



# **A spatial approach to the analysis of individual health care expenditures: the case of the Italian region of Friuli-Venezia Giulia**

Luca Grassetti and Laura Rizzi

Ottobre 2014

4 / 2014

Sezione di Statistical Applicata

# A spatial approach to the analysis of individual health care expenditures: the case of the Italian region of Friuli-Venezia Giulia

Luca Grassetto and Laura Rizzi

Dept. of Economics and Statistics - University of Udine, Italy

E-mail: luca.grassetto@uniud.it

## Abstract

Evidence of spatial autocorrelation in health data has been confirmed in a set of well-known studies. This paper aims at investigating the determinants of individual health care expenditures in the Region of Friuli-Venezia Giulia. This phenomenon is examined by considering a cross-sectional individual expenditure dataset exhibiting a hierarchical structure, due to the fact that patients are nested in their general practitioners (GPs). Individual expenditures regarding drug prescriptions, inpatient care and outpatient care are observed in a dataset from 2010. The model specification considers expenditures as a function of patient characteristics (age, gender and the presence of co-morbidities), contextual variables (population size of the municipality and the administrative area of residence) and characteristics of general practitioners (age, experience and type of physician). The behaviour of macro-units (general practitioners) is studied by introducing a random error term in the model specification. Finally, spatial correlation is included in the model. Given the size of the dataset (around 1 million patients), a feasible way to model spatial correlation is to introduce a deterministic term (the neighbourhood average expenditure measure). Moreover, the model must also take into account the typical zero inflation issue. To this end, a feasible two-stage Heckit method is adopted to introduce both the spatial component and the specific hierarchical structure of data in the sample selection model. Results showed that the spatial component presents a significant effect in both the selection and the level equations. Different weighting systems have been considered, but the model point estimates are not significantly affected by the definition of the neighbourhood. From the decision-maker's point of view, this analysis is useful to highlight the statistical significance of macro and micro-economic determinants of health care expenditures. Moreover, by analysing the three kinds of health care services separately, it is possible to focus on the determinants of each single health care expenditure.

**Keywords:** Health Care Expenditures, Health Econometrics, Microdata, Multi-level Models, Sample Selection, Spatial Econometrics

## 1 Introduction

During the last decades, the relevance of health care expenditure (HCE hereafter) has substantially increased, leading, where health care is public, to growing pressure on public budgets and to sustainability problems. For these reasons, monitoring the trends in health care expenditures and the analysis of HCE growth and health care demand drivers are fields of very active research (*Horizon 2020*). Even if per-capita health expenditures show increasing trends in the OECD, there is evidence of substantial heterogeneity of level of spending as a proportion of GDP across the countries as stated in Baltagi and Moscone (2010). Literature on these research interests has followed manifold directions, dealing with the study of demand (use) of health care services or with the analysis of HCE itself, both at the macro and micro levels. In the former case, factors affecting the probability of use and/or the amount of expenditure at the individual level has usually been evaluated (see Angulo *et al.*, 2011; Albouy *et al.*, 2010). In the latter case, determinants of health expenditure rise have been explored at the aggregate (Lopez-Casasnovas *et al.*, 2005) or micro level, both pointing out the relevance of income, age, education and health status. Moreover, since the seminal paper by Newhouse (1992), the determinants assessment has been a matter of extensive policy debates and the availability of macro and microdata on health care has driven studies on the effects of different factors, such as income (see Gerdtham and Jönsson, 2000; Baltagi and Moscone, 2010), prices, aging (see Werblow *et al.*, 2007), time, new health care profiles and technological progress.

Several studies focused on macrodata analysis and cross-country frameworks to identify the determinants of HCE (see, for instance, Gerdtham *et al.*, 1992; Baltagi and Moscone, 2010), considering the effects of gross domestic product, demographic structure and institutional factors. Other macrodata studies have focused on single countries, comparing jurisdictions or regions (see Giannoni and Hitiris, 2002; Di Matteo and Di Matteo, 1998) to disentangle heterogeneity. However, spatial spillovers are not often examined, even if the potential impact of decentralization designs have led to the examinations of spatial patterns of health expenditures, introducing spatial econometric models in the health care analysis (see Moscone and Knapp, 2005; Rosenberger *et al.*, 2005; Costa-Font and Pons-Novell, 2007). Aggregate data on HCE have also been considered to disentangle income effects in a long-run framework and income elasticity, using panel data from OECD countries (Baltagi and Moscone, 2010) and the US (Freeman, 2003), and then analyzing the stationarity and cointegration aspects. On the other side, the growing availability of administrative databases and the presence of even more efficient information systems have enabled easier development of micro-economic studies in this field of

analysis (e.g. Deb *et al.*, 2006).

There are the two main streams in the literature relative to the health economics studies on HCE microdata. Individual data are often collected to study the effect of patient social-economic characteristics, together with pathology, illness severity and general health status on personal health expenditures or demand (as in Wong *et al.*, 2011). Typically these studies regard specific pathologies and are developed in controlled experiment frameworks (e.g. Madden, 2008). Sometimes, individual data are analysed at the level of area to evaluate spatial heterogeneity in health care needs and demands (as in Wang, 2009). This study can be considered an intermediate case of the previous two. The aim of this work is to develop an empirical analysis on the evaluation of HCE determinants to address the evidence-based decision making process and policy. Disaggregated data enables us to disentangle the extent to which sociodemographic variables, heterogeneity of health care inputs and health status proxies explain spatial differences in health care expenditures. Our dataset was collected by considering the administrative health databases on the resident population of the Italian region of Friuli-Venezia Giulia (FVG). Information on about 1 million regional patients was collected. This population is also grouped by the 219 regional municipalities and the 1,092 general practitioners (GPs hereafter). The analysis of individual expenditures has been also widely treated in the econometric literature, where, for instance, Two-Part model, sample selection and latent class models have been adopted to study microdata with dependent variables presenting an excess of zeros (see, for instance, Buntin and Zaslavsky, 2004; Madden, 2008; Deb and Trivedi, 2002, respectively, for the three approaches). All these alternatives define individual expenditures as the combination of two stochastic processes (one for the presence of expenditure and another for the amount of the positive health care costs). Mixture models, where data are considered a combination of a zero mass distribution and the common cost function, have also played a relevant role in this sort of analysis with individual expenditures (as explained in Deb and Trivedi, 1997). Finally, copula bivariate probit models (Winkelmann, 2012) represent another recent modeling strategy adopted to analyse HCE patterns. To take into account of the complex hierarchical data structure and due to the computational issues connected with the huge dataset size, we, finally, decided to adopt the Heckman sample selection approach.

