Società italiana di economia pubblica

XIV CONFERENZA

IL FUTURO DEI SISTEMI DI
WELFARE NAZIONALI TRA INTEGRAZIONE EUROPEA E DECENTRAMENTO REGIONALE
coordinamento, competizione, mobilità
Pavia, Università, 4 - 5 ottobre 2002

PUBLIC VS. PRIVATE HEALTH CARE SERVICES DEMAND IN ITALY. A COUNT DATA ANALYSIS ON SHAW DATA

DANIELE FABBRI and CHIARA MONFARDINI

pubblicazione internet realizzata con contributo della COMPAGNIA di San Paolo

PUBLIC VS. PRIVATE HEALTH CARE SERVICES DEMAND IN ITALY

A COUNT DATA ANALYSIS ON SHAW DATA

Daniele Fabbri

Chiara Monfardini

Dept. of Economics - University of Bologna

Preliminary

Abstract:

Understanding the underlying process of the demand for health services is a key to a better assessment of the forces that increase health care expenditure. The Grossman model and the agency perspective on patient-physician relationship provide different, despite complementary views on this process. In the Grossman tradition, as far as the demand for health care is essentially seen as the result of patients intertemporal utility maximisation, utilisation is the product of individual preferences. In the agency approach, physicians play an active role in assessing the amount of services that patients should consume. Therefore in the analysis of health services consumption the role played by different types of provider can not be ignored. The importance of such an issue has been largely neglected in the literature.

In this paper we make use of the new Survey on Health Aging and Wealth (SHAW) data to analyse health care services utilisation explicitly acknowledging the existence of two different classes of providers: public and private. We consider visits by a specialist physician as the measure of individual health services utilisation. In the time span of the survey (year 2000), individuals can consume this service going public, private or both. In order to investigate on the determinants of these health service utilisation measures we estimate some alternative count data regression models, of which we discuss the relative advantages and disadvantages and the entailed different interpretation of the results.

JEL: C34, C35, C51, D12, I11

This version: July 2002

1 INTRODUCTION

Understanding the underlying process of the demand for health services is a key to a better assessment of the forces that increase health care expenditure. The Grossman model and the agency perspective on patient-physician relationship provide different, despite complementary views on this process. In the Grossman tradition, as far as the demand for health care is essentially seen as the result of patients intertemporal utility maximisation, utilisation is primarily patient determined, though conditioned by the health-care delivery system. In the agency approach, physicians play an active role in assessing the amount of services that patients should consume, up to the point of distorting demand according to their own preferences.

These two perspectives lead to two different streams of econometric modeling traditions: one-step models in the Grossman tradition [see Duan et al. (1983) and Cameron et al. (1988)] and two-step models in the agency tradition [see Manning et al. (1981) and Pohlmeier and Ulrich (1995)]. We review this econometric literature later in the paper. A common feature in these empirical analyses is the assumption of separability of demand functions for different services. Basically the demand for, say, general practitioners visits is assumed to be independent from the demand for specialists physicians. The only exception is Gurmu and Elder (2000).

We focus here on the problem of product differentiation. Existing econometric models perform aggregate demand analysis, i.e. model the overall counts of physician visits or specialists visits consumed by individuals as explained by covariates like income, out-of-pocket payments, coinsurance rates, health conditions. In case patients, within an health-care delivery system, could receive the same service by two different classes of providers, say public vs. private, major problems arise in performing aggregate demand estimation (have to accommodate for joint demand modeling). In the Grossman tradition, given the prominence attributed to patients' preferences, as far as the two classes of providers differ systematically in terms of unobservable characteristics and pricing policies, estimating an aggregate demand model may introduce major distortions in empirical inference. In particular, public providers tipically impose lower out-of-pocket payments but higher waiting times so that inference from aggregate models on behavioural coefficients, like price and income elasticities, may be possibly biased. In the agency perspective, given that the two classes of providers respond to different incentives structures, modeling utilization neglecting the systematic difference among providers may produce inconsistent and uninterpretable results.

In this paper we make use of the new italian Survey on Health Aging and Wealth (SHAW), conducted in the year 2001, data to analyse health care services utilisation explicitly acknowledging the existence of two different classes of providers: public and private. We consider visits by a specialist physician as the measure of individual health services utilisation. In the year before the survey (year 2000), individuals can consume this service going public, private or both. This health service utilisation measure is modelled by some alternative count data regression models, of which we discuss the relative advantages and disadvantages and the entailed different interpretation of the results.

The paper is organized as follows. In the next section we qualitatively review the existing econometric literature on health-care services utilization. In section 3 we present at length our empirical strategy by discussing the negative binomial model, the bivariate negative binomial and the hurdle model we apply to our data. Section 4 describes the data and specification adopted. The major empirical results are reported in section 5. Section 6 concludes the paper with suggestions for future research.

2 MODELS FOR HEALTH SERVICE UTILIZATION

The Grossman model and the physician-patient agency model provide complementary explanations for the demand for health care. We look at them in sequence.

2.1 THE GROSSMAN MODEL

The Grossman model emphasizes the role played by patients' choice looking at health and wealth as two interrelated assets the values of which are optimally controlled over time by the individual. In the case of health, the marginal utility of holding a marginal unit of stock has a consumption and an investment component, which together must always be equal to its marginal user cost. This consists of the interets rate, health capital depreciation and a possible change in the value of the health capital over time.

In this context the demand for health care services is a derived demand, in that services are not consumed per se but serve to maintain or improve upon a certain health status. The typical form of the individual demand function for health care services that emerges from the Grossman model is given by:

$$M(t) = f[H(t), w(t), p_m(t), age(t), E(t), X(t)]$$

The demand for health care services (for simplicity we call them medical services) at time *t*, *M(t)*, is endogenously codetermined[1] with the latent variable "health status", *H(t)*, and it is affected by the wage rate, *w(t)*, a price vector for medical services, $p_m(t)$, individual age, *age(t)*, the level of education, *E(t)*, and a vector of environmental effects, *X(t)*.

An higher wage lowers the marginal incentive to hold health as an asset for consumption use, thus depressing the demand for medical care. By way of contrast it increases the opportunity cost of sick time, hence reinforcing the incentive to hold health as an asset. Assessing the impact of wage on medical service demand is therefore an empirical matter. The impact of prices is negative like that of better education. This last one should lower the demand for investment in health because it contributes to lower health stock depreciation. Demand for medical care should increase with ageing, because it is not optimal to let health stock decline in step with depreciation.

2.2 THE AGENCY APPROACH

In the agency approach, physicians play an active role in assessing the amount of services that patients should consume as far as they typically act a double role: performing checks on the status of patient's health stock and, conditional on checks, supplying treatments aimed at restoring health stock to a desired level. Significant information asymmetry may provide physicians the opportunity to influence demand through their role as health evaluators. This informational advantage is exploited provided that physician's objective function differs from patient's. In this respect it is common to assume that physicians do not only follow Hippocratic oath (for example maximizing individual

[1] The Grossman model is deterministic, so that desidered health stock always equal actual health stock, given constraints. Therefore the demand for health services, which adjust existing health stock net of depreciation, is positively linked, one-to-one, with endogenous health stock.

health), but derive utility also from income and leisure. Therefore, when income or leisure are tailored to specific procedures and/or services, physicians will distort demand to perform more remunerative, or less time consuming, procedures/services, if the marginal benefit of a specific procedure outweighs the associated marginal costs.

In this framework a large body of empirical research is devoted to test the so called supplier inducend demand (SID) hypothesis. The SID hypothesis states that [McGuire and Pauly (1991)] in the face of negative income shocks, physicians may exploit their agency relationship with patients by providing excessive care. Income shocks examined in the literature arise from three different sources. A first source is variation in the physician/population density across areas: increased density lowers the income of existing stock of physicians, and it will lead to increased utilisation of medical procedures in an inducement-type model. Income shocks may also emerge as the consequence of an exogenous change in demand due to epidemiological shifts, evolution of needs, variation in tastes. However the most common source is variation in fees paid to physicians, generally by government payers. The inducement model has traditionally been tested by assessing how these three alternative changes in the environment facing physicians affect the utilisation of medical procedures[2]. Despite each of these testing strategies face important problems they are quite convergent in suggesting that physicians, to some extent, do actually manage demand according to economic incentives.

