Microeconometrics and Asymmetric Information: Applications to Health Care Utilization

A thesis presented

by

Ángel Marcos Vera Hernández

to

The Departament d'Economia i d'Història Econòmica

in partial fulfillment of the requirements

for the degree of

Doctor in Economics

under the supervision of

Dr. Michael Creel

Universitat Autònoma de Barcelona

Bellaterra, March 2001

A Ingrid, por un sueño.

El secreto de la felicidad es éste: que tus intereses sean los más amplios posibles y que tus reacciones hacia cosas y personas interesantes sean amistosas en vez de ser hostiles.

En "La Conquista de la Felicidad", Bertrand Russel. Traducción de J. L. Arangurén.

Contents

[1]This is a joint work with Andrés Romeu.

Acknowledgements

Mientras he hecho la tesis he disfrutado de esa experiencia algo irracional que es inventar problemas y perseverar en el proceso, considerablemente no lineal, de intentar buscarles una solución. Sin embargo, hubiese resultado demasiado arduo si no hubiera disfrutado en el camino. Es aquí donde las personas que están a tu alrededor juegan un papel fundamental. Ha resultado una experiencia muy grata, y por ello les quiero dar las gracias.

First of all, I would like to thank Michael, not only for his valuable econometric advices but also for his generosity with the time dedicated to me. I will always have in mind that his door was always open.

Mi agradecimiento también a Matilde. Con ella di los primeros pasos en la investigación y durante el camino siempre ha estado bien cerca. También quiero agradecer a Carmen las horas que pasamos en su oficina con las curvas de tour-operadores y hoteleros. Nunca pensé que las curvas de un gráfico se pudieran mover tan rápido.

Haciendo el tercer capítulo de la Tesis con Andrés aprendí mucho, tuvimos discusiones muy interesantes y nos divertimos todavía más. Le agradezco la paciencia que tuvo con mis manías, pues seguro que duplicaron a las suyas.

A todos y cada uno de mis compañeros de clase les agradezco los ratos que hemos pasado juntos, los ánimos que me han dado, los cortados que nos hemos tomado, las discusiones que hemos mantenido, así como hacerme cómplice de "Consistency of the Creative Destruction of the Capacity Utilization of the Generic Drugs". I enjoyed very much the company and support of many IDEA students. Many of them suffered GAUSS windows open in the computers. It would be unfair to give names but I particularly thank my office mates for all the times that they had to come to the computer room to tell me I had a call in the phone.

The main asset of this Deparment is the people that are in it. I have mantained a closer contact to Pau, Xavier Martínez, Xavier Jarque, Inés and Belén. I thank them for very interesting discussions and several advices.

Quiero también mostrar mi agradecimiento a Begoña Álvarez por sus consejos cuando revisé el primer artículo, sus correos tan interesantes y por ser la primera persona que se ha atrevido a citarme.

I would also like to thank those people that I have met at conferences for their suggestions. In particular, I would like to thank Frank Windmeijer for several comments to my first two chapters.

El último año ha sido más exigente que los anteriores, por ello agradezco a Belén, Pedro Álvarez, Fuen, Xavier Jarque, Nico, Begoña y Ariadna que hayan soportado mis quebraderos de cabeza con el último artículo. I would also like to thank Inés, Carmen and Pedro Barros for this thing that they know. También a Carmen y Aleix por aplazar un joint project para que pudiera acabar la tesis a tiempo.

Francesc Ortega és el màxim responsable del meu catalá. Li agreixo molt la seva paciència peró sento dir-li que difícilment separaré la llet del gel. Pot-ser ho aconseguiré si fem tants papers junts com tenim al cap.

Quiero dar las gracias a mis padres, abuelos y hermanos por la confianza que mantienen en mí. El amor que me han enseñado no viene en ningún libro, pero es la enseñanza más importante.

Mi agradecimiento más sincero a Ingrid, no solo por las ayudas del final, ni por mantenerme en la tierra, ni por tus comentarios, sino sobretodo por hacer de estos cinco años un precioso sendero. A ti te la dedico, porque siempre he creído que era un poco la tesis de los dos.

Voldria agrair el suport financer del Comissionat de Universitats i Recerca per la beca nùmero 1997FI-436, així com la DGICYT PB96-1160 del Ministerio de Educación. També voldria agrair al Servei Catalá de la Salut per facilitar las dades del segon capitol d'aquesta tesi.

Someone might be left, please feel yourself included in the error term.

Chapter 1

Foreword

During the last thirty years, the existence of asymmetric information among agents has been very important in the development of modern economic theory. The concepts of moral hazard and adverse selection have strongly influenced equilibrium analysis, regulation, optimal contracting, insurance theory, industrial organization, and health economics. In addition, new concepts have emerged: commitment, renegotiation and incomplete contracting. Unfortunately, this advance in economic theory has not been followed by empirical research. According to Chiappori and Salanié (1997):

"*While we understand the features that should characterize the optimal contracts under moral hazard or adverse selection, the number of papers providing a quantitative estimation of the importance of either phenomenon in a given context, or even exhibiting clear econometric evidence for their mere existence is surprisingly low.*"

The purpose of this thesis is to contribute to fill this gap in a field where informational asymmetries have been traditionally considered relevant by the literature, the utilization of health care services (Arrow 1963, Pauly 1968).

1.1 Preliminary discussion

Moral hazard refers to actions that the agent can take during the relation but the principal cannot verify or observe. When the agent has more relevant information than the principal, before the relation has started, then the problem is called adverse selection. For them to be a problem, the asymmetric information should affect contrarily the gains that both principal and agent expect to obtain from the relation.

Moral hazard, as well as adverse selection, are basically hypotheses about agent's informational asymmetries. It could happen that agents have the same *relevant* information, or that they do not. These concepts are not a theory, but an assumption of the theory. However, the existence or not of these asymmetries is important for the prediction that the theory makes about real world.

In what follows, we will describe how the empirical literature has dealt with moral hazard in the utilization of health care services. We will also comment on issues related to adverse selection.

1.2 The elasticity approach to moral hazard testing

Motivation

Real world contracts are not contingent on many variables that significantly affect the value of the contract. Health insurance contracts are not contingent on individual's life habits or in the severity of the illness to receive the benefit of the insurance. Therefore, moral hazard -defined in a narrow sense- is a prevalent feature of the market. However, the hypothesis to be tested is if the incentives embedded in the contract affect agents' actions or behavior. We will see why this is relevant. Consider a demand curve for visits to doctors (see figure above), as a function of the copayment. Let $\overset{*}{s}$ be the first best level of health care consumption. If there is no moral hazard, then the optimal contract specifies a zero copayment contingent on the level $\overset{*}{s}$ of utilization. However, if health care consumption is not contractible and consequently there is moral hazard then the contract specifying a zero level of copayment yields $\overset{**}{s}$ of health care utilization and therefore a positive copayment is needed to obtain a utilization level closer to the first best. However, if we had a very inelastic demand curve -possibly because health care consumption is more influenced by health status, education and culture, rather than by incentives- then the level $\overset{**}{s}$ will be very close to $\overset{*}{s}$, and the moral hazard problem would be very small. That is why the literature pays a special attention to the elasticity of health care choices with respect to incentives. Chiappori, Durand and Geoffard (1998) summaries this idea in the following paragraph:

"*The key aspect of the discussion is, in essence, an empirical one: to what extent can the introduction of a copayment be expected to significantly reduce each agent's level of health expenditures ? If moral hazard is an important feature of the medical consumption,*

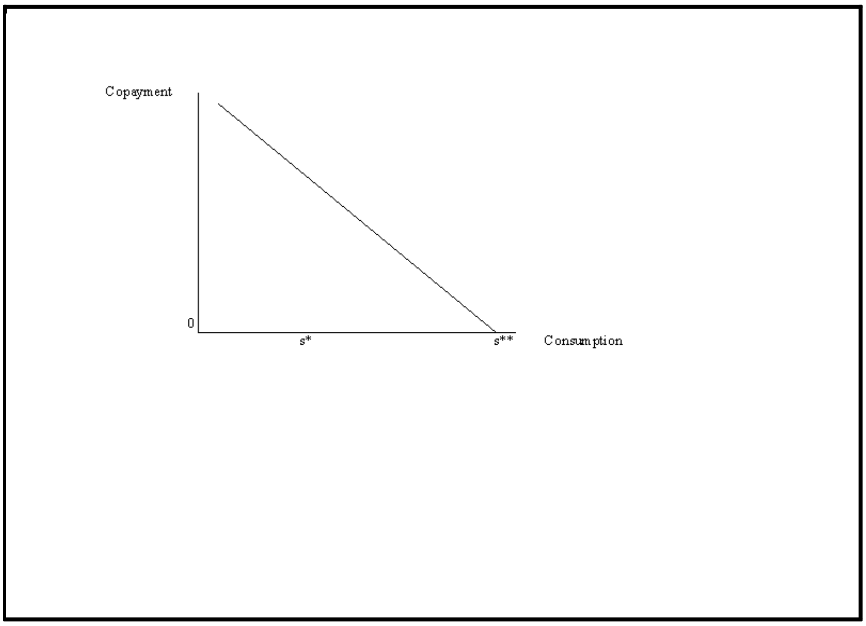

Figure 1.1: Demand function.

then not only copayments are socially acceptable, but they improve welfare. In the opposite case, however, imposing to everyone some minimum copayment is inefficient, because it reduces the scope of mutually beneficial insurance contracts without any gain in terms of aggregate risk."

Although it is interesting to find out how health care expenditures respond to incentives, the information that it provides is limited. Firstly, a downward sloping demand curve does not necessarily imply that there is moral hazard, because there is a price effect unrelated to information asymmetries. Suppose that there is no moral hazard, so the quantity of health care consumption is contractible. If for whatever reason the contract is not designed optimally and specifies a positive copayment, then the higher the copayment, the less people will consume. If health care consumption is contractible, there is nothing that makes it different from a standard good with its downward sloping demand curve. Secondly, we do not know what is the first-best level of consumption. The first level occurs in an unrealistic world of symmetric information, while our estimated demand curve is obtained from the actual world of asymmetric information. As a consequence, we do not know what of the quantity consumed is due to moral hazard since this would imply to compare the unrealistic world of perfect information with the actual one, but we have information about the former. A complete theoretical model would be needed to go from one to the other. Due to these limitations this approach is useful to give general policy directions but it would be hard to obtain specific policy measures from it. As we have mentioned, the analysis is based on estimating an elasticity which is a function of individual's preferences and agent's constraints and therefore it can be regarded as a reduced form approach.