Due to the potential spillovers at the municipality level and the geographical heterogeneity of unobservable risk factors, the model estimated here takes into consideration a spatial approach. However, given the hierarchical structure of the our dataset, a feasible way to introduce the spatial integration could take into account a modified version of the neighborhood effect defined by Moscone *et al.* (2007b) and Costa-Font and Moscone (2008). Alternatively spatial correlation could be included in our model specification by means of a spatial autoregressive (SAR) error term; however, this would lead to a complex error correlation structure (the variance-

covariance matrix is a full and sparse matrix) when dealing with disaggregated hierarchical structure. Furthermore, the idea to adopt the SAR specification only for the municipalities error term appears unfeasible under a likelihood approach, but estimation of SAR models could be accomplished by adopting a Bayesian approach (as in Eibich and Ziebarth, 2013). In all these approaches, the model estimation results would be affected by the choice of the spatial weights matrix. In our empirical approach, we tested two different paradigms: the geographical distances considering, for instance, contiguity of order 1, euclidean or Cliff and Ord distances (Cliff and Ord, 1981), and the economic distances (e.g. Case *et al.*, 1993; Boarnet, 1998) applied to health care data. Individual HCE comprises different types of required health care services: inpatient hospitalizations, outpatient care services, home care services, drug prescriptions, emergency care services and so on. Since these different sources are strictly connected, a multivariate model should be considered. Our economic evaluation of individual HCE is focused on the counterpart related to drug prescription, inpatient care and outpatient care services only. This limitation is due to the availability and the reliability of the administrative data collected in the regional health databases. However, these three expenditures sources cover the greatest part of the National and Regional Health System budgets, and they present quite relevant heterogeneity at the individual and spatial levels. We will describe their distributional characteristics in detail in Section 4. Examined data on individual expenditures refer to year 2010 and have been derived by the integrated regional data-warehouse. Other information on resident population patients and on general practitioners have been derived from personal data registries. Moreover, patients' health condition has been proxied by the information registered in the outpatient care, inpatient care, drug prescription and other health services (for instance, home care and emergency services) datasets.

The paper is organised as follows. Section 2 summarises some recent empirical evidence on the determinants of health expenditures and the subsequent public policies. Section 2.1 is devoted to empirical model specification. Section 3 presents the data with particular attention to data structure and sources. Section 4 collects the results of the model estimation. Concluding remarks and future developments for research are given in Section 6.

## 2 The analysis of the HCE

As just introduced the econometric analysis of the determinants of HCE can be addressed considering both individual and aggregated data sources. On one hand, individual data studies concerning the effects of socio-economic and health factors on health expenditure can be classified as quasi-experimental studies, epidemiological studies and observational studies regarding the analysis of the determinants of individual expenditures. On the other hand, analyses of aggregate data are adopted to study macro-variations in health needs or demand and to develop macro-area

studies. By exploring the determinants of health expenditure at the macro-level, Wang (2009) and Baltagi and Moscone (2010) assessed that gross state products, proportion of people over 65 years in the population, level of urbanisation and deprivation, along with the number of hospital beds, have significant effects on health expenditures. At the micro-level, the analyses in Getzen (2000), Madden (2008) and Deb *et al.* (2006) identify age, gender, health status, income and educational levels as determinants of individual health care demand. In our study, some variables are collected at the individual level: age class; sex; health-related exemptions and the presence of other health costs (considered as proxies of co-morbidities). All these variables are designed to allow for individual heterogeneity. The model specification is then completed by introducing the characteristics of grouping factors, the GPs (age class, experience – time from graduation and GP classification – paediatric or not) and the municipalities (i.e. population size and the presence of hospitals).

Statistical tools for spatial analyses have been thoroughly applied to health data in the analysis of the prevalence or incidence of diseases and their related costs (see Moscone, 2011; Moscone *et al.*, 2007a), in the examination of potential expenditure spillovers across the geographical area of a country (as in Costa-Font and Pons-Novell, 2007) or across countries in an international HCE comparison (as in Baltagi and Moscone, 2010). However, studies on individual health data explored in relation to HCE with the inclusion of a spatial correlation structure are rare in the health economics literature. One of the major assumptions in spatial models is that an individual’s behaviour is affected by the characteristics and behaviours of the people in the neighbourhood. In particular, the interactions among individuals may lead to a collective behaviour, which empirically represents a spatial correlation pattern. The determinants of this spatial correlation were stated in Moscone *et al.* (2007a), where mental-health expenditure analysis was carried out at the local authority level and a set of determinants of the Local Authority’s behaviour were identified. The performance of a productive unit may affect its neighbours’ behaviours and, consequently, its expenditure patterns. This effect, called the “demonstrative effect”, justifies the introduction of the deterministic component in the spatial autocorrelation structure of the specified model. The definition of the contextual effect is based on the idea that adjacent “productive units” share common observable characteristics. For this reason, the proposed model considers some variables at the aggregate level. The spatial term included aims at explaining the outcome of common exogenous environmental conditions (such as authority policies) on HCE. Finally, the correlated effects are designed to collect the unobserved determinants of interdependence. The cited results can be adapted to a general health expenditures framework and most of the stated hypotheses on the local authorities’ behaviours are also relevant for GPs. This study aims to test for the presence of a spatial correlation component and to eventually consider it in the cost prediction function. In particular, the model considers the hypothesis of si-

multaneous determination of local and neighbouring expenditures as in Brueckner (2003) or Anselin (2002). However, the application of this kind of specification to disaggregated data is not straightforward. The distance between individuals is not directly available and its potential use involves an infeasible model specification. However, the adoption of spatial correlation at the municipality level seems to be more suitable. The weighted sum of neighbouring effects can be defined by considering the average per-capita expenditures in each municipality, by assuming that individual expenditures are affected by the general neighbourhood behaviour. This allows us to specify the models by means of a feasible weighting system (about 1 million individuals are nested in 219 municipalities only).

To proceed with the analysis, we must face another two common issues in microeconomic-data analysis: zero-inflation and multilevel data structure (complex heterogeneity). As introduced previously, some alternative approaches have been developed in the econometric literature to deal with the former problem. To feasibly solve the computational problems due to large data size, we decided to adopt the sample selection model paradigm and to face the estimation issues by means of the Heckit correction in the two-step procedure (Heckman, 1979). These choices allowed us to include the administrative health macro-areas (AHA hereafter) dummies in the model. Their effects, in fact, were collinear to the GPs' ones in a fixed-effect framework. In order to evaluate the impact of alternative estimation procedures, considering a reduced model specification, we compared: two-step random and fixed-effect, maximum likelihood and copula based strategies.

The peculiar hierarchical structure of the data allowed us to consider two distinct non-nested levels of data aggregation: GPs and municipalities respectively. In the empirical application, the two grouping factors can be considered by adopting a two-level random-effects model with independent cross-classified effects. Random-effect probit and linear models were adopted in the development of the two-step estimation procedure.

## 2.1 A spatial HCE function specification

The HCE model specification can be accomplished by considering some subsequent steps. First, it is possible to define the hierarchical spatial model considering the cross-classified random intercept model specification

$$y_{ij,m} = \beta_{0j,m} + \beta' \mathbf{x}_{ij,m} + \epsilon_{ij,m}, \quad (1)$$

where  $i$ ,  $j$  and  $m$  identify individuals, municipalities and GPs, respectively;  $\beta_{0j,m}$  can be defined as the sum of the general intercept  $\beta_0$  and two independently distributed cross-classified random terms capturing the heterogeneity effects of municipalities and GPs respectively. The deterministic term  $\beta' X_{ij,m}$  captures the effects of covariates mentioned in the previous section, which are observed variables measured at the different levels of the hierarchical structure.

