (Murray et al., 2003). The former addresses

[*]Dipartimento di Istituzioni Pubbliche, Economia e Società, Universitá di Roma Tre, e-mail: antonello.maruotti@uniroma3.it

the effects of payments in the space of income in terms of income redistributive effects (Reynolds and Smolenski, 1977) and more recently in term of changes in levels of poverty (Aronson and Lambert, 1994); on the other hand, in the burden space the impact of financial payment on households requires a distributional measure analogous to those adopted in the income space (see e.g. Kakwani, 1977) and the fairness in financing contribution (FFC) developed by the World Health Organization (WHO) may represent a useful tool for this purpose (see e.g. Xu et al. 2003), since it measures health system payments in terms of their impact on the households.

Our aim is to use household survey data from 20 Italian Regions to analyze the consequences of household health system contributions. This empirical assessment helps to illustrate how the FFC, the change in the percentage of households below the poverty line and the fraction of households facing catastrophic health payments all capture different dimensions of health financing arrangements. Through this analysis, it is argued that different approaches for analyzing distribution of health system payments can be seen as complementary for policy review and formulation. To identify factors associated with impoverishment and with catastrophic payments, we adopt a bivariate model for binary data, specified by using socio-demographic characteristics and including individual- and outcome-specific random effects to capture heterogeneity sources that may arise when data are correlated. Typically, random intercept models are implemented by adding a Gaussian random effect into the linear predictor of a generalized linear model (or GLM, see Nelder and Wedderburn, 1972; McCullagh and Nelder, 1989), giving rise to a generalized linear mixed model (or GLMM, see Breslow and Clayton, 1993). A general family of *closed* random intercept models for binary outcomes are described in Caffo et al. (2007) (the model is closed in the sense that the distributions associated with the marginal and conditional link functions and the random effect distribution are all of the same family). However, if the likelihood is intractable, the need for numerical integration techniques can be solved by means of an EM algorithm for nonparametric maximum likelihood estimation of a mixing distribution, without assumptions on random effects distribution (Aitkin, 1999).

The plan of the paper is as follows. In section 2 we define the common conceptual underpinnings of the burden approach, the households' financial contribution and the technical challenges for computing the FFC index. In section 3 we introduce a correlated random effects model and the estimation procedure is explained in detail. In section 4 we present an application to the Italian NHS, analyzing the burden of health expenditures on households; conclusions are given in section 5.

2 Measurement of the Impact of Health Payments in the Space of Burden

The consequences of household system contributions whether using the income or burden approach may be examined by analyzing the distribution of contributions to health system financing isolated from the distribution of the benefits of the health system. In the following, we will focus on the burden approach since we are interested in measuring and identifying the determinants of households' burden due to health system payments, we refer to van Doorslaer et al. (1999) for the investigation of appropriate methodologies to measure the impact of health system payments in the space of income.

A health system where individuals have at the moment of seeking treatment to pay out of pockets for a substantial part of health services costs clearly restricts access to only those who can afford payment, and likely excludes the poorest members of society. Some important health interventions would not be financed at all if people had to pay for them; hence, fairness of financial risk protection requires the highest possible degree of separation between contributions and utilization. This is true for high cost interventions relative to the household *capacity to pay* (CTP).

Fair financing in health systems means that each household faces risks due to costs of the health system that are distributed according to ability to pay rather than to the risk of illness: a fairly financed system ensures financial protection to everyone. Health systems where individuals or households are sometimes lead to poverty for purchasing needed care, or forced to do without these services because of the corresponding cost, are unfair. This situation may characterize both poor and high income countries where at least a portion of the population may be inadequately protected from financial risks. Paying for health care can also be unfair since it can lead families to large unexpected expenses, i.e., costs that could not be foreseen and therefore have to be paid out of pocket at the moment the service is used rather than being covered by some kind of prepayment (such as general taxation). Let us define a household's financial contribution (HFC) which represents the household's financial burden due to health expenditures. Following Murray et al. (2003), we define for the i-th household

