Banco de Portugal
EUROSISTEMA

Estudos e Documentos de Trabalho

Working Papers

11 | 2009

DOUBLE COVERAGE AND HEALTH CARE UTILISATION: EVIDENCE FROM QUANTILE REGRESSION

Sara Moreira
Pedro Pita Barros

November 2009

Please address correspondence to
Sara Moreira
Economics and Research Department
Banco de Portugal, Av. Almirante Reis no. 71, 1150-012 Lisboa, Portugal;
Tel.: 351 21 313 0279, spmoreira@bportugal.pt

BANCO DE PORTUGAL

Edition

Economics and Research Department

Av. Almirante Reis, 71-6th

1150-012 Lisboa

www.bportugal.pt

Pre-press and Distribution

Administrative Services Department

Documentation, Editing and Museum Division

Editing and Publishing Unit

Av. Almirante Reis, 71-2nd

1150-012 Lisboa

Printing

Administrative Services Department

Logistics Division

Lisbon, November 2009

Number of copies

170

ISBN 978-989-8061-87-4

ISSN 0870-0117

Legal Deposit No 3664/83

Double coverage and health care utilisation: Evidence from quantile regression *†

Sara Moreira ‡ Pedro Pita Barros

Banco de Portugal, ISEG-TULisbon Universidade Nova de Lisboa

Abstract

An individual experiences double coverage when he benefits from more than one health insurance plan at the same time. This paper examines the impact of such supplementary insurance on the utilisation of health care. Its novelty is that within the context of count data modelling and without imposing restrictive parametric assumptions, the analysis is carried out for different points of the conditional distribution, not only for its mean location.

We use data for Portugal on the consumption of doctor visits, taking advantage of particular features of the public and private protection schemes on top of the statutory National Health Service. Results indicate that double coverage generates additional utilisation of health care and, even though it is present in the whole outcome distribution, by looking at different points we unveil that the effects are relatively smaller for more frequent users. Another interesting finding regards the source of supplementary insurance since although both public and private second layers of health care protection increase the utilisation of doctor visits, it adds more to the consumption when provided by private organizations.

Keywords: Demand for health services, Moral hazard, Count data, Quantile regression

JEL codes: I11, I18, C21, C25

*This article draws heavily on the master thesis of Sara Moreira in "Applied Econometrics and Forecasting", performed at Instituto Superior de Economia e Gestão. A reduced version of this study was awarded with the prize for best paper in the 11th Portuguese Health Economics Conference.

†We thank Anne Nolan and the participants of the Eighteenth European Workshop on Econometrics and Health Economics for helpful suggestions. We are also grateful to Nuno Alves, Cláudia Braz, Mário Centeno, Jorge C. Cunha, Adeline Delavande, Ferreira Machado, Pedro Portugal, João Santos-Silva and A. Marcos Vera-Hernández for their comments. All errors and omissions remain our responsibility.

‡Corresponding author: Sara Moreira. E-mail: spmoreira@bportugal.pt. Address: Economic Research Department, Banco de Portugal, Av. Almirante Reis, 71, 6^{o}, 1150-012 Lisboa, Portugal.

1 Introduction

The aim of this paper is to analyse the impact of double health insurance coverage (i.e. a situation in which an individual is covered by more than one health insurance plan)[1] on the utilisation of health care. It is well known that if the demand for health care reacts to changes to budget constraints and preferences, thus better or more extensive insurance should also have important effects because it affects price of services, income of the insured and opportunity cost of time in case of illnesses. The impact of double coverage is often associated to an aggravation of moral hazard that creates incentives for people to go to the doctor more frequently and eventually because of less severe illness.[2]

Organizational designs of health care systems may generate layers of coverage. The most common situation regards the case of the individual who benefits from compulsory public insurance and nevertheless decides to additionally purchase a private one. Such supplementary private health insurance usually overlaps the range of health care services provided by the statutory health system. With additional layers of coverage, people can increase the set of available providers, have a faster access and reduce out-of-pocket prices. Quantitatively, double coverage is not a negligible phenomenon. It can be found in all European countries. Furthermore, in the United States, the health reform is expected to increase health coverage, inclusively by allowing Americans to maintain their current insurance scheme while accessing new options. In such scenario, double coverage situations are expected to augment significantly in coming years. Research on this phenomenon can help to detect whether possible inefficiencies and inequalities, causing unnecessary and costly utilization, should be a concern.

Existing works addressing health insurance focus on mean effects. In contrast, by looking at other points of the conditional distribution we unveil if the stronger effects are found among more or less frequent users. Our findings are the result of the application of an innovative technique for estimating the quantile regression for counts. The estimates

[1] The terms "duplicate coverage", "supplementary health insurance" or "additional health insurance" are used alternatively in the literature.

[2] Moral hazard in this context is defined as the "change in health behaviour and health care consumption caused by insurance" (Zweifel and Manning 2000). Some authors criticize the direct association of double coverage with moral hazard, arguing on the existence of other important effects. For instance, Vera-Hernández (1999) refers the impact of insurance on individual' health, which will decrease the future consumption of health care. Also Coulson et al. (1995) points to the importance of supply-inducement by providers.

were computed with Portuguese data, using as source of double coverage the existing health insurance schemes beyond the National Health Service (NHS). Approximately a quarter of the Portuguese population has access to a second layer of health insurance coverage on top of the NHS, through mandatory occupation-based health schemes (usually referred to as health subsystems because of the Portuguese *subsistemas*) and voluntary private insurance. We focus our attention on the double coverage resulting from the former type, regarding both health insurance plans provided to public employees and insurance plans of special private companies/sectors. Results indicate positive effects of subsystems coverage on the utilisation of health care (especially large for private subsystems). An interesting finding, which could only be observed through the use of quantile analysis, is that these effects are relatively more relevant for the first levels of usage since for more frequent users, consumption behaviour seems to depend slightly less on the additional layer of insurance.

We measure health care utilisation through the traditional indicator – number of visits to a doctor – a non-negative integer count characterized by large proportion of zeros and positive skewness. Until recently the one-part and two-part models have dominated the empirical literature (Deb and Trivedi 2002). One-part models explain health care utilisation with a single equation as a function of a set of medical care determinants, usually within the Poisson and negative binomial frameworks. The second approach estimates two equations, one to distinguish between users and non-users and a second to explain the intensity of usage. In this framework the Hurdle models are the most extensively used. The appeal of two-part models comes from the fact that it is both well supported empirically (partly driven by the high incidence of zero usage), and also well connected to a principal-agent model (the individual decides whether to visit the doctor and once the initial contact is made, the doctor influences the decision about the level of usage, especially the referrals). More recent literature relies on a finite mixture variant of the latent class model, where individuals are usually distinguished between frequent and non-frequent users instead of the distinction users/non-users of the two part models. Proponents of these latter models argue that the distinction between frequent/non-frequent is better determined by health status, attitudes towards risk and life-style choices (Deb and Trivedi 2002). The two-part literature argues against this view stating that latent class frameworks are mainly driven by statistical reasoning and are not suitable for cross-sectional data (Jiménez-Martín et al. 2002). Estimators resulting from any of these econometric tools rely on assumptions about the functional form of the regression equation and the distribution of the error term. As a result, standard models determine entirely the distributional behaviour once the conditional mean response is known. An attractive alternative is the usage of nonparametric and semiparametric estimators. Introduced for continuous data in Koenker and Bassett (1978), quantile regression offers a complete picture of the effect of the covariates on the location, scale and shape of the distribution of the dependent variable. As a semiparametric method it assumes a parametric specification for the quantile of the conditional distribution but leaves the error term unspecified. It was first applied to continuous health data in Manning et al. (1995). As in Winkelmann (2006) and Liu (2007), we apply the approach suggested by Machado and Santos-Silva (2005) in which quantile regression is extended to count data through a "jittering" process that artificially imposes some degree of smoothness. This technique allows an important step forward to understand the utilisation behaviour of medical care, enabling, for example, the assessment of the efficiency and equity of a particular system and whether a health reform has a different effect among low and high users. This kind of econometric tool can also be used to analyse the need for adjustments on contracts provided by insurance companies.

Many authors have investigated the impact of supplementary health care protection derived from insurance plans with different levels of coverage (for example Cameron et al. 1988, Coulson et al. 1995, Vera-Hernández 1999, Lourenço 2007 and Barros et al. 2008). The usage of non-experimental data generally creates an endogeneity problem related to adverse selection since most of the times the decision to buy extra health care coverage is an individual choice that is likely to be influenced by its health status and attitudes towards risk. In such cases, the insurance parameter does not disentangle moral hazard and adverse selection effects (also called insurance and selection effect). The solution relies most of the times on finding reasonable instrumental variables. Our empirical application does not have this problem because the membership on public and private health subsystems was mandatory and based on professional category, and as such unrelated to the expected value of future health care consumption (unless we consider issues related with occupational choice). Moreover, contributions are based on income and not on risk characteristics of each individual.

Data was taken from the Portuguese Health Survey of 2005/2006, a cross sectional health dataset that provides a wide range of information at an individual level concerning socioeconomic conditions and health status indicators. After excluding individuals with private voluntary health insurance, observations were divided according to the type of health insurance. We chose three mutually exclusive groups according to their health care coverage: only the NHS, the NHS plus a public subsystem or the NHS plus a private subsystem. Despite having some common features, both public and private groups include several subsystems. To explain the number of doctor consultations besides the health insurance status variable we also control for health status, demographic and socioeconomic condition, seasonal and geographic effects. The variables were selected from questions included in the survey. The selection took into account the Grossman's health capital model (Grossman 1972) as well as the main factors influencing medical care consumption identified in the literature. When using data for Portugal we contribute to the literature on the impact oh health subsystems on the utilisation of health care. The motivation in studying the impact arising of public and private subsystems in the Portuguese health sector is partially related to efficiency and equity concerns. It would have been expected that the creation of a public health system in 1979, prompted the integration of public employees' plans into the NHS, which in fact did not happen. The non-integration of these subsystems raised some equity issues as the NHS provides a less comprehensive health protection plan than those available under public and private health schemes, which ensure a higher level of access to health care services at lower costs. When compared to previous research, our work is more comprehensive covering all the subsystems, while former contributions (e.g. Barros et al. 2008 and Lourenço 2007) focused on the effect of the most important public health scheme.

The paper is structured as follows. Section 2 summarizes some features of the Portuguese health care system. Section 3 describes the dataset and the relevant variables, and presents an exploratory analysis of the data. In Section 4 we introduce the quantile regression for counts as well as other models such as the one-part and two-part parametric count data models, which will be used as a benchmark. In the methodological part we present the techniques and discuss the empirical specification. In Section 5 we analyse the results from the quantile regression framework, along with the results from the benchmark parametric models. Section 6 presents the final remarks.

2 Overview of the Portuguese health care system

The Portuguese health system is a network of public and private health care providers and different funding schemes.[3] It is possible to identify three overlapping layers: the National Health Service (NHS)[4], special public and private employer-provided schemes (subsystems) and private voluntary health insurance. While the NHS is mainly financed by general taxation, subsystems resources come from employees and employers compulsory contributions (including, in the public schemes, State funds to ensure their balance). According to Barros and Simões (2007), in 2004 public funding represented 71.2 per cent of total health expenditure (of which 57.6 per cent is related with the NHS and 7.0 per cent cent with subsidies to public subsystems). Private expenditure is composed by co-payments and direct payments made by patients and, to a lesser extent, by private insurance premiums.

The complexity of the Portuguese health care system has some historical foundations. In 1979, with the creation of the NHS, legislation established that all residents have the right to health protection regardless of economic or social status. Until then, the State only had full responsibility for the health care of public employees. For the remaining population, the State provided limited preventive care and specific health services (such as maternity, child and mental care), and had some interventions in the control of infectious diseases. After the outset of the statutory public system, the health subsystems wer not integrated into the NHS and continued to cover public employes. Barros and Simões (2007) state that these schemes were kept because trade unions, which ran and managed some of the funds, were not willing to give up their privileges and forcefully defended their maintenance on behalf of their members. The existence of some private subsystems is also influenced by this reality because they cover employees of large companies that in the past were nationalized (in the sequence of the mid-70s revolution) and have been later privatized.

Individuals covered solely by the NHS (the majority of the population) face some

[3] This section is mostly based on Barros and Simões (2007) and Lourenço (2007). An interesting comparison between the Portuguese health system and other European systems is available in Bago-d'Uva and Jones (2009).

[4] In the autonomous regions, public health is ensured by regional health services (RHS of Azores and Madeira) following the same principles of the NHS but implemented by regional governments. Here it is not useful to treat them separately.

constraints in the access to public providers, in particular because of services excluded from the public network and difficulties of access due to time costs (long waiting lists and queuing) and geographical barriers. Lourenço (2007) among others, argues that the NHS coverage restrictions convert its normative completeness into an incomplete health insurance contract. In the access to medical consultations, the NHS is designed in a way that beneficiaries should first seek health care through their family doctor (general practitioner) in health care centres and then, if necessary, get appropriate referrals to a public specialist consultation (generally as out-patient consultations in public hospitals). This gatekeeper procedure is not strictly followed since there are households who do not have access to a family doctor and, even when they have, the time lag between the first step to obtain health care and its actual provision is frequently too long. Additionally, the requirements to obtain referrals are generally very demanding. For these reasons, some individuals have their first contact with medical care in hospitals' emergency rooms even if their condition would not require it. The NHS design contemplates a cost-share mechanism that in practice makes some patients (a large share of the population is exempt) pay a mandatory small co-payment to the public provider, usually on a fee-for-service basis. Given the constraints of the public network, the consumption of private services by NHS beneficiaries[5] is very common and, in the absence of private voluntary insurance schemes, they bear the full cost of such services, without being reimbursed afterwards.

A considerable share of the population (between 20-25 per cent) benefits from employer-provided health insurance through several subsystems, either private or public. Occupation-based schemes cover all employees both in working age and retirement, as well as spouses and descendants. Among those health protection plans, the largest public subsystem is a Government department (ADSE) acting as a health insurance provider for public employees, covering about 15 per cent of the population. Several specific schemes also exist, for example, for armed forces personnel. As to private subsystems, they are mainly set up by unions and cover employees of the historic telecommunications operator and postal services, as well as banking and associated insurance employees.

Each subsystem has a distinct array of medical care insurance arrangements to finance and provide health care. As a whole, we can say that they are organized differently

[5]In the course of the paper, when it says "NHS beneficiaries", we consider individuals covered solely by NHS. Therefore, this definition excludes the population with double coverage.

from the NHS, in particular because of the lower proportion of services directly provided. They basically make health care available through contracts with public/NHS and private institutions and reimburse patients costs for services supplied by private entities without contract. These features make these schemes more comprehensive health protection plans than NHS, representing both complementary and supplementary types of insurance (Lourenço 2007). People benefiting from additional health care schemes, either mandatory or voluntary, are not affected in their taxation and remain eligible to receive health care from the NHS. In this context, subsystems have been gradually positioned as mainly complementary to the statutory system. In fact, the main impact of benefiting from double coverage is having "less expensive" access to medical specialities (in particular, by reimbursing part os patients cost in all private providers, even those without contracts) that in practice makes the subsystems beneficiaries more heavy users of such type of consultations, vis-à-vis the general practitioner (GP) consultations pattern among NHS beneficiaries (Comissão para a Sustentabilidade do Financiamento do Serviço Nacional de Saúde 2006). Because of the double coverage and different State funding among subgroups of the population, the existence of these subsystems is often pointed as a factor of inequity within the Portuguese health care system (van Doorslaer and Jones 2004). Table A1 of Appendix A allows for a systematic comparison between the Portuguese non-voluntary health insurance schemes.

3 Data

3.1 Dataset

Data was taken from the fourth Portuguese Health Survey (PHS), a cross sectional health dataset designed to be representative of the Portuguese household population.[6] It provides a wide range of information at an individual level, namely demographic and socioeconomic conditions, type of health insurance, health-care utilisation, health status indicators (such as chronic diseases and long/short run disability), lifestyles (such as eating habits and

[6] PHS is carried out by the Portuguese Ministry of Health in collaboration with the National Health Institute Ricardo Jorge and the National Statistical Institute. Until now, four questionnaires have been made (1987, 1995/1996, 1998/1999 and 2005/2006) using representative probabilistic samples of the continental population (1st, 2nd and 3th PHS) and of both continental and autonomous regions of Azores and Madeira population (4th PHS). Here we made use of the last available questionnaire. Note that it is not a panel survey because the sample changes between surveys.

sports activity) and expenditure levels. The PHS sample reflects the geographical structure of the population according to the 2001 census, resulting from a two-stage cluster sampling that followed a complex design involving both stratification and systematic selection of clusters. The survey was collected by interviews carried out between February 2005 and January 2006, being selected a total of 19,950 households units and in each household all individuals were face-to-face interviewed.

The sample used in this paper comprises 35,308 observations and was obtained after defining the population of interest and handling the data. Firstly, we restrict our population to individuals without voluntary private health insurance and with less than 80 years old.[7] Secondly, we excluded 40 observations of individuals that did not report the number of visits to a doctor and 8 observations without answer regarding the subsystem they belong to. Finally, we deleted further 1,047 observations with missing values for any other relevant variables (according to the set of regressors chosen).

Two points should be made about the latter choices. Firstly, the exclusion of voluntary health insurance individuals can be pointed as a shortcoming. However, the inclusion of such variable could introduce endogeneity problems, which would be difficult to eliminate since there are no suitable instrumental variables (Barros et al. 2008). In this context and given the relatively small number of insured individuals (less than 8 per cent) it seems better to exclude such observations with a cost of restrict the analysis to the population exclusively insured through mandatory schemes. Regarding the exclusion of observations of persons with more than 80 years, the decision was related to the measurement of treatments effects, as explained in due course.

Secondly, the simplest way of handling missing data is to delete them and analyze only the sample of "complete observations" (although deleting observations reduces the efficiency of the estimation). The usage of a listwise deletion procedure is statistically appropriate only if the missing values are "missing completely at random" (Cameron and Trivedi 2005), which means that the probability of missing does not depend of its own value nor on the values of other variables in the dataset (the observed sample is a random subsample of the potential full sample). In our case listwise deletion is clearly acceptable because incomplete observations comprise a small percentage (less than 3 per

[7] We also excluded 145 observations of pregnant women whose visits to the doctor were related to their condition.

cent) of total observations. Moreover, among the relevant questions of our dataset that can create a sample selection problem, the one that generated more missing observations concerns the income level. However, most of the missing (around seventy per cent) does not result from a non answer but from individuals that declare not knowing the household income, which if not deliberately makes unlikely that unobserved factors influenced both the decision to respond and the value of the dependent variable.

Finally, another important feature that is worth noting is that the sample was distributed in order to ensure an adequate geographical distribution, and as a consequence the survey has a weight variable. If we ignore the weights we may obtain biased parameter estimates, certainly in designs where some categories have been oversampled (Wooldridge 2002, Cameron and Trivedi 2005). It is possible to ignore weights without affecting the parameter estimates, in particular when sampling weights are solely a function of independent variables, or when the model can be respecified (including new variables or interactions). This is more likely in the case where the weights are almost uncorrelated with the dependent variable, which occurs in our analysis. Additionally, the variables under the sample design of PHS are included as covariates in our estimation. The problem with the use of a weighted dataset is that it leads to artificially small standard errors for regression coefficients and therefore incorrect inferences on the significance of the different effects.[8]

3.2 The variables

To capture health care utilisation we use the total number of visits to doctors in the three months prior to the interview. The question in the survey was: "How many times did you visit a physician in the last three months?". The survey design conditioned the measure because it is overly aggregated, encompassing consultations to GPs and specialist

[8] The standard errors will usually be too small because there is no i.i.d. property (the independence assumption no longer holds). If one wants to take weighting into account and derives appropriate standard errors for the coefficients, we could include the stratification variable(s) in the regression as interaction terms with independent variables or in loglinear models as an offset variable. The problem arises when the interactions of the stratifying variable(s) with the independent variables are all significant. In this case it will not be a practical method as it will lead to a large number of effects, many of which may not be substantively interesting. An alternative is to derive corrected standard errors using software designed for complex survey designs in which cases with large weight values are treated as a replication of several cases and have a corresponding impact on the standard deviation, even though they are measured with the same accuracy as cases with small weight values. The weighted dataset should determine the estimates, but the unweighted dataset should determine the standard errors. For more details, see Cameron and Trivedi (2005) chapter 24.

doctors, as well as emergency episodes.[9] Another missing piece of information regarding the consultations is related to the nature of its provider, particularly whether it is public or private. Nevertheless, the number of doctor visits is certainly the best available indicator and any adjustment to that measure designed to capture the real amount of resources used would certainly be based on very strong assumptions.

Table I presents the covariates used in our analysis clustered into groups encompassing health insurance status, socioeconomic characteristics and health status. In addition, two further groups were also included to control for geographic and seasonal effects. We selected the variables among the raw data available in the database[10] according to their influence on medical care utilisation, taking into account the Grossman's health capital model of demand for health (1972) as well as the results of similar empirical studies (Cameron et al. 1988, Pohlmeier and Ulrich 1995, Vera-Hernández 1999, Deb and Trivedi 2002, Winkelmann 2004 and Lourenço 2007). It is not straightforward to understand the utilisation of health care since it is a result of both patients and doctors decisions, as well as of both demand and supply sides. It is possible, however, to find several channels through which the selected variables affect the number of doctor visits (even though it is not our goal to disentangle them). For instance, according to microeconomic theory, the main factors influencing the estimation of a demand curve should be the budget constraint and individual preferences, which are potentially captured through some PHS questions.

The health insurance status is often pointed as a very important factor of health care utilisation. We account for three possible insurance situations, that made us consider three mutually exclusive groups of observations: namely the "NHS" composed by individuals with only the statutory health system and two double coverage types, the "Public subsystems" for people with NHS plus a public subsystem and the "Private subsystems" for individuals with NHS and a private subsystem. These variables are of particular importance since the main goal of this work is to assess how a patient´s use of medical consultations is affected by double coverage. Insurance has an important link with the price of going to a doctor: the differences between health care protection plans as regards to costs to beneficiaries (such as co-payments and reimbursements practices) work as di-

[9]The survey includes a question about the type of doctor (GP or specialist) of the last visit but it does not allow to disentangle all the visits taken in the period of three months.

[10]Some information was excluded from the analysis, particularly the questions reported only by part of the sample according to the week of the interview.

rect prices and both mechanisms to control for its use and delivery systems are indirect costs of access. Since insurance influences the price faced by patients, it influences the budget constraint (together with income). In fact, when compared to the NHS, the subsystems provide more benefits to their beneficiaries by decreasing the price-per-service faced by patients, which whenever demand is elastic, increases their health care demand (Barros et al. 2008). See Section 4.3 for more details regarding the interpretation of the impact of these dummy variables.