The municipalities' roles in the HCE distribution can also be formalised by considering the spatial correlation between municipalities; in this case, we can introduce a deterministic model component, as in Moscone *et al.* (2007a) and Costa-Font and Moscone (2008). However, its specification is non-standard in the microdata context; therefore, we decided to introduce a modification of the standard deterministic component defined as

$$\gamma \sum_{k=1}^K w_{kj} \bar{y}_k, \quad (2)$$

where the weighting system  $w_{kj}$  is defined at the municipality level. Different weighting systems can be adopted to describe the neighbourhood municipality effect (see Zaccamer and Mason, 2011, for a full discussion on economic neighbourhood specification). The deterministic term defined in equation (2) elicitates the interaction between the units in the expenditure process, whereas the remainder of the formula represents the common linear effect of the covariates. Since the weighting system identifies the strength of the interactions between territorial units, we decided to test some alternative metrics, whose results were compared by means of a sensitivity analysis. First, the geographical distance paradigm was considered by obtaining the contiguity matrices (of a different order with rook or queen configurations) and the Cliff and Ord distances (also in their truncated versions). Moreover, by adopting the economic distance paradigm, we could define the Case and Boarnet distances. The knowledge of a spatial clustering system, such as that defined by local health authorities, allowed us to further define the concept of quasi-neighbourhood (as in Zaccamer and Mason, 2011). All distances were calculated by considering the municipalities as reference territorial units. Conditioned on the availability of GIS data concerning the municipalities, the distances between territorial units could be calculated. The paradigm of the economic distance between spatial units (municipalities in our example) was adapted to the health care framework. The construction of the distance functions was based on some possible general health-status proxies; in particular, instead of the typical economic measures, population size or percentage of individuals with at least one exception were considered in the Boarnet or Case formulations. The mean  $\bar{y}_k = \frac{1}{N_k} \sum_{i=1}^{N_k} y_{ik}$  represents the municipality per-capita expenditure. It is easy to note that the adoption of this aggregated solution is equivalent to the choice of an individual weighting system in which the distances between municipalities are applied by standardizing the weighting system with the sizes of the populations ( $N_k$ ). When introducing the spatial correlation in the model specification, the municipality error term can be dropped because the two effects are partially overlapping. For this reason, we will consider the simple random intercept model in the following model specification.

The model is defined by combining equations 1 and 2, which does not yet allow for the presence of a large number of zero expenditures. Therefore, the sample



selection model paradigm (as in Leung and Yu, 1996; Madden, 2008) is considered with the specification

$$y_{ij,m} = \begin{cases} \beta_{0j,m} + \gamma \sum_{k=1}^K w_{kj} \bar{y}_k + \beta \mathbf{x}_{ij,m} + \epsilon_{ij,m} & \text{for } y^d = 1 \\ 0 & \text{for } y^d = 0, \end{cases} \quad (3)$$

where the process  $y^d$  can be defined by considering a generalised linear mixed model, where the explicative variables may differ from those included in the model for the levels of expenditures. The Heckit method may be adopted to check for selection bias.

Models for the selection and the amount of individual expenditures may include the spatial term; in fact, the two spatial components can be defined at the municipality level by considering, respectively,

- the weighted average of neighbourhood percentages of positive expenditures and;
- the weighted average of neighbourhood per-capita expenditures (calculated on the positive expenditures only).

We decided to study the multivariate HCE phenomenon by estimating separated models for each single type of expenditure including the correlations between them using of a set of dummy variables.

Even in this simplified context, the estimation process presented some estimation issues due to the size of the population dataset and, for this reason, the adopted estimation algorithm was characterised by the following steps.

- Step 1 Estimation of the generalised mixed (probit) model for the presence of a positive expenditure.
- Step 2 Consideration of Heckman's correction in the specification of a linear mixed model for the positive expenditure sub-dataset. The *exclusion restriction* is respected by considering a different set of variables for the two steps of the estimation.

We used Stata (StataCorp, 2009) to obtain the estimates of the final models (`xtprobit`}, `xtreg`}, `heckman`} and `heckmancopula`} commands were considered). In particular, we estimated the copula sample selection model using the Stata program `heckmancopula`} developed by Hasebe (2013). Preliminary analyses have been developed in R (R Core Team, 2013).

### 3 The data

The dataset includes information on about 1 million individuals clustered in 1,092 GPs and 219 municipalities. Following the record linkages and the post-editing processes, 999,488 individuals were included in the study. The database

was extracted from the fully developed administrative regional health informative system (SISSR). The considered archives of SISSR belong to the following distinct systems: territorial (SIASI and SIES), hospital (PS NET, G2, NET LAB and so on), drug prescription and ambulatory (SIASA) informative systems. Such archives collect huge amounts of data on both pathologies and health care profiles. The HCE for drug prescription, outpatient and inpatient care and other health services were also registered here. Together with this information, data on the socio-demographic characteristics of all the resident patients were available, thanks to the population regional registry.

The first step in the data analysis dealt with the description of the phenomenon of interest. First, we observe a large proportion of null HCEs, and summary statistics reported in Table 1 show the relationship between the percentage of null expenditures and some explicative variables. Evidence suggests to treat this problem by considering a specific prediction model for excess of zero expenditures. A similar analysis was developed to investigate this issue at the municipality and GP levels, pointing out a relatively stable pattern with respect to these factors.

The descriptive analysis of positive expenditures is reported in Table 2. These summary statistics can be used to mark the relevance of individual, GP and municipality characteristics on HCE heterogeneity.

[Table 1 about here.]

[Table 2 about here.]

As mentioned previously, the spatial structure of the dataset is linked to individuals' residences. In particular, we considered municipalities as spatial units. Plots in Figure 1 show the territorial distribution of the multivariate phenomenon under examination. This descriptive analysis does not exhibit spatial patterns in the HCEs. Nevertheless, the spatial components seem to be significant in the model estimation.

[Figure 1 about here.]

## 4 The estimation results of the spatial HCE models

As stated in Section 2.1, the estimation procedure was carried out separately for each source of HCE. As just introduced, a two-step estimation procedure was considered to estimate the sample selection model, and Heckit correction was adopted to check for selection bias. While the first step focused on the estimation of the probit selection models (the results obtained are reported in Table 3), the HCE models were selection-bias-corrected linear mixed models estimated by considering positive expenditures only (the estimation results of the second-step are reported

in Table 4). A comparison between the estimated models and the simple linear mixed models for positive expenditures was also performed. Because of the significant changes in the estimates and in the models' goodness of fit measures, the necessity of Heckman's correction was confirmed in our analysis.

First-step mixed probit models (see Table 3) revealed a negative effect of male gender and increasing positive effects of class age. Patients older than 65 exhibit a significantly greater probability of demand for health care services, especially for drug prescriptions and outpatient care. Moreover, patients cared for at home were more likely to require the three types of care services analysed. This positive association between demands for different health care services was also confirmed by the positive coefficients of the dummy variables on the presence of some expenditures included in the linear predictor. These positive and significant associations between types of services revealed that health care services are not substitutes; that is, patients facing greater needs for health care are generally more likely to demand all types of care services. Being under the care of a paediatrician led to a greater probability of drug prescriptions and hospitalisations but a lower probability of outpatient expenditures. While the presence of some drug prescriptions did not show significant heterogeneity across AHAs, relevant effects of AHAs emerged in the probabilities of the other two types of services. Variances of second level errors are significant in all three probit models.