The elasticity approach in practice

In this section we will describe how to carry out the analysis based on the elasticity approach. As we discussed above, the relevant hypothesis to test is whether and how much incentives influences health care choices. This is usually done through considering the effect of insurance on a variable of medical costs or utilization of health care services (visits to doctors, number of inpatient hospital spells...). Therefore we need to estimate the following equation:

$$V = g(H, X, I, \upsilon), \text{ where} \tag{1.1}$$

$H =$ Individual health status.

$I =$ 1, if individual enjoys generous coverage; 0 if she enjoys low coverage.

$V =$ Utilization of health care services, i.e., visits.

$X =$ Sociodemographic variables.

$\upsilon =$ Unobservable variables.

It would be useful to consider also an equation that explains individual's insurance choice. This is the following one:

$$I = f(H, X, \varepsilon), \tag{1.2}$$

where ε is an error term related to insurance choice. Estimation of equation (1.1) must take into account whether the nature of the data to carry out the estimation is experimental or non-experimental data. We discuss this below.

The elasticity approach using non-experimental data. Non experimental data are obtained when the level of coverage has been decided by the individual. If there is adverse selection, we expect that agents with high probability of loss will be able to obtain full insurance, while those with low probability of loss will just obtain partial insurance (Rothschild and Stiglitz 1976). If the econometrician has not been able to condition on all the relevant risk variables, the error terms ε and v might be correlated and consequently, the insurance variable in (1.1) is correlated with the error term v. It is necessary to take into account this endogeneity in order to clarify the empirical puzzle of moral hazard and adverse selection.

The puzzle is the following. Both hypotheses, moral hazard and adverse selection, predict a positive association between health care utilization and better coverage. That is, if there is moral hazard, those with generous coverage use more frequently health services. This is because they face less incentives to restrict use. On the other hand, if there is adverse selection, those with generous coverage also use health services more frequently. But in this case, it is because they are the ones with a higher probability of loss -that is, the ones with poor health status-. If we do not take into account this endogeneity, we would be summing up these two effects, while to test moral hazard we are just interested in the one related to incentives.

Adverse selection testing. The framework above gives also a test for adverse selection. Adverse selection tests are usually carried out through the main theoretical prediction: individuals with high probability of loss will be able to obtain full insurance, while those with low probability of loss will just obtain partial insurance (Rothschild and Stiglitz 1976). Consequently, a natural approach to the problem is to estimate equation (1.2) and analyze whether the health determinants influence health coverage. If insurers observe H, as the econometrician does, health status will be taken into account in the premium and consequently H will not show up significantly in equation (1.2).

In order to examine adverse selection due to variables that the econometrician does not observe, the system of equations given by (1.1) and (1.2) should be considered. It will be very important to obtain conclusions about the correlation between the error terms: ε, v. These error terms pick up variables that are unobservable for the econometrician. A positive correlation means that the unobservable variables that increase the probability of loss, also increase the probability of having a generous coverage (that is, people that are more likely to use health care services, also are more likely to enjoy a generous coverage). Consider the set of unobservable variables that influence the use of health care services v, if these variables are observed by the insurer, then the insurer will adjust the premium to the expected loss and therefore these variables will not cause correlation with ε. Therefore, this correlation is only affected by variables that are unobserved to both the econometrician and the insurer, and therefore it is fair to refer it as a test of adverse selection.

The test described above is not new in the health economics literature and it has been proposed in a more general framework by Chiappori and Salanié (1997). The main criticism to this approach is that this correlation could also be caused by measurement errors in the regressors of the model or some other specification error so the econometrician should be careful with this problem[1].

Testing for moral hazard using experimental data. Experimental data refers to the situation where the level of coverage (that is, the variable I) has not been chosen by the individual. For instance, changes in I could be due to exogenous changes in the legislation. The extreme case was the RAND Corporation *Health Insurance Experiment* (Manning *et al.* 1987, Newhouse *et al.* 1993) where a group of people were randomly assigned to different levels of coverage. In these cases, equation (1.2) is useless since the level of coverage has not been decided by the individual. Therefore, there is no correlation between the error terms and the regressor I is exogenous, which simplifies the econometric estimation. Although this kind of data is usually preferred to non-experimental data, experimental data are hard to find.

1.3 The structural approach to moral hazard

In the previous section, we have concentrated on how to obtain an estimate of the elasticity of health care consumption to the incentives faced in the insurance contract. The approach did not attempt to recuperate consumer primitive parameters and consequently, it can be regarded as a reduced form approach. This has been the approach more commonly used in the literature, but a structural approach is also feasible.

The structural approach tries to obtain estimates of consumer primitives (technology and preferences). On one hand, this normally requires making strong assumptions about consumer behavior. On the other hand, it is a more powerful approach to policy analysis than the reduced form approach. Firstly, it is possible to solve for the policy

[1] Wolak (1994) gives an alternative approach based on a structural formulation.

function since one has obtained estimates for the parameters that define the theoretical problem. In our context, this means that one can solve for the first best or for the optimal copayment. Secondly, the parameters of a reduced form approach are a combination of raw parameters (technology and preferences) and agents' constraints. When a policy change takes place, the parameters that define the constraints will normally shift and the parameters of the reduced form model will also change (Lucas 1976). This will not occur using a structural approach.

1.4 Microeconometric considerations

Independently of the approach taken, the analyst will normally be more confident about the results if less assumptions are made about either the functional form of the model, either the distribution of disturbances. This is particularly important in a non-linear setting where the consequences of failing to verify either the assumptions on the disturbances or the functional form assumptions will yield, in general, inconsistent estimates.

In our setting, the models used for the analysis will be non-linear. The variables that measure health care utilization are either dichotomous (to have or not to have treatment) or integers (number of visits to doctors). Consequently, non-linear models will be necessary to incorporate this feature of the data.

1.5 The following three chapters

The following three chapters of the thesis deal with the issues mentioned above. The first two follow the elasticity approach using non-experimental data and consequently, the endogeneity of insurance coverage is taken into account. The last one follows a structural approach to model health care utilization but it does not model insurance choice since the data used is experimental. In this last chapter the variable that measures health care utilization is whether or not the individual has been treated by an illness episode which is a dichotomous variable. On the contrary, in chapter 2 and 3 health care utilization is measured by the number of visits to doctors in a given period of time. Consequently, the non-negative and integer nature of the data is taken into account. Chapter 2 and 3 differ in the strategy pursued for the estimation and in the health system analyzed. Chapter 2 concentrates on a strategy that makes relatively few assumptions about the unobservables at the price of a more inefficient estimates and less flexibility in the functional form of the model. In this chapter we analyze data from the Catalonia Health Survey 1994. In Chapter 3, we analyze data from United States of America using a count data model that is very flexible in the density of the count variable but assumptions about the distribution of unobservables are necessary in order to carry out the estimation.[2]

Chapter 2 is basically an analysis of health care utilization in general, and in particular the influence of incentives and the role of adverse selection in the context of the Catalonian Health System, where those who buy a private insurance (duplicate coverage) can easily visit a specialist (without general practitioner permission, waiting lists...). Therefore, it analyses how non financial incentives can affect health care choices. We found that private insurance increases the average of visits to specialist in approximately a 25% for non heads of households. It was relevant to consider the endogeneity (adverse selection) for the heads of households. The problem is analyzed by means of count data models. In order to take into account the endogeneity of insurance status, we rely on an Generalized Methods of Moments estimator. Its main contribution is to consider the endogeneity of private insurance in the Catalonian Health System, as well as the application of the econometric model to model endogeneity of insurance status.

Chapter 3 is a joint work from Andrés Romeu. The chapter considers a count data model with a dummy endogenous variable where flexibility is given by multiplying the count density by a polynomial. This expansion modifies the conditional mean of the count variable, allowing departures from the standard linear exponential specification. Estimation is carried out by Maximum Likelihood. It is found that the proposed model obtains a better fit to the data than standard models found in the literature. We find that health care utilization is affected by both Medicaid coverage and private insurance. An interesting result is that health care utilization of those with poor health conditions is less sensitive to insurance status. Its main contribution is the econometric method that combines flexibility to model count variables and a dummy endogenous variable.

Chapter 4 presents a model where the individual decides to have treatment or not when she suffers an illness spell. The decision is taken on the basis of comparing benefits and out-of-pocket monetary costs of treatment. The aim of the paper is to recover the estimates that conform a principal agent model in order to solve for the optimal

[2]In this chapter, we also analyze data on number of trips by households.

copayment. The framework also allows us to simulate the effects of not fair premiums in coverage decision. We find that optimal copayments in presence of a fair premium are surprisingly low. However, slight departures from the fair premium hypothesis give us copayments of the size found in real life. We highlight that previous approaches might have neglected the role of income effects when measuring welfare losses due to moral hazard. Its main contribution is to estimate the parameters of a principal-agent model of health care consumption.

Bibliography

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

[2] Chiappori, P.A. (1997), Empirical Contract Theory: The Case of Insurance Data. *European Economic Review*, 41, 943-950.

[3] Chiappori, P.A., Durand, F. and Geoffard, P.Y.(1998), Moral Hazard and the Demand for Physician Services: First Lessons from a French Natural Experiment. *European Economic Review,* 42: 499-511.

[4] Lucas, R. (1976), *Econometric policy evaluation: a critique.* Carnegie-Rochester conference series on public policy, 1, 19-46.

[5] Pauly, M. (1968), The Economics of Moral Hazard. A Comment. *The American Economic Review,* 58*, 531-537.*

[6] Rothschild, M. and Stiglitz, J. (1976), Equilibrium in Competitive Insurance Markets: an Essay on the Economics of Imperfect Information. *Quaterly Journal of Economics* 1976, 90, 629-650.

[7] Wolak, F. (1994), An Econometric Analysis of the Asymmetric Information, Regulator-Utility Interaction, *Annales d'Economie et de Statistiques,* 34, 13-69.