The underlying health status and the socioeconomic characteristics play a major role in the preferences' formation. Health status also influences the constraints limiting the pursuit of preferences since illness events usually imply a loss of income (although sometimes partially offset by sickness benefits). It can be questionable whether the health status regressors do not introduce a problem of endogeneity. The idea is that individuals usually become aware of their health status through medical consultations. We believe that this is not a problem, especially if one takes into account that the dependent variable is number of visits to doctors in the three months prior to the interview and individuals are likely to be conscious about their diseases for longer time. In the PHS, health status is only indirectly captured through some questions that reflect details about current medical conditions (e.g. sickness episodes and limited days) and the presence of chronic disease or pain (e.g. rheumatism, cancer and diabetes). Besides including such variables, the consumption of barbiturates as a proxy to the level of exposure to stress, as well as some other regressors related to attitudes with a potential impact on health, like the number of meals and a dummy variable identifying smokers/non-smokers, also play a role. Despite being crude measures, these last regressors may capture some remaining health aspects and some unobserved influences. Engagement in sports activities is an alternative proxy for good health but was only available for a small part of the sample, which would imply a substantial decrease in the size of the sample. Winkelmann (2004) and Winkelmann (2006) also include individual subjective self-assessment of health status. PHS provides that information (with the question "How well do you perceive your own health at the present time?", with responses "very good", "good", "fair", "poor" and "very poor") but we excluded its use. These variables are likely to create an endogeneity problem: the self-understanding of the health status influences the utilisation of medical care but it is also influenced by utilisation since the assessment is made after visiting the doctor. Moreover, the non-response is extremely high (around 30 per cent). As suggested by Windweijer and Santos-Silva (1997), we partially control for this subjective health evaluation by including long-term determinants of health (smoking and eating habits).

Table I: Description of the variables

Variables	Description
Health insurance status variables	
pubsub	=1 if the individual is covered by a public subsystem
privsub	=1 if the individual is covered by a private subsystem
Health status variables	
sick	=1 if the individual is being sick
limitdays	number of days with temporary (not long run) incapacity
limited	=1 if the individual is limited/handicapped
rheumatism	=1 if the individual has rheumatism
osteoporosis	=1 if the individual has osteoporosis
cancer	=1 if the individual has cancer
kidneystones	=1 if the individual has kidneystones
renalfailure	=1 if the individual has renalfailure
emphysema	=1 if the individual has emphysema
cerebralhemorrhage	=1 if the individual had a cerebral hemorrhage
infarction	=1 if the individual had an infarction
depressivedisorder	=1 if the individual has a depressive disorder
otherchronicaldisease	=1 if the individual has another chronical disease
highbloodpressure	=1 if the individual has high blood pressure
chronicpain	=1 if the individual has a chronic pain
diabetes	=1 if the individual has diabetes
asthma	=1 if the individual has asthma
stress	=1 if the individual has been taking sleeping pills or anxiety pills in the last two weeks
smoker	=1 if the individual smokes daily
meals	=1 if the individual makes at least three meals a day
Socioeconomic and demographic variables	
householdsize	household size of the individual
age	age in years
female	=1 if the individual is female
educmax	number of years of schooling completed with success of the most educated person living in the household
lincome	logarithm of equivalent monthly income (in euros)
single	=1 if the individual is single and do not cohabits
student	=1 if the individual is student or has it fist job or has a not remunerated job
retired	=1 if the individual is retired
Geographic variables	
Norte	=1 if the individual lives in the region "Norte" (NUTS II)
Lisboa	=1 if the individual lives in the region "Lisboa" (NUTS II)
Alentejo	=1 if the individual lives in the region "Alentejo" (NUTS II)
Algarve	=1 if the individual lives in the region "Algarve" (NUTS II)
Açores	=1 if the individual lives in the region "Açores" (NUTS II)
Madeira	=1 if the individual lives in the region "Madeira" (NUTS II)
Seasonal variables	
winter	=1 if the interview took place in the winter
spring	=1 if the interview took place in the spring
summer	=1 if the interview took place in the summer

The variables representing demographic and socioeconomic features of the interviewed can influence the decision to seek health care directly and indirectly through their impact on health care status. This is particularly evident when analysing the covariate "age". According to Grossman (1972), age captures the depreciation of health capital which influences the health status and is an important factor influencing individual preferences. It is expected that the rate of depreciation increases as the individual gets older, at least after some point of the life cycle, making the healthy times decrease. As a consequence, the demand for health care is expected to increase over the life cycle. At the same time, age is an extra variable that can be considered as a health status proxy since older individuals are, on average, less healthy and less efficient in producing health. We chose to control for age through a nonlinear relationship and by including variables that allow an assessment of its effect by gender type.

Amongst the socioeconomic covariates, a gender dummy was included because it is believed to influence the rate at which the health stock depreciates and the efficiency in producing healthy times. In particular, it is expected that health depends on biological differences between men and women through innate features, life styles and different attitudes towards health risk (Lourenço 2007). Accordingly, we also control for the marital status with the inclusion of the covariate "single". Besides the arguments of different life styles and attitudes toward risk, it is our understanding that some decisions when taken by more than one person benefit from advice and more information, which should influence health status and efficiency in producing healthy times.[11]

To control for education, we defined a variable with the number of schooling years of the most educated person living in the household (in line with the procedure in Lourenço (2007)). It is expected that more educated people are more productive in the market as well as in the household, therefore even if they seek for more health they need relatively less medical care (Jones et al. 2006). Further, different educational levels are associated with different opportunity costs and attitudes towards risk. This particular indicator was chosen, as an alternative to the usual number of schooling years of each individual, because we believe that the decision about the number of visits to a doctor is at least partially a decision of the household and benefiting from a better level of information.

[11]Most of the studies include a slightly different variable that assumes one if the person is married instead of single. The design of the survey and some previous results influenced the choice of this particular variable.

The variables "student" and "retired" capture occupational status which may explain some differences in the depreciation rate. It is expected that a person who does not work, presents lower opportunity costs of visiting a doctor, than an individual with a regular job (Lourenço 2007). Further, since hours of market and non-market can have different values and the stock of health determines the total amount of time to spend producing earnings and commodities, more active individuals should invest more in health capital (Grossman 1972).

Another variable included in the model is the monthly equivalent income. According to Grossman (1972) there are reasons to believe that medical utilisation increases with income: "The higher a person's wage rate, the greater the value to him of an increase in healthy time". The idea is that the cost of being ill is higher. A converse argument is that the opportunity cost of going to the doctor is higher for higher wages (Jones et al. 2006). In addition to this, income also represents the ability to pay, as a proxy of wealth. In the PHS, income is only compiled for the household as a whole through a categorical ordinal variable with ten thresholds that indicate intervals of net disposable household income in the month prior to the interview (including wages, pensions, and all sorts of social security benefits). A common procedure to control for income effects is including in the model a set of dummy variables, one for each category. Here, such alternative is not very attractive due to the fact that it would be impossible to take into account the composition of households. We chose a more flexible and parsimonious modelling strategy (although not problem-free) with the construction of a monthly income variable that, in a first stage assigns an income corresponding to the midpoint of the interval, and in a second stage interpolates grouped data by taking into account differences in the composition of households (in line with Pereira 1995). This procedure has the disadvantage of assuming that the income of the household is the midpoint of its income class and, additionally, for the open-ended category it was necessary to assume an arbitrarily value. We use €2500 but we test the robustness of this value by considering other figures, in particular, the estimate for the median of this last income bracket calculated using a Pareto distribution. To take into account the composition of households we used the square root scale, through dividing the household income by the square root of household size.[12]

[12] The "individual" income is measured with error given the way it is compiled in the survey and the modelling procedure. Concerning the latter, we tested different alternatives and we found only minor differences in the estimates.

The variables "Norte", "Lisboa", "Alentejo", "Algarve", "Açores" and "Madeira" ("Centro" being omitted)[13] represent the region of residence and were included to control for possible differences in the demand and supply of health care services. The regions encompass wide areas but nevertheless, when we compare them in terms of wealth or educational indicators we obtain huge differences, which could justifie different behaviours regarding the usage of health care services (not totally captured at the individual level). Apart from this argument, the main reason to include these variables is because they proxy different access to medical care supply, since some regional services are differently organized. Note that in the mainland, the five regions correspond to the five regional health administrations, and in the autonomous regions there are two different regional health services.

To control for the period of the year in which the interview took place we included the regressors "spring", "summer", and "winter" ("autumn" being omitted). This is important because there may be some seasonal differences in the health status of individuals.

3.3 An exploratory analysis of the data

Table II presents the empirical distribution of the dependent variable (y) and some basic statistics. As the table shows, the majority of observations are of the NHS group, followed by the public subsystem. The dependent variable used is a count variable (non-negative integer valued count $y = 0, 1, 2, \ldots$) with a large proportion of zeros (half of the sample) as well as a long right tail of individuals who make heavy use of health care. These features make the estimation particularly difficult since it will be necessary to use flexible models that accommodate them. For the whole sample, the average number of consultations is 1.01 and the average number of visits for those that have at least one visit is 2.04. Moreover, the unconditional variance is more than three times the unconditional mean.[14] When we analyse the average number of visits to a doctor by health insurance systems, it is possible to observe that private subsystems beneficiaries are more frequent users than the groups NHS and public subsystems. Indeed, a mean comparison *t-test* indicates that the unconditional probability does not differ across NHS and public subsystems but differ

[13] In accordance with NUT II classification (official territorial nomenclature for statistical analysis), Portugal is divided into seven regions. The survey includes data for all of them. Therefore, we use six dummies.

[14] This is a sign of possible overdispersion just confirmed when a conditional analysis is made.

when one compares NHS with private subsystems.

Table II: Empirical distribution of the dependent variable

	TOTAL	NHS	Public sub.	Private sub.
y		relative frequency		
0	50.31	50.88	48.82	41.91
1	26.94	26.53	28.54	29.83
2	10.78	10.61	11.37	12.61
3	6.77	6.82	6.15	8.72
4	1.99	2.02	1.69	2.84
5	1.12	1.06	1.25	2.10
6	0.98	0.95	1.17	0.95
7	0.20	0.21	0.14	0.21
8	0.22	0.23	0.18	0.42
9	0.08	0.07	0.13	0.11
10	0.19	0.17	0.30	0.11
11-15	0.25	0.28	0.15	0.11
16-20	0.04	0.04	0.06	0.11
21-25	0.06	0.06	0.06	0.00
26-30	0.06	0.07	0.02	0.00
		Observations		
	35,308	28,778	5,578	952
	100%	81.5%	15.8%	2.7%
		Mean		
	1.01	1.01	1.01	1.19
		Standard deviation		
	1.77	1.80	1.64	1.61
		P-value (Ho: $\mu_{Y_{NHS}} = \mu_{Y_{Subsystem}}$)		
	-	-	0.998	0.000

Table III presents the descriptive statistics of the explanatory variables by health insurance type. The mean comparison *t-test* indicates that most of the differences between the three types are significant, which suggest that a more complete account for them is required, so that an appropriate comparison of health care demand across groups can be made.

The dissimilarities between NHS and public and private subsystems are specially high among the socioeconomic pre-determined variables. The NHS group has relatively less years of education and lower income. On its turn, public subsystems beneficiaries are younger (on average about 4 years less than the other groups), have a greater proportion of students and singles and a smaller share of retired persons. The private subsystems group has less women and a smaller household size.

Table III: Descriptive statistics by health insurance system

	NHS		Public subsystem			Private subsystem		
	mean	st.dev	mean	st.dev	p-value	mean	st.dev	p-value
Health status variables								
sick	0.007	0.001	0.005	0.001	0.008	0.005	0.002	0.363
limitdays	0.613	0.015	0.488	0.030	0.000	0.536	0.077	0.327
limited	0.016	0.001	0.004	0.001	0.000	0.006	0.003	0.000
rheumatism	0.168	0.002	0.120	0.004	0.000	0.134	0.011	0.003
osteoporosis	0.069	0.001	0.060	0.003	0.014	0.068	0.008	0.943
cancer	0.019	0.001	0.020	0.002	0.688	0.022	0.005	0.491
kidneystones	0.048	0.001	0.051	0.003	0.473	0.058	0.008	0.224
renalfailure	0.014	0.001	0.011	0.001	0.196	0.014	0.004	0.971
emphysema	0.034	0.001	0.022	0.002	0.000	0.022	0.005	0.015
cerebralhemorrhage	0.018	0.001	0.013	0.002	0.000	0.020	0.005	0.654
infarction	0.014	0.001	0.011	0.001	0.103	0.014	0.004	0.956
depressivedisorder	0.074	0.002	0.074	0.004	0.934	0.082	0.009	0.395
otherchronicaldisease	0.319	0.003	0.297	0.006	0.001	0.317	0.015	0.928
highbloodpressure	0.221	0.002	0.178	0.005	0.000	0.222	0.013	0.977
chronicpain	0.148	0.002	0.110	0.004	0.000	0.119	0.010	0.006
diabetes	0.077	0.002	0.054	0.003	0.000	0.074	0.008	0.651
asthma	0.051	0.001	0.057	0.003	0.075	0.049	0.007	0.837
stress	0.119	0.002	0.104	0.004	0.001	0.124	0.011	0.631
smoker	0.162	0.002	0.138	0.005	0.000	0.179	0.012	0.200
meals	0.926	0.002	0.949	0.003	0.000	0.933	0.008	0.402
Socioeconomic and demographic variables								
householdsize	3.387	0.009	3.342	0.017	0.020	3.100	0.037	0.000
age	42.044	0.131	38.984	0.285	0.000	42.946	0.685	0.196
female	0.515	0.003	0.537	0.007	0.003	0.419	0.016	0.000
educmax	8.112	0.026	11.949	0.061	0.000	11.625	0.147	0.000
lincome	6.048	0.003	6.624	0.007	0.000	6.669	0.019	0.000
single	0.350	0.003	0.391	0.007	0.000	0.322	0.015	0.076
student	0.164	0.002	0.247	0.006	0.000	0.188	0.013	0.065
retired	0.185	0.002	0.171	0.005	0.012	0.256	0.014	0.000
Geographic variables								
Norte	0.161	0.002	0.093	0.004	0.000	0.104	0.010	0.000
Lisboa	0.126	0.002	0.146	0.005	0.000	0.232	0.014	0.000
Alentejo	0.136	0.002	0.166	0.005	0.000	0.120	0.011	0.133
Algarve	0.146	0.002	0.122	0.004	0.000	0.181	0.012	0.006
Açores	0.147	0.002	0.205	0.005	0.000	0.169	0.012	0.079
Madeira	0.139	0.002	0.147	0.005	0.127	0.060	0.008	0.000
Seasonality variables								
winter	0.255	0.003	0.254	0.006	0.873	0.314	0.015	0.000
Spring	0.258	0.003	0.255	0.006	0.702	0.235	0.014	0.110
Summer	0.249	0.003	0.237	0.006	0.052	0.246	0.014	0.825

Notes: The p-value indicates if the probability of the mean of each variable does not significantly differ across insurance types. The test is performed as a two-sample mean-comparison test (unpaired). For the comparison between the NHS and the public subsystem we considered H0: $\mu_{Y_{NHS}} = \mu_{Y_{Public\ subsystem}}$; and for the comparison between the NHS and the private subsystem we considered H0: $\mu_{Y_{NHS}} = \mu_{Y_{\Pr ivate\ subsystem}}$.

As regards the health status distributions of the three groups, it is possible to conclude that the major differences are found between the public subsystem and the NHS. Public employees seem to be the healthier, in particular when we analyse some variables related to physical limitations ("limited days" and "limited") and the presence of chronic diseases and pains. Moreover, frequent health problems (e.g. high blood pressure, diabetes and stress) are relatively more common in the NHS and private subsystem groups. This feature can be partially related with age, which is lower among the public subsystems group. Additionally, it is worth highlighting that public employees seem to be less exposed to stress and that the indicators related to attitudes show a smaller proportion of smokers and a higher average number of meals. The regional distribution of the groups is also unequal in the full sample: most of the NHS individuals are located in the North; the public employees are concentrated in Lisbon, Alentejo and Azores; and the private subsystem group has relatively more beneficiaries in the regions of Lisbon and Algarve.

4 Econometric framework

Econometrics of count data has its own modelling strategies in which discreteness and non-negativity are taken into account. Moreover, in the "count world" it is common that features other than location depend on the covariates, making the estimation of the conditional expectation poorer in the sense that provides very little information about the impact of the regressors on the outcome of interest. In this context it is potentially interesting to study the effect of regressors not only on the mean but also on single outcomes and in the full distribution.

Within the vast literature on count data it is possible to find two general categories of methods that allow a complete description of the conditional distribution of a count outcome. Following the early work of Hausman, Hall, and Griliches (1984), several fully parametric probabilistic models, like Poisson and negative binomial regressions, have been developed in order to describe the effect of the covariates on different points of a count distribution. These regressions allow inferences for all possible aspects of the outcome variable (including the computation of the marginal probability effects). However, to do it, they impose restrictive parametric assumptions on the way the independent variables affect the outcome variable. As a consequence, this approach usually faces a lack of robustness problem, even when flexible models like the hurdle or latent class models are applied. Given these limitations, it can be attractive to use non- or semiparametric techniques that freely approximate the conditional distribution. This can be achieved with the estimation of conditional quantile functions, a technique that has been applied in the context of continuous regression for a long time (Koenker and Bassett 1978). Following the contributions of Manski (1975), Manski (1985) and Horowitz (1992) regarding binary models, some effort is being made to extend the method to discrete data. Recently, the seminal work of Machado and Santos-Silva (2005) succeeded in applying the quantile framework to count data models.

Since our main aim is to assess the effect of the subsystems in different parts of the outcome distribution without imposing a probabilistic structure, the "Quantile for counts" regression model is a natural choice. In order to better understand its advantages (and disadvantages) we compare the implications drawn from the quantile regression approach with those from parametric count data models that have been used quite extensively in the analysis of health care. In this section we first present a brief description of such models (the minimum necessary to support the correct interpretation of results). A natural starting point is the Poisson regression model, followed by negative binomial models. Abandoning the rigid single index structure of these conventional approaches, we present two-part models, in particular the Hurdle and zero-inflated models. To model the number of visits to a doctor, the hurdle is usually set at zero. In fact, in the health care literature the zeros can be interpreted as a result of the individual decision to be user or non-user (in principle motivated by the existence of sickness episodes), and the intensity of usage is more likely to result from patient choice influenced by doctor´s opinion. Recently, several works (e.g. Lourenço 2007, Deb and Trivedi 2002 and Bago-d'Uva 2006) are applying a finite discrete mixture model (not continuous, like the negative binomial distribution), frequently called latent class model. In this framework, it is assumed that the observed data is a mixture of a finite number of subpopulations. In the case of two classes, the estimation splits the population into what the literature entitle "high" and "low" users, according to the intensity of usage. When compared with two-part models that have a clear dichotomy between users and non-users, these models are mainly driven by statistical reasoning (Jiménez-Martín, Lebeaga, and Martinez-Granado 2002) and are often a more suitable approach for panel data. This last feature made us exclude its application as a benchmark to the cross sectional data derived from the PHS.

4.1 Parametric models for counts

4.1.1 Poisson regression model[15]

> *"The Poisson regression model is the benchmark model for count data in much the same way as the normal linear model is the benchmark for real-valued continuous data."* Winkelmann (2008)

For each $i = 1, 2, 3, ..., N$, let y_i be a dependent count random variable and x_i a vector of independent variables. In a Poisson regression model it is assumed that the probability function of y_i, conditional on the vector of covariates x_i, is a Poisson distribution. Density is fully determined after the specification of the conditional mean (λ_i) which is usually parameterized as an exponential of a linear function, dependent on the vector of regressors x_i and the unknown vector of parameters (β).

$$f(y_i|x_i) = \frac{e^{-\lambda_i}\lambda_i^{y_i}}{y_i!}, \quad \lambda_i > 0 \tag{1}$$

$$E(y_i|x_i) = \lambda_i = \exp(x_i^{'}\beta) \tag{2}$$

The literature often criticizes the Poisson regression model by not incorporating the individual unobserved heterogeneity. This happens because it is characterized by the mean-variance equality, known as "equidispersion", which is seldom or never verified in empirical applications, that usually find "overdispersion" (Pohlmeier and Ulrich 1995, Winkelmann 2006). The "excess of zeros" is another problem, since a high frequency of zeros, often observed in count data, is not consistent with the Poisson regression. Here, we compute robust standard errors, following the pseudo maximum likelihood (Gourieroux, Monfort, and Trognon 1984). Thus if the mean function is correctly specified, then maximum likelihood estimation based on any distribution in the class of linear exponential families provides a consistent estimator of β even if the model is otherwise misspecified. Note that, by using the pseudo maximum likelihood it is not possible to overcome the drawbacks of goodness of fit, either "overdispersion" or "excess os zeros". Winkelmann (2006) emphasizes that the single crossing property of the Poisson regression model imposes a very restrictive probability change in response to a variation in covariates, concluding that it

[15] A complete characterization can be found in Winkelmann (2008), pp. 7-26 and pp. 63-126.

is not very well suited when the focus is on modelling the full probability response to a change in a regressor.

4.1.2 Negative binomial models[16]

A natural extension of the Poisson regression model is the continuous mixture models for unobserved heterogeneity. The negative binomial (Negbin) distribution is the most commonly used. It can be characterized as a continuous Poisson-gamma mixture, obtained by allowing the Poisson average population parameter λ_i to change randomly across the population with a gamma distribution $\Gamma(.)$. Following Winkelmann (2008), it can be established that

$$f(y_i; \lambda_i, \xi_i) = \frac{\Gamma(\xi_i + y_i)}{\Gamma(\xi_i)\Gamma(1 + y_i)} \left(\frac{\xi_i}{\lambda_i + \xi_i}\right)^{\xi_i} \left(\frac{\lambda_i}{\lambda_i + \xi_i}\right)^{y_i} \tag{3}$$

where ξ_i is a positive dispersion parameter, and λ_i is the conditional expectation, defined as before. The Negbin model is constructed by specifying both parameters in terms of exogenous variables (x_i) in which overdispersion (variance exceeds the mean) is contemplated. There are at least two different ways of introducing dispersion that conduct to two different values.

$$\text{NegbinI} : Var(y_i|x_i) = \lambda_i(1 + \xi_i^{-1}) \tag{4}$$

$$\text{NegbinII} : Var(y_i|x_i) = \lambda_i + \xi_i^{-1}\lambda_i^2 \tag{5}$$

Both Negbin models are fully parametric and the use of a gamma function to account for unobserved heterogeneity is only due to statistical convenience. Additionally, it still does not account for the "excess of zeros", while the single crossing property remains exactly like in the Poisson case.

Another increasingly popular count data model is the Poisson-log-normal. Although facing the same problems it tends to dominate the Negbin framework in what regards to the goodness of fit.