[Table 3 about here.]

The second-step linear mixed models (see Table 4) for the expenditure amounts on drug prescriptions, outpatient care and hospital services assessed the selection bias, since the effect of the inverse Mills ratios were significant. Thus, the adopted Heckman's two-step procedure is justified. The age-class (reference age-class: "less than 24 years") effect was confirmed positive and increasing because all the dummy coefficients differed significantly from zero in all three models. Instead, being under the care of a paediatrician was a factor that reduced the average amount of expenditures on all types of care services. The greatest predicted average expenditure in drug prescriptions and outpatient services was observed among patients residing in the administrative health area 4 (AHA 4), while the highest predicted expenditures for hospitalisation were among those residing in AHA 5. The population size of the municipalities, included as a weighting factor, showed a negligible effect. The associations between HCE and the presence of different kinds of expenditures appeared, instead, of more difficult interpretation. For drug prescription and outpatient care expenditures were positively affected by the presence of the other kind of HCE. On the contrary inpatient care expenditures were negatively related to them. These variables were introduced into the models as proxies of the presence of multiple pathologies, and thus, of bad health status.

[Table 4 about here.]

To justify the choice of two-step random effects models, we conducted a comparison of some alternative model estimation methods. In particular, we considered the copula based, the maximum likelihood and the two-step procedure based on fixed-effects model specification. The maximum likelihood estimation based on copulas were obtained considering the Frank copula as defined in Hasebe and Vijverberg (2012) Results of the comparison were collected in Tables 5, 6 and 7 for the drug, outpatient and hospital care expenditures respectively. The macro-area dummies were omitted here in order to obtain fully comparable models. In fact, the fixed-effect model specifications were affected by multicollinearity (GPs are clustered in the macro areas). The comparison showed that the different estimation methods brought to very similar results. The alternative specifications of endogeneity did not affect the selection and level processes. The only contradictory result was the one of Frank copula applied to hospitalisation expenditure. Computational issues were observed in the in-patient expenditure model and for this reason the estimation results under Frank copula were different from the others

[Table 5 about here.]

[Table 6 about here.]

[Table 7 about here.]

## 5 Discussion

In this paper, we focused on the assessment of HCE determinants within a microdata framework, which allowed us to combine the advantages of both individual and macro-level studies. The estimation results of the sample selection models confirmed some well-known health expenditure drivers, such as gender, age and health care need. The relevant role of individual health status was supported by the positive correlation between different sources of expenditures (e.g. Werblow *et al.*, 2007). This result also supports the idea that different types of health care services are not substitute services when the whole population demand is considered and increasing levels of expenditures for all types of services reveal higher levels of need. In our analysis, variables included as proxies of co-morbidity, such as the home assistance service factor and “other expenditures” dummies, showed positive effects in all models, except hospitalization in-patient services.

The cross-classified hierarchical structure of our database allowed us to assess positive spatial patterns, which were also identified in some literature (e.g. Moscone *et al.*, 2007a,b, for the macroeconomic-data analysis) mindful of cross-section dependence or spatial spillovers. Selection models (on probability of some expenditure) on drug prescriptions and hospitalization presented strong positive spatial component effects, probably underlying relevant roles of GP and proximity

to the hospitals. In fact, models on the amount of expenditures revealed a positive spatial effect when drug prescriptions and outpatient care expenditures were considered.

Generally, the probability of access to the health care system and the level of expenditures strongly depend on the characteristics of the population, both at the individual and area levels. In fact, the models estimation showed a significant relationship between HCEs and age-class, gender (as in Werblow *et al.*, 2007) and the size of the municipalities' population, as stated also by Costa-Font and Pons-Novell (2007). On the GP level, the paediatricians presented higher probabilities of drug prescriptions and hospitalisations and a lower probability of outpatient visits; however, their effect on the amount of expenditures are negative, as expected, considering the population in their charge. Finally, in all estimated models, the introduction of the inverse Mills ratio revealed the relevance of a selection-process consideration in this type of analysis.

## 6 Conclusions

This paper aimed to analyse individual HCE data by considering both the micro and macroeconomic determinants of the phenomenon. In particular, we studied the spatial distribution of the expenditures without considering aggregate data. To take into account all the informative aspects of our dataset, a hierarchical spatial sample selection model was proposed and evaluated. The model was applied to the HCE data of the whole population residing in the Italian region of FVG. This application of hierarchical spatial models to the individual-level analysis of HCE allowed us to examine the phenomenon through the original disaggregated data.

From the decision makers' points of view, the analysis of both the micro and macro determinants of HCE can be very relevant. The analysis of the determinants of expenditures, besides leading to a better understanding of this health economic phenomenon, shows that most of the significant effects in the disaggregated estimated models are connected with individual specific variables. However, the contextual variables (at the levels of general practitioners and of municipalities) are also influential. In fact, a small but significant heterogeneity between GPs was identified, and the GPs' specific error component is quite relevant, as shown by the analysis. Moreover, the role of the spatial deterministic component demands further investigation, since a significant positive spatial correlation can be observed throughout the regional territory.

The estimated models presented interesting results, but other drivers of HCE should also be included (i.e. more accurate measures of individual health status, social deprivation at the municipality level and GPs' performance measures). Therefore, this work should be considered a first step in the analysis of individual expenditures at the individual population level.

As discussed in Section 1, the model specifications and estimations are devel-

oped through a feasible, simplified approach. Generalisations of the model can also consider a SAR error structure. Moreover, the model can be developed in a more general sample selection framework (at least considering the fixed-effect model specifications). Finally, a proper multivariate model specification can be considered, even if preliminary analyses show that correlations between different expenditures have negligible effects on model estimation, due to the huge amount of available observations.

*Acknowledgement:* the authors would like to thank Prof. Badi H. Baltagi and all the other participants to the “IV Health Econometrics Workshop” for the precious and helpful suggestions and comments.

## References

- Albouy, V., Davezies, L. and Debrand, T. (2010) Health expenditure models: A comparison using panel data. *Economic Modelling*, **27**, 791–803.
- Angulo, A. M., Barberán, R., Egea, P. and Mur, J. (2011) An analysis of health expenditure on a microdata population basis. *Economic Modelling*, **28**, 169–180.
- Anselin, L. (2002) Under the hood issues in the specification and interpretation of spatial regression models. *Agricultural economics*, **27**, 247–267.
- Baltagi, B. H. and Moscone, F. (2010) Health care expenditure and income in the oecd reconsidered: Evidence from panel data. *Economic Modelling*, **27**, 804 – 811.
- Boarnet, M. G. (1998) Spillovers and the locational effects of public infrastructure. *Journal of Regional Science*, **38**, 381–400.
- Brueckner, J. K. (2003) Strategic interaction among governments: An overview of empirical studies. *International regional science review*, **26**, 175–188.
- Buntin, M. B. and Zaslavsky, A. M. (2004) Too much ado about two-part models and transformation?: Comparing methods of modeling medicare expenditures. *Journal of health economics*, **23**, 525–542.
- Case, A. C., Rosen, H. S. and Hines Jr, J. R. (1993) Budget spillovers and fiscal policy interdependence: Evidence from the states. *Journal of public economics*, **52**, 285–307.
- Cliff, A. and Ord, J. (1981) *Spatial process, models and applications*. Pion, London.