Chapter 2

Duplicate Coverage and Demand for Health Care

2.1 Introduction

The existence of asymmetric information in markets related to health care, rapid technological change, the concern about equity and finally, the increase in health expenditures have induced the large amount of literature produced in this field. An important phenomena related to some of these issues that has received very little attention in the literature is *Duplicate Coverage.*

An individual enjoys *duplicate coverage* when he is covered by a compulsory public medical insurance as well as a private one. This paper evaluates two important issues. First, the relation between duplicate coverage and the selection mechanism that takes place in the private insurance market. We will argue that the existence of a compulsory public insurance may be incompatible with market screening. The second issue deals with the influence of duplicate coverage on the number of visits that individuals make to the doctor. The effect of duplicate coverage on the number of visits could be ambiguous owing to two opposed effects. On the one hand, an aggravation of the moral hazard that leads people to go to the doctor more frequently and because of less severe illnesses. On the other, we could think of situations where the privately financed treatment is perceived as more efficient (in the sense of less visits for the same treatment) than the public one. This leads to a reduction on the number of visits that individuals make.

Quantitatively, duplicate coverage is not a negligible phenomenon either and can be found in several European countries, such as, Great Britain. The data used for this paper corresponds to Catalonia, where 20% of the population (about 1.2 million) had duplicate coverage in 1995.

The study presented here is important for several reasons. First, because of the relevant information it reveals to the policy makers. In particular, it may help the policy maker to understand why people with similar health conditions differ in their use of medical care. This information is important to control the expenditures in health care as well as assess the equity of the system. On one hand, it can help to detect whether possible inefficiencies causing unnecessary utilization due to moral hazard should be a concern. On the other, if people with similar health conditions differ in their health care utilization due to income or other socioeconomic variables, the equity of the system should be more carefully studied. The second relevant contribution of this paper is that it helps to understand the selection process that takes place in the private medical insurance market when a compulsory public insurance is present. Next, this study attempts to accommodate the endogeneity of the insurance choice decision and the non negativity of the number of visits in a cross section sample. Fourth the choice of the instruments necessary for the econometric work is motivated within a theoretical model.

The survey used is taken from the 1994 Catalonia Health Survey (Servei Català de la Salut 1994). The data reveals that the average number of visits to the doctor of people covered only by the public system (1.54) is smaller than the average for people that enjoy duplicate coverage (1.68). Furthermore, this relationship holds in seven out of eight health regions in Catalonia. However, one does not know if this difference is attributable to the effect of duplicate coverage or rather the effect of other variables related to duplicate coverage, such as health conditions.

Both the theoretical model and the empirical framework focus on two central issues of the analysis. First, the variable chosen to measure the demand for health care is the number of visits to specialists in the 12 months prior to the interview. The discreteness of this dependent variable as well as its non negativity calls for count data modelling. The second one is the endogeneity of the private insurance decision when studying the demand for health care.

To provide an intuition for the source of endogeneity, assume that the number of visits to doctors and the insurance decision are functions of both observable variables X and Y respectively. Furthermore, a variable η, say an allergy, influences both visits and private insurance decision. In principle, we could think that people suffering from allergy are more likely to both go more frequently to the doctor and buy an additional private insurance. Although there is data available on X and Y, η is unobservable, that is, data about η cannot be found. Therefore we cannot use it when we estimate the equations of visits and insurance. It means that the error terms $\varepsilon(\eta)$ and $\zeta(\eta)$ of these equations will capture the effect of η, that is, of the allergy. The following equations might be useful to get an insight of the issue:

$$\text{Visits} = v(\text{X}, \text{Private insurance}) + \varepsilon(\eta), \tag{2.1}$$

$$\text{Private insurance} = I(\text{Y}) + \zeta(\eta). \tag{2.2}$$

Endogeneity is the correlation between a regressor and the error term of the equation. In this case, both the regressor 'Private insurance' of equation (2.1) and $\varepsilon(\eta)$ depend upon η and therefore there is a correlation between them, causing the endogeneity. If endogeneity is ignored when estimating, it will lead to inconsistent estimates. Since endogeneity is due to unobservable variables and the variables available to us can only proxy true health status, then endogeneity is likely to be present.

Adverse selection is not only a potential problem for the insurance carrier but also for the statistician interested in estimating the true effect of the insurance in the number of visits. If we ignore endogeneity and the selection mechanism of the market is such that those with poor health conditions are the ones most likely to buy an additional private medical insurance, then we will overestimate the insurance effect.

We will argue how in presence of a compulsory public insurance, it is possible that private insurance providers are not able to offer a menu of contracts that both those with good unobservable health conditions and those with poor unobservable health conditions will accept. On the contrary, they will offer a contract that only unhealthy ones will accept. Therefore those with a private medical insurance are the most likely to be ill and we anticipate an overestimation of the insurance effect when endogeneity is ignored.

In our results we find evidence of endogeneity for the subsample of heads of household, that, if ignored, yields an upward biases in estimation. For the subsample of non-heads of household, endogeneity does not seem important, and we find a positive effect of duplicate coverage on the visits to specialists. The finding that the endogeneity of duplicate coverage is more important for heads of household is consistent with the idea that these are the ones that take the insurance decision for the whole family.

Summarizing, this paper studies the effect of duplicate coverage over both individual's demand for visits to doctors and the selection mechanism of the market. We use convenient econometric techniques that accommodates the endogeneity of the private insurance choice as well as the non negativity of the data.

The paper is organized as follows. Section 2 describes the relevant aspects and limitations of the 1994 Catalonian Health Survey and the institutional setting of the Spanish health care system. Section 3 briefly describes studies that relates the utilization of health care services and insurance. Section 4 describes a theoretical model that motivates the choice of the econometric instruments. In this section, we propose a hypothesis about the selection process of the market, we explain the different effects of insurance on individual demand for visits and we give a careful interpretation for this effect. Section 5 is devoted to the discussion of the econometric techniques. Section 6 interprets the econometric results and finally, in section 7 the main conclusions are summarized.

2.2 Institutional setting and data

The Spanish National Health Service provides free coverage for medical and hospital care (except for dental and psychological care). The coverage is almost universal, around a 97.1% of population in 1993 benefited from it (Whitaker and Sánchez 1997). Doctor visits are free of charge and non-pensioners patients only pay 40% of outpatient prescription charges costs while, pensioners get full subsidization of prescriptions.

Apart from obtaining care through the public system, the individual can obtain health care in the private sector. In this case, the individual can pay in a fee-for-service basis and/or purchase private medical insurance. People that buy a private medical insurance will not see their contribution to finance the public system affected and as a consequence they are still eligible to receive medical care from the public medical system, the latter situation is what we call *duplicate coverage.*

There are two kinds of private medical insurance. The most widely contracted, which covers 91% of people with a private medical insurance (Murillo 1992), gives the patient the right to obtain health care with negligible use of copayments. A Spanish consumers organization (OCU) published in 1997 a study about the different private insurance offered in the market. The copayments were at most 1.8 Euros. There is little difference in the coverage and price levels among the private insurance contracts that are available in the market. If an individual contracts a private medical insurance, he receives a catalogue with the list of doctors that he or she can visit under the coverage of the insurance. If the consumer visits a private doctor who is not included in the catalogue then she will have to pay the whole fee.

There are various reasons to obtain private health care. In the public system, the choice of doctor is very restricted, there are waiting lists and queuing. Timetable restrictions as well as administrative shortcomings are important. The perceived quality of service may also be different in the public system than in the private one. In the extreme, the consumer could identify a more comfortable waiting room of a private doctor as *better* care.

The empirical results are based on the 1994 Catalonia Health Survey. The survey provides socioeconomic characteristics, life styles, health indicators and utilization of health related services at the individual level, although some information about the household is reported.

Particularly important for this study is the availability of information about coverage. The total number of visits to doctors in the 12 months prior to the interview is recorded. However one does not know under which kind of coverage each visit was made, and therefore, it is not possible to know which proportion of the visits were to the public doctor, to a private insurance catalogue doctor, or even to a private one out of the catalogue. If the individual made a visit to a doctor in the 15 days prior to the interview then the coverage of this last visit is recorded.

The analysis of this information for individuals under duplicate coverage is presented in Table 2.1. The data about the visits to a general practitioner may seem surprising. People with duplicate coverage visit General Practitioner (GP) in the public system more frequently. This is because there are other reasons, different than obtaining health care, to visit a GP. It is common that people who enjoy duplicate coverage visit the public GP to get prescriptions with subsidization. Other cause to visit a public GP, different from obtaining health care, is that an individual needs the permission of the public GP to visit a specialist in the public system. People also visit the public GP to obtain the medical certificates in order to notify an illness to the employer and the administration.

The column labelled ‘Specialist’ in Table 2.1 presents the expected result that people with duplicate coverage visit private insurance specialists more frequently than other specialists. As Table 2.1 shows, people with duplicate coverage also visit public specialists and specialists out of the catalogue. This makes it clear that visits to specialists should be considered a *heterogenous good.* On one hand, an individual with duplicate coverage will find visiting a private insurance doctor more convenient and is more likely to wait less than if he visited a doctor of the public system. However, one could find very good doctors in the public system and with advanced medical equipment available to them. This may explain why an individual with duplicate coverage would visit a specialist of the public system. On the other hand, people having duplicate coverage visit also doctors who are out of the insurance catalogue. This may be explained by considering a private medical insurance as a package, that may include some doctors who are not completely satisfactory to the consumer. There are prestigious doctors who have not contracted their services with any insurer. Many visits to specialists are due to chronic diseases and therefore the consumer may appreciate a long term relation with the doctor. Apart from this, some people are covered by a private medical insurance that has been contracted by their employer, and therefore not chosen by the consumer.

2.3 Literature review

Studies that try to evaluate the influence of medical insurance on health care utilization can be divided according to the nature of the data. Manning *et al.* (1987) use data from the Rand Health Insurance Experiment where people were randomly assigned to insurance plans with different copayments. Chiappori *et al.* (1998), take advantage of a change in the copayments that some French insurers charged to their insureds. In these two studies, the individual has not chosen the copayment and therefore it is exogenous. In this case, it is said that the data are experimental. They are preferable to nonexperimental data because the analyst does not need to take into account the endogeneity of insurance choice. However experimental data are not available in many countries and samples are often unrepresentative of the general population. This justifies the use of nonexperimental data. Example of studies that use nonexperimental data and in consequence take into account endogeneity of the insurance plan are Cameron *et al.* (1988), Coulson *et al.* (1995) and Holly *et al.* (1998).

Cameron *et al.* (1988) study the influence of medical insurance on several health care utilization variables. The authors build a theoretical demand model where the interdependence between the demand for visits and the insurance decision is clear. However their econometric framework does not take into account simultaneously the endogeneity of insurance and the non negativity of the variable number of visits. Since their study takes place in Australia, duplicate coverage is not considered. In the study it is found that people with more generous coverage visit the doctor more frequently. The null hypothesis of exogeneity is strongly rejected for the subsample of people above certain income level.