$$HFC_i = \frac{HE_i}{CTP_i} \tag{1}$$

where HE_i includes all costs attributable to the household, even if it is not aware of paying, such as the share of sales or value-added taxes it pays on consumption, which governments devote to financing health services, and the contribution via insurance provided, and partly financed, by employers and CTP_i is a measure of the household's capacity to pay. A debate of what is CTP is still ongoing; however we define the CTP as the effective income (in Xu et al., 2003, household consumption expenditure is used as the proxy for effective income) minus subsistence expenditure requirements (actual food expenditure was used as a proxy for household subsistence expenditure).

The burden of health financing on a particular household is the share of its *capacity to pay* lost to cover health needs. For poor households, reasonably a large share of income goes for basic necessities, particularly food, whereas richer households have more margin for other expanses, including those for health care. Food spending represents an approximation to expenditure on basic needs; an implicit limit is that this approach does not consider e.g. the households' savings, which may be of little impact in developing countries but may lead to bias estimates in Western Countries such as Italy. The way health care is financed is perfectly fair if the ratio of total health contribution to CTP is identical for all households, independently of their income, their health status or their use of health systems. Obviously, the financing would be unfair if poor households spend a larger share than rich ones, either because they are less protected by prepayment systems and so have to pay relatively more out of pocket, or because the prepayment arrangements are regressive. A summary index of the HFC distribution should allow a comparison of fairness between countries/regions. The FFC seems to be the most appropriate summary measure (Xu et al. 2003); it is defined as

$$FFC = 1 - \sqrt{\frac{\sum_{i=1}^{n} w_i |HFC_i - HFC_0|^3}{\sum_{i=1}^{n} w_i}} \tag{2}$$

where HFC_i is defined according to (1), $HFC_0 = \frac{\sum_{i=1}^{n} HE_i}{\sum_{i=1}^{n} CTP_i}$ and w_i represents sample weights. Extending the logic swept under this approach, we analyze changes in the economic status due to catastrophic health payments (a out of pocket expenditure share grater than the 40% of the CTP) as well influences of health expenditures on the social status to evaluate the share of those households led to poverty by out of pocket expenditures.

3 The model

Let us start supposing we have recorded J binary outcomes Y_{ij}, $i = 1, \ldots, n$ and $j = 1, \ldots, J$, together with a set of p_j covariates $\mathbf{x}_{ij} = (\mathbf{x}_{ij1}, \ldots, \mathbf{x}_{ijp_j})$. To describe association among outcomes, it is reasonable to assume that they share some common unobservable features. Let us denote with u_{ij}, $i = 1, \ldots, n$, $j = 1, \ldots, J$, a set of individuals outcome-specific random effects, accounting for heterogeneity and dependence between outcomes. We assume that, conditional on the covariates and the random effects, the observed outcomes Y_{ij} are drawn from independent Bernoulli random variables:

$$Y_{ij} \mid \mathbf{x}_{ij}, u_{ij} \sim Bin(1, \pi_{ij}), \quad \text{logit}(\pi_{ij}) = \beta_{0j} + \sum_{l=1}^{p_j} x_{ijl}\beta_{jl} + u_{ij} \tag{3}$$

where $\beta_j = (\beta_{0j}, \beta_{1j}, \ldots, \beta_{p_j j})$ is an outcome-specific vector of regression parameters. The random effects $\mathbf{u}_i = (u_{i0}, \ldots, u_{iJ})$ are usually assumed to be drawn from a multivariate parametric distribution, say $\mathcal{G}(\cdot)$, and thus account for dependence between responses. Various alternatives have been proposed, see e.g. Rodriguez and Goldman (1995).