[16] A complete characterization can be found in Winkelmann (2008), pp.20-26 and pp. 127-142.

4.1.3 Two-part models[17]

It is potentially interesting to study the effect of regressors not only on the mean, but also on single outcomes,the most prominent of which are zeros. Standard count data models presented above are too restrictive to properly model zeros (problem of "excess zeros"). The literature points to two-part models as the most suitable solution to this problem. These models assume that consumption decisions are made in two steps. The framework is specially known by its empirical advantages, but there are also strong theoretical reasons in favour of this specification, related with the fact that zeros can often reflect corner solutions in economic choice models, making the zeros eventually dependent on other driving forces.

The Hurdle regression model is the best known two-part model. It was firstly proposed by Mullahy (1986), who presented a count data model with a decision structure similar to of the overall Cragg model. In most applications the hurdle is set at zero, assuming that zeros result from a different process, creating a clear distinction between no consumption and positive consumption. This model is a finite discrete mixture that combines different probability functions for zeros and positive integers, allowing underdispersion and overdispersion. Formally, the density function can be written as:

$$f(y_i|x_i) = \{ \begin{smallmatrix} f_0(0|x_i) & \text{for } y_i=0 \\ f_1(y_i|x_i)\frac{1-f_0(0|x_i)}{1-f_1(0|x_i)} & \text{for } y_i>0 \end{smallmatrix} \tag{6}$$

where $f_0(0|x_i) = \Pr(y_i = 0|x_i)$ and $f_1(y_i|x_i, y_i > 0)$ is the truncated distribution. Different probability functions for $f_0(.)$ and $f_1(.)$ lead to different specifications of the Hurdle model. For the first part the most common specifications are the Poisson, Negbin or binary models like logit and probit, while for the second part it is usually used a Poisson or a Negbin. In this work, besides the Poisson Hurdle model that uses a complementary loglog for $f_0(0|x_i)$ and a zero truncated Poisson for $f_1(y_i|x_i, y_i > 0$, we will estimate a probit and logit for $f_0(0|x_i)$ and a zero truncated Negbin I and II. These models although fully parametrized and strongly dependent on the assumption that zeros and positive values result from different processes, are more flexible than the one-part models in what regards to the single crossing property. Indeed, it can be proved that the marginal probability effects can switch signs twice (Winkelmann 2008).

[17] A complete characterization can be found in Winkelmann (2008), pp.173-193.

Another group of two-part models are the zero-inflated class of models, which also address the problem of "excess of zeros". Their difference relative to Hurdle models is that they assume zeros to be generated from two different processes, based on the idea that an individual will be willing to utilise but in a second stage, although admitting to consume strictly positive quantities decide not to do so. Sometimes literature refers to the two types of zeros as strategic and incidental. This framework uses the same kind of models for the first part and for the second part count distributions not truncated. In this work we will estimate zero-inflated Poisson and zero-inflated Negbin II. The choice between Hurdle and zero-inflated models can be made on substantive or statistical grounds (for instance, by using the Vuong test for nested models).

4.2 Quantile regression for counts

The models presented in the last subsection are not as informative as conditional quantiles. They are conceived to study the impact on the mean - the information on the distribution of the relevant outcomes is determined either by the conditional expectation or by the "arbitrarily" distribution chosen (e.g. Poisson, Negbin). Quite recently the quantile regression (Koenker and Bassett 1978) was extended to count data (Machado and Santos-Silva 2005) through the conjugation of a nondifferentiable sample objective function with a discrete dependent variable.

Let y be a count random variable and their α-quantile defined as:

$$Q_y(\alpha) = \min\left[\eta | P(y \leq \eta) \geq \alpha\right] \qquad \text{where} \qquad 0 \leq \alpha < 1 \tag{7}$$

The α-quantile has the same discrete support as y and cannot be a continuous function of the covariates (x). Machado and Santos-Silva (2005) suggested a procedure known as “jittering” to artificially impose some degree of smoothness. The basic idea is to build a continuous auxiliary variable (y^*) whose quantiles have a one-to-one known relationship with the quantiles of the count variable of interest. The y^* is obtained by adding to the count variable a uniform random variable, independent of y and x:[18]

[18] Machado and Santos-Silva (2005) showed that there is a little loss of generality in assuming that U is uniform. In fact they argue that it is possible to choose another distribution for U as long as it has a support on $[0, 1)$ and a density function bounded away from 0. The advantages of using a uniform distribution are purely algebraic and computational.

$$y^* = y + u \qquad \text{where} \qquad u \sim uniform[0,1) \tag{8}$$

The continuity problem of the dependant variable is solved but the derivatives are not continuous for integer values of y^*. Machado and Santos-Silva (2005) proved that given some regularity conditions, valid asymptotic inference is possible. Among those conditions, it is particularly relevant the existence of at least one continuously distributed covariate. The standard quantile regression is applied to a monotonic transformation of y^* that ensures that the estimated quantiles are non-negative and the transformation is linear in the parameters of a vector of regressors.

In order to implement the procedures, the authors suggest the following parametric representation of the α-quantile of y^*:

$$Q_{y^*}(\alpha|x) = \alpha + \exp\left[x'\beta(\alpha)\right], \qquad 0 \leq \alpha < 1. \tag{9}$$

The reason for adding α to the right side is that y^* is bounded from below at α due to the way it is constructed. The exponential form is traditionally assumed in count data models. We believe that this specification provides a good parsimonious approximation to the unknown conditional quantile functions. The linear transformation is specified as:

$$Q_{T(y^*;\alpha)}(\alpha|\mathbf{x}) = x'\beta(\alpha), \tag{10}$$

where $T(y^*;\alpha) = \left\{ \begin{smallmatrix} \log(y^*-\alpha) & \text{for} & y^*>\alpha \\ \log(\varepsilon) & \text{for} & y^*\leq\alpha \end{smallmatrix} \right.$, being ε a small positive number ($0 < \varepsilon < \alpha$).[19]

This is feasible because quantiles are equivariant to monotonic transformations and to censoring from below up to the quantile of interest. The vector of covariates $\beta(\alpha)$ is obtained as a solution to a standard quantile regression of a linear transformed variable by minimizing an asymmetrically weighted sum of absolute errors

$$\min \sum_{i=1}^{n} \rho_\alpha\left[T(y^*;\alpha) - x_i'\beta\right] \qquad \text{where} \qquad \rho_\alpha(v) = v\left[\alpha - I(v<0)\right]. \tag{11}$$

Machado and Santos-Silva (2005) proved that although the quantile regression is not differentiable everywhere, the estimator is consistent and asymptotically normal:

[19] We will use 1.0E-10 as Machado and Santos-Silva (2005) did.

$$\sqrt{n}\left[\widehat{\boldsymbol{\beta}}(\alpha)-\boldsymbol{\beta}(\alpha)\right] \xrightarrow{D} N(0, D^{-1}AD^{-1}) \tag{12}$$

with $A=\alpha(1-\alpha)E(XX')$ and $D=E\left[f_T(X'\beta(\alpha)|X)X'X\right]$, where f_T denotes the conditional density of $T(y^*;\alpha)$ given X.

Because "noise" has been artificially created for technical reasons, Machado and Santos-Silva (2005) suggest a Monte Carlo procedure - an "average-jittering" - which consists in obtaining an estimator that is the average of m independent "jittering" samples with the same size. The difference between samples is the dependent variable y^* because it is created as the sum of y (constant between samples) with m different draws of the uniform distribution. The main advantage of this procedure is that the resulting estimator is more efficient than the one obtained from a single draw and a misspecification-robust estimator of the covariance matrix is available.

The importance of this procedure derives from the possibility of performing inferences on the variable of interest y. Machado and Santos-Silva (2005) showed that marginal effects of the smoothed variable y^* are easily obtained and interpreted and that there is a correspondence between the two quantile functions:

$Q_y(\alpha|x)=\lceil Q_{y^*}(\alpha|x)-1\rceil$, where $\lceil a\rceil$ denotes the ceiling function (returns the smallest integer greater than, or equal to a).

Because of the monotone transformation of y^*($T(y^*;\alpha)$), the relationship between coefficient estimates $\widehat{\beta}(\alpha)$ and $y^*$and y is essentially non-linear, making it hard to interpret $\widehat{\beta}(\alpha)$ in terms of y^* and y. It is possible to test the null hypothesis that a covariate has no effect on $Q_y(\alpha|x)$ because it is equivalent to test whether the variable has no impact on the $Q_{y^*}(\alpha|x)$. The problem is when the variable is significant in $Q_{y^*}(\alpha|x)$. In such case it could be non significant in the conditional quantile of y.[20] This occurs because different quantiles of y^*correspond to the same quantiles of y. In fact, a change in x_j will affect $Q_y(\alpha|x)$ only if it is capable of changing the integer part of $Q_{y^*}(\alpha|x)$. Machado and Santos-Silva (2005) call this "magnifying glass effect" of $Q_{y^*}(\alpha|x)$.

[20] It is not possible to just look at β_j, as it becomes necessary to evaluate case by case if a given magnitude in x_j induces changes in the α-quantile of y. Inference about the partial effect of a particular variation of the regressor, given that all other variables remain fixed at $\widetilde{\mathbf{x}}$ is made through the following expression:

$\Delta_j Q_y(\alpha|\widetilde{\mathbf{x}},x_j^0,x_j^1)=Q_y(\alpha|\widetilde{\mathbf{x}},x_j^1)-Q_y(\alpha|\widetilde{\mathbf{x}},x_j^0)$

4.3 Empirical specification

The conditional quantiles used in our analysis are defined as

$$Q_{y_i^*}(\alpha|x) = \alpha + \exp\left[\beta_0(\alpha) + \beta_1(\alpha)pubsub_i + \beta_2(\alpha)privsub_i + \boldsymbol{\gamma}(\alpha)\mathbf{z}_i\right], 0 \leq \alpha < 1 \quad (13)$$

where $pubsub_i$ and $privsub_i$ indicate if a person benefit from double coverage through a "public insurance health subsystem" or a "private insurance health subsystem", respectively. The vector $\mathbf{z}_i$ includes all other characteristics that were controlled for in this regression.[21] In addition to all independent variables referred in Section 3.2, we use a third order polynomial in "age" and a third order polynomial in "age" crossed with the gender variable ($age \times female$).

Our main focus will be on the coefficients $\beta_1(\alpha)$ and $\beta_2(\alpha)$. They measure the impact of additional layers of health insurance (on top of the statutory NHS) on the utilisation of visits to a physician, by reflecting the differences in consumption patterns between NHS and public and private subsystems. This empirical work gives some evidence regarding the effects of a potential reform in the Portuguese health care system. We can state our interest as to measure the potential impact of the elimination of double coverage on the utilisation of health services, i.e. the potential decrease in doctors consultations amongst the subsystems beneficiaries due to their insurance status. When compared to the general treatment effects techniques, our analysis lacks on a certain features. For instance, when studing the impact of a reform is frequent the usage of panel data comparing the outcome before and after the reform (e.g. Winkelmann (2006)).[22] Nevertheless, we have a important positive distinctive feature. By analyzing not only the mean effect but also the impact on the whole outcome distribution, we present an improvement over past research methodology by showing that mean treatment effects can "miss a lot". The advantages of applying a quantile regression approach go further than just statistical convenience. Using this technique we are able to study the heterogeneity of the effect of double coverage, without imposing any a priori distributional structure. In fact, Bitler

[21] The vector of coefficients is now $\beta(\alpha) = [\beta_0(\alpha), \beta_1(\alpha), \beta_2(\alpha), \boldsymbol{\gamma}(\alpha)]$, being $\beta_0(\alpha)$, $\beta_1(\alpha)$, $\beta_2(\alpha)$ scalars and $\boldsymbol{\gamma}(\alpha)$ a vector.

[22] In such case, the typical empirical strategies include pre-reform/post-reform differences-in-differences where one compares the changes in the utilization between affected and unaffected sub-populations. A drawback in our analysis relative to more general approaches is that we estimate the impact in 2005-2006, which may change in case of different time paths between groups.

et al. (2006) showed that unlike what is usually done in the majority of welfare reform studies that rely on estimating mean impacts, it is necessary to allow for heterogenous treatment effects, in which the quantile regression methodology can play a very useful role. In our case, with quantile regression we can see if the impact of the health subsystems differs accordingly with the level of utilisation of doctor visits, which gives insight of what can happen in case of a reform.

Our specification is consistent with the measurement of the double coverage effect in the case of ignorability of the treatment conditional on a set of covariates ($\mathbf{z}_i$) and we arguee that there is no reasons to believe that different health insurance status have distinct distributions of unobservable determinants of health care utilisation.[23] Moreover, when selecting the variables we guarantee that treated (public and private subsystems beneficiaries) and untreated (NHS beneficiaries) groups have a common support by using only observations in the intersection of the domains. This means that it is necessary to have subpopulations in each age-state: NHS, *privsub* and *pubsub* (see Wooldridge (2002) for details). This procedure made us exclude from the population of interest individuals with more than 80 years old. Another problem that can rise is the strong correlation between some potential regressors and the treatment dummies. One of the cases that deserved special attention is the unemployment status. See robustness analysis in Appendix D.

In general, the estimation of insurance effects can be erroneous if the researcher ignores the adverse selection role in the decision to obtain health protection. In such case, it creates an endogeneity problem that results in an overestimation of the impact. The particular features of the Portuguese subsystems gather conditions to consider the variables $\beta_1(\alpha)$ and $\beta_2(\alpha)$ exogenous, i.e. not correlated with the beneficiaries health status. This happens because membership on public and private health subsystems was mandatory and based on professional category, and as such unrelated to the expected value of future health care consumption. Note that even the contributions are based on income (and not on risk characteristics of each individual) and it is very implausible that individuals want to work as public employees or in companies with private subsystems just to benefit from this additional health insurance, especially if one takes into account that, by

[23] In fact, the assumption of ignorability conditional on the set of covariates ($\mathbf{z}_i$) is naturally dependent on the inexistence of unobservable characteristics (omitted variables) with a different distribution among subsystems.

default, people are covered by a health care protection system (NHS). It is also unlikely that employers choose individuals on the basis of unobservable variables related to their health or even household health. The only requirement is that the potential employee is suitable for the job and has no infectious disease which could be controlled through our set of pre-determined variables. Previous studies using data for Portugal defended the exogeneity of the covariates indicating if the individuals benefit from a subsystem (e.g. Barros et al. 2008 and Lourenço 2007). Jones, Koolman, and van Doorslaer (2006) analysed the effect of supplementary insurance on the probability of visiting a specialist physician, allowing for potential endogeneity of the insurance variable and, for Portugal, they conclude that the increased probability of utilisation is not due to selection effects. In spite of this evidence, the appropriateness of the exogeneity assumption was tested in a sensitivity analysis exercise that consisted on running models on spouses and descendants only (see Appendix E for details).

The insurance covariates can capture two effects that underestimate the impact of double coverage. Firstly, the fact that the subsystems beneficiaries enjoy more or better treatment than NHS beneficiaries may decrease the future consumption of health care (Barros et al. 2008 and Vera-Hernández 1999). This is because over lifetime, better health care would translate into a significant accumulation of health advantages not totally captured in the other controlled variables.[24] This issue will be addressed in Section 5.3, by restricting the analysis to young beneficiaries who did not yet had time to accumulate such advantages and compare the results with those of the larger sample. Finally, another important comment of the double coverage coefficients is that they cannot be totally associated with a moral hazard behaviour but instead to a joint effect of moral hazard from the beneficiaries and supply-induced demand by the providers. The latter is related to the fact that doctors for patients with health subsystems may require more tests in order to justify more visits. With our dataset it is not possible to distinguish between the contact and frequency decisions that, if possible, would help to isolate the supply-induced demand of the suppliers from the "pure" demand from the patients. However, according to Barros et al. (2008), the payments to subsystems providers are relatively low so the magnitude of this effect will be very small. Independently of that, the important point

[24]This is likely for the Portuguese case, since before 1979, the State only covered the costs of health care for civil servants and this fact implies that the elder cohorts of subsystems beneficiaries received relatively better access to health care in comparison with the NHS counterparts.

here is to capture how much the system design increases the utilisation of resources related to consultations, being only indirect a association to demand/supply impacts or moral hazard effects.

5 Results

5.1 Preliminary results[25]

In this section we present the results of the parametric count data models mentioned in section 4.1. All models were estimated using the maximum likelihood method (or pseudo-maximum likelihood).[26] The software used was STATA. Starting with single crossing log-linear regressions group, Table B1 of Appendix B presents the results for the Poisson pseudo-likelihood, the Negbin I and II, and for the Poisson log-normal model. The estimated effects are generally similar across specifications. Health insurance semi-elasticities vary from 5.6 to 6.7 per cent and from 15.0 to 18.7 per cent in the public and private subsystems, respectively. An interesting finding is that the ratio between private and public subsystems remains almost unchanged at 2.6-2.8. Among the health status variables, only the dummy "smoker" is not significant at the five per cent level in the Poisson and Negbin II. In addition to that, estimates are similar across regressions, while the dummies "limited", "renailfailure" and "infarction" have the highest relative differences. Discrepancies between estimates are even lower in the case of socioeconomic regressors, in particular, the effect of age is amazingly identical. In that group of variables, the impacts of education and income are not significant in some models. Geographic and seasonality effects are significant (except for the difference between "summer" and "autumn", the latter being the default season) and do not change considerably among regressions.

In a second stage of this preliminary analysis we compute the two-part models. Table B2 of Appendix B shows the results for the hurdle formulation.[27] Regarding the Pois-

[25] For comparability reasons we use the same set of regressors. In the two-part models the set is included in both equations. The economic interpretation will be made in the next section with the quantile regression results.

[26] The maximum likelihood estimator is the one that maximizes the likelihood function, which is the product of the individual densities conditional of the covariates (Cameron and Trivedi 2005). The models belong to the linear exponential family can be estimated by maximum pseudo-likelihood, which does not depend on the particular densities. It just depends on the conditional mean being well specified (Gourieroux et al. 1984).

[27] Hurdle models can be estimated in two separated parts. We include all regressors to estimate both

son Hurdle model, the first impressive fact is that health subsystems dummies are not significant in the second equation (specially the public type), which means that health insurance plans influence the decision to consult a doctor but after wards they do not seem to have impact on the quantity of subsequent visits. In what concerns the health status regressors it should be mentioned that while being handicapped seems relevant explaining the number of visits, it does not influence the initial decision, whereas the smoker and eating habits may not have an impact on the number of visits, but only on the probability of visiting a doctor. Finally, most socioeconomic characteristics are not important in the second part of the model. This result seems to corroborate that after deciding to visit a doctor the quantity of consultations may depend on the doctor's opinion. Apart from the Poisson Hurdle framework, other specifications were estimated, in particular, the $f_0(0|x_i)$ was tested with a probit and logit distributions, and the $f_1(y_i|x_i, y_i > 0)$ was modeled with a zero truncated Negbin I and II. Among the different distributions, the results for the probability of having no visits to a doctor are very similar regarding the significance of the independent variables. Furthermore, the probit estimates are analogous to the ones derived from the complementary loglog ($f_0(0|x_i)$ of the Poisson Hurdle model). In what concerns the positively valued observations of the dependent variable, the significance of the regressors is also similar even though some differences exist in the estimates. Public plan remains not significant in explaining the expected number of visits when one decides to visit the doctor, while the private plan seems to have a positive effect (significant at the 5 per cent level) if the real distribution is a zero truncated Negbin I. With the same distribution, the impact of age is also significant, unlike the results obtained with zero truncated Poisson and zero truncated Negbin II.

Finally, table B3 of Appendix B presents the results using two different models: zero-inflated Poisson and zero-inflated Negbin II. Overall, the dummies controlling for the health insurance plan are not significant, however, at a confidence level of 5 per cent, the effect of private subsystem is different from zero in the inflated part (of the model first equation that determines whether the count is zero). Regarding the significance of health status variables, the results are similar to the ones of the Hurdle framework (although the values of the estimates, the results vary a lot between the two specifications). The

the binary model using indicator variable d (where $d = 1$ if $y = 0$ and $d = 0$ else) and the truncated model using the subsample of positive counts only. This feature differs from the zero-inflated framework that cannot be estimated in different parts.

socioeconomic characteristics are overall more important in explaining the inflated part since the decision on the number of visits is only influenced by the dummy "retired" and by age (but not by a third order polynomial).

Models can be compared and evaluated at different levels. Most of the models are non-nested, with the exception of the Poisson and the Negbin models. This implies that we will have to use criteria as the Bayes Information Criteria (BIC), the Akaike Information Criterion (AIC) and Consistent Akaike Information Criterion (CAIC), which are popular measures for comparing maximum likelihood models that account for the fact that the number of parameters differs across specifications. Table B4 of Appendix B shows the log-likelihood and the selection criteria for the different models. Among the one-part models, the Poisson log-normal is preferred. Note that the likelihood ratio test clearly rejects the Poisson model against the other single crossing models. Moreover, the Negbin dispersion parameter (either constant or dependent on the mean in Negbin I and II, respectively) is clearly different from zero thus allowing to reject the Poisson hypothesis of equivariance (see Table B1 of Appendix B).[28] These results corroborate the findings of other works (e.g. Winkelmann (2004) and Deb and Trivedi (2002)) that use similar dependent variable. Results from the two-part models indicate that when compared with its zero-inflated peers, the Hurdle models are preferred and the specifications that use the Poisson distribution are worse than the ones that use the Negbin specification.[29]

5.2 Quantile regression results

Estimations were performed with the *qcount* package of STATA (Miranda 2006) after some slight adjustments. Regarding the number of jittered samples used to obtain the results, preliminary experiments showed that the coefficients are not very sensitive to a particular sample of uniform random variables used to jitter the data: with 1500 samples almost no changes were detected both in coefficients and in standard deviations.[30] The decision of which quantiles to compute took into account the problem under analysis and the empirical distribution of the relevant outcome. Since the marginal quantiles are zero for all $\alpha \leqslant 0.50$, it becomes more interesting to compute conditional quantiles on the

[28] The goodness-of-fit χ^2 (Gilbert 1992) also tells us that we can reject the hypothesis that this data is Poisson distributed.

[29] To analyse a particular relationship between two Hurdle models we can also use Vuong tests.

[30] This result was no surprising due to the high number of observations of our database.

upper tail of the distribution where the effect of covariates changes rapidly. Note that in the lower tail, a variation in the conditional quantiles of the artificial outcome $Q_{y^*}(\alpha|x)$ may be mostly due to the random noise that has been added and as a result we expect to find quantiles more flat. Moreover, it is economically more interesting to look at the behaviour of individuals who make heavy use of health care. In this scenario, and despite the fact that we will still be presenting the first quartile, we will focus on quantiles above the median, and accordingly we will compute results for each decile after the median.