Coulson *et al.* (1995) estimate an empirical model to study the effect of supplemental medical insurance (Medigap policies) on the number of prescriptions filled or refilled by an individual in a designated two-week period. Duplicate coverage is again not considered, since Medigap policies cover the purchase of outpatient prescription drugs which are not covered by the Medicare program. The results show that, on one hand the insurance choice is exogenous to the decision regarding the number of prescription drugs and on the other hand supplemental insurance raises the number of drugs filled.

Holly *et al.* (1998) use the 1992-93 Swiss Health Survey to estimate a simultaneous two equation model for the probability of having at least one inpatient stay and purchasing supplemental insurance. They find that the effect of supplemental insurance plan is to increase the probability of a person having at least one inpatient stay. In Switzerland, duplicate coverage is not an issue either.

Recently Álvarez (2001) has studied the determinants of the number of visits to doctors and urgency visits in the last previous fifteen days to the interview. She uses two complementary datasets, the 93 Spanish Health Survey and the Spanish Expenditure Survey to obtain income information. Form an econometric point of view, she focuses on choosing among a variety of model specifications, rather than on the specific issue of duplicate coverage.

2.4 Theoretical framework

2.4.1 The model

The model is in the spirit of Cameron *et al.* (1988) but it models explicitly the unobservable variables and provides some insight about the kind of variables that will be used as instruments for the insurance decision in the empirical part of the paper.

Let v_{1i} denote, for individual i, the number of visits to the doctor through the public system, v_{2i} through a private medical insurance and v_{3i} through a private doctor paying the full fee. We summarize this information in the vector $v_i = (v_{1i}, v_{2i}, v_{3i})$.

The vector $s = (s_1, s_2, s_3)$ contains information about the perceived quality of the service offered by the corresponding supplier. $H_i = H(v_i|S_i, \eta_i, \varepsilon_i, s)$ is a health production function that depends on both observable (S_i) and unobservable (η_i) individual's health characteristics, a random health shock (ε_i), as well as the quality of the service.

We assume the following conditional probability distribution for ε_i:

$$F(\varepsilon_i|S_i, Z_i, \eta_i). \tag{2.3}$$

It is important to point out that both S_i and Z_i represent observable characteristics but they play a different role. The variables in S_i are directly related to the individual's health (chronic illnesses, age, sex...) while Z_i contains variables that do not determine health but the likelihood of occurrence of future health shocks. An example is illustrative. Compare two individuals, one belonging to a low social class and the other to a high one. If a heart attack happens to both of them, there is nothing intrinsic to social class to make us think that one of them is going to feel worse than the other, so social class does not belong in S_i. On the contrary, the likelihood of occurrence of a heart attack may actually differ, since they are exposed to different environmental conditions. Therefore the social class is a variable likely to belong to Z_i. Employment status is also likely to belong to Z_i.

The timing of decisions is as follows. At t=1 the individual already knows her endowments of S_i, Z_i, η_i. At this stage she decides whether to take additional private insurance, taking into account the probability distribution of the random variable ε_i conditional on the information available to her. At t=2, the random variable ε_i is realized and observed by the consumer, and finally, at t=3, the consumer decides on visits to the doctor.

We will solve the individuals optimization problem backwards. At t=3, conditional on the realized value of ε_i, that is $\widehat{\varepsilon}_i$, and the previous decision on the insurance choice, the consumer i solves

$$\underset{\{C, v_1, v_2, v_3\}}{Max} U(C_i, H_i) \tag{2.4}$$

$$\text{s.t.} \begin{cases} H_i = H(v_i|S_i, \eta_i, \varepsilon_i, s) \\ p_3 v_3 + C_i = Y_i - I \times P_i \\ I_i \in \{0, 1\}; \\ v_1, v_2, v_3 \in N. \end{cases} \tag{2.5}$$

So the consumer chooses a vector of number of visits $\overline{v}$ and the quantity consumed of other goods, C_i , in order to maximize a utility function $U(.)$ which is a function of H_i and C_i. Y_i stands for consumer's total income and P_i for the insurance premium. Since visits to the public doctor are free of charge and copayments within the private insurance schemes are negligible then we only introduce v_{3i} in the budget constraint, with its corresponding price per visit p_3. I_i can only take value 0 or 1. If it takes value 0 then the individual has chosen in t=1 not to take an additional private insurance and therefore the parameters s_2,and P_i will not play any role when deciding the demand for visit and v_2 will be automatically zero. On the contrary $I_i = 1$ means that the individual has bought an additional private insurance and therefore enjoys duplicate coverage. In this case, v_{2i}, s_2, and P_i do matter when deciding the demand for visits and the individual is not constrained to $v_{2i} = 0$. Therefore the demand functions are:

$$\overline{v}_{j\,i} = v_j\left(S_i, \widehat{\varepsilon}_i, \eta_i, Y_i - I_i * P_i, s, p_3\right) \quad for\ j = 1, 2, 3 \tag{2.6}$$

where both η_i and $\widehat{\varepsilon_i}$ are unobservable to the econometrician. Since Z_i only affects ε_i stochastically, then Z_i does not show up in equation (2.6) because the demand for visits is taken once the realization of ε_i has been observed.

Once the consumer knows the demand for visits to the doctor conditional on having purchased private medical insurance and on each possible realization of the random variable ε_i, then the decision whether to contract or not a private insurance will be taken by comparison of the expected utilities. So by substituting the demand functions in (2.4) for $I_i = 1$ and integrating with respect to the conditional density of ε_i then the expected utility of purchasing private insurance in t=1 is given by

$$E\left[U_i | I = 1\right] = \int U(C_i, H_i(\overline{v_i} | S_i, \eta_i, \varepsilon_i, s))\ \partial F(\varepsilon_i / S_i, Z_i, \eta_i). \tag{2.7}$$

which can be expressed in reduced form as

$$E\left[U_i | I = 1\right] = \overset{*}{U}\left(S_i, Z_i, \eta_i, s, Y_i - P, p_3\right). \tag{2.8}$$

Equivalently for the case of $I = 0$

$$E\left[U_i | I = 0\right] = \overset{**}{U}\left(S_i, Z_i, \eta_i, s, Y_i, p_3\right). \tag{2.9}$$

Finally the consumer will decide to buy private medical insurance if the difference

$$\Delta U_i = \overset{**}{U}_i - \overset{*}{U}_i = \Delta U(S_i, Z_i, \eta_i, s, Y_i, P_i, p_3) \tag{2.10}$$

is negative.

From this model we can conclude that the variables Z_i are likely to be good instruments since they influence the insurance decision but not the demand for visits directly. That is to say, they only influence the demand for visits through the insurance decision. The reason is that, when the insurance decision is taken, the consumer considers the likelihood of occurrence of health shocks. However the demand for visits depends on the actual health shocks ($\widehat{\varepsilon}_i$), which is why Z_i does not show up in the demand for visits. Therefore the decision whether to purchase medical private insurance or not depends on S_i as well as Z_i but the demand for visits only depends on S_i.

A second conclusion of the theoretical model is that endogeneity seems a reasonable hypothesis since both the demand for visits, given by equation (2.6) and the decision whether to contract a private medical insurance, summarized in (2.10), depend upon η_i which is an unobservable variable and consequently it will be captured by the error terms of the econometric model.

So far, we have neglected the agency relationship that exists between the doctor and the patient, which may give rise to supplier induced demand. Unfortunately, the type of Survey data we have available is not suited to take these effects into consideration The reason is that one cannot identify the first contact to a physician for a new illness spell among visits reported by individuals in the survey. Since sample period and illness spells are not the same, it is not necessarily the case that the first visit reported by the individual corresponds to a first contact to a physician because of a new illness spell. The first visit within the sample period could be due to an illness that requires regular attention and therefore the same doctor may have been visited before the sample period began. Furthermore, it could be the case that the second visit reported in the sample period corresponds to a new illness spell.

Another shortcoming of the model is that it assumes that after the health shock has occurred, the consumer plans the whole series of visits. It may be more realistic to assume that several shocks could occur and the number of visits could depend on the number of shocks, the kind of shocks and possibly on the time between shocks.

2.4.2 Expected influence of duplicate coverage on individuals demand for visits

Up to this point, a hypothesis about the effect of duplicate coverage on the demand for visits has not been established. A complete model to derive this effect is difficult because of the number of relevant elements: risk issues and heterogeneity of visits. In order to get some insight one could try to study the two issues separately.

First concentrate on the effect of risk shifting associated with duplicate coverage. Visiting a public doctor may cause some losses to the consumer, such as travel costs, waiting time, and poor perceived quality of care. These losses could be seen as a form of copayment and therefore the compulsory public insurance would not be a complete insurance. A risk averse consumer will benefit from contracting private medical insurance that reduces her losses when visiting the doctor. Since private medical insurance will be closer a complete insurance, then moral hazard problems would be exacerbated and consumers who contract a private medical insurance would have a larger demand for doctor visits.

The preceding paragraph is useful to understand risk issues associated with duplicate coverage and its implications for demand for health care, but it does not deal with the heterogeneity of visits. One could think of situations where the heterogeneity of visits play an important role in determining the total number of visits. Consider the case of a patient without duplicate coverage who has already visited the public specialist. If she is disappointed by the quality of the service obtained, she could additionally visit a private specialist, paying the full fee. On the contrary, patients that enjoy duplicate coverage can choose their preferred specialist within a closed list. In this case, he could be satisfied with just one visit. In this example, the patient that enjoys duplicate coverage makes a smaller number of visits.

2.4.3 Can we interpret the effect of insurance on visits as moral hazard?

For the purpose of this section, assume that the number of visits to doctor is given by the following linear function:

$$v_i = \theta X_i + \phi I_i, \tag{2.11}$$

where X_i is some individual characteristic and $I_i = 1$ if the individual enjoys a private insurance and $I_i = 0$ otherwise.

In the empirical literature the coefficient (ϕ) associated with the dummy for insurance has been interpreted as moral hazard. However this interpretation does not seem to be entirely correct.

Insurance has two different effects: a desirable one that is owing to risk reduction and an undesirable one, caused by moral hazard. As Meza (1993) shows, both effects will usually lead to increased consumption. It is well known that there is moral hazard when an insured individual consumes more than the first best level. That is, moral hazard should be measured as the difference in consumption between the actual world where there is asymmetric information after the signature of the contract and the unrealistic world of perfect information. However the literature sometimes neglects the desirable insurance effect, the one caused by risk reduction. The risk reduction effect implies an income transfer from the healthy states to the ill states. If health care services are a normal good, this would imply an increase in health care consumption with respect an uninsured situation. In other words, if there are income effects the hicksian demand curves is the relevant curve to measure price or incentives effects. The uncompensated demand curve will overestimate the price effect since it will take both the price and the income effect. See Meza (1983) and Belsey (1988) for a more formal discussion.