- Costa-Font, J. and Moscone, F. (2008) The impact of decentralization and inter-territorial interactions on spanish health expenditure. *Empirical Economics*, **34**, 167–184.
- Costa-Font, J. and Pons-Novell, J. (2007) Public health expenditure and spatial interactions in a decentralized national health system. *Health economics*, **16**, 291–306.
- Deb, P., Munkin, M. K. and Trivedi, P. K. (2006) Bayesian analysis of the two-part model with endogeneity: application to health care expenditure. *Journal of Applied Econometrics*, **21**, 1081–1099.
- Deb, P. and Trivedi, P. K. (1997) Demand for medical care by the elderly: a finite mixture approach. *Journal of applied econometrics*, **12**, 313–336.
- Deb, P. and Trivedi, P. K. (2002) The structure of demand for health care: latent class versus two-part models. *Journal of health economics*, **21**, 601–625.
- Di Matteo, L. and Di Matteo, R. (1998) Evidence on the determinants of canadian provincial government health expenditures: 1965–1991. *Journal of Health Economics*, **17**, 211–228.
- Eibich, P. and Ziebarth, N. (2013) Examining the structure of spatial health effects using hierarchical bayes models. Tech. rep., Verein für Socialpolitik/German Economic Association.
- Freeman, D. G. (2003) Is health care a necessity or a luxury? pooled estimates of income elasticity from us state-level data. *Applied Economics*, **35**, 495–502.
- Gerdtham, U.-G. and Jönsson, B. (2000) International comparisons of health expenditure: theory, data and econometric analysis. *Handbook of health economics*, **1**, 11–53.
- Gerdtham, U.-G., Sjøgaard, J., Andersson, F. and Jönsson, B. (1992) An econometric analysis of health care expenditure: a cross-section study of the oecd countries. *Journal of health economics*, **11**, 63–84.
- Getzen, T. E. (2000) Health care is an individual necessity and a national luxury: applying multilevel decision models to the analysis of health care expenditures. *Journal of Health Economics*, **19**, 259 – 270.
- Giannoni, M. and Hitiris, T. (2002) The regional impact of health care expenditure: the case of italy. *Applied Economics*, **34**, 1829–1836.
- Hasebe, T. (2013) Copula-based maximum-likelihood estimation of sample-selection models. *The Stata journal*, **13**, 547–573.

- Hasebe, T. and Vijverberg, W. P. (2012) A flexible sample selection model: A gtl-copula approach. Tech. rep., Discussion Paper Series, Forschungsinstitut zur Zukunft der Arbeit.
- Heckman, J. J. (1979) Sample selection bias as a specification error. *Econometrica: Journal of the econometric society*, 153–161.
- Leung, S. F. and Yu, S. (1996) On the choice between sample selection and two-part models. *Journal of econometrics*, **72**, 197–229.
- Lopez-Casasnovas, G., Costa-Font, J. and Planas, I. (2005) Diversity and regional inequalities in the spanish “system of health care services”. *Health Economics*, **14**, S221–S235.
- Madden, D. (2008) Sample selection versus two-part models revisited: The case of female smoking and drinking. *Journal of Health Economics*, **27**, 300–307.
- Moscone, F. (2011) Geographical variations in expenditure of learning disability services in england. *Applied Economics*, **43**, 2997–3005.
- Moscone, F. and Knapp, M. (2005) Exploring the spatial pattern of mental health expenditure. *Journal of mental health policy and economics*, **8**, 205.
- Moscone, F., Knapp, M. and Tosetti, E. (2007a) Mental health expenditure in england: A spatial panel approach. *Journal of Health Economics*, **26**, 842 – 864.
- Moscone, F., Tosetti, E. and Knapp, M. (2007b) Sur model with spatial effects: an application to mental health expenditure. *Health economics*, **16**, 1403–1408.
- Newhouse, J. P. (1992) Medical care costs: how much welfare loss? *The Journal of Economic Perspectives*, 3–21.
- R Core Team (2013) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
- Rosenberger, R. S., Sneh, Y., Phipps, T. T., Gurvitch, R. *et al.* (2005) A spatial analysis of linkages between health care expenditures, physical inactivity, obesity and recreation supply. *Journal of Leisure Research*, **37**, 216–235.
- StataCorp (2009) *Stata Statistical Software: Release 11*. College Station, TX.
- Wang, Z. (2009) The determinants of health expenditures: evidence from us state-level data. *Applied Economics*, **41**, 429–435.
- Werblow, A., Felder, S. and Zweifel, P. (2007) Population ageing and health care expenditure: a school of ‘red herrings’? *Health Economics*, **16**, 1109–1126.



- Winkelmann, R. (2012) Copula bivariate probit models: with an application to medical expenditures. *Health Economics*, **21**, 1444–1455.
- Wong, A., van Baal, P. H., Boshuizen, H. C. and Polder, J. J. (2011) Exploring the influence of proximity to death on disease-specific hospital expenditures: a carpaccio of red herrings. *Health economics*, **20**, 379–400.
- Zaccomer, G. and Mason, P. (2011) A new spatial shift-share decomposition for the regional growth analysis: a local study of the employment based on italian business statistical register. *Statistical Methods & Applications*, **20**, 329–356.