Table IV presents the parameter estimates of the quantiles regressions (the corresponding standard errors are shown in Table C1 of the Appendix C). As we can see, quantile regression does not restrict the way regressors affect different regions of the distribution, allowing the assessment of whether health insurance systems have significant and variable impacts over the different outcomes. The signs of the regressors do not switch across the different quantiles (except for the dummy "summer", whose effect, albeit highly insignificant, is positive in the lower tail and becomes negative in the upper quantiles). All variables are significant in at least one quantile and when compared with the previous one-part models, we conclude that the regressors that were significant in those ones are also significant in the quantile regressions. In the group of health status regressors, the covariates that control for current medical conditions are highly significant as expected. Among the chronic diseases dummies, only the cerebral hemorrhage effect is not significant in quantiles above the $0.7y^*-$ quantile. Concerning the indicators related to attitudes with impact on health status, we find that both the number of meals and smoking habits are insignificant in the upper tail of the distribution. In the case of socioeconomic characteristics, the statistical significance is, to a large extent, lower in both tails of the distribution. Most of the variables related to the region of residence and to seasonality effects have a significant impact on the consumption of visits to doctors.

Table IV: Quantile regression results: coefficients

	$\widehat{\beta(0.25)}$	$\widehat{\beta(0.50)}$	$\widehat{\beta(0.60)}$	$\widehat{\beta(0.70)}$	$\widehat{\beta(0.80)}$	$\widehat{\beta(0.90)}$
Health insurance status variables						
pubsub	0.078	0.088	0.095	0.096	0.073	0.055†
privsub	0.200	0.229	0.247	0.232	0.185	0.148
Health status variables						
sick	0.680	0.602	0.590	0.601	0.547	0.772
limitdays	0.071	0.073	0.076	0.074	0.071	0.073
limited	0.136†	0.205†	0.247	0.321	0.335	0.368
rheumatism	0.134	0.140	0.139	0.140	0.148	0.150
osteoporosis	0.282	0.207	0.182	0.152	0.115	0.091
cancer	0.468	0.464	0.430	0.386	0.403	0.525
kidneystones	0.149	0.154	0.175	0.188	0.221	0.211
renalfailure	0.167‡	0.220	0.212	0.226	0.260	0.234
emphysema	0.090‡	0.210	0.222	0.227	0.232	0.238
cerebralhemorrhage	0.133†	0.135†	0.134†	0.163	0.191	0.189
infarction	0.228	0.327	0.343	0.341	0.290	0.217
depressivedisorder	0.187	0.231	0.247	0.253	0.246	0.248
otherchronicaldisease	0.435	0.451	0.471	0.458	0.384	0.352
highbloodpressure	0.407	0.382	0.367	0.322	0.260	0.208
chronicpain	0.172	0.197	0.220	0.230	0.221	0.224
diabetes	0.449	0.368	0.340	0.316	0.293	0.292
asthma	0.290	0.325	0.339	0.340	0.275	0.230
stress	0.441	0.360	0.342	0.305	0.293	0.250
smoker	-0.205	-0.176	-0.168	-0.154	-0.095	-0.034‡
meals	0.188	0.158	0.129	0.114	0.081†	0.070†
Socioeconomic and demographic variables						
householdsize	-0.063	-0.060	-0.060	-0.060	-0.039	-0.017†
age	-1.072	-1.014	-1.048	-1.071	-0.727	-0.559
age^2	0.234	0.222	0.231	0.241	0.160	0.121
age^3	-0.015	-0.014	-0.014	-0.015	-0.010	-0.007
age*female	0.558	0.580	0.641	0.750	0.490	0.335
(age*female)^2	-0.120	-0.129	-0.146	-0.181	-0.116	-0.078
(age*female)^3	0.007‡	0.008	0.009	0.012	0.007	0.005†
female	-0.321†	-0.321	-0.345	-0.357	-0.216	-0.091‡
educmax	0.010	0.014	0.015	0.015	0.010	0.005†
lincome	0.069	0.058	0.060	0.060	0.053	0.030‡
single	-0.218	-0.198	-0.202	-0.218	-0.164	-0.116
student	-0.252	-0.246	-0.272	-0.253	-0.179	-0.172
retired	0.168	0.149	0.134	0.115	0.120	0.143
Geographic variables						
Norte	-0.052‡	-0.036‡	-0.045‡	-0.043‡	-0.055‡	-0.095
Lisboa	-0.087†	-0.066†	-0.081†	-0.086	-0.092	-0.105
Alentejo	-0.273	-0.221	-0.221	-0.193	-0.157	-0.142
Algarve	-0.248	-0.204	-0.196	-0.168	-0.145	-0.166
Açores	-0.371	-0.343	-0.363	-0.382	-0.338	-0.352
Madeira	-0.534	-0.514	-0.561	-0.612	-0.518	-0.483
Seasonality						
winter	0.171	0.173	0.178	0.172	0.137	0.146
spring	0.093	0.093	0.099	0.092	0.071	0.068
summer	0.048‡	0.037‡	0.024‡	0.000‡	-0.020‡	-0.002‡
Constant	-1.018	-0.406	-0.165‡	0.126‡	0.270†	0.657

Notes: Coefficients marked with ‡ and †are not significant at a 5 and 1 per cent level, respectively. Standard errors are available in Table C1 of Appendix C.

We now turn to the analysis of the magnitude of the effects under analysis. The direct interpretation of Table IV may suggest some misleading conclusions. Note that $\widehat{\beta(\alpha)}$ is a vector of linear partial effects on $Q_{T(y^*;\,\alpha)}(\alpha|x)$. To fully understand the impacts, the analysis should be made through $Q_{y^*}(\alpha|x)$, which is not so easily computed due to its non-linearity as well as to the fact that it is a function of $\alpha-$quantile. Given its non-linearity, the parameter provides an incomplete picture of the covariates' effects on the shape of the distribution. Furthermore, the fact that it is a function of α implies, for instance, that a variable with the same estimated coefficient in all quantiles will have a proportional effect that varies with $\alpha-$quantile. A possible way to take into account the non-linearity is to compute partial effects for specific individuals, say $\widetilde{\mathbf{x}}$. Inference for the marginal effect of a dummy x_j given that all other variables remain fixed at $\widetilde{\mathbf{x}}$ is made through $Q_{y^*}(\alpha|\widetilde{\mathbf{x}}, x_j = 1) - Q_y^*(\alpha|\widetilde{\mathbf{x}}, x_j = 0) = \left[\exp(\gamma_j(\alpha)) - 1\right][Q_{y^*}(\alpha|\widetilde{\mathbf{x}}) - \alpha]$ and for a continuous variable x_l is $\gamma_l(\alpha)\,[Q_{y^*}(\alpha|\widetilde{\mathbf{x}}) - \alpha]$.[31] To facilitate the comparison of effects across different α we also compute the semi-elasticities of $Q_{y^*}(\alpha|\widetilde{\mathbf{x}})$, which is done simply by taking the ratio of the partial effect to $Q_{y^*}(\alpha|\widetilde{\mathbf{x}})$. Table V shows the results for a specific individual, say the "default" individual, who is defined by setting the continuous variables at the sample median and the dummy variables equal to zero.[32]

Results from the parametric models are not directly comparable with the results from the quantile regressions. First, with parametric models it is estimated the expected values and it is possible to compute some outcomes, whereas the quantile regression provides results across quantiles. Another important difference is that the marginal effects are slightly different due to the fact that we work with $Q_{y^*}(\alpha|x)$ which is a function of a constant (α) that changes across quantiles. The comparison between models must be indirect (e.g. through the measurement of the impact of one covariate in terms of the effect of other variable). Amongst parametric models, the ones more easily comparable with the quantile regression models are the two-part models, since they allow at least

[31] Table C2 of Appendix C presents the results. The marginal effects of some covariates are calculated in a different way. This is the case of the income that is computed as $\gamma_{lincome}(\alpha) * \left[1/\overline{income}\right][Q_{y^*}(\alpha|\widetilde{\mathbf{x}}) - \alpha]$, the "age when male" that is set as $[\gamma_{age}(\alpha) + 2\gamma_{age^2}(\alpha) * \overline{age} + 3\gamma_{age^3}(\alpha) * \overline{age}^2] * [Q_{y^*}(\alpha|\widetilde{\mathbf{x}}) - \alpha]$, and the "age when female" that is $[\gamma_{age}(\alpha) + \gamma_{agexfemale}(\alpha) + 2(\gamma_{age^2}(\alpha) + \gamma_{(agexfemale)^2}(\alpha)) * \overline{age} + 3(\gamma_{age^3}(\alpha) + \gamma_{(agexfemale)^3}(\alpha)) * \overline{age}^2] * [Q_{y^*}(\alpha|\widetilde{\mathbf{x}}) - \alpha]$.

[32] The "default" individual is a healthy man with a household of 3 persons, 9 years of schooling, €500 of monthly income, not single or retired, living in the Centre region of Portugal and interviewed in autumn. Also note that, the vector $\widetilde{\mathbf{x}}$ is set with the dummies *pubsub* and *privsub* equal to zero, so the "default" individual has the NHS insurance plan (belongs the control group). It is worth mentioning that setting the continuous variables at the median or at the mean produces very similar marginal effects.

more than one crossing.[33]

Table V: Quantile regression results: semi-elasticities (SE)

	SE(0.25)	SE(0.50)	SE(0.60)	SE(0.70)	SE(0.80)	SE(0.90)
Health insurance status variables						
pubsub	0.027	0.029	0.033	0.037	0.034	0.031
privsub	0.072	0.082	0.093	0.096	0.093	0.087
Health status variables						
sick	0.316	0.264	0.266	0.302	0.332	0.633
limitdays	0.024	0.024	0.026	0.028	0.033	0.041
limited	0.047	0.073	0.093	0.139	0.181	0.242
rheumatism	0.047	0.048	0.049	0.055	0.072	0.088
osteoporosis	0.106	0.073	0.066	0.060	0.056	0.052
cancer	0.194	0.189	0.178	0.172	0.226	0.376
kidneystones	0.052	0.053	0.063	0.076	0.112	0.128
renalfailure	0.059	0.079	0.078	0.093	0.135	0.143
emphysema	0.031	0.075	0.082	0.093	0.119	0.146
cerebralhemorrhage	0.046	0.046	0.048	0.065	0.096	0.113
infarction	0.113	0.123	0.135	0.149	0.153	0.131
depressivedisorder	0.067	0.083	0.092	0.105	0.127	0.153
otherchronicaldisease	0.177	0.182	0.199	0.213	0.213	0.229
highbloodpressure	0.163	0.149	0.147	0.139	0.135	0.125
chronicpain	0.061	0.070	0.081	0.095	0.113	0.137
diabetes	0.184	0.142	0.134	0.136	0.155	0.184
asthma	0.109	0.122	0.134	0.148	0.144	0.140
stress	0.180	0.138	0.135	0.131	0.155	0.154
smoker	-0.060	-0.052	-0.051	-0.052	-0.041	-0.018
meals	0.067	0.055	0.045	0.044	0.039	0.039
Socioeconomic characteristics variables						
householdsize	-0.021	-0.019	-0.020	-0.022	-0.018	-0.009
age when male*	0.005	0.004	0.005	0.006	0.004	0.003
age when female*	0.001	0.001	0.000	0.000	-0.001	-0.001
female	0.152	0.147	0.164	0.201	0.171	0.172
educmax	0.003	0.005	0.005	0.006	0.004	0.003
income**	0.005	0.004	0.004	0.005	0.005	0.004
single	-0.064	-0.057	-0.060	-0.072	-0.069	-0.060
student	-0.072	-0.070	-0.079	-0.082	-0.075	-0.086
retired	0.059	0.051	0.047	0.045	0.058	0.084
Geographic variables						
Norte	-0.016	-0.011	-0.014	-0.015	-0.024	-0.049
Lisboa	-0.027	-0.020	-0.026	-0.030	-0.040	-0.054
Alentejo	-0.078	-0.063	-0.066	-0.064	-0.066	-0.072
Algarve	-0.071	-0.059	-0.059	-0.057	-0.061	-0.083
Açores	-0.101	-0.093	-0.101	-0.116	-0.131	-0.161
Madeira	-0.134	-0.128	-0.142	-0.168	-0.184	-0.208
Seasonality variables						
winter	0.060	0.060	0.065	0.069	0.067	0.086
spring	0.032	0.031	0.034	0.035	0.034	0.038
summer	0.016	0.012	0.008	0.000	-0.009	-0.001

Notes: Semi-elasticities marked with * or ** were multiplied by 10 or 100, respectively.

[33] The comparison with one-part models should be done by looking at the quantiles near the mean.

As we have already mentioned, in the quantile regression framework it is possible that a significant coefficient of a variable on $y^*_\alpha-$ quantile does not affect a particular conditional $y_\alpha-$ quantile. Nevertheless, when it is found that the $y^*_\alpha-$ quantile depends on the covariate for several quantiles, then it should be possible to detect a subpopulation for which the semi-elasticity on $y_\alpha-$ quantile is different from zero (Machado and Santos-Silva (2005) and Miranda (2008)). For clarity take the following example, if we consider the median and compute the $Q_{y^*}(0.50|x = \widetilde{\mathbf{x}})$ we obtain 0.73 and as a consequence the $Q_y(0.50|x = \widetilde{\mathbf{x}})$ is equal to zero consultations. When considering an individual equal to the "default" except that he has diabetes, then Q_{y^*} is higher (0.84), but the estimated median for that individual is still zero. Hence the marginal effect of "diabetes" on the $y_\alpha-$ quantile is zero, even though it has a significant positive effect on the $y^*_\alpha-$ quantile. Conversely, if the sixth decile is used the $Q_{y^*}(0.60|x = \widetilde{\mathbf{x}})$ is equal to 0.90 and as a consequence $Q_y(0.60|x = \widetilde{\mathbf{x}})$ is also equal to zero consultations. But now, a diabetic has Q_{y^*} equal to 1.03, making Q_y equal to one consultation which means that the estimated impact of diabetes on the sixth decile of the outcome distribution is one additional visit to a doctor.

Starting with the impact of double coverage, it is visible that the insurance coefficients do not change a lot across the estimated quantiles, but it is possible to find a pattern: both public and private subsystems have an increasing positive effect on the number of doctor visits until the sixth/seventh decile and a decreasing positive effect thereafter (Table V). It can be concluded that having extra insurance through health subsystems is important to determine whether or not to visit a doctor and is slightly less relevant in explaining further visits to a doctor. Moreover, the similarities between the patterns of both subsystems are clear when we compute the ratio between them across quantiles, since it remains almost unchanged. In fact, the effect of private subsystem insurance plans is between 2.6 and 2.9 times higher than the impact of the public employees' insurance plans. This result shows that health insurance double coverage does lead to further use of health care and its public/private nature is also quite important, as private subsystems double coverage induces much more consumption than public subsystems double coverage. This evidence may be related to some differences between subsystems regarding the reimbursing policy. Overall, the positive effect of double coverage is consistent with the results obtained for the expected number of visits using the traditional parametric models. (see B1 of Appendix

B). We tested the exogeneity of the double coverage variables with the construction of a sub-sample that potentially includes only indirect beneficiaries of public subsystems (spouses and descendants), by considering population covered by public health subsystems that do not work in general government. The results are presented and discussed in Appendix E.

To better understand the effect of health subsystems on health care utilisation we used the point estimates to predict the $y_\alpha-$ quantile (note that here we consider the relevant outcome) for each observation in a simulation exercise in which all variables are set equal to their actual values, except the health insurance status. The latter assumes three distinct possible outcomes: only NHS, NHS plus a public subsystem and NHS plus a private subsystem. Results measured by relative frequencies are presented in Table VI.

Table VI: Frequencies of estimated quantiles for the number of visits to a doctor

	0	1	2	3	4	5	6	7	8	9	$\geqslant 10$
						NHS					
$Q_y\widehat{(0.25}\|x)$	89.4	8.3	1.4	0.4	0.2	0.1	0.1	0.0	0.0	0.0	0.0
$Q_y\widehat{(0.50}\|x)$	58.2	32.8	5.5	1.7	0.7	0.4	0.2	0.1	0.1	0.1	0.2
$Q_y\widehat{(0.75}\|x)$	1.3	69.3	17.9	5.6	2.5	1.1	0.7	0.5	0.2	0.2	0.7
$Q_y\widehat{(0.90}\|x)$	0.0	23.4	46.3	15.1	6.2	3.1	1.8	1.1	0.7	0.5	1.8
						Public subsystem					
$Q_y\widehat{(0.25}\|x)$	87.9	9.4	1.6	0.5	0.3	0.1	0.1	0.0	0.0	0.0	0.0
$Q_y\widehat{(0.50}\|x)$	54.0	35.7	6.3	2.0	0.9	0.5	0.2	0.2	0.1	0.1	0.2
$Q_y\widehat{(0.75}\|x)$	0.7	65.7	20.1	6.5	2.9	1.4	0.8	0.5	0.3	0.2	0.8
$Q_y\widehat{(0.90}\|x)$	0.0	19.5	47.2	16.6	6.7	3.5	1.9	1.2	0.8	0.6	2.0
						Private subsystem					
$Q_y\widehat{(0.25}\|x)$	83.6	12.3	2.4	0.8	0.3	0.2	0.1	0.1	0.1	0.1	0.1
$Q_y\widehat{(0.50}\|x)$	46.8	40.3	7.5	2.6	1.2	0.6	0.3	0.2	0.1	0.1	0.3
$Q_y\widehat{(0.75}\|x)$	0.2	60.0	23.4	7.6	3.6	1.8	0.9	0.7	0.5	0.3	1.0
$Q_y\widehat{(0.90}\|x)$	0.0	13.2	47.7	19.5	7.7	4.0	2.3	1.5	1.0	0.7	2.4

Notes: Estimates are based on a simulation exercise that start by predicting the $y^*_\alpha-$ quantile for all 35,308 individuals setting all control variables in their actual values except the health insurance status, which is set in the three possible cases. After that, the $y_\alpha-$ quantiles are computed applying $Q_y(\alpha|x) = \lceil Q_{y^*}(\alpha|x) - 1\rceil$ and tabulated according to their frequencies.

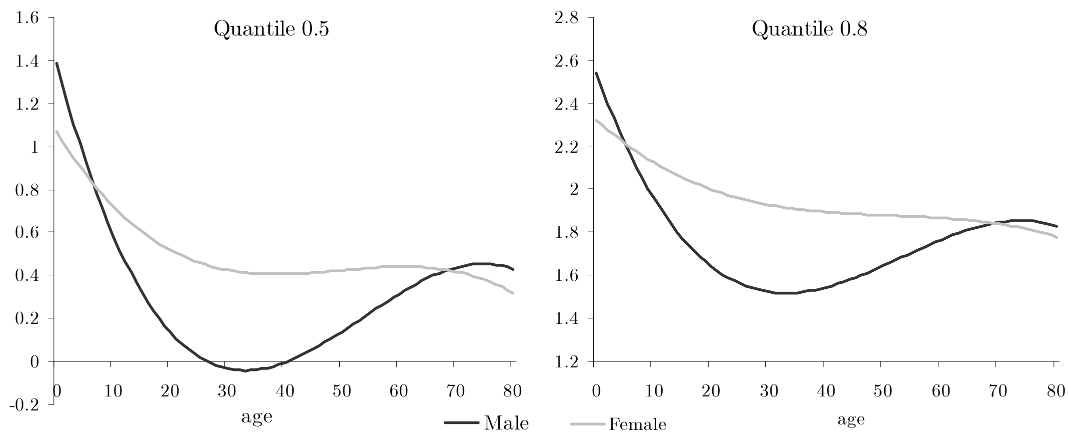
Given that half of the sample has zero visits, it is not surprising that the first conditional quartile is zero for almost all observations. When the estimates from different quantiles are compared, we have the perception that the distribution changes differently across the health insurance plans. For instance, the proportion of individuals with a predicted quantile of zero or one consultation is always lower with double coverage, but these relative effects change between quantiles. In particular, when covered solely by NHS, the proportion is 91.0, 70.7 and 23.4 per cent for the $0.50y-$, $0.75y-$, $0.90y-$quantile, respectively, while with a private subsystem on top of the NHS the proportion is 87.1, 60.2 and 13.2 per cent for the $0.50y-$, $0.75y-$, $0.90y-$quantile, respectively. This means that being double covered causes a decreasing path in the difference of proportion of individuals with a certain (increasing) number of visits that is steeper from the $0.50y-$quantile to the $0.75y-$quantile than from the $0.75y-$quantile to the $0.90y-$quantile.

Regarding the effects of health status variables as a whole, it is visible that most of the regressors have a positive effect that increases with α, which means that the distribution of the number of doctor consultations for individuals that are sicker is displaced upwards and more spread-out. As expected, having an episode of illness, seems especially important to determine whether or not the individual visits a doctor and, taking into consideration the results of the last decile, it is even more important in explaining the subsequent visits. This last finding is in accordance with the results derived from the two-part models (Table B2 and B3 of Appendix B). The effect of long term incapacity becomes gradually more relevant, as for the first quantiles it is not significant and for higher levels of consumption it has a very important impact. Amongst the chronic diseases we found evidence of a positive increasing effect along the estimated quantiles, except for the dummy "osteoporosis" that has a decreasing impact, and for "infarction", "otherchronicaldisease", "highbloodpressure", "diabetes" and "asthma" that have a constant effect in the different parts of the distribution. When compared with the parametric models, we conclude that the signs of the coefficients are equal and the magnitude of the relative effects between variables present minor differences. The proxy for the level of exposure to stress has an effect that does not vary much across quantiles, and the other regressors related to attitudes towards health care have decreasing effects. The negative and decreasing impact of being a smoker contrasts with the results of Lourenço (2007), which although using a slightly different variable found positive effects on the consumption of visits to a doctor. Another interesting result is that having the habit of eating several times a day has also a positive impact. These results show that individuals that take better care of their health by not smoking and having a higher number of meals also complement their care by being more pro-active in visiting a doctor. These attitudes towards health care seem to more than offset the impact of the improved health stemming from non-smoking and having a higher number of meals. The comparison between these findings and the ones derived from the two-part models validates this conclusion. When we study the impact of being a smoker on the decision to visit or not to visit the doctor we found a positive impact, which means that the probability of having zero visits is higher for the smoker than for the non-smoker. Regarding the second part of the model (the choice on the number of consultations), we found that in similar conditions a smoker decides, on average, to consume more doctor visits than a non-smoker (the sign is positive, although not always significant).

Socioeconomic characteristics seem to have a similar impact across quantiles. Concerning the household size effect, the results indicate that an individual consumes on average less consultations if the number of members of his/her household is larger. These findings are in accordance with the ones found in Winkelmann (2006). A possible economic explanation for this effect is the presence of "economies of experience" within the family due to the fact that decisions taken by more than one person benefit from more in-depth information, which on its turn influence health status and efficiency in producing healthy times. It is also plausible that scale economies play a role if it is true that when visiting a doctor patients often also ask for symptoms of diseases of their relatives in order to prevent further visits.