2.3 THE BASIC FRAMEWORK FOR MODELS OF VISITS' COUNTS

The class of econometric models of health service demand we consider here is that concerned with discrete counts of medical visits. In this case excess zeroes is the most relevant modeling issue[3]. From a purely statistical viewpoint basically the problem consists in building enough flexibility into the econometric model to account for the excess probability mass concentrated in the zero counts. Tackling the problem has major econometric and economic implications. For the sake of expositional clarity we will focus here on the economic issues, introducing econometric issues but leaving details on them to section 4.

In general terms the problem of built-in-flexibility can be addressed in either a single process perspective (determining both null and positive counts), or in a double process perspective (one generating the zeroes vs. the positives and one determining the positives provided that a positive has been already generated). In the context of our problem this amounts to say that in a single process approach all the visits counts, zeroes included, are driven by the same process. On the other hand, when a double process is envisaged contact process (to access to medical treatment or not?) is distinguished from utilization (given that the first answer is YES, how much to consume?).

From an economic viewpoint the double process perspective has a natural appeal in the health economics literature as far as it distinguishes the two-part character of the decisionmaking process in health care demand [Stoddart and Barer (1981)]. While at the first stage it is the patient who decides whether or not she needs medical attention and

---

[2] Representative studies that use physician density changes to proxy for income shocks are Fuchs (1978) and Cromwell and Mitchell (1986). Gruber and Owings (1996) use exogenous demand changes, while Rice (1984) and Yip (1998) examine fee changes.

[3] Similar methodological problems arise while considering continuous demand measures like expenditure [see Newhouse and The Insurance Experiment Study Group (1993)].

therefore to access a physician (contact analysis), in the second stage the health care providers toghether with the patient determine the intensity of the treatment (frequency analysis). This modeling approach has, given certain conditions, a sound structural interpretation [see Santos Silva and Windmeijer (2001)] which motivated its braod adoption in the empirical studies. Moreover it provides a unifying empirical framework for the two abovementioned theories of health care demand. A Grossman-like interpretation might be called for explaining the contact decision, while an agency perspective could be invoked for the interpretation of the frequency decision.

This theoretical partition underpins the choice and interpretation of typical regressors' coefficients introduced in each of the two-part components. Take for instance the paper by Pohlmeier and Ulrich (1995). They estimate two distinct two-part models for general practitioners visits and for specialists visits on a sample of 5.000 employed germans. They control for sex, income, age, education, chronic conditions, physician density in place of residence, plus a set of other covariates. It is interesting here to notice the results on physician density. The two-part model estimates show that physician density does not affect the contact choice while it has a positive impact on the frequency decision. The authors note that "while physician density proxies an availability effect for the patient at the first stage, it captures both demand and supplier response at the second stage. … we are inclined to interpret this finding as some evidence of supplier-induced demand". Similarily also other common covariates to the two parts are given different interpretation in the contanct and frequency analysis.

A common feature of models for visit counts is the lack of control for medical services' prices. Pohlmeier and Ulrich (1995) is not an exception. This is due to unavailability of detailed data on single visits outlays. As far as surveys are designed to gather total number of visits per time period no data are available on each visit payment[4]. Therefore, monetary opportunity costs are typically captured by private insurance status variables [like in Pohlmeier and Ulrich (1995) and Deb and Trivedi (1997)] or, more precisely, by individual coinsurance rate [like in Deb and Trivedi (2002)].[5] The availability of private insurance is found to positively affect contact choice but not frequency choice [Pohlmeier and Ulrich (1995)]. Similar effects are found for copayment rates: higher copayment rates result in a lower probability of contact while frequency is unaffected [Deb and Trivedi (2002)]. These results are coherent with a Grossman interpretation but less so with an agency perspective.

Coming to the results concerning other typical regressors in models for visits' counts we see that some predictions of the Grossman model are frequently contradicted by empirical evidence[6]. In particular good health status is found to be negatively related to the number of visits. This results is coherently consistent across all the papers we reviewed despite differences in econometric specification. Education, typically measured as years of schooling, is usually found to increase visits counts [see Deb and Trivedi (1997, 2002)]. Pohlmeier and Ulrich (1995) show that higher education reduces contact decision for GPs visits while increases it for specialists, in both cases unaffecting

[4] Pohlmeier and Ulrich (1995) argue that the impact of prices may be neglected given that for many western health care systems the direct price of medical services is close to zero.

[5] Introducing insurance status variables raises endogeneity problems. See Cameron et al. (1988), Windemeijer and Santos Silva (1997), Vera Hernandez (1999).

[6] Wagstaff concludes that "the majority of the model's structural parameters are in fact of the 'wrong sign'" [Wagstaff (1986), p. 216].

frequency. Santos Silva and Windmeijer find (2001) that education positively affects contacts and negatively affects frequency for specialists visits. Evidence concerning the impact of income and age tends to be more coherent with the theory.

For the sake of completeness it has to be noticed that, on a purely statistical ground, there is no clear evidence that econometric models based on the two process approach should be preferred to those relying on a single process approach. Actually it has been shown [Deb and Trivedi (1997, 2002)] that sufficiently flexible specification, based on latent class analysis, let single process models better fit the emprical distribution of visits' counts. We will return on this issue later on.

2.4 THE AIMS OF OUR ANALYSIS

In the following we develop a count data analysis of specialist visits in Italy. A remarkable feature of the market for medical professional consultancy in Italy is the presence of two broad distinguishable class of providers: public, highly regulated, specialists, and private, less regulated, ones. We want to account for this peculiarity in our analysis, an issue which has been largely neglected in the literature. Substitution/complementarity relationships between these two classes of providers both arises from the demand and the supply side suggesting that they cannot be separately examined.

The fact private consultancy is typically of higher accuracy, implies lower waiting times at the cost of higher out-of-pocket payment comparing to public ones raises the issue of demand side joint determination of both counts quite obviously.

On the supply side it is pretty relevant to realize that the role of physician incentives affect utilization. Indirect evidence of this is provided by **Table 1**. In countries where general practitioners (GPs) are payed fee-for-service, per-capita consultations are slightly more than in countries where they are payed according to capitation, but are almost double than in countries where GPs are salaried. This provides a strong, additional rationale for our analysis of health service utilization in which we will emphasize the role played by different types of provider.

3 DATA AND INSTITUTIONAL SETTING

Our data source is the new Survey on Health Aging and Wealth (SHAW) collected in 2001. The survey focusses on individuals aged 50 or more. The dataset includes a wide range of microlevel information on socioeconomic characteristics of individuals and households, including specific variables on working and living conditions as well as variables on health condition and health care utilization. We restrict our attention on householder, either male or female. We preferred not to use observations on householders' partners since demand interdependence through family relationships could emerge [see Deb (2001)]. Given the structure of the survey our choice affects the composition of our sample in that we have a larger incidence of male householders comparing to the universe of people aged 50 or more. The total sample consists of 1050 individuals.

We model, as a dependent variable, the number of visits to a specialist physician. These include opticians, dentists and any other physician specialised in a certain field. In performing our analysis of visit counts separately for public and private specialists we had to drop 35 observations with missing values in counts for public specialists visits and 40 in counts for public specialists visits. Joint non missing values are available for 1002 observations. **Table 2** shows the tabulations for the separate counts in our dependent variables. Zero counts are approximately 60% of both distributions; alternatively, partecipation rates are similarly around 40%. 408 and 415 individuals are observed with at least one visit to a public and a private specialist respectively. Private consultations are more frequent on our sample due to larger incidence of higher counts - 3.7 vs. 3 on positive counts -. Contact decision process leads to similar sample means partecipation rates across providers' types, while the second stage process differentiates conditional frequencies of visits across types. This provides a first evidence that the process underlying the contact decision is different from the second stage process.