In countries where duplicate coverage is not an issue, studies try to estimate an equation for utilization of health services where one explanatory variable is insurance coverage. Usually the presence of insurance increases consumption. Can we say that this increase in medical services utilization is due to moral hazard? The answer is no. That is, moral hazard is when an insured individual consumes more in the real world than in the *unrealistic* world of perfect information where an insurance contract has been signed. On the contrary, ϕ measures the difference in consumption between two real words, one with insurance and the other without insurance. Therefore ϕ will typically overestimate moral hazard effects. A positive sign of ϕ is consistent with moral hazard, but does not imply it.

With duplicate coverage, things are more complex. The first best solution can be defined as the solution that maximizes the welfare of the agent (consumer) given a level of welfare for both principals (public and private insurance carriers). So, the first best solution will be an allocation specifying the level of consumption from the public and the private insurance. So any deviation from this, caused by the ex-post asymmetry of information, would be moral hazard. Therefore, it seems that comparing two real situations (consumption with duplicate coverage and consumption with only public coverage) says little about the comparison among the allocation of the real world and the one of perfect information. However, it is likely that in the duplicate coverage case ϕ is more close to measure moral hazard, since most of the risk reduction effect should be already included in the public compulsory insurance.

In this study ϕ will be interpreted as the effect of additional private insurance on individual consumption but not related to moral hazard. One must say that some motivation is lost, since interpreting ϕ as moral hazard gives an indication of inefficiency in the market.

2.4.4 Effect of public insurance on market selection

In presence of adverse selection, insurers will try to screen the market, that is, offer a menu of contracts to encourage different types of agents to self-select. The insurer will offer complete insurance (a large premium and zero copayment) for high risks and incomplete insurance (low premium and a large copayment) for low risks (Rothschild and Stiglitz 1976).

However if there is duplicate coverage, both type of agents already enjoy a compulsory public insurance. In this situation, the low risks may not find very attractive to buy a private insurance, since it would be incomplete insurance and they already have incomplete coverage (the public insurance) for free. Moreover payments to the private insurance will have to cover all the treatment costs, not only the part that is not covered by the compulsory public insurance. As so, the existence of a compulsory public insurance could be incompatible with both types of agents buying private insurance and therefore private insurers could be forced to take only bad risks, for which the optimal contract is based on a high premium. These seems consistent with the Spanish evidence where copayments are small and a large variety of contracts is not available. As a consequence, those that enjoy private insurance in a duplicate coverage setting are those more likely to be ill. As we argued in the introduction, this will yield overestimates of the duplicate coverage effect on the demand for visits when using a econometric model that ignores endogeneity.

Two last comments might contribute to understand the argument above. First, this reasoning holds for the same level of perceived risk. Therefore, it is consistent with the fact that older people do not generally enjoy private insurance. Second, it is important to realize that this is a ceteris paribus argument. That is we could find an unhealthy individual who does not enjoy duplicate coverage owing to the fact that he is poor, and a rich one, who is healthy and who has duplicate coverage. Fortunately, in an econometric model one can control, at least partially, this ceteris paribus assumption,

2.5 Econometric modelling

This section discusses the different ways to estimate the econometric model. As pointed out in the introduction, the econometric model should take into account that the dependent variable, number of visits, takes only non-negative integer values. An appropriate framework is *count data modelling* (Hausman *et al.* 1984, Cameron and Trivedi 1986, Winkelmann and Zimmermann 1995).

In order to go from a theoretical model to an econometric one, the functional forms of the equations need to be specified. The theoretical model is taken just as a guide and therefore the functional forms of the theoretical model are not specified. On one hand, the number of visits to specialists is a non-negative integer and therefore in order to solve the consumers problem integer optimization techniques should be applied for which analytical solutions are hard to find. On the other, non-negativity of the data already imposes restrictions on the functional form of the econometric model. Cameron *et al.* (1988) have commented on the difficulties of estimating a structural model of this form.

Another issue that should be considered when specifying the econometric model is the endogeneity of the insurance choice. This endogeneity shows up in equations (2.6) and (2.10) through the common element η. If I_i represents the individual's insurance choice and u_{1i} the error term for the econometric equation for the number of visits then $E\left[I_i \times u_{1i}\right] \neq 0$ since both I_i and u_{1i} depend upon η_i. When this endogeneity exists, standard methods of estimation such as Least Squares or Maximum Likelihood lead to inconsistent estimates. If individuals with poor health conditions are the ones that hold private insurance, then endogeneity will lead to overestimates of the effect of insurance on visits.

As we have already mentioned, information regarding the system that the consumer has used to visit the doctor (public, private insurance and private out of catalogue) is not available. So the analysis will be done for the total number of visits. In the following, we will use v_i to denote the total number of visits by individual i.

2.5.1 Absence of endogeneity

As a benchmark and in order to be able to perform endogeneity tests, estimation under the no endogeneity hypothesis should be considered. If private information η_i, which is unknown to the econometrician, is also unknown to the consumer in t=1, when contracting the private medical insurance, then the expected utility computed in equations (2.8) and (2.9) would have been taken with respect to the joint probability distribution of ε_i and η_i, that is $\partial F(\varepsilon_i, \eta_i|S_i, Z_i)$ and therefore (2.10) would not depend on η_i. In this case the decision about contracting private medical insurance I_i would be an *exogenous regressor* in equations regarding the number of visits to the doctor.

In order to test the exogeneity of the insurance decision, using a Hausman test, one should obtain an estimator that is consistent and efficient under the null hypothesis of no endogeneity and inconsistent under the alternative. This will be done by maximum likelihood, assuming that the dependent variable, number of visits to doctors, is drawn from a specific discrete probability function. In this way, not only the non negativity of the variable but also its discreteness will be taken into account.

The Poisson distribution is the benchmark in count data applications. This assumes that the conditional mean of the dependent variable is equal to the conditional variance. In our data, as is common in health care utilization data, the conditional variance is bigger than the mean, a situation called overdispersion. In presence of overdispersion, the Poisson model is no longer efficient. Since we are interested in performing a Hausman test, we need a model that accommodates overdispersion. In this situation, it is very common to use the Negbin-2 (Cameron and Trivedi 1986), which assumes a Negative Binomial distribution for the dependent variable and a quadratic relationship between conditional mean and variance. Its density function is

$$f(v_i|\alpha,\lambda_i) = \frac{\Gamma\left(\alpha^{-1}+v_i\right)}{\Gamma\left(\alpha^{-1}\right)\Gamma\left(v_i+1\right)}\left(\frac{\alpha^{-1}}{\lambda_i+\alpha^{-1}}\right)^{\alpha^{-1}}\left(\frac{\lambda_i}{\lambda_i+\alpha^{-1}}\right)^{v_i}, \tag{2.12}$$

where

$$E\left[v_i|\lambda_i,\alpha\right] = \lambda_i \text{ and } Var\left[v_i|\lambda_i,\alpha\right] = \lambda_i + \alpha\lambda_i^2. \tag{2.13}$$

In order to ensure the positiveness of the conditional mean, λ_i is usually parametrized as $\lambda_i = \exp(X_i\beta)$ where $X_i = [\overset{*}{X}_i, I_i]$. $\overset{*}{X}_i$ denotes a k-1 row vector including the variables that influence visits to specialists, independently of the insurance choice. That is $\overset{*}{X}_i$ includes the following variables of the theoretical model: S_i, $\widehat{\varepsilon}_i$, Y_i, P_i, p_3. The variable I_i will take the value of 0 if the individual does not enjoy duplicate coverage and 1 if she does. β denotes a k column vector. Maximum likelihood estimation of this model yields consistent and efficient estimations, provided the assumption underlying the model, specifically the functional forms and the distributional assumptions, are true.

2.5.2 Estimation in the presence of endogeneity

When endogeneity arises, that is, $E\left[I_i\eta_i|\overset{*}{X}_i, Z_i\right] \neq 0$, Maximum Likelihood estimation of the equation of the number of visits would yield inconsistent estimates. The theoretical model described above shows that one should expect

$$E\left[I_i\eta_i|\overset{*}{X}_i, Z_i\right] \neq 0, \tag{2.14}$$

since the unobserved heterogeneity η_i is present in both equations (2.6) and (2.10)

This section describes how to obtain consistent estimates for the demand for visits taking into account the endogeneity. The equations of the number of visits will be estimated using the Generalized Method of Moments (GMM) (Hansen 1982). These techniques has been applied before by Mullahy (1997) in a model of cigarette smoking behavior and by Windmeijer and Santos Silva (1997), in a model of GP visits. Windmeijer and Santos Silva study the number of visits in the previous 30 days to the interview. They consider the endogeneity of the short term self reported health index[1], while the long term health variables are considered exogenous. Since we study the number of visits in a year, we do not consider short term health measures. As them we consider exogenous the long term health index[2].

The demand for visits to doctors, given by (2.6) suggests the following econometric model for the total number visits:

$$v_i = \exp\left(\overset{*}{X}_i\overset{*}{\beta} + I_i\phi\right) + u_{1i}. \tag{2.15}$$

where the exponential function is used to ensure the positiveness of the demand for visits and u_{1i} refers to the error term of the econometric specification that includes unobservable variables. Mullahy (1997) has suggested a multiplicative specification for the error term u_{1i}. We will comment on this later.

[1]Their argument is that people's perception about their short term health index might be affected by their recent visits to the doctor.

[2]It would be desirable to treat all the health variables as potentially endogenous. However, this a difficult task given the problem of weak instruments that arises even in panels of moderate temporal dimension when regressors are highly persistent. See Blundell and Bond (1998).

Equation (2.10) describes the insurance choice as a discrete choice problem. Assuming linearity of the function

$$\Delta U(S_i, Z_i, \eta_i, s, Y, P, p_3) = X_i\alpha + Z_i\gamma + \eta_i \tag{2.16}$$

the econometric model for the insurance choice will be

$$\begin{aligned} \overset{*}{I_i} &= X_i\alpha + Z_i\gamma + u_{2i} \\ I_i &= 1 \;\; for \quad \overset{*}{I_i} \geq 0 \\ I_i &= 0 \qquad otherwise, \end{aligned} \tag{2.17}$$

where u_{2i} also captures η_i. A problem with the estimation of (2.17) is the fact that an important variable P_i, the insurance premium, is only available for those who have private insurance. Since this equation is not the main aim of our analysis, we will assume a reduced form for it.