Given model assumptions the marginal likelihood function can be written as follows:

$$\begin{aligned} L(\cdot) &= \prod_{i=1}^{n} \left\{ \int_{\mathcal{U}} f(\mathbf{y}_i \,|\mathbf{x}_i, \mathbf{u}_i) d\mathcal{G}(\mathbf{u}_i) \right\} = \\ & \prod_{i=1}^{n} \left\{ \int_{\mathcal{U}} \prod_{j=1}^{J} f(y_{ij} \,|\mathbf{x}_{ij}, u_{ij}) d\mathcal{G}(\mathbf{u}_i) \right\} \end{aligned} \tag{4}$$

where $\mathcal{U}$ is the support for $\mathcal{G}(\cdot)$.

Under Gaussian assumptions on $\mathbf{u}_i$, the marginal likelihood can not be written in closed form; however, if random effects are assumed to follow a known distribution, the likelihood function can be computed directly using numerical methods. Where there are only a few random effects, numerical integration methods (e.g Gaussian Quadrature or Adaptive Gaussian Quadrature, if random effects are Gaussian random variables) could be applied; on the other hand, for a large number of random effects simulation methods (e.g. Monte Carlo Expectation Maximization, MCEM) seem to be more feasible.

Alfò and Trovato (2004) propose a semiparametric model with unspecified density $\mathcal{G}(\cdot)$ for random effects developing the model in a finite mixture context. The choice of a flexible specification is preferred to parametric alternatives, as suggested by Heckman and Singer (1984). Since parametric alternatives could often result in oversmoothing (Knorr-Held and Raßer, 2000) and the marginal maximization through numerical approximation or simulation methods can be very intensive (Gueorguieva, 2001).

As long as the likelihood is bounded, it is maximized with respect to $\mathcal{G}(\cdot)$ by a discrete distribution with at most $K \leq n$ support points. Let us suppose that this discrete distribution puts masses p_k on locations $\mathbf{u}_k = (u_{k1}, \ldots, u_{kJ})$, $k = 1, \ldots, K$. The likelihood function in (4) can be written as:

$$L(\cdot) = \prod_{i=1}^{n} \left\{ \sum_{k=1}^{K} \left[\prod_{j=1}^{J} f(y_{ij} \mid \mathbf{x}_{ij}, u_{kj}) \right] p_k \right\} \tag{5}$$

where $p_k = \Pr(\mathbf{u}_k) = \Pr(u_{k1}, \ldots, u_{kJ})$, $k = 1, ..., K$, represents the prior probability of locations $\mathbf{u}_k$. The use of finite mixtures have several significant advantages over parametric mixture models; since it may help to classify subjects in clusters characterized by homogeneous values of random effect parameters. This is interesting in health sciences, where components can be considered as groups with similar propensity to event of interest.

Estimation may be accomplished by a standard EM algorithm. For a given K we assume that the hidden component indicator vector $\mathbf{z}_i$ is a multinomial random variable with weights p_k, $k = 1, \ldots, K$. Thus, the likelihood for the complete data $(\mathbf{y}_i, \mathbf{z}_i)$, $i = 1, \ldots, n$, is given by:

$$L_c(\cdot) = \prod_{i=1}^{n} \prod_{k=1}^{K} (f_{ik} p_k)^{z_{ik}} = \prod_{i=1}^{n} \prod_{k=1}^{K} \left(\prod_{j=1}^{J} f_{ijk} p_k \right)^{z_{ik}} \tag{6}$$

where $f_{ijk} = f(y_{ij} | \mathbf{x}_{ij}, u_{kj})$ represent *component-specific* univariate conditional densities.
The log-likelihood is written as:

$$\begin{aligned} \ell_c(\cdot) &= \sum_{i=1}^{n} \sum_{k=1}^{K} z_{ik} \{\log(p_k) + \log f_{ik}\} = \\ &= \sum_{i=1}^{n} \sum_{k=1}^{K} z_{ik} \left\{ \log(p_k) + \sum_{j=1}^{J} \log(f_{ijk}) \right\}. \end{aligned} \tag{7}$$