A first indication of overdispersion in the data is obtained when the sample variance of the dependent count variable is found to be greater than its sample mean. After inclusion of regressors, the Poisson model sample conditional variance will decrease with respect to the sample variance, while the sample average of the conditional mean will be equal to the sample mean if a constant is included among the regressors. Cameron and Trivedi point out that if the sample variance is more than twice the sample mean –this is true in our data for both public and private visits - the data are likely to exhibit overdispersion even after inclusion of regressors, as in cross-section data regressions usually explain less than half of the variation of the dependent variable.

Table 3 contains a cross-tabulation of the two kinds of visits which reflects the view that the two phenomena are jointly determined. It can be noticed that the two count variables display an excess of frequency of the pair (0,0) – about 36% of the total number of the observed pairs of counts- in their joint distribution. We performed the Pearson Chi-square test on the corresponding contingency table, and found strong evidence of dependence between the two count variables.

Explanatory variables are conventional predisposing variables and variables capturing the access to medical services. **Table 5** contains a description of the variables used in this piece of empirical work. We tried to keep our specification as parsimonious as possible, while mimicking similar specification in the literature. In this respect our specification is very close to Deb and Trivedi (1997) and quite similar to Pohlmeier and Ulrich (1995) thus allowing us to make useful comparisons.

It should be noticed that public specialists are payed according to administered prices, while private ones are free to set prices according to competitive pressures coming from close substitutes. This feature would suggest that controlling for out-of-pocket payments would be quite relevant in our case study. SHAW collects informations on total amount paid out-of-pocket for the cumulative count of visits, both specialist and generic, in each type of provider. However no-response rate was quite large (23% for public and 17% for private visits). Moreover averaging outlays across multiple visits could severely distort results. We preferred, at this stage, not to use payments information in the modeling stage.

4 ECONOMETRIC MODELS FOR COUNT DATA

4.1 UNIVARIATE MODELS

We model the demand for physician services by measuring it as counts of utilization, i.e. number of visits, resulting from an underlying discrete probability function. The simplest model for count data is based on the Poisson distribution, which is characterized by a single parameter m. Having available a sample of N independent observations

(y_i, x_i), where y_i denote the count variable of interest and x_i a set of covariates, the Poisson regression model is defined by the conditional density:

$$f^P(y_i \mid x_i; b) = \frac{e^{-m_i} m_i^{\,y_i}}{y_i!} \quad y_i = 0,1,2,..... \tag{1}$$

where: $m_i = \exp(x_i^{'} b)$, $m_i > 0$.

The Poisson distribution implies the property of equidispersion:

$$E(y_i \mid x_i) = V(y_i \mid x_i) = m_i$$

which appears to be very restrictive in most empirical applications, where the conditional variance exceeds the conditional mean. The standard parametric model accounting for overdispersion is based on the Negative Binomial (NB) distribution. This can be derived as a compound Poisson process where the parameter of the Poisson distribution includes a gamma distributed random variable reflecting individual heterogeneity: $y_i \sim Poisson(m_i n_i)$ with $n_i \sim Gamma(a, l)$[7], with $a = l$, and the negative binomial distribution is obtained by integrating over n_i:

$$\begin{aligned} f^{NB}(y_i \mid x_i; a, b) &= \int_0^\infty \frac{e^{-(m_i n_i)}(m_i n_i)^{y_i}}{y_i!} g(n_i) dn_i \\ &= \frac{\Gamma(y_i + a)}{\Gamma(a)\Gamma(y_i + 1)} \left(\frac{a}{m_i + a} \right)^a \left(\frac{m_i}{m_i + a} \right)^{y_i} \end{aligned} \tag{2}$$

where $m_i = \exp(x_i^{'} b)$ as above, and the conditional mean and variance are given by:

$$E(y_i \mid x_i) = m_i$$

$$V(y_i \mid x_i) = m_i + f m_i^2$$

where $f = a^{-1} > 0$ is an overdispersion parameter, making the variance greater than the mean, as observed in many data sets. The parameters (a, b) can be estimated by the maximizing numerically the log-likelihood function corresponding to the density above (estimation is automatically implemented in some statistical packages, like STATA). This is the most common implementation of the Negative Binomial Model, NB2 in the terminology of Cameron and Trivedi (1998). The additional parameter characterizing the NB distribution makes it more flexible than the Poisson, to which it reduces when $f = 0$. In most applications, NB regression models are likely to provide more efficient estimators than those based on Poisson distribution, as failure of the assumption of equidispersion has similar consequences to failure of the homoskedasticity assumption in the linear regression model (Cameron and Trivedi, 1998).

---

[7] The density function for the positive continuous variable n_i is given by: $g(n_i) = \dfrac{n_i^{a-1} l^a}{\Gamma(a)} \exp(-l n_i)$, where $l > 0, a > 0$ and $\Gamma(a) = \int_0^\infty e^{-t} t^{a-1} dt = (a-1)!$, $a > 0$.

An alternative way of dealing with the "excess zeros" displayed by most count variables is represented by the hurdle model. This modification of the basic model was firstly introduced by Mullahy (1986), and thereafter received a great deal of attention in the empirical analysis of the usage of medical services. The hurdle model can be interpreted as a two part model, in which a binary model for the decision of use, determining the probability of crossing a zero threshold, is combined with a truncated count data model on positive counts, explaining the extent of use conditionally to some use. To illustrate the hurdle model, define a dummy variable describing the non use of a doctor in a given period: i.e. $d_i = 1$ if $y_i = 0$. The probability function is then given by:

$$f^H(y_i \mid x_i; J_1, J_2) = f_1(0 \mid x_i; J_1)^{d_i} \left[(1 - f_1(0 \mid x_i; J_1)) f_{trunc}(y_i \mid x_i, y_i > 0; J_2)\right]^{(1-d_i)} \tag{3}$$

where:

$$f_1(0 \mid x_i; J_1) = pr(y_i = 0 \mid x_i; J_1)$$

$$f_{trunc}(y_i \mid x_i, y_i > 0; J_2) = \frac{f_2(y_i \mid x_i; J_2)}{1 - f_2(0 \mid x_i; J_2)}$$

The model specifies a binary probability determining whether the count has a zero realization. If the realization is positive, the hurdle is crossed and the conditional distribution is described by a truncated count model. The two processes can be driven by the same explanatory variables, but the interpretation of parameters will be different depending on the considered stage.

The log-likelihood functions corresponding to (3) factors in two components, which can be separately maximized on the whole sample and on the positive observations respectively:

$$\ln L(J_1, J_2) = \sum_i d_i \ln f_{1i}(J_1) + (1 - d_i)\ln(1 - f_{1i}(J_1)) + \sum_{i \mid d_i = 0} \ln f_{2i}(J_2) - \ln(1 - f_{2i}(0))$$

Estimation of the parameters requires some choice for the two density functions. In our application we use a probit model for the binary outcome, and a truncated negative binomial density for the intensity of use part of the model.

4.2 THE MULTIVARIATE APPROACH

The application of multivariate non-linear non-Gaussian models as those arising when the aim of the analysis is the joint explanation of a given number of count variables is still relatively rare. This is true despite the kind of event counts typically examined in the health economics literature is often represented by different measures of health care utilization like number of doctor consultation, either general practitioner or specialist, non-doctor health professional visits, prescription drug use etc. These measures are likely to be jointly dependent and their interrelation can be described in an analogous way to the seemingly unrelated regression model. The Poisson bivariate model is the most popular model in this context. As illustrated by Cameron and Trivedi (1998), this model can be obtained by the so-called trivariate reduction technique (Kocherlakota and Kocherlakota, 1993), consisting in the convolution of independent random counts with a common component in the sum. The main features of this model are the following. The marginal distributions are both Poisson, and the marginal model, if correctly specified, give consistent but inefficient estimates with respect to joint estimation. The correlation coefficient implied by the joint distribution is individual specific and fully describes the dependence structure of the variables, but it is not very “flexible”, as it is bound to be non-negative. Finally, the model imposes the restriction of equidispersion on each count variables.