In what follows, we consider the estimation of (2.15). Since the unobservable variable η_i is captured by both u_{1i} and u_{2i} then $E\left[u_{1i}u_{2i}\right] \neq 0$. A possible solution could be to assume that u_{1i} has a multiplicative effect on the exponential term of (2.15) and then take natural logarithms of both sides of the equation. Then, we could apply linear instrumental variables estimation. However this solution is ruled out by the fact that there are many observations with no visits to doctors, and the value of the logarithmic is not defined.

In order to estimate (2.15) by GMM, one needs to specify a moment condition. In this case, it is natural to choose

$$E[(v_i - \exp(X_i\beta))\,|W_i] = 0, \tag{2.18}$$

where W_i is a vector of valid instruments which are correlated with insurance but not with the unobservable variables η_i. The validity of this instruments can be tested using a overidentifying restriction test. The GMM method consists on minimizing a weightened distance of the unconditional moments conditions based on (2.18). More information about this method can be obtained from standard textbooks in econometrics (Davidson and Mackinnon 1993, Hamilton 1994).

2.5.3 Elasticities

Since the equation for the number of visits is a nonlinear one, the coefficients of the model require careful interpretation (Cameron *et al.* 1988). The interpretation is slightly different for dummy variables, such as duplicate coverage, and for continuous variables. Consider how to interpret the dummy variable of duplicate coverage, I_i. Since $E[v_i] = \exp\left(\overset{*}{X}_i\overset{*}{\beta} + I_i\phi\right)$, then

$$E[v_i|I_i = 1]/E[v_i|I_i = 0] = \exp(\phi) \simeq 1 + \phi, \text{for } \phi \text{ small enough.}$$

so ϕ can be interpreted as the proportionate increase in the mean of the visits owing to the duplicate coverage effect. That is to say, if ϕ were 0.2, then on average individuals with duplicate coverage would have twenty per cent more visits. For continuous variables such as, age or family size, the following relation holds (Mullahy 1998):

$$\frac{dv_i}{dx_{ik}}\frac{x_{ik}}{v_i} = \beta_k x_{ik}$$

that is, the elasticity of v_i with respect to x_{ik} is linear with coefficient β_k.

2.5.4 Discussion of variables

Table 2.2 shows the definitions of the variables that have been used in the estimation. According to equation (2.6) of the theoretical model, the determinants of the number of visits are variables that affect individual's health (S_i), the health shock, income, the unobservable variable η_i, the perceived quality of the service and the private doctors fee. Our data do not include measures of the last two. Instead we use variables that refer to the health region and town size where the individual lives. Although the health shock is mostly unobservable to the econometrician, we can try to control much of it using dummies variables for accidents and limited activity episodes due to chronic illnesses. Individual's health perception is measured by a self assessed health index, dummy variables for individuals with chronic conditions and disability, and a variable for the number of chronic conditions. Since income is represented by a categorical variable, we also included the number of members in the household. Since one third of the sample did not answer the income question, we have not deleted these observations. We have included a dummy variable for individuals that has not answered the income question. We have also included age, sex, education, head of household, and employment status (self employed, employee...) as determinants of the number of visits.

The variables that will be used as instruments are very important. These are the variables that affect health care utilization only indirectly, that is through the health insurance choice. The variables Z_i in the theoretical model satisfy this condition. They only affect the likelihood of occurrence of health shock. We have chosen reported social class and occupation (manager, administrative, blue collar...) as candidates to be included in Z_i. We think that these variables do not affect utilization, once one has controlled for other variables, specially insurance, health status, education and income. We have also included as instruments some interaction terms and the prediction of the duplicate coverage variable using a logit model (Windmeijer and Santos Silva 1997).

2.6 Econometric results

The models were estimated using a subsample of those aged between 18 and 59. People above 60 have problems to buy private medical insurance since insurers will normally reject them. After deleting observations corresponding to people that did not answer relevant questions the subsample size is 7281 individuals.

It is interesting to compare GMM estimates with Negbin-2, since the former is consistent under endogeneity while the last would be inconsistent. Table 2.3 shows the results for the estimation using the whole sample. We see that the estimation of the duplicate coverage effect is quite different between the models (0.28 for the Negbin-2 and -0.03 for the GMM). This is consistent with endogeneity bias. However, given the large standard error of the GMM estimation, a Hausman test based on this coefficient would not reject the null hypothesis of exogeneity. On the contrary, the Hausman test for all the coefficients did reject exogeneity with a P-value of 1.44*1e-12.

It is interesting to split the sample between heads of household and non-heads of household. Results are presented in Tables 2.4 and 2.5. These show how the evidence of endogeneity is much stronger for the subsample of heads of household. The coefficient for the duplicate coverage effect diverges quite a lot between the Negbin-2 (0.359) and the GMM (-0.514) for the heads of household -which is also consistent with endogeneity bias- but they are almost the same for the subsample of non-heads of household (0.25 for GMM and 0.27 for Negbin-2). Again, due to the large standard error for the GMM estimates, a Hausman test for the duplicate coverage coefficient does not reject exogeneity. However, the Hausman test for all the coefficients for the subsample of heads of household take a P-value of 2.52*1e-9, rejecting exogeneity, while for the subsample of non-heads of household does not reject it at the 80% of confidence, since the P-value is 0.23. The estimate for duplicate coverage is not the only one that differs between GMM and Negbin-2 for the heads of household, other differences include NumChron (1.35 vs 0.47), Sex (0.6 vs 0.12) and Incapacity (0.69 vs -0.00).

The P-values for the test of overidentifying restrictions of the GMM models are 0.62, 0.45, 0.89. This is a general test for the specification of the model, and therefore the validity of instruments. The hypothesis of correct specification is not rejected, which suggests that the models are reasonably well specified, and the instruments are valid. We have also estimated by GMM models where the error term u_{1i} is multiplicative in the moment condition (2.15) as has been suggested by Mullahy (1997). In this case the P-values of the overidentifying restriction test is 0.14 for the non-heads of household and 0.048 for heads of household. Therefore, the specification of the multiplicative model for heads of household is rejected, while it is not with the additive specification. Windmeijer and Santos Silva (1997) have argued that this may also be taken as an indication of endogeneity.

The results above show that endogeneity is much more important for heads of household than others. The explanation for this is very intuitive. The endogeneity occurs because the same unobservable variables that affected the insurance choice, also influenced the number of visits. However it is very likely that a non-head's insurance status is determined by the head of household's unobservable variables. This will occur when the head of household takes the insurance decision for all the members of the family and either other members' unobservable variables are unknown or are not take in into account.

From the results it seems that endogeneity is not important for non-heads of household. In this case, it seems plausible to accept the results given by the Negbin-2 for the subsample of non-heads of household.

These results show that the duplicate coverage effect implies an increase of about 27% in the average of number of visits to specialists, and this effect is statistically different from zero at usual confidence levels. All the variables related to health such as, chronic illnesses and self assessed health have an important effect on the demand for visits. The variables related to health shocks, such as accidents or limitation in usual activities due to chronic illnesses are also very important. Women go more often to the doctor that men. When interpreting the negative effect of age on the number of visits, it should be taken into account that we have already taken a large set of health related variables into account, therefore age seems to measure an experience effect, which is likely to affect the choice between GP and specialist. Those who have undertaken university studies go more often to the specialist. This could be due to their general knowledge which could enable them to identify the specialist they need. Income seems to be important. People with higher income go more often to the specialist. However, we have to interpret this result with caution due to the categorical nature of the variable and the missing values. The number of members in the household has a negative impact on visits to specialists, perhaps reflecting the importance of income-per capita. Although we do not present the results for regions and size of the town, their effect is also important when determining the number of visits to specialists.

Given that endogeneity seems an important problem for the subsample of heads of household, the estimates for the Negbin-2 could be inconsistent. On the other hand, the GMM estimates have large standard errors. In this case, only some of the estimates for variables related to health conditions are statistically significant.

Although they are not the main goal of the paper, we will briefly comment on the discrete choice model for the duplicate coverage decision. We have estimated a logit model that has been used as instrument for the GMM. According to the theoretical model, the determinants of the insurance choice decision are the same as the determinants of the number of visits (except those regarding the health shocks as limitation in activity or accidents, which are unknown when buying the insurance) and in addition the variables in Z_i, that determine the likelihood of occurrence of health shocks. As we have argued above, social class and occupation are our candidates. The basic conclusion is that insurance decision is more determined by socioeconomic variables as income, education, social class, occupation and employment status than by health related variables. This is consistent with the results found by Cameron *et al.* (1988) for Australia. Although it seems to imply the absence of adverse selection, this is only for observable variables, while the overestimation of insurance coverage found above gives evidence, for the subsample of heads of household, of adverse selection on unobservable variables. Ettner (1997) also finds evidence of adverse selection on unobservable variables, when logit results were inconclusive for observable variables. The effect of higher income, education, social class and to be self employed increases the probability of a person having duplicate coverage. For the subsample of non-heads of household, those individuals with a self employed head of household with university education have a larger probability of having duplicate coverage. This shows the importance of the head of household's characteristics in determining the whole family's insurance coverage.

2.7 Extensions

As we have acknowledged in the paper, the incentives faced to visit a specialist differ according to individual's insurance status. That is why we focused on studying visits to specialists. However, an analysis of utilization of emergency care, GP visits and hospitalization episodes might provide new insights about the incentives of the system. Another issue that merits attention is how complementary or substitutive are the different types of medical care available to the individual. People without free direct access to the specialist (for instance, people without duplicate coverage) might use emergency care to obtain care from an specialist without visiting a GP.

2.8 Conclusions

In this paper we have investigated how duplicate coverage influences both the demand for visits to doctors and the selection mechanism of the market.

The specification of the econometric framework has been motivated by a theoretical model of consumer demand. Since the unobservable information was explicitly modelled, we could get some insights about the endogeneity of the insurance choice decision and the role of the instruments for the econometric specification.

We have established a hypothesis about the selection mechanism in the private insurance market. We argue that the existence of a compulsory public insurance may prevent screening by private insurers. Therefore, people that buy a private insurance are likely to be, ceteris paribus, the ones with poor unobservable health conditions.