## List of Figures

- 1 Spatial distribution of total health care and its components. . 18

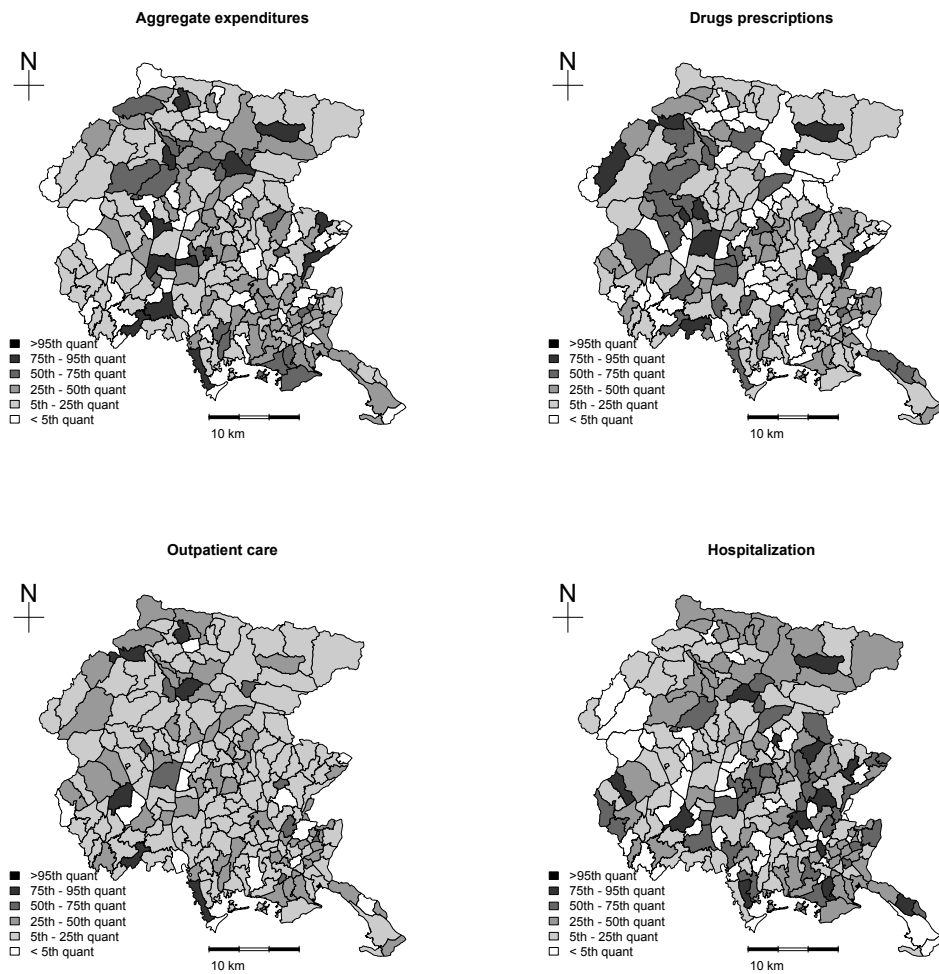


Figure 1: Spatial distribution of total health care and its components.

## List of Tables

1	Conditional percentages of positive health care expenditures in year 2010. Data refers to 999,488 individuals clustered in 219 municipalities and in 1092 GPs. . . . .	20
2	Conditional means of positive health care expenditures in year 2010. Data refers to 999,488 individuals clustered in 219 municipalities and in 1092 GPs. . . . .	21
3	First-step probit models estimation results. . . . .	22
4	Second-step linear mixed models estimation results. . . . .	23
5	The comparison of some estimation strategies for sample-selection model applied to drug expenditures. Sample-selection two-steps estimation strategies considering both fixed and random-effects models are compared with the maximum likelihood and the copula-based ones. . . . .	24
6	The comparison of some estimation strategies for sample-selection model applied to outpatient care expenditures. Sample-selection two-steps estimation strategies considering both fixed and random-effects models are compared with the maximum likelihood and the copula-based ones. . . . .	25
7	The comparison of some estimation strategies for sample-selection model applied to hospital care expenditures. Sample-selection two-steps estimation strategies considering both fixed and random-effects models are compared with the maximum likelihood and the copula-based ones. . . . .	26

Table 1: Conditional percentages of positive health care expenditures in year 2010. Data refers to 999,488 individuals clustered in 219 municipalities and in 1092 GPs.

Variables	Percentage of individual Expenditures			Number of observations
	Drug prescriptions	Outpatient care	Inpatient care	
<i>General</i>	66.13%	64.95%	10.00%	999, 488
<i>Gender</i>				
Female	71.16%	70.01%	10.86%	518, 675
Male	60.70%	59.49%	9.07%	480, 813
<i>Age Class</i>				
Under 25	46.64%	47.74%	7.82%	200, 089
25-44	53.23%	56.50%	8.34%	250, 921
45-64	69.85%	67.39%	8.25%	303, 550
Over 65	90.64%	84.65%	15.65%	244, 928
<i>GP type</i>				
General Practitioner	67.34%	67.13%	9.79%	915, 959
Paediatrician	52.80%	41.11%	12.35%	83, 529
<i>Home assistance level</i>				
0	63.41%	62.48%	8.65%	902, 717
1	88.21%	85.39%	18.92%	46, 884
2	94.47%	90.41%	25.98%	49, 887
<i>Administrative area</i>				
1	67.05%	65.75%	10.19%	196, 941
2	67.05%	66.44%	10.29%	115, 279
3	64.62%	65.03%	10.12%	57, 278
4	64.82%	64.41%	10.14%	288, 300
5	66.84%	65.98%	9.75%	88, 575
6	66.57%	63.90%	9.62%	253, 115

Table 2: Conditional means of positive health care expenditures in year 2010. Data refers to 999,488 individuals clustered in 219 municipalities and in 1092 GPs.

Variables	Mean values of individual Expenditures		
	Drug prescriptions	Outpatient care	Inpatient care
<i>General</i>	337.82	391.26	4931.28
<i>Gender</i>			
Female	324.12	391.46	4577.11
Male	355.15	391.01	5388.47
<i>Age Class</i>			
Under 25	68.89	189.80	2164.24
25-44	138.16	297.47	3393.94
45-64	306.42	369.44	5240.85
Over 65	600.98	569.75	6697.95
<i>GP type</i>			
General practitioner	358.49	404.70	5291.43
Paediatrician	48.69	150.52	1801.07
<i>Home assistance level</i>			
0	280.75	347.27	4082.79
1	599.38	653.44	7381.19
2	801.49	708.73	8368.06
<i>Administrative area</i>			
1	365.89	432.99	5096.43
2	359.27	395.24	5133.62
3	317.10	354.10	4960.98
4	332.43	377.48	4944.47
5	331.70	387.76	5101.52
6	318.65	381.61	4613.43

Table 3: First-step probit models estimation results.

	Presence of positive Health Care Expenditures					
	Drug prescriptions		Outpatient Care		Inpatient care	
<i>Intercept</i>	-2.393	(0.625)	0.340	(0.459)	-3.656	(0.234)
<i>Male</i>	-0.171	(0.003)	-0.175	(0.003)	-0.008	(0.004)
<i>Age 25 – 44</i>	0.250	(0.005)	-0.000	(0.005)	0.241	(0.008)
<i>Age 45 – 64</i>	0.640	(0.005)	0.171	(0.005)	0.105	(0.008)
<i>Age ≥ 65</i>	1.255	(0.006)	0.484	(0.005)	0.247	(0.008)
<i>Home Assist. Lev.2</i>	0.348	(0.009)	0.298	(0.008)	0.332	(0.007)
<i>Home Assist. Lev.3</i>	0.533	(0.011)	0.359	(0.009)	0.532	(0.007)
<i>Paediatrician</i>	0.354	(0.015)	-0.444	(0.013)	0.655	(0.012)
<i>Admin. Area 2</i>	0.011	(0.018)	0.058	(0.016)	0.061	(0.013)
<i>Admin. Area 3</i>	-0.011	(0.022)	0.053	(0.020)	0.041	(0.017)
<i>Admin. Area 4</i>	-0.001	(0.014)	0.035	(0.015)	0.060	(0.011)
<i>Admin. Area 5</i>	-0.005	(0.020)	0.033	(0.018)	-0.002	(0.016)
<i>Admin. Area 6</i>	0.050	(0.018)	0.012	(0.019)	0.073	(0.013)
<i>Spatial component</i>	2.783	(0.656)	-0.962	(0.704)	12.024	(2.306)
<i>Population size</i>	-0.002	(0.002)	0.003	(0.001)	0.008	(0.002)
<i>Dummy Drugs Exp.</i>	–	–	0.783	(0.003)	0.333	(0.005)
<i>Dummy Outp. Exp.</i>	0.777	(0.003)	–	–	0.698	(0.005)
<i>Dummy Hospital.</i>	0.378	(0.006)	0.810	(0.006)	–	–
$\sigma_u$	0.137	(0.003)	0.119	(0.003)	0.080	(0.003)

Table 4: Second-step linear mixed models estimation results.