Similarly to the univariate framework, the bivariate Negative Model represents an useful tool to handle overdispersed count data. Marshall and Olkin (1990) generate a bivariate negative binomial mixture beginning with two marginal Poisson distributions whose parameters contain a common gamma-distributed heterogeneity term. This approach is also followed by Gurmu and Elder (2000), who generalize the bivariate negative binomial distribution by postulating a first-degree polynomial expansion of the unobserved heterogeneity term, based again on a gamma density. This amounts to the introduction of a further parameter in the joint distribution, and makes the bivariate negative binomial a testable model nested in the generalized one. We present hereafter the bivariate negative binomial model whose estimation results are described in the next section. Using the same notation as in the univariate case, let the two joint count variables be Poisson distributed as follows: $y_{1i} \sim Poisson(m_{1i}n_i)$, $y_{2i} \sim Poisson(m_{2i}n_i)$, with $m_{ji} = \exp(x_{ji}^{'} b_j)$, *j=1,2*, and n_i is a common unobserved heterogeneity term with gamma density $g(n_i)$. Similarly to the univariate case, the joint density is derived by integrating over the heterogeneity term:

$$
\begin{aligned}
f^{BIVNB}(y_{1i}, y_{2i} \mid x_i; a, b_1, b_2) &= \int_0^{\infty} \prod_{j=1}^{2} \left[\frac{e^{-(m_{ji}n_i)} (m_{ji}n_i)^{y_{ji}}}{y_{ji}!} \right] g(n_i) dn_i \\
&= \left[\prod_{j=1}^{2} \left(\frac{m_{ji}^{y_{ji}}}{\Gamma(y_{ji}+1)} \right) \right] \frac{\Gamma(y_{.i}+a)}{\Gamma(a)} a^{-y_{.i}} \left(1+\frac{m_{.i}}{a}\right)^{-(a+y_{.i})}
\end{aligned} \tag{4}
$$

where $m_{.i} = m_{1i} + m_{2i}$, $y_{.i} = y_{1i} + y_{2i}$.

The marginal distributions of this model are still negative binomial, and the correlation between the two count variables (conditional to the covariates) is individual specific, being a function of the m_{ji}, and constrained to be non-negative:

$$
Corr(y_{i1}, y_{i2} \mid x_i) = \frac{m_{1i}m_{2i}}{a} \Big/ \sqrt{(m_{1i} + m_{1i}^2/a)(m_{2i} + m_{2i}^2/a)} \tag{5}
$$

5 RESULTS

5.1 THE NEGATIVE BINOMIAL ESTIMATES

We start our empirical analysis by estimating two univariate NB models on the number of specialist public and private consultations respectively. This approach ignores the joint nature of the two health care demand determination processes, and takes into account the excess-zeros pattern by specifying a more general statistical distribution than the Poisson. The Maximum Likelihood estimation results[8] reported in **Table 7** reveal that

[8] The estimation has been obtained using STATA 7.

the Poisson distribution is indeed rejected by the data, as the “nesting” parameter f is found to be significantly different from zero. This confirms the stylized facts on overdispersion of the data emerged by the descriptive analysis.

The main findings concerning the role of the inserted explanatory variables are the following. Family income appears to be an important determinant of the number of private consultations, with higher income families tending to increase their utilization of private health care. Also, the level of schooling has not a significant impact on private services demand. On the contrary, the demand of public specialist visits is not affected by the family income variable, while it positively reacts to an increase in the years of schooling. The education effect result agrees with the conventional reason that education makes individuals more informed consumers of medical care services, and signals that more educated people are oriented towards a more frequent use of the services offered inside the public sector.

The possession of a private health insurance increases the consultation of private specialist. This is a common result in the applied literature which is coherent with fours stories. The first one relates to price elasticites (being double insured allows to access private health care at lower out-of-pocket payments). According to the second explanation, this could also be the effect of an adverse selection process making the frequent health services users to look for supplementary coverage and cost reimbursment.[9] A third key of interpretation is represented by moral hazard where incentives by the patient and the physicians for over-treatement aligne against the insurer. The last possible explanation has to do with supplier induced demand in a wide sense Pohlmeier and Ulrich find no evidence of such behaviour as the private insurance dummy is only significant in the firs stage – i.e. contact decision- of their hurdle model.

Turning to the demographic variables, we find that individual’s age play no role in both equations. The effect of this variable is usually found to be negative until some age (which varies from 33 to 52 in different studies), and increasing thereafter. We observe coherent coefficient signs, but these parameters are not enough precisely estimated. Women appear to seek more medical care than men, as usually evidenced in empirical studies. In our context, this is true both for private and public specialist consultations.

The health status measures display the usual empirical link with the degree of utilization of medical care. This increases when chronic conditions or physical limitations are present, the level of self-perceived health is poor and in presence of eyesight troubles (for private visits), and decreases with excellent self-assessed health (public visits). Individuals who never smoked seek less both public and private medical consultations. Customary consumers of super-alchoolic drinks use more public specialist services and less private doctor visits.

Regional-specific unobservable factors make the demand for public doctor consultation in central and southern Italy lower than in northern Italy. The effect of the size of the community of residence, aimed at proxying the opportunity costs of visiting a physician, turns out not to be significant. Finally, the variables which proxy the accessibility to the two kind of medical services show the expected sign, with the ratio of physicians per bed in private providers exhibiting a negative effect on the number of visits

[9] Following this interpretation, a problem of endogeneity of the private insurance variable can be envisaged.

demanded from public physicians and the amount of per-capita public expenditure increasing the number of public specialist consultations.

5.2 THE BIVARIATE NEGATIVE BINOMIAL ESTIMATES

Table 8 displays the Maximum Likelihood Estimation[10] results we get when the number of public and private visits is allowed to be generated by a joint process represented by the bivariate negative distribution. This modelling framework acknowledges both features emerged from the descriptive analysis: overdispersion and dependency of the two health utilization variables. Taking into account their joint determination will provide more efficient estimation of the parameters. Some relevant differences with respect to the univariate estimation results are observed in the magnitude and significance of the estimated coefficients (notice, in particular, that the private insurance indicator looses its explanatory power in the private equation), while the direction of the analysed effects keeps generally the same.

The most interesting result is represented by the conditional correlation value, which is obtained as the after-estimation sample average of the correlation coefficient in (5). This average measure of correlation is quite high (although it should be accompanied by a precision of estimation measure). Gurmu and Elder (2000) find a value of 0.3 in their application to counts of doctor and non-doctor consultations, that they claim to be strongly dependent also according to some independency tests proposed by Cameron and Trivedi (1998). Two comments on this observed correlation measure are due. First, the cross tabulation in **Table 3** shows that the two count variables exhibit a pattern that we can call "joint overdispersion", meaning that they not only exhibit an excess of zero values in their marginal distributions, but also an excess of frequency of the pair (0,0) – about 36% of the total number of the observed pairs of counts- in their joint bivariate density function. If we adopt the two-part interpretation of the univariate hurdle model and generalize it to a bivariate setting, this group represents the non-users of medical services, deciding not to contact any kind of physicians. The composition of this group is likely to distort the correlation coefficient. If, for example, this group is mainly generated by the population of the healthy people, this will induce a higher correlation coefficient for the two counts.

A second point concerns the possibility of a particular interpretation of the estimated correlation, in case the conditional (to the observable regressors) correlation can be interpreted as the correlation between the unobservable part of the non linear regression model. In our example this is mainly represented by out-of-pocket payments for the private doctor visits. These will enter both equation of the model, with negative sign in the private equation. Finding a positive correlation after estimation could then be interpreted as evidence that private services prices affect negatively also the demand for public ones, showing that the two goods as complements. This interpretation seems not appropriate in our present study, in the light of what appears a major limitation of the NB bivariate model: the correlation is constrained to be non-negative.

Nevertheless, the above considerations push our future analysis in the direction of a bivariate hurdle model[11], in which the first part is aimed at explaining the zero-pairwise observations (no-contact with any medical services providers), while the second is conditional to some contact (either with public or private provider). Generalizing the

[10] We made use of the GAUSS routines kindly made available by Gurmu and Elder (2000).