Regarding the effect of age, from Figure 1 we see that the utilisation of health care is very high in the first years of life and decreases until $30 - 40$ years old, more for men than for women, and thereafter it increases for men while remaining fairly constant for women. These results seem intuitive and are consistent with the literature: the initial decreasing path may be related to the fact that children often require more health care (having therefore periodic doctor appointments); and after some point in the life cycle it is expected an increasing use of health services both if we consider that age is a health status proxy or a indicator of the depreciation rate (Grossman 1972). Most of the applications studying health care demand consider that age has a quadratic relationship with the utilisation of medical services (Pohlmeier and Ulrich 1995, Winkelmann 2006 and Lourenço 2007). We tried to introduce the quadratic relation in our specification, but both coefficients did not appear significant and we found that a third order polynomial allows a much better fit to the data. Additionally, we modelled the ageing and gender effects together. Note that in our specification, it makes little sense to interpret the dummy "female" alone. The advantages of assessing the ageing effect by gender type are clear from Figure 1: men tend to consume less while women's behaviour towards health demand is smoother over the life cycle. Comparing the effects of age on the median to the ones on the $0.80y-$ quantile, we observe that the pattern of the effects is similar, but the impact of age is less pronounced in explaining high levels of visits to a doctor. This last result is very much in line with Winkelmann (2006) that shows that age in the upper tail of the distribution of the number of visits has an insignificant effect.

Figure 1: Effect of age in the 0.5y*– quantile and 0.8y*–quantile



Note: The curves are computed as $Q_{y^*}(\alpha|\widetilde{\mathbf{x}})+\gamma_{female}+\left[\gamma_{age}(\alpha)+\gamma_{femaleage}(\alpha)female\right]*age+\left[\gamma_{age^2}(\alpha)+\gamma_{femaleage^2}(\alpha)female\right]*age^2+\left[\gamma_{age^3}(\alpha)+\gamma_{femaleage^3}(\alpha)female\right]*age^3$ being the set of covariates equal $\widetilde{\mathbf{x}}$ except for age that is variable.

The level of income has a positive but negligible effect on the utilisation of health care, constant across the different quantiles. Conceptually, it is possible to find at least two channels of income influences. The first is derived from the Grossman's model (1972), in which the income determines the budget constraint and, therefore, the ability to pay for health care. The second channel is related to the fact that different levels of income can explain differences in the opportunity cost of being ill and in the cost of visiting the doctor, especially if we closely relate income with the wage rate. In Portugal, the first channel may not be important because of the features of its health care system. This is broadly applicable to both private and public subsystems and to NHS beneficiaries, although to the latter in a minor extent. Direct costs of beneficiaries are relatively small as most of the cost of a consultation is borne by the health care system, which is financed predominantly by general taxation or by employers and employees compulsory contributions. In this context, the second channel can be more relevant and it is consistent with the estimated small effect of income over all the outcome distribution. Also the educational level has a small positive impact that does not change significantly across the estimated quantiles. This appears to indicate that individuals with high educational levels face a higher opportunity cost of being ill and this more than offset the opportunity cost of visiting the doctor. Moreover, there is no evidence supporting the idea that more educated people are able to improve health more efficiently generating fewer doctors' consultations. The previous empirical evidence of Pohlmeier and Ulrich (1995), Winkelmann (2006) and Lourenço (2007) also found small positive effects for both income and education variables.

Concerning the effect of marital status, results point out that single people visit doctors less often. These findings may indicate that they are less risk-averse regarding their health. As to the occupational status, the estimated semi-elasticities are positive for retired individuals and negative for students, meaning that the utilisation of health care increases over the life cycle, being lower when we study, higher when we work and much higher when we retire. In the interpretation of the results we should be aware that these particular variables may capture to some extent Grossman's income and age effects: traditionally, students are the youngest in the database and retirees the oldest.

Finally, the coefficients related with the area of residence indicate that individuals from "Centro" utilise more consultations, followed by "Norte" and "Lisboa". Individuals from the autonomous regions consume much less health care services than the ones from mainland. Regarding the estimated effects in the different quantiles, we conclude that they are more or less constant for most of the cases and growing in the case of "Norte" and "Lisboa". The seasonality variables indicate that individuals consume less visits to doctors in the summer while in the autumn their doctor's consultations reach a peak.

5.3 Further results: cumulative health effects of double coverage

As mentioned in Section 4.3, some individuals may have enjoyed health insurance double coverage for a long period of time which may generate cumulative health benefits from a hypothetical better medical follow-up accumulated over time. If this occurs, the difference in the number of consultations between the "treated" and "control" groups could decrease with age. The idea is that recent beneficiaries of a health subsystem (more likely the younger generations) did not have time to accumulate such health benefits, whereas the older beneficiaries (more likely the older generations) had time to do so, and that will make them relatively healthier when compared with "untreated" individuals. If this behaviour is not fully controlled by the health status variables, the estimated double coverage effect can be positively biased. Following Barros et al. (2008), we estimate our specification in different age groups, within the quantile regression framework. To do that we worked with three subsamples: individuals with more than 18 years old, a younger cohort with people between 18 to 45 years old and a older cohort with people between 45 to 80 years old. Table VII presents the double coverage coefficients (and standard errors).

Table VII: Double coverage in different age groups

	$\widehat{\beta(0.25)}$	$\widehat{\beta(0.50)}$	$\widehat{\beta(0.60)}$	$\widehat{\beta(0.70)}$	$\widehat{\beta(0.80)}$	$\widehat{\beta(0.90)}$
age $\geq$ 18						
pubsub	0.070	0.081	0.084	0.077	0.054	0.032
	(0.035)	(0.031)	(0.031)	(0.030)	(0.026)	(0.027)
privsub	0.174	0.213	0.224	0.198	0.140	0.103
	(0.071)	(0.060)	(0.057)	(0.050)	(0.045)	(0.054)
18 $\leq$ age $\leq$ 45 (N= 12637)						
pubsub	0.220	0.209	0.193	0.177	0.114	0.029
	(0.060)	(0.056)	(0.054)	(0.056)	(0.051)	(0.046)
privsub	0.176	0.284	0.345	0.384	0.300	0.255
	(0.137)	(0.129)	(0.123)	(0.119)	(0.093)	(0.111)
45 $\leq$ age $\leq$ 80 (N=16637)						
pubsub	0.002	0.031	0.034	0.036	0.048	0.042
	(0.044)	(0.037)	(0.035)	(0.032)	(0.032)	(0.033)
privsub	0.193	0.213	0.184	0.152	0.091	0.022
	(0.077)	(0.065)	(0.058)	(0.050)	(0.044)	(0.058)

Notes: "N" means the number of observations in each subsample. Values in brackets correspond to the standard errors. Results for other variables can be found from Table C3 to Table C8 in Appendix C.

The most important fact is that the effects of both public and private subsystems are higher for the younger generations and this occurs in the whole distribution. When we restrict the analysis to individuals with more than eighteen years old, thus rising the average age, both insurance coefficients decrease (slightly more in the upper tail of the distribution), whereas the younger cohort (individuals with more than eighteen and less than forty five) has the largest estimated treatment effects. The differences are very expressive, especially for public employees. This is consistent with Barros et al. (2008) findings. For different levels of visits to a doctor, beneficiaries from private subsystems and public subsystems now behave in a different way. Regarding the public subsystems, quantile regression results show that the effect of supplementary insurance of the younger cohort decreases considerably across the distribution, which is points to a double coverage impact relatively lower among young high users. Note that this is clearly more marked than the findings resulting from the full sample. For the private subsystem, the estimated impact of the younger group increases until the $0.70y^*$- quantile and decreases thereafter. This path is similar to that one obtained for the full sample. The results seem to confirm the suspicion that the estimated effects for the elder groups are lower, possibly reflecting accumulated health benefits from the existence of the subsystems.[34] In this context the best indicator of moral hazard would be one obtained from the sample of individuals that possibly did not have time to incorporate such benefits. The caveat is the reduction of the sample, in particular of the "treated" individuals.

6 Conclusions

This paper analyses the impact of additional coverage on the utilisation of doctor consultations at different levels of the outcome distribution, contributing to the empirical literature on moral hazard and equity in the health care sector. Using a recent quantile regression method for count data, we overcome a limitation of traditional parametric count data models by investigating the effect of supplementary insurance on the whole outcome distribution without imposing restrictive assumptions. The application to the Portuguese case, allows us to discard the selection bias problem by using only individuals who benefit from health insurance double coverage (through subsystems) on a mandatory basis and by analysing its impact in different age cohorts.

Our results show that the additional insurance is very important in explaining the utilisation of doctor visits in all levels of usage, being slightly higher for medium-intensity users. That is, double coverage leads to a relatively higher increase in the utilisation of

[34] The final report of a commission that recently assessed the sustainability of the NHS funding concluded, however, that there is no evidence that the additional consumption of health care services have impact on the self-assessment of health status.

visits to a doctor for regular users of the health system vis-à-vis heavy users. When the effects of public and private health insurance plans are compared it seems that double coverage derived from private health insurance is much higher than the one derived from the health insurance plan of public employees. The analysis for the youngest cohort shows that the estimated effects of both public and private health insurance on top of the NHS are higher than the ones for the full sample, possibly reflecting accumulated health benefits.

To explain the differences in the utilisation of health care between the different health insurance status we control for several demographic, socioeconomic and health status variables, as well as geographic and seasonal effects. Not surprisingly, results indicate that the existence of chronic disease or pain is extremely relevant in explaining doctor visits, especially for high users. Among the demographic and socioeconomic characteristics, age (also as proxy of health status) assumes a unique role, especially when combined with gender. In the first years of living the consumption of health care is very high and it decreases until 30-40 years old, more for men than for women, and thereafter it increases for men and remains fairly constant for women. Education and income present significant positive effects (constant over the whole distribution) although less important than those of other regressors. Results from quantile regression are similar to those from previous literature in terms of the significance of key covariates, but the combination of age and gender is novel in the literature.

In short, health insurance double coverage generates additional utilisation of health care. This additional consumption effect is probably only slightly higher for medium-intensity users than for heavy users. Another interesting piece of evidence is the large difference in impact according to the type of health insurance double coverage. The second layer of health insurance coverage adds more to utilisation when provided by private organizations than when obtained from government financed entities.

References

Bago-d'Uva, T. (2006). Latent class models for utilisation of health care. *Health Economics 15*, 329–343.

Bago-d'Uva, T. and A. Jones (2009). Health care utilisation in Europe: New evidence from the ECHP. *Journal of Health Economics 28*(2), 265–279.

Barros, P., M. Machado, and A. Galdeano (2008). Moral hazard and the demand for health services: A matching estimator approach. *Journal of Health Economics 27*(4), 1006–1025.

Barros, P. P. and J. Simões (2007). *Portugal: Health systems review*, Volume 9. Health systems in Transition.

Bitler, M., J. Gelbach, and H. Hoynes (2006). What mean impacts miss: Distributional effects of welfare reform experiments. *The American Economic Review 96*(4), 987–1012.

Cameron, A. and P. Trivedi (2005). *Microeconometrics: methods and applications.* Cambridge.

Cameron, A., P. Triverdi, F. Milne, and J. Piggott (1988). A microeconometric model of the demand for health care and health insurance in Australia. *The Review of Economic Studies 55*(1), 85–106.

Comissão para a Sustentabilidade do Financiamento do Serviço Nacional de Saúde (Ed.) (2006). *Relatório* F*inal.*

Coulson, N., J. Terza, C. Neslusan, and B. Stuart (1995). Estimating the moral-hazard effect of supplemental medical insurance in the demand for prescription drugs by the elderly. *The American Economic Review 85*(2), 122–126.

Deb, P. and P. Trivedi (2002). The structure of demand for health care: latent class versus two-part models. *Journal of Health Economics 21*, 601–625.

Gourieroux, C., A. Monfort, and A. Trognon (1984). Pseudo maximum likelihood methods: applications to Poisson models. *Econometrica 52*, 701–721.

Grossman, M. (1972). On the concept of health capital and the demand for health. *The Journal of Political Economy 80*(2), 223–255.

Hausman, J., B. Hall, and Z. Griliches (1984). Econometric models for count data with an application to the patents RD relationship. *Econometrica 52*, 909–938.

Horowitz, J. L. (1992). A smooth maximum score estimatior for the binary response model. *Econometrica 60*, 505–531.

Jiménez-Martín, S., J. Lebeaga, and M. Martinez-Granado (2002). *Latent class versus two-part models in the demand for physician services across the European Union*, Section II7, pp. 101–116. Econometric Analysis of Health Data. Wiley.

Jones, A., X. Koolman, and E. van Doorslaer (2006). The impact of having supplementary private health insurance on the use of specialists. *Annales DtÉconomie et de Statistique 83/84*, 251–275.

Jones, A., N. Rice, and P. Contoyannis (2006). *The Dynamics of health.* Elgar Companion to Health Economics.

Koenker, R. and G. Bassett (1978). Regression quantiles. *Econometrica 46*, 33–50.

Liu, C. (2007). Utilization of general practitioners' services in Canada and the United States: a quantile regression for counts analysis. *mimeo*.

Lourenço, O. D. (2007). *Unveiling health care consumption groups.* PhD dissertation, Faculdade de Economia da Universidade de Coimbra.

Machado, J. and J. Santos-Silva (2005). Quantiles for counts. *Journal of the American Statistical Association 100*, 1226–1237.

Manning, W., L. Blumberg, and L. Moulton (1995). The demand for alcohol: The differential response to price. *Journal of Health Economics 14*(2), 123–148.

Manski, C. F. (1975). Maximum score estimation of the stochastic utility model of choice. *Journal of Econometrics 3*, 205–228.

Manski, C. F. (1985). Semiparametric analysis of discrete response: Asymptotic properties of the maximum score estimatior. *Journal of Econometrics 27*, 313–333.

Miranda, A. (2006). Qcount: Stata program to fit quantile regression models for count data. *Statistical Software Components S456714, Boston College Department of Economics*.

Miranda, A. (2008). Planned fertility and the family background: a quantile regression for count analysis. *Journal of Polulation Economics 21*, 67–81.

Pereira, J. (1995). Equity, health and health care: an economic study with reference to Portugal. *Department of Economics and Related Studies of the University of York 341*.

Pohlmeier, W. and V. Ulrich (1995). An econometric model of the two-part decisionmaking process in the demand for health care. *The Journal of Human Resources 30*(2), 339–361.

van Doorslaer, E. and A. Jones (2004). Income-related inequality in health and health care in the European Union. *Health Economics 13*, 605–608.

Vera-Hernández, A. (1999). Duplicate coverage and demand for health care: The case of Catalonia. *Health Economics 8*, 123–148.

Windweijer, F. and J. Santos-Silva (1997). Endogeneity in count data models: An application to demand for health care. *Journal of Applied Econometrics 12*(3), 281–294.

Winkelmann, R. (2004). Health care reform and the number of doctor visits: An econometric analysis. *Journal of Applied Econometrics 19*, 455–472.

Winkelmann, R. (2006). Reforming health care: evidence from quantile regressions for counts. *Journal of Health Economics 25*, 131–145.

Winkelmann, R. (2008). *Econometric analysis of count data*. Springer.

Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data*. MIT press.

Zweifel, P. and W. G. Manning (2000). *Moral hazard and consumer incentives in health care*, Volume 1A of *Handbook of Health Economics*, Section I6, pp. 409–459. Elsevier Science.

7 Appendix

A Portuguese health case system

Table A1: Comparison between the NHS and the health subsystems

	NHS	Public subsystems	Private subsystems
Membership	Health care system for all residents in the continent regardless of economic and social background (universal coverage). Regional Health Systems with the same characteristics exist in the Azores and Madeira.	Compulsory and based on professional or occupational career (independent of the risk of falling sick). The beneficiaries are public employees (and their spouses and dependants) and public pensioners. There are specific schemes for some ministries/professions, like the military personnel who have their own subsystem. **Examples**: ADSE (*Assistência na Doença dos Servidores do Estado*), SSMJ (*Serviços Sociais do Ministério da Justiça*), ADMA (*Assistência na Doença aos Militares da Armada*), ADME (*Assistência na Doença aos Militares do Exercito*), ADFA (*Assistência na Doença aos Militares da Força Aérea*), SAD PSP (*Assistência na Doença da Polícia de Segurança Pública*), SAD GNR (*Serviços de Assistência à Doença à GNR*).	Some are compulsory and based on professional or occupational career (independent of the risk of falling sick). The beneficiaries are workers (and their spouses and dependants) and pensioners of sectors of activity or important private companies (in general large companies) that have their own health insurance schemes. **Examples**: SAMS (*Serviços de Assistência Médico-Social*) for banking employees and pensioners (and their dependants); PT-ACS (*Portugal Telecom- Associação de Cuidados de Saúde*) for the employees and pensioners (and their dependants) of the main telecomunications operator.
Delivery systems	Provides health care directly: on NHS providers (very limited in certain services, as dental care).	NHS providers: with equal rights and obligations relative to NHS counterparts (same costs sharing mechanism and access problems); Private providers with contract; Provides health care directly (not in all subsystems and services).	NHS providers: with equal rights and obligations relative to NHS counterparts (same costs sharing mechanism and access problems); Private providers with contract; Provides health care directly (not in all subsystems and services).
Direct costs to beneficiaries	NHS providers (acute hospital care, general practice and mother and child care): with low co-payments and an important part of beneficiaries exempted; Private (specialist and dental consultations, diagnosis services, renal dialysis and physiotherapy treatments) with full payment. Health expenditure supported by beneficiaries is subsidised by the State through a tax credit at a rat of 30% (without upper limits).	NHS providers: with equal copayments between NHS beneficiaries and subsystems beneficiaries. Law requires that the subsystems pay the NHS for the medical care; Subsystems providers and private providers with contract: available with low copayments; Private providers without contract: paying the full cost of the service which aftwards is partially reimbursed (75-80% up to a certain limit, dependent of the type of service). Health expenditure supported by beneficiaries is subsidised by the State through a tax credit at a rat of 30% (without upper limits).	NHS providers: with equal copayments between NHS beneficiaries and subsystems beneficiaries. Law requires that the subsystems pay the NHS for the medical care; Subsystems providers and private providers with contract: available with low copayments; Private providers without contract: paying the full cost of the service which aftwards is partially reimbursed (75-80% up to a certain limit, dependent of the type of service). Health expenditure supported by beneficiaries is subsidised by the State through a tax credit at a rat of 30% (without upper limits).
Financing sources	Predominantly by general taxation.	By employers and employees compulsory contributions (a percentage of 1.5% of gross wage in 2007, up from a historical contribution of 1%, regardless of the health status and the number of members of the household) topped with money from general government budget. The employees expenditure represents only 10% of total expenditure.	By employers and employees compulsory payroll contributions (a percentage of the remuneration, regardless of the health status and the number of members of the household).

B Results from preliminary model

Table B1: Results from single-crossing models

	Poisson pseudo-ml		Negbin I		Negbin II		Poisson log-normal	
	Coef.	S.E	Coef.	S.E	Coef.	S.E	Coef.	S.E.
Health insurance status variables								
pubsub	$0.056^{\dagger}$	(0.024)	0.061	(0.020)	$0.061^{\dagger}$	(0.025)	0.067	(0.022)
privsub	0.150	(0.043)	0.159	(0.037)	0.168	(0.044)	0.187	(0.044)
Health status variables								
sick	0.538	(0.079)	0.501	(0.067)	0.627	(0.081)	0.611	(0.067)
limitdays	0.056	(0.003)	0.052	(0.002)	0.068	(0.003)	0.066	(0.002)
limited	0.272	(0.082)	0.181	(0.059)	0.329	(0.089)	0.259	(0.058)
rheumatism	0.089	(0.023)	0.101	(0.018)	0.101	(0.022)	0.120	(0.020)
osteoporosis	0.120	(0.029)	0.136	(0.022)	0.154	(0.032)	0.163	(0.027)
cancer	0.440	(0.056)	0.354	(0.043)	0.515	(0.059)	0.455	(0.043)
kidneystones	0.123	(0.032)	0.124	(0.027)	0.161	(0.031)	0.168	(0.030)
renalfailure	0.256	(0.079)	0.139	(0.052)	0.393	(0.097)	0.240	(0.051)
emphysema	0.121	(0.040)	0.107	(0.034)	0.176	(0.040)	0.166	(0.037)
cerebralhemorrhage	$0.124^{\dagger}$	(0.051)	0.122	(0.041)	0.128	(0.048)	0.137	(0.047)
infarction	$0.126^{\dagger}$	(0.051)	0.180	(0.046)	0.179	(0.047)	0.226	(0.053)
depressivedisorder	0.168	(0.028)	0.160	(0.023)	0.212	(0.029)	0.213	(0.026)
otherchronicaldisease	0.326	(0.019)	0.315	(0.015)	0.364	(0.019)	0.366	(0.016)
highbloodpressure	0.214	(0.022)	0.232	(0.017)	0.249	(0.022)	0.265	(0.019)
chronicpain	0.158	(0.023)	0.145	(0.018)	0.201	(0.023)	0.196	(0.020)
diabetes	0.270	(0.029)	0.259	(0.021)	0.314	(0.029)	0.309	(0.025)
asthma	0.243	(0.035)	0.224	(0.027)	0.275	(0.038)	0.263	(0.030)
stress	0.279	(0.026)	0.273	(0.020)	0.305	(0.026)	0.310	(0.022)
smoker	$-0.039^{\ddagger}$	(0.033)	-0.099	(0.023)	$-0.041^{\ddagger}$	(0.032)	-0.077	(0.024)
meals	$0.089^{\dagger}$	(0.036)	0.096	(0.028)	0.099	(0.036)	0.104	(0.030)
Socioeconomic and demographic variables								
householdsize	-0.023	(0.009)	-0.035	(0.006)	-0.031	(0.009)	-0.035	(0.006)
age	-0.753	(0.072)	-0.729	(0.050)	-0.752	(0.069)	-0.739	(0.054)
age^2	0.166	(0.019)	0.160	(0.014)	0.163	(0.019)	0.160	(0.015)
age^3	-0.010	(0.001)	-0.010	(0.001)	-0.010	(0.001)	-0.010	(0.001)
age*female	0.489	(0.098)	0.460	(0.069)	0.459	(0.095)	0.454	(0.074)
$(age*female)^2$	-0.113	(0.026)	-0.104	(0.019)	-0.107	(0.026)	-0.104	(0.020)
$(age*female)^3$	0.007	(0.002)	0.006	(0.001)	0.007	(0.002)	0.006	(0.002)
female	$-0.252^{\dagger}$	(0.103)	-0.240	(0.067)	$-0.224^{\dagger}$	(0.099)	-0.219	(0.072)
educmax	$0.004^{\ddagger}$	(0.002)	0.007	(0.002)	$0.006^{\dagger}$	(0.002)	$0.007^{\dagger}$	(0.002)
lincome	$0.012^{\ddagger}$	(0.018)	$0.029^{\dagger}$	(0.013)	$0.025^{\ddagger}$	(0.018)	0.034	(0.014)
single	-0.142	(0.035)	-0.147	(0.027)	-0.150	(0.033)	-0.150	(0.028)
student	-0.205	(0.040)	-0.194	(0.028)	-0.185	(0.039)	-0.194	(0.030)
retired	0.147	(0.028)	0.120	(0.021)	0.160	(0.029)	0.147	(0.025)
Geographic variables								
Norte	-0.110	(0.028)	-0.075	(0.023)	-0.090	(0.028)	-0.080	(0.026)
Lisboa	-0.108	(0.029)	-0.087	(0.023)	-0.100	(0.029)	-0.094	(0.027)
Alentejo	-0.173	(0.029)	-0.167	(0.024)	-0.176	(0.030)	-0.179	(0.027)
Algarve	-0.127	(0.032)	-0.147	(0.024)	-0.140	(0.032)	-0.159	(0.027)
Açores	-0.306	(0.034)	-0.298	(0.025)	-0.302	(0.034)	-0.331	(0.028)
Madeira	-0.378	(0.040)	-0.412	(0.027)	-0.396	(0.040)	-0.448	(0.000)
Seasonality variables								
winter	0.125	(0.024)	0.120	(0.019)	0.145	(0.024)	0.143	(0.021)
spring	0.079	(0.025)	0.069	(0.019)	0.079	(0.024)	0.077	(0.021)
summer	$0.036^{\ddagger}$	(0.025)	$0.024^{\ddagger}$	(0.019)	$0.036^{\ddagger}$	(0.026)	$0.026^{\ddagger}$	(0.022)
Constant	$0.303^{\dagger}$	(0.145)	$0.193^{\ddagger}$	(0.104)	$0.165^{\ddagger}$	(0.139)	$-0.217^{\dagger}$	(0.111)
α			0.790	(0.031)	0.662	(0.025)	0.768	(0.008)

Notes: Coefficients marked with ‡ and †are not significative at a 5 and 1 per cent level, respectively. The S.E. for Poisson, NegBinI and NegBin II are robust standard errors. Coefficient α is the dispersion parameter (for the Negbin I, the $Var(y|x) = \lambda + \alpha$, and for the Negbin II, the $Var(y|x) = \lambda(1 + \alpha)$) and if is equal 0 corresponds to dispersion equal one, and it is simply a Poisson.