We have estimated econometric models of count data by Maximum Likelihood and Generalized Method of Moments to account for the non-negativity of the number of visits. GMM estimation also allows us to control the endogeneity of the insurance choice decision. We have split the sample in a subsample of heads of household and others.

The results show that endogeneity is important for heads of household, while it is not a serious problem for others. This is consistent with the idea that heads of household take the insurance decision for all the members of the family.

The model for non-heads of household shows evidence of a positive effect of duplicate coverage on visits to specialists. However, contrary to the empirical literature in this field, we do not interpret this effect as moral hazard. The results show that variables related to health status (acute illness, disabilities, chronic illnesses...) have an important impact in determining the number of visits to doctors, but not the insurance choice. Income, and education are important in determining both demand for visits and the insurance decision choice. Other socioeconomic variables such as social class and employment status are important to determine the insurance choice decision. The GMM estimates have large standard errors and only estimates of health related variables are statistically different from zero at a 95% confidence level.

2.9 Tables

Table 2.1. Visits to doctors by people that enjoys Duplicate Coverage in the last 15 days of the interview

General Practitioner			Specialists*		
Private insurance	Public system	Private not in catalogue	Private insurance	Public system	Private not in catalogue
23.8 %	67.3 %	8.9 %	55.92 %	22.72 %	21.36 %

Note: * specialist do not include either dentists or psychologists.

Table 2.2. Definitions of variables

Endogenous variables	
Visits	Number of visits to a specialist doctor, within the twelve months previous to the interview
Dc	1= for those that enjoys both public and private insurance, 0 for those with only public insurance
Exogenous variables	
Health Status	
Hea	Self assessed long term health. The dummy variables are Hea1 (very good), Hea2 (good), Hea3 (average), Hea4 (bad). Excluded category: Excellent health.
Chron	1 for those with chronic condition(s), 0 otherwise.
NumChron	Number of reported chronic conditions, divided by 10.
Limact1	1 for those limited in activity to work due to a chronic condition, 0 otherwise.
Acc	1 for those who have suffered an accident, 0 otherwise.
Disab	1 for those with some disability, 0 otherwise.
Socioeconomic	
Age	Age divided by 100.
Sex	1 for females, 0 for males.
Pri	1 for heads of household, 0 otherwise.
Edu	Maximum level of education completed. The dummy variables are Edu1 (primary), Edu2 (secondary) Edu3 (university). Excluded category: Without studies
Self	1 for self employed, 0 otherwise.
Employee	1 for employee, 0 otherwise.
Housewife	1 for housewife, 0 otherwise.
Incapacity	1 for those with incapacity to work, 0 otherwise.
Size	Number of members in the household, divided by 10.
Inc	Annual gross household income. The dummy variables Inc1 (between 6024 and 12048 euros), Inc2 (between 12048 and 18072 euros), Inc3 (between 18072 and 30120 euros), Inc4 (between 30120 and 42168 euros), Inc5 (more than 42168 euros). Excluded category: less than 6024 euros
Incmiss	1 for not reported income, 0 otherwise.

Continuation Table 2.2

LabRisk	1 for those who have reported to suffer labour risk, 0 otherwise
Class	Self reported social class. The dummy variables are Class1 (low-medium), Class2(medium), Class3 (medium-high or high). Exluded category: low social class.
Occup	Head of household's last occupation. The dummy variables are Occup1 (executive,technical), Occup2 (administrative, sales), Occup3 (low level administrative), Occup4 (Head of blue collars). Excluded category: blue collars, farmers.
PriEdu	Maximum level of education completed by the head of household. The dummy variables are PriEdu1 (primary), PriEdu2 (secondary), PriEdu3 (university).Excluded category: Without studies.
PriSelf	1 if the head of household is self employed, 0 otherwise.
PriEmpl	1 if the head of household is an employee, 0 otherwise.
PriHous	1 if the head of household is housewife, 0 otherwise.
PriInca	1 for those with household principal with incapacity to work, 0 otherwise.
ToInh	Number of inhabitants in the town. Dummy variables for ToInh1 (between 2000 and4999 inhabitants), ToInh2 (between 5000 and 9999 inhabitants), ToInh3 (between10000 and 24999) ToInh4 (between 25000 and49999 inhabitants), ToInh5 (between 50000 and 99999),ToInh6 (between 100000 and 500000 inhabitants) and ToInh7 (more than 500000 inhabitants). Exluded category: less than 2000 inhabitants.
Regions	6 Dummy variables for each Health Region.

Table 2.3. Estimates for the whole sample

	Negbin-2 for visits	**Gmm for visits**	**Logit for Dc**
Constant	-0.9709* 0.200	-0.6966* 0.257	-2.3439 0.364
Dc	0.2874* 0.045	-0.0377 0.326	- -
Hea1	0.0504 0.0968	0.1003 0.111	0.0526 0.137
Hea2	0.2409* 0.0882	0.2545* 0.097	-0.0527 0.125
Hea3	0.5643* 0.0998	0.6078* 0.113	-0.1156 0.152
Hea4	0.7240* 0.1445	0.8402* 0.170	-0.4475 0.292
Chron	0.3529* 0.0527	0.3712* 0.068	0.0086 0.083
Numchron	0.8532* 0.1445	0.5166* 0.163	0.0418 0.258
Limact1	0.2066* 0.089	0.1327 0.122	- -
Limact2	0.8068* 0.081	0.7168* 0.121	- -
Acc	0.4914* 0.053	0.2714* 0.063	- -
Disab	0.3231* 0.081	0.2509* 0.103	0.0844 0.159
Age	-1.2942* 0.215	-0.9257* 0.349	2.3236 0.358
Sex	0.5904* 0.052	0.4160* 0.070	0.0615 0.091
Pri	0.0283 0.059	0.0673 0.085	-0.0733 0.100
Edu1	0.0176 0.081	0.0778 0.127	0.6258* 0.020
Edu2	0.0238 0.093	0.1087 0.173	1.0544* 0.225
Edu3	0.0745 0.108	0.2636 0.240	0.8552* 0.257

Continuation Table 2.3

	Negbin-2 for visits	**Gmm for visits**	**Logit for Dc**
Self	-0.0078 0.072	-0.1115 0.098	0.2140 0.142
Employee	-0.0005 0.056	-0.1250 0.071	0.1273 0.094
Housewife	0.0923 0.069	-0.0418 0.092	-0.0996 0.123
Incapacity	0.7807* 0.131	0.4283* 0.187	-0.0471 0.331
Size	-0.4361* 0.152	-0.3427 0.183	-1.5151* 0.255
Inc1	0.0360 0.075	0.0386 0.102	0.3364* 0.155
Inc2	0.0534 0.082	-0.0173 0.115	0.6017* 0.163
Inc3	0.1933 0.099	0.1858 0.139	1.0534* 0.181
Inc4	0.1068 0.165	0.1133 0.171	1.0881* 0.232
Inc5	0.4253* 0.165	0.3800 0.201	1.4452* 0.270
Incmiss	0.0179 0.078	0.0676 0.109	0.7652* 0.157
LabRisk	0.0137 0.043	0.0496 0.056	-0.0507 0.071
Class1	- -	- -	0.2563 0.189
Class2	- -	- -	0.6994* 0.179
Class3	- -	- -	1.0085* 0.223
Occup1	- -	- -	0.5987* 0.160
Occup2	- -	- -	0.4433* 0.143
Occup3	- -	- -	0.3186* 0.101
Occup4	- -	- -	0.2248 0.158

Continuation Table 2.3

	Negbin-2 for visits	**Gmm for visits**	**Logit for Dc**
PriEdu1	- -	- -	0.0386 0.164
PriEdu2	- -	- -	0.3567 0.184
PriEdu3	- -	- -	0.5186* 0.211
PriSelf	- -	- -	0.6621* 0.114
PriEmpl	- -	- -	-0.1956* 0.098
PriHous	- -	- -	0.1627 0.217
PriInca	- -	- -	-0.2865 0.238
Incmis x Edu3	- -	- -	0.4861* 0.197
Self x Sex	- -	- -	-0.1609 0.174
Alfa	1.6019* 0.0488	- -	- -
Sample Size	7281	7281	7281
LogL	-11371.65	-	-3412.37
R2	0.22	-	-
Overidentifying Test	-	Test= 14.57 DF =17 P-value = 0.6263	-

Note. * denotes the estimate is statistically significantly different from zero at a confidence level of 95%. Standard errors are at the bottom row of each cell. City and regions were also included in the estimation, results are available from the author. R^2 is based on deviance residuals Cameron and Windmeijer (1996).

Table 2.4. Estimates for subsample heads of household

	Negbin-2 for visits	**Gmm for visits**	**Logit for Dc**
Constant	-1.4222* 0.392	-0.6287 1.706	-2.4439* 0.577
Dc	0.3597* 0.088	-0.5141 1.032	- -
Hea1	-0.0270 0.200	-0.0301 0.480	0.1709 0.235
Hea2	0.2143 0.180	0.1604 0.490	0.0878 0.211
Hea3	0.5075* 0.200	0.4983 0.494	0.0900 0.251
Hea4	0.4952 0.282	0.5232 0.566	-0.2239 0.467
Chron	0.4644* 0.107	0.5097* 0.131	0.1878 0.135
Numchron	1.3512* 0.286	0.4715 0.323	-0.3724 0.427
Limact1	0.4756* 0.160	0.3495 0.199	- -
Limact2	0.7509* 0.142	0.7439* 0.163	- -
Acc	0.8510* 0.099	0.5163* 0.134	- -
Disab	0.6382* 0.140	0.5950* 0.178	0.2487 0.239
Age	-0.9632* 0.444	-0.5178 1.057	1.9208* 0.589
Sex	0.6058* 0.123	0.1283 0.130	0.1943 0.193
Edu1	-0.0527 0.142	-0.1396 0.221	0.6268* 0.267