Within the E-step, the presence of missing data is handled by taking the conditional expectation of the log-likelihood for complete data given the observed data $\mathbf{y}_i$ and the current ML parameter estimates, say $\boldsymbol{\theta}^{(r)} = \{\boldsymbol{\delta}^{(r)}, \boldsymbol{\pi}^{(r)}\} = \{\boldsymbol{\beta}^{(r)}, \mathbf{u}^{(r)}, \boldsymbol{\pi}^{(r)}\}$. Using Bayes's rule, we replace z_{ik} with the conditional expectation:

$$E(z_{ik} \mid \mathbf{y}_i, \boldsymbol{\theta}^{(r)}) = w_{ik}^{(r)} = \frac{f_{ik} p_k}{\sum_{l=1}^{K} f_{il} p_l} = \frac{\left(\prod_{j=1}^{J} f_{ijk} \right) p_k}{\sum_{l=1}^{K} \left(\prod_{j=1}^{J} f_{ijl} \right) p_l}, \tag{8}$$

which represents the posterior probability that the $i-$th unit belongs to the $k-$th component of the finite mixture, $i = 1, \ldots, n$, $k = 1, \ldots, K$. The conditional expectation of the complete log likelihood given the data vector $\mathbf{y}_i$ is defined accordingly:

$$\begin{aligned} Q(\cdot) &= E_{\boldsymbol{\theta}^{(r)}} \{\ell_c(\cdot) | \mathbf{y}_i\} = \sum_{i=1}^{n} \sum_{k=1}^{K} w_{ik} \{\log(p_k) + \log f_{ik}\} = \\ &= \sum_{i=1}^{n} \sum_{k=1}^{K} w_{ik} \left\{ \log(p_k) + \sum_{j=1}^{J} \log(f_{ijk}) \right\} \end{aligned} \tag{9}$$

which corresponds to a finite mixture of K multivariate (J-dimensional) distributions with common weights w_{ik}. Maximizing $Q(\cdot)$ with respect to $\boldsymbol{\theta}^{(r)}$ we obtain updated ML parameter estimates $\hat{\boldsymbol{\theta}}^{(r+1)}$ given the posterior weights $w_{ik}^{(r)}$. The estimated parameters are the solution of the following M-step equations:

$$\frac{\partial Q}{\partial p_k} = \sum_{i=1}^{n} \left\{ \frac{w_{ik}}{p_k} - \frac{w_{iK}}{p_K} \right\} = 0, \tag{10}$$

$$\frac{\partial Q}{\partial \boldsymbol{\delta}} = \sum_{i=1}^{n} \sum_{k=1}^{K} w_{ik} \frac{\partial}{\partial \boldsymbol{\delta}} \log(f_{ik}) = \sum_{i=1}^{n} \sum_{k=1}^{K} w_{ik} \frac{\partial}{\partial \boldsymbol{\delta}} \left(\sum_{j=1}^{J} \log(f_{ijk}) \right). \tag{11}$$

which are weighted sums of K likelihood equations for standard GLMs with common weights w_{ik}. Solving the first equation we obtain:

$$\hat{p}_k^{(r)} = \sum_{i=1}^{n} \frac{w_{ik}^{(r)}}{n}, \tag{12}$$

which represents a well known result from ML in finite mixtures. The E- and M-steps are repeatedly alternated until the log-likelihood relative difference changes by an arbitrarily small amount.

4 Empirical application: fairness of the NHS in Italy

Data refer to 2002 household survey on family consumptions, recorded by ISTAT and summarized in the annual sampling analysis on household consumptions. Recorded variables refer mostly to household (total consumption, number of family members, food expenditure and out of pocket expenditure). In the following, we show the FFC as a benchmark for allowing comparisons among Italian Regions and other European Countries and stress impoverishment and catastrophic payments ideas since they may be taken as proxies of the NHS failures. The main limit of the indicators proposed by WHO lies in their sensitivity to the choice of the subsistence threshold.