Table B2: Results from Hurdle count data models

	Poisson Hurdle model				$Pr(y=0)$		$E(y\|y>0)$	
	$Pr(y=0)$		$E(y\|y>0)$		Logit	Probit	ZT NegBin I and II	
	Coef.	S.E.	Coef.	S.E.	Coef.	Coef.	Coef.	Coef.
Health insurance status variables								
pubsub	-0.068	(0.024)	0.039‡	(0.034)	-0.094	-0.059	0.096‡	0.044‡
privsub	-0.209	(0.054)	0.101‡	(0.057)	-0.287	-0.176	0.231†	0.138‡
Health status variables								
sick	-1.054	(0.215)	0.423	(0.079)	-1.338	-0.782	0.558	0.637
limitdays	-0.149	(0.010)	0.045	(0.003)	-0.168	-0.087	0.064	0.068
limited	-0.046‡	(0.081)	0.284	(0.097)	-0.069‡	-0.039‡	0.357	0.443
rheumatism	-0.160	(0.030)	0.073†	(0.029)	-0.190	-0.110	0.176	0.115
osteoporosis	-0.427	(0.051)	0.056‡	(0.036)	-0.520	-0.296	0.110†	0.106†
cancer	-0.666	(0.093)	0.407	(0.062)	-0.809	-0.456	0.470	0.580
kidneystones	-0.207	(0.049)	0.114	(0.037)	-0.253	-0.147	0.164	0.207
renalfailure	-0.421	(0.113)	0.298	(0.087)	-0.490	-0.263	0.106‡	0.587
emphysema	-0.229	(0.059)	0.141	(0.046)	-0.265	-0.137	0.184†	0.244
cerebralhemorrhage	-0.245	(0.092)	0.121†	(0.061)	-0.269†	-0.138†	0.199†	0.168†
infarction	-0.526	(0.109)	0.076‡	(0.055)	-0.622	-0.333	0.222†	0.135‡
depressivedisorder	-0.285	(0.045)	0.161	(0.033)	-0.333	-0.186	0.268	0.251
otherchronicaldisease	-0.422	(0.019)	0.209	(0.026)	-0.569	-0.340	0.380	0.290
highbloodpressure	-0.435	(0.027)	0.082	(0.029)	-0.563	-0.333	0.160	0.121
chronicpain	-0.233	(0.030)	0.136	(0.029)	-0.293	-0.172	0.212	0.233
diabetes	-0.600	(0.045)	0.158	(0.036)	-0.760	-0.441	0.201	0.225
asthma	-0.324	(0.042)	0.189	(0.043)	-0.435	-0.255	0.335	0.222
stress	-0.568	(0.040)	0.178	(0.032)	-0.706	-0.410	0.293	0.245
smoker	0.111	(0.022)	0.089‡	(0.047)	0.192	0.116	0.041‡	0.124†
meals	-0.126	(0.031)	0.015‡	(0.049)	-0.224	-0.131	0.012‡	0.020‡
Socioeconomic and demographic variables								
householdsize	0.041	(0.006)	0.016‡	(0.013)	0.068	0.041	0.014‡	0.004‡
age	0.631	(0.053)	-0.351	(0.112)	1.000	0.618	-0.826	-0.416
age^2	-0.126	(0.015)	0.072†	(0.029)	-0.204	-0.127	0.191	0.079†
age^3	0.007	(0.001)	-0.004†	(0.002)	0.012	0.008	-0.012†	-0.005‡
age*female	-0.293	(0.076)	0.279‡	(0.148)	-0.483	-0.299	1.020	0.300†
(age*female)^2	0.062	(0.022)	-0.065‡	(0.038)	0.105	0.064	-0.243	-0.071†
(age*female)^3	-0.004†	(0.002)	0.004‡	(0.003)	-0.006†	-0.004†	0.015†	0.004‡
female	0.162†	(0.072)	-0.129‡	(0.161)	0.256†	0.163†	-0.606‡	-0.106‡
educmax	-0.013	(0.002)	-0.006‡	(0.003)	-0.018	-0.011	-0.008‡	-0.006‡
lincome	-0.075	(0.015)	-0.033‡	(0.025)	-0.102	-0.061	-0.036‡	-0.033‡
single	0.141	(0.027)	-0.070‡	(0.051)	0.219	0.134	-0.206†	-0.096‡
student	0.151	(0.026)	-0.135†	(0.067)	0.230	0.145	-0.897	-0.135†
retired	-0.191	(0.035)	0.126	(0.037)	-0.251	-0.148	0.184	0.175
Geographic variables								
Norte	0.017‡	(0.031)	-0.152	(0.038)	0.028‡	0.019†	-0.293	-0.174
Lisboa	0.015‡	(0.032)	-0.127	(0.039)	0.036‡	0.024†	-0.241	-0.162
Alentejo	0.183	(0.031)	-0.113	(0.038)	0.276	0.169	-0.220	-0.137
Algarve	0.182	(0.030)	-0.037‡	(0.043)	0.271	0.166	-0.097‡	-0.057‡
Açores	0.240	(0.030)	-0.259	(0.051)	0.361	0.222	-0.724	-0.318
Madeira	0.402	(0.030)	-0.202	(0.062)	0.584	0.362	-0.628	-0.238
Seasonality variables								
winter	-0.138	(0.022)	0.084†	(0.034)	-0.199	-0.120	0.155	0.130
spring	-0.072	(0.022)	0.062‡	(0.035)	-0.111	-0.066	0.111†	0.065†
summer	-0.035‡	(0.022)	0.027‡	(0.036)	-0.055‡	-0.032‡	0.015‡	0.020‡
Constant	-0.198‡	(0.117)	0.718	(0.210)	0.115‡	0.037‡	0.143‡	-0.159‡
α							1.309	2.181

Notes: Coefficients marked with ‡ and †are not significative at a 5 and 1 per cent level, respectively. The S.E. are robust standard error.

Table B3: Results from Zero-Inflated count data models

	Zero-inflated Poisson				Zero-inflated Negbin II			
	y		Inflated		y		Inflated	
	Coef.	S.E	Coef.	S.E	Coef.	S.E	Coef.	S.E
Health insurance status variables								
pubsub	0.036‡	(0.030)	-0.121‡	(0.076)	0.038‡	(0.028)	-0.195‡	(0.117)
privsub	0.094‡	(0.051)	-0.384†	(0.170)	0.109†	(0.047)	-0.745†	(0.322)
Health status variables								
sick	0.455	(0.074)	-1.923†	(1.016)	0.541	(0.072)	-35.821	(1.919)
limitdays	0.044	(0.003)	-15.416	(1.588)	0.058	(0.003)	-35.596	(0.234)
limited	0.202†	(0.092)	0.148‡	(0.24)	0.229	(0.089)	0.357‡	(0.477)
rheumatism	0.063†	(0.025)	-0.398	(0.115)	0.081	(0.022)	-1.281	(0.338)
osteoporosis	0.089	(0.031)	-1.235	(0.285)	0.146	(0.029)	-2.513†	(1.090)
cancer	0.377	(0.060)	-1.150	(0.424)	0.442	(0.054)	-38.326	(0.622)
kidneystones	0.116	(0.034)	-0.299‡	(0.175)	0.147	(0.030)	-0.962†	(0.477)
renalfailure	0.281	(0.082)	-0.827‡	(0.531)	0.361	(0.086)	-37.050	(0.631)
emphysema	0.110	(0.043)	-0.519†	(0.206)	0.146	(0.04)	-1.302	(0.454)
cerebralhemorrhage	0.101‡	(0.055)	-0.825†	(0.408)	0.119†	(0.048)	-2.089‡	(1.431)
infarction	0.085‡	(0.049)	-1.945	(0.73)	0.159	(0.044)	-16.444	(4.085)
depressivedisorder	0.135	(0.030)	-0.658	(0.197)	0.169	(0.027)	-13.955	(2.893)
otherchronicaldisease	0.167	(0.022)	-0.841	(0.071)	0.203	(0.020)	-1.799	(0.175)
highbloodpressure	0.084	(0.026)	-1.084	(0.125)	0.152	(0.022)	-3.086	(0.904)
chronicpain	0.124	(0.026)	-0.395	(0.109)	0.151	(0.023)	-1.407	(0.325)
diabetes	0.179	(0.031)	-1.331	(0.202)	0.247	(0.029)	-2.849	(1.043)
asthma	0.160	(0.040)	-0.602	(0.137)	0.169	(0.038)	-1.310	(0.302)
stress	0.202	(0.028)	-1.194	(0.189)	0.269	(0.024)	-20.554	(0.901)
smoker	0.039‡	(0.044)	0.296	(0.080)	-0.005‡	(0.038)	0.229‡	(0.117)
meals	0.072‡	(0.045)	-0.122‡	(0.112)	0.113	(0.039)	0.106‡	(0.166)
Socioeconomic and demographic variables								
householdsize	0.007‡	(0.011)	0.091	(0.019)	-0.005‡	(0.010)	0.109	(0.026)
age	-0.338	(0.097)	1.045	(0.175)	-0.429	(0.090)	1.414	(0.270)
age^2	0.063†	(0.026)	-0.225	(0.050)	0.082	(0.024)	-0.293	(0.080)
age^3	-0.004‡	(0.002)	0.014	(0.004)	-0.005	(0.002)	0.018	(0.007)
age*female	0.240‡	(0.124)	-0.436‡	(0.234)	0.238†	(0.113)	-0.670‡	(0.349)
$(age*female)^2$	-0.053‡	(0.032)	0.113‡	(0.070)	-0.052‡	(0.030)	0.163‡	(0.108)
$(age*female)^3$	0.003‡	(0.002)	-0.009‡	(0.006)	0.003‡	(0.002)	-0.013‡	(0.010)
female	-0.118‡	(0.133)	0.202‡	(0.220)	-0.100‡	(0.122)	0.385‡	(0.321)
educmax	-0.006†	(0.003)	-0.045	(0.007)	-0.004‡	(0.003)	-0.077	(0.012)
lincome	-0.022‡	(0.022)	-0.212	(0.050)	-0.001‡	(0.019)	-0.270	(0.074)
single	-0.085‡	(0.045)	0.201†	(0.089)	-0.088†	(0.039)	0.362	(0.127)
student	-0.115†	(0.057)	0.189†	(0.090)	-0.098‡	(0.052)	0.310†	(0.129)
retired	0.124	(0.031)	-0.286†	(0.123)	0.155	(0.028)	-0.253‡	(0.248)
Geographic variables								
Norte	-0.141	(0.033)	-0.168‡	(0.094)	-0.109	(0.03)	-0.141‡	(0.164)
Lisboa	-0.124	(0.034)	-0.227†	(0.100)	-0.109	(0.031)	-0.311‡	(0.180)
Alentejo	-0.135	(0.034)	0.181†	(0.092)	-0.143	(0.032)	0.324†	(0.160)
Algarve	-0.059‡	(0.039)	0.283	(0.091)	-0.074†	(0.035)	0.479	(0.154)
Açores	-0.247	(0.043)	0.189†	(0.095)	-0.236	(0.038)	0.438	(0.154)
Madeira	-0.168	(0.052)	0.643	(0.094)	-0.159	(0.048)	1.233	(0.160)
Seasonality variables								
winter	0.064†	(0.029)	-0.257	(0.070)	0.087	(0.027)	-0.420	(0.108)
spring	0.060†	(0.030)	-0.089‡	(0.069)	0.063†	(0.027)	-0.139‡	(0.102)
summer	0.027‡	(0.031)	-0.053‡	(0.070)	0.021‡	(0.028)	-0.106‡	(0.105)
Constant	0.699	(0.179)	0.444‡	(0.369)	0.360†	(0.158)	-0.365‡	(0.560)
α					0.491	(0.046)		

Notes: Coefficients marked with ‡ and †are not significative at a 5 and 1 per cent level, respectively. The S.E. are robust standard error. The inflated distribution is a probit in the fist model and a logit in the second.

Table B4: Results from Zero-Inflated count data models

	parameters	log-likelihood	BIC	AIC	CAIC
One-part models					
Poisson pseudo-ml	44	-49,123	98,706	49,211	98,750
Negbin I	45	-45,122	90,715	45,212	90,760
Negbin II	45	-45,274	91,020	45,364	91,065
Poisson log-normal	45	-44,586	89,643	44,676	89,688
Hurdle models					
Poisson Hurdle	88	-47,100	95,122	47,276	95,210
Logit-Negbin I	89	-44,314	89,559	44,492	89,648
Probit-Negbin II	89	-44,238	89,408	44,416	89,497
Zero-inflated models					
Zero-inflated Poisson	88	-47,010	94,942	47,186	95,030
Zero-inflated Negbin II	89	-44,565	90,062	44,743	90,151

Note: The BIC, AIC and CAIC are popular measures for comparing maximum likelihood models.

C Results from quantile regression

Table C1: Quantile regression results: standard errors

	$\widehat{\beta(0.25)}$	$\widehat{\beta(0.50)}$	$\widehat{\beta(0.60)}$	$\widehat{\beta(0.70)}$	$\widehat{\beta(0.80)}$	$\widehat{\beta(0.90)}$
Health insurance status variables						
pubsub	0.032	0.029	0.028	0.028	0.024	0.025
privsub	0.070	0.055	0.053	0.051	0.046	0.053
Health status variables						
sick	0.097	0.095	0.081	0.106	0.120	0.097
limitdays	0.004	0.003	0.003	0.003	0.003	0.004
limited	0.094	0.081	0.080	0.078	0.077	0.073
rheumatism	0.032	0.025	0.024	0.023	0.023	0.024
osteoporosis	0.039	0.031	0.030	0.029	0.028	0.032
cancer	0.065	0.050	0.042	0.047	0.056	0.068
kidneystones	0.045	0.039	0.038	0.038	0.035	0.038
renalfailure	0.075	0.069	0.064	0.065	0.059	0.067
emphysema	0.060	0.049	0.043	0.043	0.046	0.050
cerebralhemorrhage	0.076	0.055	0.054	0.061	0.051	0.062
infarction	0.079	0.063	0.063	0.053	0.050	0.065
depressivedisorder	0.042	0.035	0.032	0.030	0.030	0.034
otherchronicaldisease	0.025	0.021	0.020	0.019	0.017	0.019
highbloodpressure	0.030	0.024	0.022	0.021	0.020	0.022
chronicpain	0.031	0.026	0.025	0.023	0.022	0.024
diabetes	0.037	0.028	0.027	0.028	0.028	0.030
asthma	0.047	0.038	0.036	0.035	0.035	0.037
stress	0.035	0.027	0.024	0.024	0.025	0.025
smoker	0.035	0.031	0.033	0.034	0.028	0.028
meals	0.044	0.041	0.040	0.039	0.036	0.036
Socioeconomic and demographic variables						
householdsize	0.010	0.009	0.008	0.009	0.008	0.007
age	0.080	0.067	0.066	0.071	0.063	0.059
age^2	0.022	0.019	0.019	0.020	0.018	0.017
age^3	0.002	0.002	0.002	0.002	0.001	0.001
age*female	0.112	0.095	0.093	0.092	0.082	0.085
(age*female)^2	0.031	0.026	0.026	0.026	0.023	0.023
(age*female)^3	0.002	0.002	0.002	0.002	0.002	0.002
female	0.109	0.092	0.089	0.083	0.074	0.085
educmax	0.003	0.003	0.003	0.003	0.002	0.002
lincome	0.021	0.018	0.019	0.018	0.016	0.017
single	0.041	0.037	0.038	0.040	0.034	0.034
student	0.042	0.039	0.039	0.042	0.034	0.034
retired	0.037	0.029	0.027	0.026	0.028	0.027
Geographic variables						
Norte	0.037	0.033	0.032	0.031	0.029	0.030
Lisboa	0.040	0.033	0.032	0.031	0.030	0.033
Alentejo	0.039	0.034	0.034	0.034	0.031	0.033
Algarve	0.039	0.036	0.035	0.033	0.029	0.032
Açores	0.039	0.035	0.035	0.035	0.032	0.031
Madeira	0.042	0.037	0.036	0.039	0.036	0.033
Seasonality variables						
winter	0.031	0.026	0.027	0.027	0.023	0.024
spring	0.031	0.027	0.027	0.027	0.024	0.023
summer	0.031	0.027	0.027	0.027	0.024	0.025
Constant	0.166	0.144	0.143	0.141	0.121	0.130

Table C2: Quantile regression results: marginal effects

	ME(0.25)	ME(0.50)	ME(0.60)	ME(0.70)	ME(0.80)	ME(0.90)
Health insurance status variables						
pubsub	0.012	0.027	0.037	0.051	0.059	0.069
privsub	0.033	0.075	0.104	0.130	0.158	0.194
Health status variables						
sick	0.147	0.240	0.298	0.412	0.567	1.412
limitdays	0.011	0.022	0.029	0.038	0.057	0.092
limited	0.022	0.066	0.104	0.189	0.310	0.540
rheumatism	0.022	0.044	0.055	0.075	0.124	0.196
osteoporosis	0.049	0.067	0.074	0.082	0.095	0.116
cancer	0.090	0.172	0.199	0.235	0.386	0.839
kidneystones	0.024	0.048	0.071	0.104	0.192	0.285
renalfailure	0.027	0.072	0.088	0.127	0.231	0.320
emphysema	0.014	0.068	0.092	0.127	0.203	0.326
cerebralhemorrhage	0.021	0.042	0.053	0.089	0.163	0.253
infarction	0.052	0.112	0.152	0.203	0.262	0.293
depressivedisorder	0.031	0.076	0.104	0.144	0.216	0.341
otherchronicaldisease	0.082	0.166	0.223	0.291	0.364	0.512
highbloodpressure	0.076	0.135	0.164	0.190	0.231	0.280
chronicpain	0.028	0.063	0.091	0.129	0.192	0.305
diabetes	0.086	0.129	0.150	0.186	0.264	0.411
asthma	0.051	0.112	0.150	0.202	0.246	0.314
stress	0.084	0.126	0.151	0.178	0.264	0.344
smoker	-0.028	-0.047	-0.057	-0.071	-0.070	-0.040
meals	0.031	0.050	0.051	0.060	0.066	0.088
Socioeconomic and demographic variables						
householdsize	-0.010	-0.017	-0.022	-0.030	-0.030	-0.021
age when male	0.002	0.003	0.005	0.007	0.006	0.007
age when female	0.000	0.000	0.000	-0.001	-0.002	-0.002
female	0.074	0.141	0.194	0.292	0.308	0.402
educmax	0.002	0.004	0.006	0.008	0.007	0.007
lincome	0.002	0.003	0.004	0.005	0.007	0.006
single	-0.030	-0.052	-0.068	-0.098	-0.118	-0.133
student	-0.034	-0.064	-0.088	-0.112	-0.128	-0.192
retired	0.028	0.047	0.053	0.061	0.099	0.186
Geographic variables						
Norte	-0.008	-0.010	-0.016	-0.021	-0.042	-0.110
Lisboa	-0.013	-0.019	-0.029	-0.041	-0.068	-0.121
Alentejo	-0.036	-0.058	-0.073	-0.088	-0.113	-0.161
Algarve	-0.033	-0.054	-0.066	-0.077	-0.105	-0.185
Açores	-0.047	-0.085	-0.113	-0.159	-0.223	-0.360
Madeira	-0.063	-0.117	-0.159	-0.229	-0.314	-0.465
Seasonality variables						
winter	0.028	0.055	0.072	0.094	0.114	0.191
spring	0.015	0.028	0.039	0.048	0.057	0.086
summer	0.007	0.011	0.009	0.000	-0.016	-0.003

Notes: Marginal effects marked with * or ** were multiplied by 10 or 100, respectively.