[11] To our knowledge, the only application of bivariate hurdle models is given by Hellstrom on number of leisure trips and total number of overnight stays on Swedish tourism data.

bivariate truncated distribution to a Poisson-lognormal (or NB-lognormal), model as proposed by Winkelmann (2001) in the univariate case, would allow a flexible correlation pattern between the two count variables. This would relax both the non-negativity constraint of the conditional correlation coefficient and one of the basic constraining feature of the conventional hurdle model, i.e. the independence between the hurdle step and the truncated distribution.

5.3 THE HURDLE MODEL ESTIMATES

In the present analysis we limit our attention to the univariate hurdle approach, as the development of the bivariate model requires a deeper methodological investigation. The single equation modelling exercise neglects the joint mechanism determining the demand of the considered services, but can nevertheless shed some light on the opportunity of separately modelling the two subsequent stages corresponding to contact and frequency decisions. The Maximum Likelihood estimation results of the two parts of the model (probit at the first stage, truncated negative binomial at the second one) are contained in **Tables 9** and **10**[12].

A first look at both tables reveals that the first stage model exhibits a better fit than the second stage one. As Pohlmeier and Ulrich point out, household data are better suited to quantify the determinants of the contact decision, while the frequency of use also depends on supply side factors on which observable information is limited. Also, the number of observations is considerably reduced in the second part of the model. Despite this, there is a number of relevant comments concerning differences between the parameters across the two stages and, more interestingly, with the univariate NB model, which does not distinguish between the two parts.

The variables included as regressors exert on the modelled probability of non-contacting a public/private specialist a similar effect to what was found in the single equation NB model. To higher family income corresponds higher probability of contacting a private specialist. The income variable is now significant also in determining a less probable contact with a public specialist. Consistently with our previous findings, more educated individual tend to have higher probability of contacting a public physician. It has to be noticed that this set of variables turns out not to be relevant in the second stage model. Pohlmeier and Ulrich find the same result on both the counts of general practitioner and specialist visits. This means that once the kind of provider is chosen, income and education do not affect the frequency behaviour.

The female dummy, health status variables, the regional dummies and the number of physicians per bed in private hospitals have the same sign effect in both parts of the model, and this is still consistent with the interpretation we put forward for the univariate NB model. But the hurdle model allows to disentangle their coefficients on the contact decision and the number of visits respectively. These parameters are mostly significant at both stages and have different magnitudes. Public per-capita expenditure only affects positively the decision to contact of a public specialist, but not the number of referrals. The second measure of accessibility, represented by the number of doctors per bed in public hospitals is now significant in the second part of the model and negatively related to the number of visits provided by private specialists..Finally, an interesting remark has to do with the role of the possession of a private health insurance. This has no importance

[12] In order to implement estimation with the truncated negative binomial distribution we resorted to the STATA ado file provided by Hilbe (1999) on the Stata Technical Bulletin.

in the contact of either kind of specialists, but is positively affecting the frequency of both private and public specialist visits. This last evidence is plausably due to an adverse selection effect, with the frequent users being doubly insured.

6 CONCLUSIONS

In the present paper we develop a count data analysis of specialist visits in Italy. A remarkable feature of the market for medical professional consultancy in Italy is the presence of two broad distinguishable class of providers: public, highly regulated, specialists, and private, less regulated, ones. We want to account for this peculiarity in our analysis, an issue which has been largely neglected in the literature.

Existing econometric models perform aggregate demand analysis, i.e. model the overall counts of physician visits or specialists visits consumed by individuals as explained by covariates like income, out-of-pocket payments, coinsurance rates, health conditions. In case patients, within an health-care delivery system, could receive the same service by two different classes of providers, say public vs. private, major problems arise in performing aggregate demand estimation.

In this paper we make use of the new italian Survey on Health Aging and Wealth (SHAW), conducted in the year 2001, data to analyse health care services utilisation explicitly acknowledging the existence of two different classes of providers: public and private. We consider visits by a specialist physician as the measure of individual health services utilisation. In the year before the survey (year 2000), individuals can consume this service going public, private or both. This health service utilisation measure is modelled by some alternative count data regression models.

From the univariate Negative Binomial model estimates we derived empirical evidence coherent with common findings in this stream of literature. Moreover we received a strong confirmation of the importance of modelling the two counts as driven by different, despite non necessarily, indenpendent processes. This conclusion is further supported by the results from the bivariate Negative Binomial estimate. The hurdle model indicates the importance of a further dimension, arising by separate consideration of the contact and frequency decision processes.

Therefore our first explorative analysis, despite not conclusive, points out the major features of a devisable model for our peculiar case study. Accordingly our future research should move towards a bivariate hurdle model, in which the first part is aimed at explaining the zero-pairwise observations, while the second is conditional to some contact. A requirement of this model is a flexible conditional correlation, between the two count variables, allowing for possibly negative values.

An alternative approach[13] is a bivariate count model with Latent Classes. In a single equation framwork Deb and Trivedi (2002) suggests that two-part models are dominated by Latent Class Models (LCM). The variation in demand for health care is explained

[13] The major limitation in the literature based on hurdle model is due to the possibly misconceived assumption that zeroes reflect the choice of not contacting a physician. Actually both zeroes and positives might be the product of two related process: emergence of need and service utilization. Individuals may decide not to contact a physician either because they don't need it or because they prefer not to do it even if they need. Similarly we may observe low levels of utilization either because of a low level of need or because of a low preference for treatment.

relatively more by individual intrinsic characteristics than by physician or supply factors. Despite the Principal-Agent framework seems relevant in this context, however, from a statistical point of view, LCM models provide better performance than classical TPM as far as "it is better to permits mixing with respect to both zeros *and* positives" (different processes describing for example healtyh/infrequent users and ills/frequent users can generate both zeros and positives counts)[14]. A careful comparison between bivariate hurdle model and a bivariate count model with latent classes should be developed on both statistical and economic interpretation grounds.

[14] This is coherent with analogous results in hospital profiling literature [see Silber, Rosenbaum and Ross (1995)] where it is shown that individual predictors explain more than 80% of the variation in individual medical outcome.

7 REFERENCES

Cameron, A.C., P.K. Trivedi, 1998, "Regression Analysis of Count Data", Cambridge University Press, Cambridge.

Cameron, A.C., P.K. Trivedi, F. Milne and J. Piggott, 1988, "A Microeconometric Model of the Demand for Health Care and Health Insurance in Australia", Review of Economic Studies; 55(1), pp. 85-106.

Deb, P. and A.M. Holmes, 1998, "Substitution of Physicians and Other Providers in Outpatient Mental Health Care", Health Economics; 7(4), pp. 347-61.

Deb, P. and P. Trivedi, 1997, "Demand for medical care by the elderly: a finite mixture approach", Journal of Applied Econometrics, 12, pp. 313-336.

Dustmann, C. and F. Windmeijer, 1999, "Wages and the Demand for Health-A Life Cycle Analysis", Institute for Fiscal Studies Working Paper: W99/20.

Grossman, M., 1982, "The demand for health after a decade", Journal of Health Economics, 1(1), pp. 1-3.

Gurmu, S. and J. Elder, 2000, "Generalized Bivariate Count Data Regression Models", Economics Letters; 68(1), pp. 31-36.

Gurmu, S., 1997, "Semi-Parametric Estimation of Hurdle Regression Models with an Application to Medical Utilization", Journal of Applied Econometrics, 12, pp. 225-242.

Hellstrom, J., 2001(?), "An Endogenously Stratified Bivariate Count Data Model for Household Tourism Demand", mimeo Dept. of Economics, Umea University.

Hilbe, J (1999), "Zero-truncated Poisson and negative binomial regression", Stata Technical Bulletin 1/99, pp.37-40; Reprints Vol 8, pp.233-236.

Karatzas, G., 1992, "On the Effect of Income and Relative Price on the Demand for Health Care--The EEC Evidence: Comment", Applied Economics; 24(11), pp. 1251-53.

Kocherlakota, S. and K. Kocherlakota, 1993, Bivariate Discrete Distributions, New York, Marcel Dekker.