	<b>Health Care Expenditures</b>					
	<b>Drug prescriptions</b>		<b>Outpatient Care</b>		<b>Inpatient care</b>	
<i>Intercept</i>	2.029	(0.839)	0.677	(1.239)	9.334	(1.380)
<i>Male</i>	0.217	(0.004)	-0.041	(0.004)	0.070	(0.006)
<i>Age 25 – 44</i>	0.186	(0.009)	0.265	(0.005)	0.095	(0.014)
<i>Age 45 – 64</i>	0.837	(0.013)	0.369	(0.006)	0.414	(0.014)
<i>Age <math>\geq</math> 65</i>	1.593	(0.018)	0.577	(0.008)	0.474	(0.015)
<i>Paediatrician</i>	-0.560	(0.019)	-0.323	(0.015)	-0.951	(0.020)
<i>Admin. Area 2</i>	0.001	(0.022)	0.051	(0.019)	0.026	(0.014)
<i>Admin. Area 3</i>	-0.031	(0.027)	0.015	(0.032)	0.006	(0.019)
<i>Admin. Area 4</i>	0.047	(0.019)	0.072	(0.021)	0.019	(0.011)
<i>Admin. Area 5</i>	-0.009	(0.026)	0.012	(0.024)	0.043	(0.016)
<i>Admin. Area 6</i>	0.018	(0.027)	-0.034	(0.030)	-0.019	(0.017)
<i>Spatial component</i>	0.446	(0.179)	0.801	(0.242)	0.074	(0.171)
<i>Population size</i>	0.006	(0.002)	0.005	(0.001)	-0.007	(0.003)
<i>Dummy Drugs Exp.</i>	–	–	0.174	(0.012)	-0.096	(0.010)
<i>Dummy Outp. Exp.</i>	0.067	(0.013)	–	–	-0.402	(0.016)
<i>Dummy Hospital.</i>	0.238	(0.007)	0.533	(0.009)	–	–
<i>Inverse Mills Ratio</i>	-1.178	(0.028)	-0.570	(0.026)	-1.023	(0.021)
$\sigma_u$	0.152		0.107		0.041	
$\sigma_e$	1.372		1.105		0.887	



Table 5: The comparison of some estimation strategies for sample-selection model applied to drug expenditures. Sample-selection two-steps estimation strategies considering both fixed and random-effects models are compared with the maximum likelihood and the copula-based ones.

<b>Heckman Models: Selection Equations</b>				
	<i>Two-Step Random-Effects</i>	<i>Two-Step Fixed-Effects</i>	<i>Maximum Likelihood</i>	<i>Frank Copula</i>
<i>Intercept</i>	-3.333 (0.320)	-1.901 (0.632)	-1.824 (0.605)	-1.856 (0.613)
<i>Male</i>	-0.171 (0.003)	-0.171 (0.003)	-0.168 (0.003)	-0.166 (0.003)
<i>Age 25 – 44</i>	0.251 (0.005)	0.251 (0.005)	0.256 (0.005)	0.263 (0.005)
<i>Age 45 – 64</i>	0.640 (0.005)	0.641 (0.005)	0.651 (0.005)	0.677 (0.005)
<i>Age ≥ 65</i>	1.255 (0.006)	1.257 (0.006)	1.195 (0.006)	1.235 (0.006)
<i>Paediatrician</i>	0.356 (0.015)	0.414 (0.071)	0.416 (0.070)	0.416 (0.071)
<i>Home Assist. Lev. 2</i>	0.348 (0.009)	0.349 (0.009)	0.423 (0.008)	0.419 (0.008)
<i>Home Assist. Lev. 3</i>	0.532 (0.011)	0.534 (0.011)	0.674 (0.010)	0.659 (0.010)
<i>Dummy Outp. Exp.</i>	0.777 (0.003)	0.777 (0.003)	0.769 (0.003)	0.353 (0.006)
<i>Dummy Hosp. Exp.</i>	0.378 (0.006)	0.378 (0.006)	0.351 (0.006)	0.773 (0.003)
<i>Spatial component</i>	4.220 (0.480)	1.913 (0.953)	1.802 (0.912)	1.826 (0.924)
<i>Pop. Size</i>	-0.001 (0.002)	-0.001 (0.002)	-0.001 (0.002)	0.000 (0.002)
<i>Pseudo R<sup>2</sup></i>		0.191		
<i>Correctly Classified</i>		0.731		
<b>Heckman Models: Expenditures Equations</b>				
<i>Intercept</i>	2.951 (0.499)	1.817 (1.182)	1.481 (1.171)	1.779 (1.169)
<i>Male</i>	0.217 (0.004)	0.202 (0.004)	0.184 (0.004)	0.165 (0.004)
<i>Age 25 – 44</i>	0.185 (0.009)	0.216 (0.008)	0.258 (0.008)	0.250 (0.007)
<i>Age 45 – 64</i>	0.836 (0.013)	0.906 (0.012)	0.999 (0.008)	1.013 (0.007)
<i>Age ≥ 65</i>	1.591 (0.018)	1.702 (0.017)	1.845 (0.009)	1.910 (0.008)
<i>Paediatrician</i>	-0.560 (0.019)	-0.288 (0.125)	-0.406 (0.097)	-0.319 (0.095)
<i>Dummy Outp. Exp.</i>	0.065 (0.013)	0.144 (0.012)	0.244 (0.006)	0.301 (0.005)
<i>Dummy Hosp. Exp.</i>	0.237 (0.006)	0.262 (0.006)	0.293 (0.006)	0.269 (0.006)
<i>Pop. Size</i>	0.007 (0.002)	0.008 (0.003)	0.008 (0.003)	0.007 (0.003)
<i>Spatial component</i>	0.249 (0.109)	0.431 (0.256)	0.444 (0.254)	0.359 (0.253)
<i>R<sup>2</sup></i>		0.355		
<b>Dependence parameters</b>				
<i>Inv.Mills</i>	-1.181 (0.028)	-1.008 (0.027)		
$\rho$			-0.525 (0.006)	
$\theta$				-2.952 (0.046)

Table 6: The comparison of some estimation strategies for sample-selection model applied to outpatient care expenditures. Sample-selection two-steps estimation strategies considering both fixed and random-effects models are compared with the maximum likelihood and the copula-based ones.