Table C3: Quantile regression results: estimated coefficients when age>=18

	$\widehat{\beta(0.25)}$	$\widehat{\beta(0.50)}$	$\widehat{\beta(0.60)}$	$\widehat{\beta(0.70)}$	$\widehat{\beta(0.80)}$	$\widehat{\beta(0.90)}$
Health insurance status variables						
pubsub	$0.070^{\dagger}$	0.081	0.084	0.077	$0.054^{\dagger}$	$0.032^{\ddagger}$
privsub	$0.174^{\dagger}$	0.213	0.224	0.198	0.140	$0.103^{\ddagger}$
Health status variables						
sick	0.746	0.664	0.648	0.625	0.582	0.805
limitdays	0.065	0.067	0.068	0.067	0.064	0.066
limited	$0.165^{\ddagger}$	0.234	0.271	0.334	0.352	0.379
rheumatism	0.141	0.150	0.149	0.145	0.149	0.148
osteoporosis	0.294	0.222	0.198	0.166	0.128	0.104
cancer	0.461	0.454	0.424	0.375	0.382	0.494
kidneystones	0.159	0.168	0.187	0.201	0.228	0.214
renalfailure	$0.151^{\dagger}$	0.193	0.197	0.198	0.238	0.215
emphysema	$0.052^{\ddagger}$	0.152	0.169	0.165	0.184	0.213
cerebralhemorrhage	$0.135^{\ddagger}$	0.136	$0.128^{\dagger}$	$0.143^{\dagger}$	0.191	0.195
infarction	0.303	0.347	0.347	0.342	0.297	0.209
depressivedisorder	0.201	0.241	0.256	0.253	0.249	0.234
otherchronicaldisease	0.379	0.393	0.407	0.391	0.340	0.320
highbloodpressure	0.417	0.390	0.372	0.321	0.259	0.209
chronicpain	0.175	0.198	0.221	0.223	0.211	0.212
diabetes	0.449	0.364	0.336	0.307	0.289	0.293
asthma	0.232	0.268	0.283	0.278	0.241	0.193
stress	0.454	0.368	0.345	0.310	0.286	0.245
smoker	-0.209	-0.182	-0.175	-0.157	-0.095	$-0.036^{\ddagger}$
meals	0.189	0.160	0.130	0.107	0.092	$0.089^{\dagger}$
Socioeconomic and demographic variables						
householdsize	-0.049	-0.048	-0.049	-0.048	-0.031	$-0.012^{\ddagger}$
age	$-0.365^{\ddagger}$	$-0.319^{\ddagger}$	$-0.359^{\ddagger}$	$-0.139^{\ddagger}$	$0.177^{\ddagger}$	$-0.061^{\ddagger}$
age^2	$0.094^{\ddagger}$	$0.084^{\ddagger}$	$0.096^{\ddagger}$	$0.062^{\ddagger}$	$-0.021^{\ddagger}$	$0.022^{\ddagger}$
age^3	$-0.006^{\ddagger}$	$-0.005^{\ddagger}$	$-0.006^{\ddagger}$	$-0.004^{\ddagger}$	$0.001^{\ddagger}$	$-0.001^{\ddagger}$
age*female	$-0.078^{\ddagger}$	$0.032^{\ddagger}$	$0.226^{\ddagger}$	$0.133^{\ddagger}$	$-0.108^{\ddagger}$	$0.285^{\ddagger}$
$(\text{age*female})^2$	$0.007^{\ddagger}$	$-0.022^{\ddagger}$	$-0.068^{\ddagger}$	$-0.066^{\ddagger}$	$0.004^{\ddagger}$	$-0.070^{\ddagger}$
$(\text{age*female})^3$	$-0.001^{\ddagger}$	$0.001^{\ddagger}$	$0.004^{\ddagger}$	$0.005^{\ddagger}$	$0.000^{\ddagger}$	$0.004^{\ddagger}$
female	$0.660^{\ddagger}$	$0.548^{\ddagger}$	$0.357^{\ddagger}$	$0.683^{\ddagger}$	$0.701^{\ddagger}$	$0.011^{\ddagger}$
educmax	$0.004^{\ddagger}$	0.010	0.011	0.010	$0.006^{\dagger}$	$0.002^{\ddagger}$
lincome	0.068	0.057	0.058	0.054	0.046	$0.020^{\ddagger}$
single	-0.215	-0.189	-0.192	-0.202	-0.147	-0.100
student	$0.039^{\ddagger}$	$0.016^{\ddagger}$	$0.014^{\ddagger}$	$0.052^{\ddagger}$	$0.064^{\ddagger}$	$0.033^{\ddagger}$
retired	0.190	0.166	0.152	0.135	0.137	0.163
Geographic variables						
Norte	$-0.078^{\dagger}$	$-0.060^{\ddagger}$	$-0.068^{\dagger}$	$-0.070^{\dagger}$	$-0.082^{\dagger}$	-0.115
Lisboa	$-0.098^{\dagger}$	$-0.078^{\dagger}$	-0.099	-0.107	-0.119	-0.130
Alentejo	-0.308	-0.256	-0.261	-0.228	-0.195	-0.167
Algarve	-0.245	-0.205	-0.200	-0.173	-0.161	-0.171
Açores	-0.372	-0.348	-0.372	-0.376	-0.356	-0.357
Madeira	-0.494	-0.482	-0.532	-0.554	-0.492	-0.468
Seasonality variables						
winter	0.161	0.161	0.165	0.151	0.119	0.126
spring	0.099	0.098	0.101	0.089	0.065	$0.064^{\dagger}$
summer	$0.076^{\dagger}$	$0.061^{\dagger}$	$0.044^{\ddagger}$	$0.015^{\ddagger}$	$-0.017^{\ddagger}$	$0.000^{\ddagger}$
Constant	-2.076	-1.457	-1.206	-1.280	$-1.039^{\dagger}$	$-0.033^{\ddagger}$

Note: The subsample has 28,736 observations. Coefficients marked with ‡ and †are not significative at a 5 and 1 per cent level, respectively.

Table C4: Quantile regression results: estimated coefficients when 18<=age<=45

	$\widehat{\beta(0.25)}$	$\widehat{\beta(0.50)}$	$\widehat{\beta(0.60)}$	$\widehat{\beta(0.70)}$	$\widehat{\beta(0.80)}$	$\widehat{\beta(0.90)}$
Health insurance status variables						
pubsub	0.220	0.209	0.193	0.177	0.114†	0.029‡
privsub	0.176‡	0.284†	0.345	0.384	0.300	0.255†
Health status variables						
sick	0.254‡	0.515†	0.686†	0.732	0.634	0.651†
limitdays	0.142	0.134	0.132	0.130	0.120	0.105
limited	-0.044‡	0.092‡	0.117‡	0.218‡	0.347‡	0.492
rheumatism	0.239	0.203†	0.229	0.263	0.267	0.276
osteoporosis	0.497‡	0.524	0.501	0.469†	0.380	0.301‡
cancer	0.556	0.438†	0.422	0.312†	0.276‡	0.482‡
kidneystones	0.136‡	0.163‡	0.270‡	0.406	0.444	0.458
renalfailure	0.046‡	0.071‡	0.009‡	0.309‡	0.366‡	0.420‡
emphysema	0.040‡	0.105‡	0.218‡	0.389†	0.400	0.344†
cerebralhemorrhage	0.170‡	0.014‡	0.371‡	0.702‡	0.722‡	0.657†
infarction	0.971‡	0.882‡	0.814‡	0.872‡	0.857‡	0.494‡
depressivedisorder	0.370	0.421	0.425	0.438	0.393	0.365
otherchronicaldisease	0.507	0.552	0.613	0.631	0.521	0.437
highbloodpressure	0.435	0.535	0.560	0.524	0.425	0.294
chronicpain	0.218	0.270	0.341	0.376	0.368	0.360
diabetes	0.465	0.368	0.392	0.415	0.372	0.380
asthma	0.217†	0.250	0.257	0.311	0.273	0.278
stress	0.808	0.730	0.724	0.664	0.560	0.427
smoker	-0.136	-0.140‡	-0.124	-0.105†	-0.072‡	-0.012‡
meals	0.178†	0.142	0.130‡	0.135‡	0.118‡	0.057‡
Socioeconomic and demographic variables						
householdsize	-0.036†	-0.041	-0.046	-0.053	-0.045	-0.024‡
age	-1.326‡	-1.285‡	-1.263‡	-2.047‡	-2.445‡	-1.387‡
age^2	0.372‡	0.350‡	0.343‡	0.597‡	0.765‡	0.441‡
age^3	-0.033‡	-0.030‡	-0.030‡	-0.056‡	-0.077‡	-0.046‡
age*female	1.639‡	2.061‡	1.851‡	2.295‡	2.049‡	1.084‡
(age*female)^2	-0.552‡	-0.671‡	-0.605‡	-0.744‡	-0.690‡	-0.373‡
(age*female)^3	0.054‡	0.065‡	0.059‡	0.072‡	0.070‡	0.039‡
female	-0.989‡	-1.471‡	-1.242‡	-1.610‡	-1.391‡	-0.648‡
educmax	0.022	0.023	0.024	0.024	0.017	0.008‡
lincome	0.120	0.116	0.113	0.100†	0.068	0.034‡
single	-0.185	-0.177	-0.194	-0.237	-0.234	-0.166
student	-0.063‡	-0.068‡	-0.065‡	-0.040‡	0.003‡	0.018‡
retired	0.216‡	0.260‡	0.287‡	0.433‡	0.345‡	0.221‡
Geographic variables						
Norte	0.016‡	0.020‡	-0.003‡	-0.022‡	-0.028‡	-0.055‡
Lisboa	0.110‡	0.117‡	0.101‡	0.081‡	0.036‡	-0.004‡
Alentejo	-0.221	-0.243	-0.277	-0.317	-0.227	-0.188
Algarve	-0.257	-0.260	-0.297	-0.308	-0.201	-0.174
Açores	-0.263	-0.270	-0.313	-0.387	-0.379	-0.346
Madeira	-0.573	-0.568	-0.625	-0.739	-0.771	-0.537
Seasonality variables						
winter	0.256	0.252	0.250	0.250	0.199	0.172
spring	0.127†	0.136†	0.138†	0.132†	0.103‡	0.075‡
summer	0.063‡	0.076‡	0.076‡	0.051‡	0.022‡	0.070‡
Constant	-1.775‡	-1.039‡	-0.801‡	0.336‡	1.392‡	1.160‡

Notes: The subsample has 12,637 observations. Coefficients marked with ‡ and †are not significant at a 5 and 1 per cent level, respectively.

Table C5: Quantile regression results: estimated coefficients when 45<=age<=80

	$\widehat{\beta(0.25)}$	$\widehat{\beta(0.50)}$	$\widehat{\beta(0.60)}$	$\widehat{\beta(0.70)}$	$\widehat{\beta(0.80)}$	$\widehat{\beta(0.90)}$
Health insurance status variables						
pubsub	0.002‡	0.031‡	0.034‡	0.036‡	0.048‡	0.042‡
privsub	0.193	0.213	0.184	0.152	0.091†	0.022‡
Health status variables						
sick	0.722	0.656	0.628	0.556	0.445	0.742
limitdays	0.053	0.051	0.051	0.050	0.049	0.048
limited	0.107‡	0.211†	0.252	0.268	0.308	0.309
rheumatism	0.147	0.156	0.154	0.144	0.152	0.151
osteoporosis	0.317	0.249	0.216	0.192	0.149	0.110
cancer	0.475	0.456	0.413	0.364	0.380	0.447
kidneystones	0.165	0.170	0.192	0.194	0.201	0.178
renalfailure	0.201	0.250	0.259	0.240	0.246	0.234
emphysema	0.098‡	0.156	0.181	0.186	0.189	0.220
cerebralhemorrhage	0.146†	0.171	0.142	0.136	0.181	0.196
infarction	0.314	0.372	0.362	0.334	0.286	0.189
depressivedisorder	0.142	0.176	0.193	0.188	0.179	0.152
otherchronicaldisease	0.322	0.316	0.307	0.282	0.264	0.246
highbloodpressure	0.404	0.352	0.324	0.271	0.230	0.173
chronicpain	0.164	0.171	0.179	0.181	0.173	0.170
diabetes	0.446	0.354	0.319	0.287	0.278	0.264
asthma	0.219	0.243	0.236	0.216	0.192	0.132
stress	0.387	0.307	0.279	0.254	0.230	0.194
smoker	-0.291	-0.263	-0.271	-0.216	-0.130	-0.081†
meals	0.201	0.167	0.148	0.126	0.108	0.113
Socioeconomic and demographic variables						
householdsize	-0.035†	-0.033	-0.030	-0.021‡	-0.005‡	0.004‡
age	-0.857‡	-0.175‡	0.755‡	1.489‡	0.550‡	-1.802‡
age^2	0.169‡	0.050‡	-0.090‡	-0.213‡	-0.080‡	0.307‡
age^3	-0.010‡	-0.003‡	0.004‡	0.011‡	0.004‡	-0.017‡
age*female	1.925‡	1.759‡	0.622‡	-0.536‡	1.306‡	4.042‡
(age*female)^2	-0.324‡	-0.308‡	-0.149‡	0.035‡	-0.243‡	-0.682‡
(age*female)^3	0.017‡	0.016‡	0.009‡	0.000‡	0.014‡	0.037‡
female	-3.286‡	-2.842‡	-0.182	2.154‡	-1.890‡	-7.495‡
educmax	-0.002‡	0.003‡	0.004‡	0.003‡	0.000‡	-0.001‡
lincome	0.055†	0.041‡	0.046†	0.043†	0.041†	0.017‡
single	-0.302	-0.232	-0.209	-0.152	-0.111†	-0.072‡
student	0.820	0.484‡	0.580†	0.424‡	0.154‡	0.094‡
retired	0.215	0.190	0.177	0.156	0.154	0.175‡
Geographic variables						
Norte	-0.114†	-0.106	-0.120	-0.116	-0.125	-0.156
Lisboa	-0.145	-0.145	-0.174	-0.175	-0.195	-0.190
Alentejo	-0.340	-0.272	-0.266	-0.221	-0.200	-0.169
Algarve	-0.224	-0.175	-0.163	-0.139	-0.157	-0.162
Açores	-0.391	-0.369	-0.378	-0.356	-0.348	-0.341
Madeira	-0.418	-0.413	-0.451	-0.410	-0.402	-0.397
Seasonality variables						
winter	0.127	0.131	0.128	0.111	0.108	0.103
spring	0.081†	0.088	0.086	0.076	0.065†	0.044‡
summer	0.075‡	0.054‡	0.028‡	-0.009‡	-0.035‡	-0.043
Constant	-0.798‡	-1.347‡	-3.134‡	-4.205‡	-1.706‡	3.552‡

Notes: The subsample has 16,637 observarions. Coefficients marked with ‡ and †are not significant at a 5 and 1 per cent level, respectively.

D Further results: the regressor "unemployed"

As already mentioned, when selecting the variables, we guaranteed that "treated" and "untreated" groups have a common support by using only observations in the intersection of the domains. Another problem that can emerge is the strong correlation between some potential regressors and the double coverage dummies. One of the cases that deserves special attention is the dummy "unemployed". Similar studies include this regressor (see for example Winkelmann (2006)). Theoretically, it is reasonable to include it because it captures the occupational status that influences the opportunity cost of visiting a doctor, the income effect and the depreciation rate of health. Clearly, the individuals that do not work present lower opportunity costs (both time and monetary related) when visiting a doctor when compared to individuals that have a job. In accordance, we anticipate that these covariates tend to increase the utilisation of doctor visits. Moreover, unemployed individuals may also present a higher rate of health depreciation because being unemployed can induce a high level of stress. In our analysis its inclusion causes some problems. Table D1 shows that there is a clear different proportion of unemployed between subsystems, which is related to the different working conditions, specially in the case of public sector employees vis-à-vis private employees. As such, it can happen that this variable (if included in the specification) may capture some treatment effects.

Table D1: Unemployment status by subsystem

	NHS		Public subsystems		Private subsystems	
unemployed	Freq	%	Freq.	%	Freq.	%
= 0	27,423	95.3	5,552	99.5	932	97.9
= 1	1,355	4.7	26	0.5	20	2.1

The solution chosen for this issue was to exclude this variable from the conditional quantile specification. To analyse the robustness of this option, we re-estimate the selected specification excluding the unemployed individuals from the sample. As a whole, the results are very similar and, in particular, the health insurance coefficients lead to the same conclusion, although with a level sightly higher (see Table D2).

Table D2: Quantile regression results: estimated coefficients without unemployed individuals

	$\widehat{\beta(0.25)}$	$\widehat{\beta(0.50)}$	$\widehat{\beta(0.60)}$	$\widehat{\beta(0.70)}$	$\widehat{\beta(0.80)}$	$\widehat{\beta(0.90)}$
Health insurance status variables						
pubsub	0.087	0.092	0.100	0.101	0.075	$0.056^{\dagger}$
privsub	0.226	0.246	0.262	0.243	0.192	0.155
Health status variables						
sick	0.693	0.619	0.601	0.607	0.547	0.775
limitdays	0.071	0.073	0.075	0.074	0.071	0.073
limited	$0.144^{\ddagger}$	0.215	0.254	0.325	0.339	0.373
rheumatism	0.143	0.142	0.143	0.144	0.152	0.151
osteoporosis	0.259	0.186	0.167	0.143	0.108	$0.086^{\dagger}$
cancer	0.470	0.464	0.433	0.387	0.400	0.509
kidneystones	0.152	0.157	0.176	0.188	0.228	0.220
renalfailure	$0.151^{\ddagger}$	0.211	0.208	0.229	0.267	0.233
emphysema	$0.091^{\ddagger}$	0.205	0.220	0.224	0.226	0.229
cerebralhemorrhage	$0.135^{\ddagger}$	$0.128^{\dagger}$	$0.127^{\dagger}$	0.162	0.197	0.184
infarction	0.290	0.323	0.339	0.336	0.292	0.222
depressivedisorder	0.177	0.223	0.238	0.244	0.238	0.245
otherchronicaldisease	0.437	0.452	0.473	0.463	0.383	0.351
highbloodpressure	0.412	0.385	0.368	0.321	0.256	0.203
chronicpain	0.178	0.208	0.231	0.236	0.224	0.222
diabetes	0.430	0.350	0.319	0.295	0.278	0.293
asthma	0.320	0.351	0.361	0.355	0.284	0.239
stress	0.433	0.357	0.339	0.301	0.289	0.246
smoker	-0.202	-0.173	-0.165	-0.152	-0.091	$-0.039^{\ddagger}$
meals	0.186	0.160	0.131	$0.116^{\dagger}$	0.079	$0.066^{\dagger}$
Socioeconomic characteristics variables						
householdsize	-0.064	-0.061	-0.059	-0.059	-0.039	-0.019
age	-1.058	-1.008	-1.041	-1.069	-0.733	-0.574
age^2	0.229	0.219	0.228	0.241	0.162	0.124
age^3	-0.014	-0.013	-0.014	-0.015	-0.010	-0.008
age*female	0.532	0.562	0.619	0.727	0.477	0.332
$(age*female)^2$	-0.113	-0.125	-0.141	-0.176	-0.113	-0.076
$(age*female)^3$	$0.006^{\dagger}$	0.007	0.008	0.011	0.007	$0.004^{\dagger}$
female	-0.315	-0.312	-0.335	-0.349	-0.212	$-0.093^{\ddagger}$
educmax	0.010	0.015	0.015	0.015	0.010	$0.006^{\dagger}$
lincome	0.070	0.058	0.062	0.062	0.052	$0.030^{\ddagger}$
single	-0.221	-0.202	-0.201	-0.217	-0.163	-0.119
student	-0.251	-0.249	-0.271	-0.252	-0.173	-0.166
retired	0.174	0.156	0.138	0.118	0.120	0.144
Demographic variables						
Norte	$-0.060^{\ddagger}$	$-0.044^{\ddagger}$	$-0.048^{\ddagger}$	$-0.047^{\ddagger}$	$-0.057^{\ddagger}$	-0.104
Lisboa	$-0.099^{\dagger}$	$-0.073^{\dagger}$	$-0.083^{\dagger}$	-0.089	-0.093	-0.111
Alentejo	-0.284	-0.233	-0.228	-0.201	-0.161	-0.154
Algarve	-0.259	-0.214	-0.202	-0.177	-0.148	-0.167
Açores	-0.372	-0.341	-0.358	-0.379	-0.336	-0.353
Madeira	-0.526	-0.509	-0.552	-0.603	-0.511	-0.483
Seasonality variables						
winter	0.174	0.181	0.188	0.183	0.144	0.146
spring	0.089	0.093	0.100	0.096	0.072	0.065
summer	$0.044^{\ddagger}$	$0.035^{\ddagger}$	$0.025^{\ddagger}$	$0.005^{\ddagger}$	$-0.017^{\ddagger}$	$-0.006^{\ddagger}$
Constant	-1.009	-0.402	$-0.184^{\ddagger}$	$0.107^{\ddagger}$	$0.274^{\dagger}$	0.677

Notes: The subsample has 34,356 observations. Coefficients marked with ‡ and †are not significant at a 5 and 1 per cent level, respectively.

E Further results: adequacy of the exogeneity assumption on the subsystems dummies

The features of the Portuguese health subsystems are unique (in particular, the fact that the membership is mandatory and based on professional category and contributions are related to income and not to risk characteristics of the individuals). The literature usually assumes that the dummies indicating if an individual benefits from those protection plans gather conditions to be interpreted as exogenous (e.g. Barros et al. 2008 and Lourenço 2007). The basic idea underlying this assumption is that the health risk characteristics of the individuals do not influence the probability of benefiting from a subsystem. It can be argued, however, that more risk averse people are more likely to be public employees. The main argument behind this position is that employment decisions can be based in all the features of the job, including health benefits.[35] The objective of the following exercise is to provide further evidence on this issue, focusing on the public subsystems.

Among the population that benefits from employer provided health insurance through public schemes, there are direct beneficiaries, i.e. workers, and indirect beneficiaries, such as spouses and dependants. Our strategy was to study the impact of public subsystems using a sub-sample of individuals that are potentially "immune" to the above mentioned problem, i.e. a sample of indirect beneficiaries of subsystems.[36] The PHS dataset does not provide this information directly. It is, however, possible to use a proxy to disentangle between direct and indirect beneficiaries. Since we want to exclude individuals that are currently workers of the public sector, we use a question that indicates for the employed individuals what is the sector of activity (according to the Portuguese classification of economic activities) in which they are currently working. In particular, public employees are mostly classified in the sectors L (public administration, defense and compulsory social security), M (education) and N (health). Therefore, we created a subsample of beneficiaries of public subsystems that excludes observations of individuals that are employed in

[35] In fact, although in Portugal all residents have the right to health protection through NHS, it can be argued that as these subsystems are more beneficial to their members, their complementary and supplementary coverage can be taken into account when choosing their job.

[36] Note that only direct beneficiaries are more likely to have chosen their professional activities based on their risk features. Cohabitation decisions are usualy driven by other factors. However, it can be argued that the direct beneficiaries based their employment decisions not only on their own health but also on the health of their households. We believe that this is not an important consideration, because the government recruitment usualy takes place at the beginning of the employment spell.

these three sectors. This procedure is the best possible approximation to the sample of spouses and descendants of public employees. Nevertheless, two minor drawbacks should be borne in mind: we are also excluding individuals of the private sector (the ones that work in the sectors M and N) and a reduced number of public employees may still remain in the sample (the ones employed in the other sectors). Table E1 shows some differences between this subsample and the one previously used.