Lopez N.A., 1998, "Unobserved Heterogeneity and Censoring in the Demand for Health Care", Health Economics; 7(5), pp. 429-37.

Manning et al., 1981, "A Two-Part Model of the Demand for Medical Care: Preliminary Study from the Health Insurance Study", in Sheffler R.M. and L.F. Rossiter, eds., Health, Economics, and Health Economics, Amsterdam, North-Holland, pp. 103-123.

Marshall, A.W. and I. Olkin, 1990, "Multivariate Distributions Generated from Mixtures of Convolution and Product Families" in H.W. Block, A.R. Sampson and T.H. Davids eds., Topics in Statistical Dependence, IMS Lecture Notes-Monograph Series, volume 16, 371-393.

Mullahy, J., 1986, "Specification and testing of some modified count data models", Journal of Econometrics, 33 (3), pp. 341-365.

Mullahy, J., 1998, "Much Ado About Two: Reconsidering Retransformation and the Two-Part Model in Health Econometrics", Journal of Health Economics, 17(3), pp. 247-281.

Muurinen, J. M., 1982, "Demand for health - A generalised Grossman model", Journal of Health Economics, 1(1), pp. 5-28.

Newhouse and The Insurance Experiment Study Group, 1993, Free for All? Lessons from the RAND Health Insurance Expenditure, Cambridge, Harvard University Press.

Pohlmeier, W. and V. Ulrich, 1995, "An Econometric Model of the Two-Part Decisionmaking Process in the Demand for Health Care", Journal of Human Resources; 30(2), pp. 339-61.

Propper, C., 2000, "The Demand for Private Health Care in the UK", Journal of Health Economics; 19(6), pp. 855-76.

Santos Silva, J.M.C. and F. Windmeijer, 2001, "Two-Part Multiple Spell Models for Health Care Demand", Journal of Econometrics; 104(1), pp. 67-89.

Schellhorn, M., 2001, "The Effect of Variable Health Insurance Deductibles on the Demand for Physician Visits", Health Economics; 10(5), pp. 441-56.

Schellhorn, M., et al.., 2000, "Health Services Utilization of Elderly Swiss: Evidence from Panel Data, Health Economics"; 9(6), pp. 533-45.

Silber, J.H., P.R. Rosenbaum and R.N. Ross, 1995, "Comparing the Contributions of Groups of Predictors: Which Outcomes Vary with Hospital rather than Patients Characteristics?", Journal of the American Statistical Association, 90(429), pp. 7-18.

Stoddart, G.L. and M.L. Barer, 1981, "Analyses of the Demand and Utilization Through Episodes of Medical Care", in van der Gaag, J. and M. Perelman, eds., Health, Economics and Health Economics, Amsterdam, North-Holland.

Vera Hernandez, A.M., 1999, "Duplicate Coverage and Demand for Health Care. The Case of Catalonia", Health Economics; 8(7), pp. 579-98.

Wagstaff, A., 1986, "The Demand for Health - Some New Empirical Evidence", Journal of Health Economics, 5(3), pp. 195-233.

Windmeijer, F. and J.M.C. Santos Silva, 1997, "Endogeneity in Count Data Models: An Application to Demand for Health Care", Journal of Applied Econometrics; 12(3), pp. 281-94.

Winkelmann, R. (2001), "Health Care Reform and the Number of Doctor Visits – An Econometric Analysis", CEPR Discussion Paper Series, No. 3021.

Zweifel, P. and F. Breyer, 1997, Health Economics, Oxford, Oxford University Press.

Table 1: Per-capita general practitioners' consultations across some European countries

	1990-1997	Subgroup mean
	Fee-for-service	
Belgium	7.9	
France	6.2	
Germany	6.0	6.7
	Capitation	
Italy	6.7	
Netherlands	5.7	
United Kingdom	5.8	6.1
	Salary	
Finland	4.0	
Iceland	4.9	
Norway	3.8	
Portugal	3.2	
Sweden	2.9	3.8

Source: Our elaboration on OECD Health Data '99, OECD, Paris, 1999

Table 2: Tabulations of specialists visits in our sample

		PUBLIC			PRIVATE		
	Count	Freq.	Percent	Cum.	Freq.	Percent	Cum.
	0	607	59.8	59.8	595	58.9	58.9
	1	138	13.6	73.4	109	10.8	69.7
	2	105	10.3	83.7	95	9.4	79.1
	3	53	5.2	89.0	54	5.4	84.5
	4	35	3.5	92.4	41	4.1	88.5
	5	23	2.3	94.7	39	3.9	92.4
	6	19	1.9	96.6	19	1.9	94.3
	7	8	0.8	97.3	13	1.3	95.5
	8	5	0.5	97.8	9	0.9	96.4
	9	1	0.1	97.9	2	0.2	96.6
	10	8	0.8	98.7	13	1.3	97.9
	11	3	0.3	99.0	5	0.5	98.4
	12	3	0.3	99.3	3	0.3	98.7
	13	3	0.3	99.6	1	0.1	98.8
	14	1	0.1	99.7	2	0.2	99.0
	15	3	0.3	100.0	3	0.3	99.3
	16				1	0.1	99.4
	17				2	0.2	99.6
	18						
	19						
	20				4	0.4	100.0
Total		1015			1010		
Positives		408			415		

	Mean	Variance	St. dev.	Mean	Variance	St. dev.
Full Sample	1.210	5.058	2.249	1.537	8.205	2.864
Positive counts	3.010	7.167	2.677	3.740	11.730	3.425
Partecipation rate	0.402	0.240	0.490	0.411	0.242	0.492

Table 3: Cross-tabulation of PUBLIC vs PRIVATE specialists visits in our sample

		PUBLIC VISITS												
	COUNT	0	1	2	3	4	5	6	7	8	9	10	+10	tot
PRIVATE VISITS	0	358	87	61	28	18	15	6	3	3	1	5	7	592
	1	62	22	11	6	4	0	2	1	0	0	0	0	108
	2	67	7	7	5	5	0	1	1	1	0	1	0	95
	3	29	8	7	5	3	1	0	0	0	0	0	0	53
	4	25	4	6	1	3	1	0	0	0	0	0	1	41
	5	25	5	3	0	0	2	4	0	0	0	0	0	39
	6	8	1	2	1	1	1	4	0	0	0	0	0	18
	7	6	0	3	1	0	1	1	0	0	0	0	1	13
	8	6	0	0	1	0	0	0	1	0	0	0	1	9
	9	1	0	0	1	0	0	0	0	0	0	0	0	2
	10	5	1	3	1	0	1	0	1	0	0	1	0	13
	+10	12	2	1	1	0	0	0	0	1	0	0	2	19
	tot	604	137	104	51	34	22	18	7	5	1	7	12	1002

Table 4: Sample moments of joint PUBLIC-PRIVATE specialist visits distribution

	Partecipation rate	Mean number of visits
Public+Private conditional on joint positives	64.3%	4.171
Public conditional on zero private	39.5%	2.850
Private conditional on zero public	40.7%	3.528
Public conditional on positive private	40.0%	1.239
Private conditional on positive public	41.2%	1.616
Public+Private conditional on positive private		4.924
Public+Private conditional on positive public		4.568

Table 5: Description of variables

Variable	Description
Dependent	
Public specialist visits	Number of visits to a public specialist in the year before survey (2000)
Private specialist visits	Number of visits to a private specialist in the year before survey (2000)
Explanatory	
Family income	Monthly family income, net of income taxes and social insurance rates
Education	Number of year of education
Unemployed	=1 if the person is unemployed
Female	=1 if the person is female
Single	=1 if the person is unmarried or widow
Age	Age in years
Chronic conditions	=1 if the person suffers from chronic conditions
Physical limitations	=1 if the person has a condition that limits activities of daily life
Poor self-perceived health	=1 if self-perceived health is poor
Excellent self-perceived health	=1 if self-perceived health is excellent
Hearing troubles	=1 if the person suffers from hearing troubles
Eyesight troubles	=1 if the person suffers from eye troubles
Never smoked	=1 if the person never smoked in his life
Alchool consumption	=1 if the person consumes alchool regularily
Private health insurance	=1 if the person is covered by private health insurance
Central region	=1 if the person lives in central regions
Southern region	=1 if the person lives in southern regions
Public exp. per-capita	Public expenditure per capita in the residing Local Health Authority
Availability of private hospitals	=1 if private hospitals are present in the residing Local Health Authority area
Physicians per bed in private	Ratio of physician per bed in private hospitals operating in the residing Local Health Authority area
Physicians per bed in public	Ratio of physician per bed in public hospitals operating in the residing Local Health Authority area
Population	Total population in place of residence (in thousands of inhabitants)