The out of pocket expenditure has been evaluated by considering expenses for hospitalization, first aid, specialist examinations and relevant assistance services. Food consumption has been taken out of the *purchase booklet* which consider all food expenditures, alcoholics and tobacco excluded. By using equation (2), the value of the FFC index for out of pocket payments in Italy in 2002 is equal to 0.880, suggesting a low level of fairness in the out of pocket payment component of the Italian health care financing system as compared with those of other western Countries (for a not self-contained comparison see Murray et al., 2003). The FFC index is based on the idea that households should be required to pay for health proportionally to their CTP; hence it has a value less than one when there is inequality in health care payments. Such inequality may be due to horizontal or vertical inequity, or both (Wagstaff, 2002). If we consider the FFC index in terms of horizontal and vertical inequity, it can be assumed to reflect changes in proportions for health care expenditure between households with similar CTP (i.e. horizontal inequity) and, simultaneously, with different CTP (i.e. vertical inequity) or a combination of the two.

Our aim is to study in depth the fairness among Italian Regions to capture differences due to the federalist health system introduced in 2001; thus, we evaluate the FFC index for the 20 Italian Regions (see Table 1). Deep differences can be pointed out by the analysis: lower values of FFC may arise because the better-off of those who pay a larger proportion of their CTP or viceversa. Since FFC focuses on inequalities, it cannot capture the effect of health expenditures on the social life of households; hence, indices of catastrophic events and impoverishment have to be considered.

A pure descriptive analysis of such indexes may be not informative enough, since *policy makers* should be interested in identifying socio-demographic determinants that may lead households to the poverty line and inducing a higher proportion of out of pocket expenditures. For this purpose, we suggest to adopt the GLMM framework described in section 3. We define two variables: catastrophic payments and impoverishment due to health expenditures. A catastrophic payment happens when more than the 40% of the CTP goes to health expenditures; impoverishment means to live under the subsistence line once health expenditures have been provided. Here we consider the subsistence line as the absolute poverty line, as defined in Table 2, since subsistence (for developed Countries such as Italy) is slightly different from the need of food and includes other primary needs that have to be considered in the definition of the absolute poverty line (ISTAT, 2002).

To simplify the analysis and dissemination of results, we summarize the percentage of catastrophic payments and impoverishment in Table 3. From a pure descriptive analysis we may notice that impoverished households represent the 1.74% of all households with highest values in the southern Regions where the maximum is equal to 4.10% in Sicily and 3.57% in Calabria, with all the southern Regions showing higher values. Obviously, also northern Regions are subject to impoverishment; the highest values in the north are those of Piemonte (1.53%) and Trentino Alto Adige (1.57%).

The estimate of the number of households subject to catastrophic payments is 2.30% for the whole Italian territory. As for impoverishment, catastrophic payments are more frequent in the South: Sicily has a catastrophic value twice that of Italy value. In most northern Regions such ratio is lower the the Italian mean; Lombardia is an exception with 2.74%.

However, we are interested in identifying determinants of these two phenomena; hence, we propose a bivariate correlated random effect model in a finite mixture context introducing socio-demographic variables in the linear predictor. Variables definition is provided in Table 4.

We assume that CATA and IMPOOR are, conditionally on the covariates and on outcome-specific, correlated, random effects, Bernoulli random variables.

Difference in the regression parameters can be found for the OLD variable parameter estimate: in the IMPOOR model its effect is *positive*, while the presence of an old man in the family reduce the probability of being subject to catastrophic payments; OLD and CHILD are the only variables which have a significant effect in the IMPOOR equation, while the others are not significant at any level. The family structure and particularly the presence of children (CHILD, NOCHILD) may affect the probability of CATA as well as OLD (Table 5). Adopting the semiparametric approach described above, the estimate $\hat{\mathcal{G}}$ is a two-point distribution with masses (0.69, 0.31) on locations [(-0.9, 4.46); (-0.61, 1.33)], for the random effects in the CATA and IMPOOR equations, respectively. The first location corresponds to people who have a low propensity to fall down into poverty and a low to be subject to catastrophic payments. The second component groups individuals with a very high propensity to being subject to both, but higher catastrophic payments. Random effects are positively correlated; the strong positive correlation between these sources of heterogeneity is consistent with general economic and behavioral theories.