<b>Heckman Models: Selection Equations</b>				
	<i>Two-Step Random-Effects</i>	<i>Two-Step Fixed-Effects</i>	<i>Maximum Likelihood</i>	<i>Frank Copula</i>
<i>Intercept</i>	0.141 (0.274)	1.231 (0.553)	1.002 (0.519)	1.090 (0.533)
<i>Male</i>	-0.175 (0.003)	-0.174 (0.003)	-0.177 (0.003)	-0.176 (0.003)
<i>Age 25 – 44</i>	0.000 (0.005)	0.001 (0.005)	0.006 (0.005)	-0.003 (0.005)
<i>Age 45 – 64</i>	0.171 (0.005)	0.172 (0.005)	0.166 (0.005)	0.165 (0.005)
<i>Age ≥ 65</i>	0.484 (0.005)	0.485 (0.005)	0.469 (0.005)	0.476 (0.005)
<i>Paediatrician</i>	-0.444 (0.013)	-0.039 (0.071)	-0.047 (0.071)	-0.042 (0.071)
<i>Home Assist. Lev. 2</i>	0.297 (0.008)	0.298 (0.008)	0.356 (0.007)	0.357 (0.008)
<i>Home Assist. Lev. 3</i>	0.358 (0.009)	0.359 (0.009)	0.410 (0.008)	0.414 (0.008)
<i>Dummy Drugs Exp.</i>	0.782 (0.003)	0.783 (0.003)	0.778 (0.003)	0.777 (0.003)
<i>Dummy Hosp. Exp.</i>	0.810 (0.006)	0.811 (0.006)	0.797 (0.006)	0.808 (0.006)
<i>Spatial component</i>	-0.566 (0.429)	-2.400 (0.867)	-2.037 (0.814)	-2.167 (0.836)
<i>Pop. Size</i>	0.000 (0.002)	0.000 (0.002)	0.000 (0.002)	0.000 (0.002)
<i>Pseudo R<sup>2</sup></i>		0.158		
<i>Correctly Classified</i>		0.720		
<b>Heckman Models: Expenditures Equations</b>				
<i>Intercept</i>	1.960 (0.469)	2.529 (1.811)	2.689 (1.800)	3.483 (1.775)
<i>Male</i>	-0.042 (0.004)	-0.052 (0.003)	-0.048 (0.003)	-0.038 (0.003)
<i>Age 25 – 44</i>	0.265 (0.005)	0.266 (0.005)	0.265 (0.005)	0.238 (0.005)
<i>Age 45 – 64</i>	0.370 (0.006)	0.383 (0.006)	0.376 (0.005)	0.337 (0.005)
<i>Age ≥ 65</i>	0.578 (0.008)	0.608 (0.008)	0.595 (0.006)	0.520 (0.006)
<i>Paediatrician</i>	-0.323 (0.015)	-0.189 (0.078)	-0.188 (0.079)	-0.155 (0.079)
<i>Dummy Drugs Exp.</i>	0.176 (0.013)	0.230 (0.012)	0.205 (0.005)	0.130 (0.004)
<i>Dummy Hosp. Exp.</i>	0.535 (0.009)	0.570 (0.008)	0.554 (0.005)	0.502 (0.005)
<i>Pop. Size</i>	0.005 (0.002)	0.004 (0.002)	0.004 (0.002)	0.005 (0.002)
<i>Spatial component</i>	0.555 (0.094)	0.388 (0.361)	0.366 (0.358)	0.243 (0.354)
<i>R<sup>2</sup></i>		0.175		
<b>Dependence parameters</b>				
<i>Inv.Mills</i>	-0.592 (0.026)	-0.478 (0.025)		
$\rho$			-0.457 (0.006)	
$\theta$				-4.043 (0.042)

Table 7: The comparison of some estimation strategies for sample-selection model applied to hospital care expenditures. Sample-selection two-steps estimation strategies considering both fixed and random-effects models are compared with the maximum likelihood and the copula-based ones.

<b>Heckman Models: Selection Equations</b>				
	<i>Two-Step Random-Effects</i>	<i>Two-Step Fixed-Effects</i>	<i>Maximum Likelihood</i>	<i>Frank Copula</i>
<i>Intercept</i>	-3.286 (0.154)	-3.433 (0.451)	-3.401 (0.446)	-3.259 (0.365)
<i>Male</i>	-0.008 (0.004)	-0.007 (0.004)	-0.007 (0.004)	0.002 (0.004)
<i>Age 25 – 44</i>	0.241 (0.008)	0.240 (0.008)	0.239 (0.008)	0.229 (0.008)
<i>Age 45 – 64</i>	0.105 (0.008)	0.105 (0.008)	0.107 (0.008)	0.104 (0.008)
<i>Age ≥ 65</i>	0.248 (0.008)	0.246 (0.008)	0.260 (0.008)	0.252 (0.008)
<i>Paediatrician</i>	0.659 (0.012)	1.065 (0.083)	1.071 (0.083)	1.035 (0.082)
<i>Home Assist. Lev. 2</i>	0.329 (0.007)	0.336 (0.007)	0.296 (0.007)	0.330 (0.006)
<i>Home Assist. Lev. 3</i>	0.527 (0.007)	0.539 (0.007)	0.463 (0.007)	0.495 (0.006)
<i>Dummy Drugs Exp.</i>	0.334 (0.005)	0.335 (0.005)	0.337 (0.005)	0.338 (0.005)
<i>Dummy Outp. Exp.</i>	0.699 (0.005)	0.702 (0.005)	0.702 (0.005)	0.703 (0.005)
<i>Spatial component</i>	9.244 (1.549)	10.615 (4.350)	10.309 (4.305)	8.889 (3.510)
<i>Pop. Size</i>	0.002 (0.002)	0.004 (0.003)	0.004 (0.003)	0.004 (0.003)
<i>Pseudo R<sup>2</sup></i>		0.092		
<i>Correctly Classified</i>		0.900		
<b>Heckman Models: Expenditures Equations</b>				
<i>Intercept</i>	7.869 (0.652)	7.025 (4.205)	6.493 (4.003)	1.333 (3.895)
<i>Male</i>	0.070 (0.006)	0.070 (0.006)	0.100 (0.007)	0.067 (0.005)
<i>Age 25 – 44</i>	0.097 (0.014)	0.093 (0.014)	0.005 (0.016)	0.308 (0.013)
<i>Age 45 – 64</i>	0.416 (0.014)	0.410 (0.014)	0.371 (0.015)	0.544 (0.013)
<i>Age ≥ 65</i>	0.477 (0.015)	0.468 (0.015)	0.371 (0.016)	0.885 (0.013)
<i>Paediatrician</i>	-0.949 (0.020)	-1.109 (0.186)	-1.409 (0.141)	-0.216 (0.110)
<i>Dummy Drugs Exp.</i>	-0.093 (0.011)	-0.098 (0.011)	-0.163 (0.010)	0.260 (0.008)
<i>Dummy Outp. Exp.</i>	-0.399 (0.016)	-0.406 (0.016)	-0.544 (0.012)	0.313 (0.010)
<i>Pop. Size</i>	-0.011 (0.002)	-0.002 (0.005)	0.000 (0.005)	-0.001 (0.004)
<i>Spatial component</i>	0.259 (0.080)	0.350 (0.523)	0.495 (0.498)	0.610 (0.484)
<i>R<sup>2</sup></i>		0.226		
<b>Dependence parameters</b>				
<i>Inv.Mills</i>	-1.017 (0.021)	-1.034 (0.021)		
$\rho$			-0.884 (0.002)	
$\theta$				3.598 (0.045)