Continuation Table 2.4

	Negbin-2 for visits	**Gmm for visits**	**Logit for Dc**
Edu2	0.0127 0.169	0.5177 0.240	1.2818* 0.288
Edu3	-0.0781 0.196	-0.1241 0.260	1.1096* 0.333
Self	-0.1891 0.138	-0.2235 0.212	0.6843* 0.201
Employee	-0.0969 0.123	-0.2874 0.194	-0.0748 0.182
Housewife	-0.1825 0.242	0.0227 0.332	-0.2404 0.402
Incapacity	0.6965* 0.203	-0.0091 0.273	-0.7022 0.454
Size	0.1120 0.292	-0.0073 0.339	-1.2316* 0.405
Inc1	0.0981 0.137	0.1335 0.174	0.3018 0.219
Inc2	0.0738 0.151	0.0449 0.204	0.6517 0.233
Inc3	0.1764 0.179	0.3378 0.265	1.1218* 0.260
Inc4	0.0314 0.261	-0.4649 0.295	0.8241* 0.334
Inc5	0.3387 0.282	0.6038 0.473	1.5297* 0.385
Incmiss	0.1481 0.145	0.2553 0.196	0.6368* 0.232
LabRisk	-0.0845 0.077	-0.0034 0.097	-0.0643 0.104

Continuation Table 2.4

	Negbin-2 for visits	**Gmm for visits**	**Logit for Dc**
Class1	- -	- -	0.0353 0.271
Class2	- -	- -	0.7100* 0.258
Class3	- -	- -	1.2622* 0.341
Occup1	- -	- -	0.5501* 0.254
Occup2	- -	- -	0.4328 0.225
Occup3	- -	- -	-0.0202 0.166
Occup4	- -	- -	-0.0448 0.251
Incmis x Edu3	- -	- -	0.1279 0.350
Self x Sex	- -	- -	-0.1859 0.380
Alfa	2.0667* 0.115	- -	- -
Sample Size	2688	2688	2688
LogL	-3589.27	-	-1327.35
R2	0.28	-	-
Overidentifying Test	-	Test= 13.97 DF =14 P-value = 0.451	-

Note. * denotes the estimate is statistically significantly different from zero at a confidence level of 95%. Standard errors are at the bottom row of each cell. City and regions were also included in the estimation, results are available from the author. R^2 is based on deviance residuals Cameron and Windmeijer (1996).

Table 2.5. Estimates for subsample of non-heads of household

	Negbin-2 for visits	**Gmm for visits**	**Logit for Dc**
Constant	-0.8507* 0.224	-0.7570* 0.292	-2.3201* 0.496
Dc	0.2762* 0.053	0.2565 0.310	– –
Hea1	0.0830 0.109	0.1217 0.124	-0.0087 0.172
Hea2	0.2667* 0.100	0.2522* 0.112	-0.1221 0.158
Hea3	0.6344* 0.114	0.6044* 0.132	-0.2218 0.194
Hea4	0.9928* 0.168	0.8953* 0.196	-0.5422 0.381
Chron	0.3060* 0.060	0.2964* 0.078	-0.1007 0.108
Numchron	0.7119* 0.164	0.6076* 0.183	0.3210 0.330
Limact1	0.1267 0.111	0.0603 0.111	– –
Limact2	0.7596* 0.102	0.7451* 0.106	– –
Acc	0.3281* 0.064	0.1519* 0.072	– –
Disab	-0.0625 0.099	-0.0628 0.117	-0.0466 0.218
Age	-1.4781* 0.257	-1.0688* 0.3689	2.4542* 0.478
Sex	0.5893* 0.061	0.4970* 0.090	0.1225 0.113
Edu1	0.1186 0.097	0.2038 0.107	0.7107* 0.257
Edu2	0.1247 0.109	0.1891 0.1479	1.2240* 0.276
Edu3	0.2613* 0.129	0.3903* 0.163	1.0650* 0.319

Continuation Table 2.5

	Negbin-2 for visits	Gmm for visits	Logit for Dc
Self	0.1686 0.092	0.0199 0.107	0.8228* 0.257
Employee	0.0343 0.062	-0.0492 0.0789	0.1138 0.110
Housewife	0.1843* 0.073	0.0385 0.0915	-0.0441 0.140
Incapacity	0.6286* 0.184	0.4796* 0.210	0.3116 0.434
Size	-0.7433* 0.181	-0.6332* 0.242	-1.8097* 0.341
Inc1	0.0186 0.092	0.0368 0.114	0.3826 0.223
Inc2	0.0325 0.099	-0.0291 0.118	0.5778* 0.233
Inc3	0.2411* 0.120	0.1766 0.148	1.0385* 0.258
Inc4	0.2130 0.174	0.1462 0.208	1.3351* 0.328
Inc5	0.4337* 0.212	0.2610 0.232	1.3603* 0.389
Incmiss	-0.0237 0.092	0.0006 0.114	0.8518* 0.222
LabRisk	0.0210 0.052	0.0416 0.064	-0.0869 0.099
Class1	– –	– –	0.2276 0.265
Class2	– –	– –	0.7083* 0.251
Class3	– –	– –	0.8633* 0.303
Occup1	– –	– –	0.6111* 0.210
Occup2	– –	– –	0.4520* 0.189
Occup3	– –	– –	0.5238* 0.129
Occup4	– –	– –	0.4074* 0.2063

Continuation Table 2.5

	Negbin-2 for visits	**Gmm for visits**	**Logit for Dc**
PriEdu1	- -	- -	-0.0351 0.178
PriEdu2	- -	- -	0.3209 0.202
PriEdu3	- -	- -	0.6147* 0.232
PriSelf	- -	- -	0.7440* 0.129
PriEmpl	- -	- -	-0.2239 0.115
PriHous	- -	- -	0.2598 0.271
PriInca	- -	- -	-0.1309 0.261
Incmis x Edu3	- -	- -	0.5755* 0.250
Self x Sex	- -	- -	-0.7717* 0.288
Alfa	1.3588* 0.048	- -	- -
Sample Size	4593	4593	4593
LogL	-7689.14	-	-2058.65
R2	0.19	-	-
Overidentifying Test	-	Test= 10.28 DF =17 P-value = 0.89	-

Note. * denotes the estimate is statistically significantly different from zero at a confidence level of 95%. Standard errors are at the bottom row of each cell. City and regions were also included in the estimation, results are available from the author. R^2 is based on deviance residuals Cameron and Windmeijer (1996).

Bibliography

[1] Álvarez, B (2001), La Demanda Atendida de Consultas Médicas y Servicios Urgentes en España. *Investigaciones Económicas*, 25.

[2] Belsey, T.J. (1988), Optimal Reimbursement Health Insurance and the Theory of Ramsey Taxation. *Journal of Health Economics*, 7, 321-336.

[3] Blundell, R. and Bond, R. (1998), Initial Conditions and ;oment Restrictions in Dynamic Panel Data models. *Journal of Econometrics, 87, 115-143.*

[4] Cameron, A.C. and Trivedi, P.K. (1986), Econometric Models Based on Count Data: Comparisons and Applications of some Estimators and Tests *Journal of Applied Econometrics*,1, 29-53.

[5] Cameron, A.C., Trivedi, P.K., Milne, F, and Pigott, J. (1988), A Microeconometric Model of the Demand for Health Care and Health Insurance in Australia. *Review of Economic Studies,* 55, 85-106.

[6] Cameron, A.C. and Windmeijer, F.A.G. (1996), R-Squared Measures for Count Data Regression Models with Applications to Health Care Utilization. *Journal of Business and Economics Statistics*, 14, 209-220.

[7] Coulson, N.E., Terza, J., Neslulan, C.A. and Stuart, C.B. (1995), Estimating the Moral-Hazard Effect of Supplemental Medical Insurance in the Demand for Prescription Drugs by the Elderly. *American Economic Association Papers and Proceedings,* 85, 122-126.

[8] Chiappori, P.A., Durand, F. and Geoffard, P.Y.(1998), Moral Hazard and the Demand for Physician Services: First Lessons from a French Natural Experiment. *European Economic Review,* 42: 499-511.

[9] Davidson, R. and J. G. Mackinnon. (1993), *Estimation and Inference in Econometrics,* Oxford: Oxford University Press.

[10] Ettner, S. L. (1997), Adverse Selection and the Purchase of Medigap Insurance by the Elderly. *Journal of Health Economics,* 16, 543-562.

[11] Grogger, J. A (1990), Simple Test for Exogeneity in Probit, Logit and Poisson Regression Models. *Economic Letters,* 33, 329-332.

[12] Hamilton, J.D. (1994) *Time Series Analysis,* New Jersey: Princeton University Press.

[13] Hansen, L. P. (1982), Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica,* 50, 1029-54.

[14] Hausman, J.A., Hall B. H. and Griliches, Z. (1984), Econometric Models for Count Data with Applications to the Patents-R&D Relationship. *Econometrica,* 52, 909-38.

[15] Heckman, J. (1978), Dummy Endogenous Variables in a Simultaneous Equation System, *Econometrica*, 46, 931-959.

[16] Holly, A., Gardiol, L., Domenighetti, G., Bisig, B. (1998), An Econometric Model of Health Care Utilization and Health Insurance in Switzerland. *European Economic Review,* 42, 513-522.

[17] Machado, M. Dollars and Performance: Cost-Effectiveness of Substance Abuse Treatment in Maine. *Journal of Health Economics.* Forthcoming.

[18] Manning, W.G., Newhouse, J. P. *et al.* (1987), Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment. *The American Economic Review,* 77, 251-277.

[19] Meza, D. (1983), Health Insurance and the Demand for Health Care. *Journal of Health Economics,* 2, 47-54.

[20] Mullahy, J. (1997), Instrumental-Variable Estimation of Count Data Models: Applications to Models of Cigarette Smoking Behavior. *The Review of Economics and Statistics,* 79, 586-593.

[21] Mullahy, J. (1998), Much Ado about Two: Reconsidering Retransformation and the Two-Part Model in Health Economics. *Journal of Health Economics*, 17, 247-281.

[22] Murillo, C. (1992), La Industria del Seguro Privado y los Seguros Sanitarios. *Revista de Economía,* 13, 84-90.

[23] Rothschild, M. and Stiglitz, J. (1976), Equilibrium in Competitive Insurance Markets: an Essay on the Economics of Imperfect Information. *Quaterly Journal of Economics* 1976, 90, 629-650.

[24] Servei Català de la Salut (1994), *Enquesta de Salut de Catalunya.* Barcelona

[25] Whitaker, D. and Sánchez, P. L. (1997), *Sistema Sanitario Español: Alternativas para su Reforma.* Madrid: National Economic Research Associates.

[26] Windmeijer, F.A.G. and Santos Silva, J.M.C. (1997), Endogeneity in Count Data Models: An Application to Demand for Health Care. *Journal of Applied Econometrics,* 12, 281-294.

[27] Winkelmann, R. and Zimmermann, K.F.(1995), Recent Developments in Count Data Modelling: Theory and Application. *Journal of Economic Surveys,* 9, 1-23.