As can be easily observed, the estimated correlation between the random effects is close to unit in magnitude, which is the boundary of the parameter space. In the proposed semiparametric approach this represent a drawback: in fact, when a discrete mixing distribution is adopted, the correlation coefficient is estimated only on a small (K) number of points and when the correlation is high this can lead to a set of K points that are almost aligned. Thus, when correlation is high, the estimated parameter tends towards the bounds of the interval (-1; 1) as pointed out by Smith and Moffatt (1999).

The choice of the number of components in the finite mixture is based on the BIC criterion.

5 Conclusion

We can draw some conclusions: first we may observe the existence of a core of social iniquity made up of impoverished households, also for non voluntary catastrophic payments. We have to add the percentage of impoverished households due to health expenditures to the percentage of poor Italian households; this means that health expenditures increase poverty. Moreover, despite of the NHS role of social insurance and global coverage of population health needs, there is an additional proportion of households which have to spend more than the threshold fixed by the WHO (40% of the capacity to pay) for health care services. In absolute terms, more than 1.200.000 households are involved. Among these we find the most fragile and those which are not able (or renounce) to call for their rights to receive a global coverage; we believe that policy makers should pay particular attention to these households.

The Regions where poverty incidence is higher (ISTAT, 2002) do not correspond to those where there is a larger share of impoverished households; this demonstrates that poverty and impoverishment are two different phenomena, and also that impoverishment depends on adopted health policies. A final argument is that government choices, related to health policies, are not neutral in terms of equity, both on a regional and national basis. It is important to underline that while regional funding takes into consideration different degrees of care needs (with capitation weighted for age), it does not take into account socio-economic differences. This means that different levels of households exempted by co-payments are not considered, although we have showed they have equity impacts. In conclusion, we propose that co-payments will be tuned with appropriateness of consumption, while exemption levels should be regionally established, constraining (a citizen right) the financial impact on household budget (or saying better, on the capacity to pay).

References

Aitkin, M. (1999). A general maximum likelihood analysis of variance components in generalized linear models. *Biometrics*, **55**, 117–128.

Alfó, M. Trovato, G. (2004) Semiparametric mixture models for multivariate count data, with application. *Econometrics Journal*, **7**, 1–29.

Aronson, J.R. and Lambert, P.J. (1994). Decomposing the Gini coefficient to reveal the vertical, horizontal and reranking effects of income taxation. *National Tax Journal*, **47**, 273–294.

Breslow, N.E. and Clayton, D.G. (1993) Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association*, **88**, 9-25.

Caffo, B., An, M.W. and Rohde, C. (2007). Flexible random intercept models for binary outcomes using mixtures of normals. *Computational Statistics and Data Analysis*, 51:5220-5235.

Gueorguieva, R. (2001) A multivariate generalized linear mixed model for joint modelling of clustered outcomes in the exponential family. *Statistical Modelling*, 1:177–193.

Heckman, J.J. Singer, B. (1984) A method for minimizing the impact of distributional assumptions in econometric models of duration, *Econometrica*, **52**, 271–320.

ISTAT (2002). La povertá relativa in Italia.

Kakwani, N.C. (1977). Measurement of Tax Progressivity: An International Comparison. *Economics Journal*, **87**, 71–80.

Knorr-Held, L. and Raßer, S. (2000). Bayesian Detection of Clusters and Discountinities in Disease Maps. *Biometrics*, **56**, 13–21.

McCullagh, P. and Nelder, J.A. (1989). *Generalized Linear Models*. Chapman & Hall, New York.

Munkin, M.K. Trivedi, P.K. (1999) Simulated maximum likelihood estimation of multivariate mixed-Poisson regression models, with application. *Econometrics Journal*, **2**:29–48.

Murray et al. (2003). Assessing the distribution of household financial contribution to the health system: concepts and empirical application. In *Health System Performance Assessments*. Ginevra. WHO.