Table 6: Descriptive statistics for the regressors

Variable	FULL SAMPLE				CONDITIONAL ON POSITIVE PUBLIC COUNT				CONDITIONAL ON POSITIVE PRIVATE COUNT			
	Mean	St. Dev	Min	Max	Mean	St. Dev.	Min	Max	Mean	St. Dev.	Min	Max
Family income	3.145	2.341	0.30	25.00	3.040	2.616	0.30	25.00	3.503	2.609	0.30	25.00
Family income_sq	15.366	38.236	0.09	625.00	16.065	52.845	0.09	625.00	19.062	43.390	0.09	625.00
Education	7.748	4.737	-	21.00	7.142	4.323	-	21.00	8.566	5.037	-	21.00
Education_sq	82.440	89.295	-	441.00	69.657	75.677	-	441.00	98.687	101.248	-	441.00
Unemployed	0.714	0.452	-	1.00	0.792	0.407	-	1.00	0.672	0.470	-	1.00
Female	0.462	0.499	-	1.00	0.549	0.498	-	1.00	0.489	0.500	-	1.00
Single	0.297	0.457	-	1.00	0.297	0.457	-	1.00	0.292	0.455	-	1.00
Age	63.7	9.4	50.0	91.0	65.1	9.5	50.0	91.0	62.8	9.1	50.0	91.0
Age_sq	4145.1	1233.8	2500.0	8281.0	4323.5	1260.8	2500.0	8281.0	4030.7	1192.7	2500.0	8281.0
Chronic conditions	0.345	0.476	-	1.00	0.493	0.501	-	1.00	0.390	0.488	-	1.00
Physical limitations	0.193	0.395	-	1.00	0.306	0.462	-	1.00	0.214	0.411	-	1.00
Poor self-perceived health	0.134	0.341	-	1.00	0.208	0.407	-	1.00	0.164	0.371	-	1.00
Excellent self-perceived health	0.574	0.495	-	1.00	0.441	0.497	-	1.00	0.559	0.497	-	1.00
Hearing troubles	0.062	0.241	-	1.00	0.074	0.261	-	1.00	0.063	0.243	-	1.00
Eyesight troubles	0.121	0.326	-	1.00	0.162	0.369	-	1.00	0.149	0.357	-	1.00
Never smoked	0.547	0.498	-	1.00	0.559	0.497	-	1.00	0.516	0.500	-	1.00
Alchool consumption	0.018	0.133	-	1.00	0.017	0.130	-	1.00	0.012	0.109	-	1.00
Private health insurance	0.059	0.236	-	1.00	0.049	0.216	-	1.00	0.075	0.263	-	1.00
Central region	0.199	0.399	-	1.00	0.191	0.394	-	1.00	0.178	0.383	-	1.00
Southern region	0.360	0.480	-	1.00	0.319	0.467	-	1.00	0.359	0.480	-	1.00
Public exp. per-capita	1.935	0.415	0.92	3.38	1.995	0.425	0.92	3.38	1.924	0.403	0.92	3.38
Availability of private hospitals	0.828	0.377	-	1.00	0.838	0.369	-	1.00	0.819	0.385	-	1.00
Physicians per bed in private	0.208	0.132	-	0.49	0.202	0.130	-	0.49	0.208	0.135	-	0.49
Physicians per bed in public	0.427	0.103	0.18	0.65	0.424	0.103	0.18	0.65	0.426	0.101	0.18	0.65
Population	251	602	0.337	2653	270	613	0.337	2653	222	546	0.337	2653
Population_sq	425	1503	0.000	7040	448	1501	0.000	7040	347	1305	0.000	7040

Table 7: Estimates of the negative binomial model

	PUBLIC				PRIVATE			
	Coef.	Std. Err.	z		Coef.	Std. Err.	z	
Family income	-0.0150	0.0604	-0.250		0.1325	0.0534	2.480	**
Family income_sq	0.0034	0.0027	1.260		-0.0059	0.0025	-2.400	**
Education	0.0884	0.0422	2.100	**	0.0536	0.0429	1.250	
Education_sq	-0.0047	0.0024	-1.960	**	0.0009	0.0021	0.400	
Unemployed	0.1340	0.1571	0.850		-0.0261	0.1484	-0.180	
Female	0.5016	0.1323	3.790	***	0.3799	0.1253	3.030	***
Single	-0.3127	0.1469	-2.130	**	0.0819	0.1358	0.600	
Age	-0.0618	0.0841	-0.740		0.1008	0.0817	1.230	
Age_sq	0.0005	0.0006	0.790		-0.0008	0.0006	-1.360	
Chronic conditions	0.5243	0.1195	4.390	***	0.2908	0.1420	2.050	**
Physical limitations	0.5555	0.1389	4.000	***	-0.1342	0.1662	-0.810	
Poor self-perceived health	0.1887	0.1679	1.120		0.6551	0.2016	3.250	***
Excellent self-perceived health	-0.4988	0.1320	-3.780	***	0.0508	0.1425	0.360	
Hearing troubles	0.2119	0.2062	1.030		0.0568	0.2252	0.250	
Eyesight troubles	0.1490	0.1558	0.960		0.5495	0.1884	2.920	***
Never smoked	-0.2181	0.1294	-1.690	*	-0.3677	0.1186	-3.100	***
Alchool consumption	0.7600	0.5200	1.460		-0.7896	0.4621	-1.710	*
Private health insurance	0.2467	0.2547	0.970		0.4802	0.2297	2.090	**
Central region	-0.5653	0.1801	-3.140	***	-0.0870	0.1819	-0.480	
Southern region	-0.2271	0.1267	-1.790	*	0.0250	0.1370	0.180	
Public expenditure per-capita	0.4402	0.1332	3.310	***	0.0058	0.1436	0.040	
Availability of private hospitals	0.8841	0.2148	4.120	***	0.0047	0.2168	0.020	
Physicians per bed in private	-2.6632	0.6084	-4.380	***	-0.1227	0.6239	-0.200	
Physicians per bed in public	1.0603	0.6668	1.590		-0.7492	0.6420	-1.170	
Population/100	0.0038	0.0346	0.110		-0.0476	0.0370	-1.290	
Population/100_sq	0.4960	1.3950	0.360		0.9880	1.4900	0.660	
Constant	-0.0385	2.8387	-0.010		-3.3173	2.6925	-1.230	
Ln(alpha)	0.5761	0.0969			0.9911	0.0773		
Alpha	1.7791	0.1723			2.6941	0.2084		
Number of observations	1015				1010			
Wald $chi^2(26)$	231.84				140.84			
Prob > chi^2	0.0000				0.0000			
Log likelihood	-1373.51				-1544.72			
Pseudo R^2	0.0626				0.0293			