Nelder, J.A. and Wedderburn, R.W.N. (1972). Generalized linear models. *Journal of the Royal Statistical Society - Series A*, **135**, 370–384.

Reynolds, M. and Smolensky, E. (1977). *Public expenditures, taxes and the distribution of income: the Unites States 1950, 1960, 1970*. New York Academic Press. New York.

Smith, M.D. and Moffatt P.G. (1999). Fisher's information on the correlation coefficient in bivariate logistic models. *Australian and New Zeland Journal of Statistics*, **41**:315–330.

van Doorslaer, E., Wagstaff, A., van der Burg, H., Christiansen, T. and Citoni, G. (1999). The redistributive effect of health care finance in twelve OECD countries. *Journal of Health Economics*, **18**: 291–313.

van Ophem, H. (2000) Modeling selectivity in count data models. *Journal of Business and Economic Statistics*, **18**:503–510.

Wagstaff, A. (2002). Reflections on and alternatives to WHO's fairness of financial contribution index. *Health Economics*, **11**, 103–115.

Xu et al. (2003). Household health system contributions and capacity to pay: definitional, empirical and technical challanges. In *Health Systems Performance Assessments*. Ginevra. WHO

Regions	***FFC***
Emilia Romagna	0.864
Lombardia	0.871
Basilicata	0.872
Liguria	0.872
Puglia	0.874
Campania	0.875
Calabria	0.876
Trentino Alto Adige	0.879
ITALIA	**0.880**
Sicilia	0.880
Molise	0.881
Friuli Venezia Giulia	0.881
Veneto	0.882
Piemonte - Valle d'Aosta	0.885
Marche	0.888
Sardegna	0.891
Abruzzo	0.893
Umbria	0.904
Lazio	0.907
Toscana	0.907

Table 1: Fairness in Financing Contribution (FFC) index - Italian Regions

# of family members	*Poverty line (euros)*
1	494.07
2	823.02
3	1095.19
4	1342.22
5	1564.56
6	1778.65
7 or more	1976.28

Table 2: Poverty line in 2002. Source: ISTAT

Regions	*CATA*	*IMPOOR*
Italia	2.30	1.74
Piemonte- Valle d'Aosta	1.67	1.53
Lombardia	2.74	0.75
Trentino Alto Adige	2.40	1.57
Veneto	1.89	0.78
Friuli Venezia Giulia	1.82	0.74
Liguria	1.36	1.56
Emilia Romagna	2.68	1.37
Toscana	1.04	1.12
Umbria	1.32	1.48
Marche	1.21	0.93
Lazio	2.47	1.07
Abruzzo	1.49	3.15
Molise	2.19	2.39
Campania	1.78	3.39
Puglia	2.73	3.27
Basilicata	1.94	2.48
Calabria	1.41	3.57
Sicilia	4.79	4.10
Sardegna	1.97	2.43

Table 3: Percentage frequence of household with catastrophic and impoverishment payments - Italian Regions - 2002

Variable	**Definition**
CATA	1 if more than 40% of the CTP is used for health care payments
IMPOOR	1 if household goes under the absolute poverty line after health care payments
CITY	1 if the household is resident in a city
OLD	1 if an elderly lives in the family
SINGLE	1 if the household is a single
NOCHILD	1 if the household is married but has no children
CHILD	1 if the household is married and has children

Table 4: Variable definitions

Variable	**CATA**		**IMPOOR**	
	Coef.	Std. Err.	Coef.	Std. Err.
CITY	0.145	0.112	-0.100	0.130
OLD	-0.226	0.110	0.933	0.129
SINGLE	0.175	0.147	-0.046	0.167
NOCHILD	0.297	0.136	0.010	0.170
CHILD	0.372	0.164	0.341	0.162
CONSTANT	-5.991	3.919	-5.029	3.295
log-likelihood	-5955.967			

Table 5: Multivariate mixed effects logit model parameter estimates