Table 8: Estimates of the bivariate negative binomial model

	PUBLIC			PRIVATE		
	Coef.	St. Err.	t-stat	Coef.	St. Err.	t-stat
Family income	-0.037	0.083	-0.444	0.142	0.078	1.816
Family income_sq	0.005	0.004	1.197	-0.006	0.005	-1.211
Education	0.124	0.058	2.131	0.037	0.076	0.482
Education_sq	-0.006	0.003	-1.855	0.125	0.407	0.307
Unemployed	0.201	0.245	0.822	-0.026	0.090	-0.291
Female	0.343	0.144	2.378	0.381	0.129	2.944
Single	-0.06	0.299	-0.199	0.214	0.228	0.939
Age	-0.105	0.100	-1.055	0.100	0.101	0.992
Age_sq	0.001	0.001	1.066	-0.001	0.001	-1.086
Chronic conditions	0.547	0.157	3.497	0.268	0.145	1.849
Physical limitations	0.469	0.145	3.231	-0.185	0.205	-0.903
Poor self-perceived health	0.203	0.269	0.756	0.608	0.280	2.173
Excellent self-perceived health	-0.657	0.316	-2.077	0.031	0.313	0.100
Hearing troubles	0.085	0.196	0.432	0.083	0.253	0.329
Eyesight troubles	0.146	0.183	0.800	0.568	0.185	3.076
Never smoked	-0.287	0.192	-1.490	-0.360	0.138	-2.617
Alchool consumption	0.922	0.524	1.761	-0.527	0.499	-1.057
Private health insurance	0.025	0.123	0.204	0.355	0.224	1.581
Central region	-0.673	0.261	-2.575	-0.052	0.190	-0.275
Southern region	-0.286	0.133	-2.155	0.071	0.119	0.596
Public expenditure per-capita	0.411	0.171	2.405	0.025	0.096	0.256
Availability of private hospitals	1.134	0.228	4.977	0.009	0.109	0.078
Physicians per bed in private	-3.232	0.708	-4.566	-0.091	0.166	-0.546
Physicians per bed in public	0.516	0.854	0.604	-0.832	0.727	-1.144
Population/100	0.019	0.081	0.238	-0.024	0.014	-1.791
Population/100_sq	-0.037	0.295	-0.126	0.020	0.064	0.316
Constant	1.745	3.511	0.497	-3.262	3.150	-1.035
Ln(alfa)	-0.197	0.076	-2.603			
Conditional mean		1.226			1.531	
Conditional variance		4.997			5.666	
Conditional correlation			0.544			
Number of observations			1002			
Log likelihood			-3166.76			

Table 9: Estimates of the double hurdle model: first stage

	PUBLIC				PRIVATE			
	Coef.	Std. Err.	z		Coef.	Std. Err.	z	
Family income	0.0670	0.0441	1.520		-0.1243	0.0424	-2.930	***
Family income_sq	-0.0071	0.0023	-3.120	***	0.0045	0.0022	2.030	**
Education	-0.0829	0.0329	-2.520	**	-0.0120	0.0305	-0.390	
Education_sq	0.0052	0.0017	2.980	***	-0.0010	0.0016	-0.650	
Unemployed	-0.1477	0.1232	-1.200		0.1235	0.1193	1.040	
Female	-0.3727	0.1019	-3.660	***	-0.2168	0.0997	-2.170	**
Single	0.3205	0.1100	2.910	***	-0.0698	0.1048	-0.670	
Age	0.0642	0.0647	0.990		-0.0372	0.0636	-0.580	
Age_sq	-0.0006	0.0005	-1.150		0.0003	0.0005	0.650	
Chronic conditions	-0.4154	0.1071	-3.880	***	-0.2338	0.1036	-2.260	**
Physical limitations	-0.3935	0.1360	-2.890	***	0.0530	0.1335	0.400	
Poor self-perceived health	-0.1047	0.1630	-0.640		-0.3274	0.1596	-2.050	**
Excellent self-perceived health	0.2942	0.1091	2.700	***	-0.0151	0.1077	-0.140	
Hearing troubles	-0.0382	0.1880	-0.200		-0.0898	0.1768	-0.510	
Eyesight troubles	0.0459	0.1445	0.320		-0.2722	0.1425	-1.910	*
Never smoked	0.1375	0.0946	1.450		0.1643	0.0923	1.780	*
Alchool consumption	-0.0861	0.3088	-0.280		0.3954	0.3514	1.130	
Private health insurance	0.0826	0.1893	0.440		-0.1075	0.1791	-0.600	
Central region	0.3191	0.1346	2.370	**	0.1486	0.1311	1.130	
Southern region	0.2227	0.1074	2.070	**	0.0927	0.1051	0.880	
Public expenditure per-capita	-0.4098	0.1192	-3.440	***	0.1260	0.1108	1.140	
Availability of private hospitals	-0.5967	0.1740	-3.430	***	0.1790	0.1624	1.100	
Physicians per bed in private	2.1196	0.5167	4.100	***	-0.0890	0.4698	-0.190	
Physicians per bed in public	-0.1779	0.4999	-0.360		-0.4292	0.4890	-0.880	
Population/100	-0.0498	0.0284	-1.750	*	0.0286	0.0281	1.020	
Population/100_sq	1.4000	1.1230	1.250		-0.3770	1.1020	-0.340	
Constant	-0.3882	2.1361	-0.180		1.6116	2.0960	0.770	
Number of obs	1015				1010			
Wald $chi^2(26)$	159.67				74.8			
Prob > chi^2	0.000				0.000			
Log likelihood	-592.309				-646.71			
Pseudo R^2	0.1339				0.0545			

Table 10: Estimates of the double hurdle model: second stage

	PUBLIC				PRIVATE			
	Coef.	Std. Err.	z		Coef.	Std. Err.	z	
Family income	0.0689	0.0623	1.100		0.0205	0.0574	0.360	
Family income_sq	-0.0034	0.0032	-1.060		-0.0021	0.0028	-0.750	
Education	0.0042	0.0473	0.090		0.0356	0.0437	0.820	
Education_sq	0.0009	0.0027	0.350		0.0006	0.0022	0.280	
Unemployed	-0.0208	0.1847	-0.110		0.0317	0.1530	0.210	
Female	0.2672	0.1512	1.770	*	0.2491	0.1318	1.890	*
Single	-0.1393	0.1548	-0.900		-0.0315	0.1433	-0.220	
Age	-0.0091	0.0886	-0.100		0.0942	0.0888	1.060	
Age_sq	0.0000	0.0007	0.020		-0.0008	0.0007	-1.160	
Chronic conditions	0.2201	0.1359	1.620		0.0859	0.1400	0.610	
Physical limitations	0.3255	0.1582	2.060	**	-0.1085	0.1702	-0.640	
Poor self-perceived health	0.2278	0.1736	1.310		0.3442	0.1996	1.720	*
Excellent self-perceived health	-0.3490	0.1585	-2.200	**	0.0308	0.1443	0.210	
Hearing troubles	0.3023	0.2198	1.370		0.0502	0.2370	0.210	
Eyesight troubles	0.2828	0.1674	1.690	*	0.4048	0.1774	2.280	**
Never smoked	-0.1115	0.1397	-0.800		-0.2518	0.1233	-2.040	**
Alchool consumption	0.8598	0.4154	2.070	**	-0.4857	0.5383	-0.900	
Private health insurance	0.4447	0.2648	1.680	*	0.5313	0.2161	2.460	**
Central region	-0.4664	0.1964	-2.370	**	0.0490	0.1794	0.270	
Southern region	-0.0659	0.1503	-0.440		0.1724	0.1457	1.180	
Public expenditure per-capita	0.0894	0.1479	0.600		0.1302	0.1521	0.860	
Availability of private hospitals	0.8088	0.2243	3.610	***	0.2100	0.2203	0.950	
Physicians per bed in private	-1.4148	0.6434	-2.200	**	-0.0959	0.6306	-0.150	
Physicians per bed in public	1.0837	0.7294	1.490		-1.6624	0.6634	-2.510	**
Population/100	-0.0401	0.0373	-1.080		-0.0255	0.0389	-0.650	
Population/100_sq	1.6540	1.4820	1.120		1.0990	1.5480	0.710	
Constant	-0.1787	2.9754	-0.060		-2.0919	2.9220	-0.720	
lnalpha								
constant	-0.4233	0.2283	-1.850	*	-0.2858	0.1917	-1.490	
alpha	0.6549				0.7514			
LR test against Poisson, chi2(1)	222.103				301.201			
P	0.000				0.000			
Number of obs	408				415			
Model $chi^2(26)$	82.17				58.74			
Prob > chi^2	0.000				0.0002			
Log Likelihood	-738.307				-871.305			
Pseudo R^2	0.0527				0.0326			