Table E1: Differences between direct and indirect beneficiaries of public subsystems

	Beneficiaries of public subsystems	
	direct and indirect	indirect
Observations	5,578	2,347
(% in total observations)	15.8%	7.5%
$\overline{y}$	1.01	0.90
$\overline{\sigma_y}$	1.64	1.54

Using the sample of individuals covered solely by the NHS and this sample of indirect beneficiaries of public subsystems on top of NHS, we estimate the double coverage effects using the following conditional quantiles specification

$$Q_{y_i^*}(\alpha|x) = \alpha + \exp\left[\beta_0(\alpha) + \beta_{1ind}(\alpha) pubsub\_ind_i + \boldsymbol{\gamma}(\alpha)\mathbf{z}_i\right], 0 \leq \alpha < 1$$

where $pubsub\_ind_i$ indicates if a person benefits indirectly from double coverage through a "public insurance health subsystem". The vector $\mathbf{z}_i$ includes all the regressors previously used. Table E2 shows the estimated coefficients. As expected, the results regarding the covariates for which we are controlling for are very similar to the ones obtained with the full sample. In particular, health status variables as a whole have a positive effect that increases with α and socioeconomic characteristics seem to have a similar impact across quantiles. Regarding the impact of double coverage, it is now even more clear that public subsystem beneficiaries consume a high number of doctor visits when compared with individuals covered solely by NHS, and this positive effect slightly decreases with α. Note that the effect is more expressive than the one obtained for all public subsystems beneficiaries. Since the estimated coefficientes resulting from the subsample of indirect beneficiaries were not lower than the ones resulting from the sample with both direct and indirect beneficiaries, this sensitivity analysis exercise gives us further evidence on supporting the hypothesis that there are no adverse selection effects in our main dataset.

Table E2: Quantile regression results: estimated coefficients for indirect beneficiaries of public subsystems

	$\widehat{\beta(0.25)}$	$\widehat{\beta(0.50)}$	$\widehat{\beta(0.60)}$	$\widehat{\beta(0.70)}$	$\widehat{\beta(0.80)}$	$\widehat{\beta(0.90)}$
Health insurance status variables						
pubsub_ind	0.167	0.165	0.169	0.165	0.124	0.098
Health status variables						
sick	0.676	0.639	0.621	0.618	0.555	0.802
limitdays	0.070	0.073	0.076	0.075	0.072	0.073
limited	0.150	0.212	0.246	0.320	0.331	0.371
rheumatism	0.134	0.138	0.138	0.131	0.133	0.142
osteoporosis	0.305	0.216	0.188	0.155	0.111	0.080†
cancer	0.486	0.484	0.442	0.403	0.441	0.560
kidneystones	0.153	0.154	0.182	0.202	0.231	0.229
renalfailure	0.165	0.210	0.207	0.202	0.243	0.229
emphysema	0.072	0.171	0.181	0.194	0.198	0.206
cerebralhemorrhage	0.109	0.114†	0.111‡	0.139†	0.162	0.180
infarction	0.267	0.277	0.282	0.296	0.264	0.195
depressivedisorder	0.167	0.221	0.237	0.244	0.234	0.243
otherchronicaldisease	0.443	0.457	0.480	0.471	0.395	0.361
highbloodpressure	0.447	0.397	0.374	0.327	0.264	0.213
chronicpain	0.152	0.183	0.206	0.217	0.211	0.216
diabetes	0.460	0.376	0.349	0.334	0.312	0.305
asthma	0.284	0.345	0.367	0.371	0.302	0.247
stress	0.431	0.363	0.350	0.317	0.305	0.255
smoker	-0.241	-0.202	-0.193	-0.175	-0.107	-0.039‡
meals	0.162	0.126	0.102†	0.094†	0.074†	0.071‡
Socioeconomic characteristics variables						
householdsize	-0.068	-0.062	-0.060	-0.059	-0.039	-0.014‡
age	-1.062	-1.019	-1.056	-1.095	-0.770	-0.586
age^2	0.230	0.221	0.231	0.246	0.170	0.127
age^3	-0.014	-0.014	-0.014	-0.016	-0.011	-0.008
age*female	0.535	0.568	0.629	0.761	0.532	0.363
(age*female)^2	-0.114	-0.125	-0.142	-0.181	-0.125	-0.084
(age*female)^3	0.006	0.007	0.008	0.011	0.008	0.005†
female	-0.310	-0.322	-0.352	-0.381	-0.254	-0.127‡
educmax	0.011	0.015	0.015	0.015	0.010	0.005‡
lincome	0.074	0.062	0.068	0.069	0.058	0.032‡
single	-0.240	-0.219	-0.224	-0.244	-0.193	-0.140
student	-0.278	-0.272	-0.295	-0.281	-0.201	-0.193
retired	0.174	0.156	0.143	0.124	0.120	0.142
Demographic variables						
Norte	-0.047	-0.037‡	-0.047‡	-0.054‡	-0.067†	-0.111
Lisboa	-0.066	-0.077†	-0.095	-0.102	-0.111	-0.132
Alentejo	-0.282	-0.236	-0.229	-0.199	-0.162	-0.152
Algarve	-0.259	-0.233	-0.228	-0.199	-0.167	-0.184
Açores	-0.355	-0.342	-0.372	-0.400	-0.356	-0.362
Madeira	-0.540	-0.530	-0.579	-0.640	-0.551	-0.502
Seasonality variables						
winter	0.155	0.166	0.175	0.169	0.129	0.138
spring	0.072	0.075	0.083	0.076	0.055†	0.054†
summer	0.045	0.026‡	0.019‡	-0.005‡	-0.029‡	-0.011‡
Constant	0.000	-0.353†	-0.140‡	0.154‡	0.322†	0.706

Notes: The subsample has 31,125 observations. Coefficients marked with ‡ and †are not significant at a 5 and 1 per cent level, respectively.

WORKING PAPERS

2000

1/00 UNEMPLOYMENT DURATION: COMPETING AND DEFECTIVE RISKS
— *John T. Addison, Pedro Portugal*

2/00 THE ESTIMATION OF RISK PREMIUM IMPLICIT IN OIL PRICES
— *Jorge Barros Luís*

3/00 EVALUATING CORE INFLATION INDICATORS
— *Carlos Robalo Marques, Pedro Duarte Neves, Luís Morais Sarmento*

4/00 LABOR MARKETS AND KALEIDOSCOPIC COMPARATIVE ADVANTAGE
— *Daniel A. Traça*

5/00 WHY SHOULD CENTRAL BANKS AVOID THE USE OF THE UNDERLYING INFLATION INDICATOR?
— *Carlos Robalo Marques, Pedro Duarte Neves, Afonso Gonçalves da Silva*

6/00 USING THE ASYMMETRIC TRIMMED MEAN AS A CORE INFLATION INDICATOR
— *Carlos Robalo Marques, João Machado Mota*

2001

1/01 THE SURVIVAL OF NEW DOMESTIC AND FOREIGN OWNED FIRMS
— *José Mata, Pedro Portugal*

2/01 GAPS AND TRIANGLES
— *Bernardino Adão, Isabel Correia, Pedro Teles*

3/01 A NEW REPRESENTATION FOR THE FOREIGN CURRENCY RISK PREMIUM
— *Bernardino Adão, Fátima Silva*

4/01 ENTRY MISTAKES WITH STRATEGIC PRICING
— *Bernardino Adão*

5/01 FINANCING IN THE EUROSYSTEM: FIXED VERSUS VARIABLE RATE TENDERS
— *Margarida Catalão-Lopes*

6/01 AGGREGATION, PERSISTENCE AND VOLATILITY IN A MACROMODEL
— *Karim Abadir, Gabriel Talmain*

7/01 SOME FACTS ABOUT THE CYCLICAL CONVERGENCE IN THE EURO ZONE
— *Frederico Belo*

8/01 TENURE, BUSINESS CYCLE AND THE WAGE-SETTING PROCESS
— *Leandro Arozamena, Mário Centeno*

9/01 USING THE FIRST PRINCIPAL COMPONENT AS A CORE INFLATION INDICATOR
— *José Ferreira Machado, Carlos Robalo Marques, Pedro Duarte Neves, Afonso Gonçalves da Silva*

10/01 IDENTIFICATION WITH AVERAGED DATA AND IMPLICATIONS FOR HEDONIC REGRESSION STUDIES
— *José A.F. Machado, João M.C. Santos Silva*

2002

1/02 QUANTILE REGRESSION ANALYSIS OF TRANSITION DATA
— *José A.F. Machado, Pedro Portugal*

2/02 SHOULD WE DISTINGUISH BETWEEN STATIC AND DYNAMIC LONG RUN EQUILIBRIUM IN ERROR CORRECTION MODELS?
— *Susana Botas, Carlos Robalo Marques*

3/02 MODELLING TAYLOR RULE UNCERTAINTY
— *Fernando Martins, José A. F. Machado, Paulo Soares Esteves*

4/02 PATTERNS OF ENTRY, POST-ENTRY GROWTH AND SURVIVAL: A COMPARISON BETWEEN DOMESTIC AND FOREIGN OWNED FIRMS
— *José Mata, Pedro Portugal*

5/02 BUSINESS CYCLES: CYCLICAL COMOVEMENT WITHIN THE EUROPEAN UNION IN THE PERIOD 1960-1999. A FREQUENCY DOMAIN APPROACH
— *João Valle e Azevedo*

6/02 AN “ART”, NOT A “SCIENCE”? CENTRAL BANK MANAGEMENT IN PORTUGAL UNDER THE GOLD STANDARD, 1854 -1891
— *Jaime Reis*

7/02 MERGE OR CONCENTRATE? SOME INSIGHTS FOR ANTITRUST POLICY
— *Margarida Catalão-Lopes*

8/02 DISENTANGLING THE MINIMUM WAGE PUZZLE: ANALYSIS OF WORKER ACCESSIONS AND SEPARATIONS FROM A LONGITUDINAL MATCHED EMPLOYER-EMPLOYEE DATA SET
— *Pedro Portugal, Ana Rute Cardoso*

9/02 THE MATCH QUALITY GAINS FROM UNEMPLOYMENT INSURANCE
— *Mário Centeno*

10/02 HEDONIC PRICES INDEXES FOR NEW PASSENGER CARS IN PORTUGAL (1997-2001)
— *Hugo J. Reis, J.M.C. Santos Silva*

11/02 THE ANALYSIS OF SEASONAL RETURN ANOMALIES IN THE PORTUGUESE STOCK MARKET
— *Miguel Balbina, Nuno C. Martins*

12/02 DOES MONEY GRANGER CAUSE INFLATION IN THE EURO AREA?
— *Carlos Robalo Marques, Joaquim Pina*

13/02 INSTITUTIONS AND ECONOMIC DEVELOPMENT: HOW STRONG IS THE RELATION?
— *Tiago V.de V. Cavalcanti, Álvaro A. Novo*

2003

1/03 FOUNDING CONDITIONS AND THE SURVIVAL OF NEW FIRMS
— *P.A. Geroski, José Mata, Pedro Portugal*

2/03 THE TIMING AND PROBABILITY OF FDI: AN APPLICATION TO THE UNITED STATES MULTINATIONAL ENTERPRISES
— *José Brandão de Brito, Felipa de Mello Sampayo*

3/03 OPTIMAL FISCAL AND MONETARY POLICY: EQUIVALENCE RESULTS
— *Isabel Correia, Juan Pablo Nicolini, Pedro Teles*

4/03 FORECASTING EURO AREA AGGREGATES WITH BAYESIAN VAR AND VECM MODELS
— *Ricardo Mourinho Félix, Luís C. Nunes*

5/03 CONTAGIOUS CURRENCY CRISES: A SPATIAL PROBIT APPROACH
— *Álvaro Novo*

6/03 THE DISTRIBUTION OF LIQUIDITY IN A MONETARY UNION WITH DIFFERENT PORTFOLIO RIGIDITIES
— *Nuno Alves*

7/03 COINCIDENT AND LEADING INDICATORS FOR THE EURO AREA: A FREQUENCY BAND APPROACH
— *António Rua, Luís C. Nunes*

8/03 WHY DO FIRMS USE FIXED-TERM CONTRACTS?
— *José Varejão, Pedro Portugal*

9/03 NONLINEARITIES OVER THE BUSINESS CYCLE: AN APPLICATION OF THE SMOOTH TRANSITION AUTOREGRESSIVE MODEL TO CHARACTERIZE GDP DYNAMICS FOR THE EURO-AREA AND PORTUGAL
— *Francisco Craveiro Dias*

10/03 WAGES AND THE RISK OF DISPLACEMENT
— *Anabela Carneiro, Pedro Portugal*

11/03 SIX WAYS TO LEAVE UNEMPLOYMENT
— *Pedro Portugal, John T. Addison*

12/03 EMPLOYMENT DYNAMICS AND THE STRUCTURE OF LABOR ADJUSTMENT COSTS
— *José Varejão, Pedro Portugal*

13/03 THE MONETARY TRANSMISSION MECHANISM: IS IT RELEVANT FOR POLICY?
— *Bernardino Adão, Isabel Correia, Pedro Teles*

14/03 THE IMPACT OF INTEREST-RATE SUBSIDIES ON LONG-TERM HOUSEHOLD DEBT: EVIDENCE FROM A LARGE PROGRAM
— *Nuno C. Martins, Ernesto Villanueva*

15/03 THE CAREERS OF TOP MANAGERS AND FIRM OPENNESS: INTERNAL VERSUS EXTERNAL LABOUR MARKETS
— *Francisco Lima, Mário Centeno*

16/03 TRACKING GROWTH AND THE BUSINESS CYCLE: A STOCHASTIC COMMON CYCLE MODEL FOR THE EURO AREA
— *João Valle e Azevedo, Siem Jan Koopman, António Rua*

17/03 CORRUPTION, CREDIT MARKET IMPERFECTIONS, AND ECONOMIC DEVELOPMENT
— *António R. Antunes, Tiago V. Cavalcanti*

18/03 BARGAINED WAGES, WAGE DRIFT AND THE DESIGN OF THE WAGE SETTING SYSTEM
— *Ana Rute Cardoso, Pedro Portugal*

19/03 UNCERTAINTY AND RISK ANALYSIS OF MACROECONOMIC FORECASTS: FAN CHARTS REVISITED
— *Álvaro Novo, Maximiano Pinheiro*

2004

1/04 HOW DOES THE UNEMPLOYMENT INSURANCE SYSTEM SHAPE THE TIME PROFILE OF JOBLESS DURATION?
— *John T. Addison, Pedro Portugal*

2/04 REAL EXCHANGE RATE AND HUMAN CAPITAL IN THE EMPIRICS OF ECONOMIC GROWTH
— *Delfim Gomes Neto*

3/04 ON THE USE OF THE FIRST PRINCIPAL COMPONENT AS A CORE INFLATION INDICATOR
— *José Ramos Maria*

4/04 OIL PRICES ASSUMPTIONS IN MACROECONOMIC FORECASTS: SHOULD WE FOLLOW FUTURES MARKET EXPECTATIONS?
— *Carlos Coimbra, Paulo Soares Esteves*

5/04 STYLISED FEATURES OF PRICE SETTING BEHAVIOUR IN PORTUGAL: 1992-2001
— *Mónica Dias, Daniel Dias, Pedro D. Neves*

6/04 A FLEXIBLE VIEW ON PRICES
— *Nuno Alves*

7/04 ON THE FISHER-KONIECZNY INDEX OF PRICE CHANGES SYNCHRONIZATION
— *D.A. Dias, C. Robalo Marques, P.D. Neves, J.M.C. Santos Silva*

8/04 INFLATION PERSISTENCE: FACTS OR ARTEFACTS?
— *Carlos Robalo Marques*

9/04 WORKERS' FLOWS AND REAL WAGE CYCLICALITY
— *Anabela Carneiro, Pedro Portugal*

10/04 MATCHING WORKERS TO JOBS IN THE FAST LANE: THE OPERATION OF FIXED-TERM CONTRACTS
— *José Varejão, Pedro Portugal*

11/04 THE LOCATIONAL DETERMINANTS OF THE U.S. MULTINATIONALS ACTIVITIES
— *José Brandão de Brito, Felipa Mello Sampayo*

12/04 KEY ELASTICITIES IN JOB SEARCH THEORY: INTERNATIONAL EVIDENCE
— *John T. Addison, Mário Centeno, Pedro Portugal*

13/04 RESERVATION WAGES, SEARCH DURATION AND ACCEPTED WAGES IN EUROPE
— *John T. Addison, Mário Centeno, Pedro Portugal*

14/04 THE MONETARY TRANSMISSION N THE US AND THE EURO AREA: COMMON FEATURES AND COMMON FRICTIONS
— *Nuno Alves*

15/04 NOMINAL WAGE INERTIA IN GENERAL EQUILIBRIUM MODELS
— *Nuno Alves*

16/04 MONETARY POLICY IN A CURRENCY UNION WITH NATIONAL PRICE ASYMMETRIES
— *Sandra Gomes*

17/04 NEOCLASSICAL INVESTMENT WITH MORAL HAZARD
— *João Ejarque*

18/04 MONETARY POLICY WITH STATE CONTINGENT INTEREST RATES
— *Bernardino Adão, Isabel Correia, Pedro Teles*

19/04 MONETARY POLICY WITH SINGLE INSTRUMENT FEEDBACK RULES
— *Bernardino Adão, Isabel Correia, Pedro Teles*

20/04 ACOUNTING FOR THE HIDDEN ECONOMY: BARRIERS TO LAGALITY AND LEGAL FAILURES
— *António R. Antunes, Tiago V. Cavalcanti*

2005

1/05 SEAM: A SMALL-SCALE EURO AREA MODEL WITH FORWARD-LOOKING ELEMENTS
— *José Brandão de Brito, Rita Duarte*

2/05 FORECASTING INFLATION THROUGH A BOTTOM-UP APPROACH: THE PORTUGUESE CASE
— *Cláudia Duarte, António Rua*

3/05 USING MEAN REVERSION AS A MEASURE OF PERSISTENCE
— *Daniel Dias, Carlos Robalo Marques*

4/05 HOUSEHOLD WEALTH IN PORTUGAL: 1980-2004
— *Fátima Cardoso, Vanda Geraldes da Cunha*

5/05 ANALYSIS OF DELINQUENT FIRMS USING MULTI-STATE TRANSITIONS
— *António Antunes*

6/05 PRICE SETTING IN THE AREA: SOME STYLIZED FACTS FROM INDIVIDUAL CONSUMER PRICE DATA
— *Emmanuel Dhyne, Luis J. Álvarez, Hervé Le Bihan, Giovanni Veronese, Daniel Dias, Johannes Hoffmann, Nicole Jonker, Patrick Lünnemann, Fabio Rumler, Jouko Vilmunen*

7/05 INTERMEDIATION COSTS, INVESTOR PROTECTION AND ECONOMIC DEVELOPMENT
— *António Antunes, Tiago Cavalcanti, Anne Villamil*

8/05 TIME OR STATE DEPENDENT PRICE SETTING RULES? EVIDENCE FROM PORTUGUESE MICRO DATA
— *Daniel Dias, Carlos Robalo Marques, João Santos Silva*

9/05 BUSINESS CYCLE AT A SECTORAL LEVEL: THE PORTUGUESE CASE
— *Hugo Reis*

10/05 THE PRICING BEHAVIOUR OF FIRMS IN THE EURO AREA: NEW SURVEY EVIDENCE
— *S. Fabiani, M. Druant, I. Hernando, C. Kwapil, B. Landau, C. Loupias, F. Martins, T. Mathä, R. Sabbatini, H. Stahl, A. Stokman*

11/05 CONSUMPTION TAXES AND REDISTRIBUTION
— *Isabel Correia*

12/05 UNIQUE EQUILIBRIUM WITH SINGLE MONETARY INSTRUMENT RULES
— *Bernardino Adão, Isabel Correia, Pedro Teles*

13/05 A MACROECONOMIC STRUCTURAL MODEL FOR THE PORTUGUESE ECONOMY
— *Ricardo Mourinho Félix*

14/05 THE EFFECTS OF A GOVERNMENT EXPENDITURES SHOCK
— *Bernardino Adão, José Brandão de Brito*

15/05 MARKET INTEGRATION IN THE GOLDEN PERIPHERY – THE LISBON/LONDON EXCHANGE, 1854-1891
— *Rui Pedro Esteves, Jaime Reis, Fabiano Ferramosca*

2006

1/06 THE EFFECTS OF A TECHNOLOGY SHOCK IN THE EURO AREA
— *Nuno Alves , José Brandão de Brito , Sandra Gomes, João Sousa*

2/02 THE TRANSMISSION OF MONETARY AND TECHNOLOGY SHOCKS IN THE EURO AREA
— *Nuno Alves, José Brandão de Brito, Sandra Gomes, João Sousa*

3/06 MEASURING THE IMPORTANCE OF THE UNIFORM NONSYNCHRONIZATION HYPOTHESIS
— *Daniel Dias, Carlos Robalo Marques, João Santos Silva*

4/06 THE PRICE SETTING BEHAVIOUR OF PORTUGUESE FIRMS EVIDENCE FROM SURVEY DATA
— *Fernando Martins*

5/06 STICKY PRICES IN THE EURO AREA: A SUMMARY OF NEW MICRO EVIDENCE
— *L. J. Álvarez, E. Dhyne, M. Hoeberichts, C. Kwapil, H. Le Bihan, P. Lünnemann, F. Martins, R. Sabbatini, H. Stahl, P. Vermeulen and J. Vilmunen*

6/06 NOMINAL DEBT AS A BURDEN ON MONETARY POLICY
— *Javier Díaz-Giménez, Giorgia Giovannetti , Ramon Marimon, Pedro Teles*

7/06 A DISAGGREGATED FRAMEWORK FOR THE ANALYSIS OF STRUCTURAL DEVELOPMENTS IN PUBLIC FINANCES
— *Jana Kremer, Cláudia Rodrigues Braz, Teunis Brosens, Geert Langenus, Sandro Momigliano, Mikko Spolander*

8/06 IDENTIFYING ASSET PRICE BOOMS AND BUSTS WITH QUANTILE REGRESSIONS
— *José A. F. Machado, João Sousa*

9/06 EXCESS BURDEN AND THE COST OF INEFFICIENCY IN PUBLIC SERVICES PROVISION
— *António Afonso, Vítor Gaspar*

10/06 MARKET POWER, DISMISSAL THREAT AND RENT SHARING: THE ROLE OF INSIDER AND OUTSIDER FORCES IN WAGE BARGAINING
— *Anabela Carneiro, Pedro Portugal*

11/06 MEASURING EXPORT COMPETITIVENESS: REVISITING THE EFFECTIVE EXCHANGE RATE WEIGHTS FOR THE EURO AREA COUNTRIES
— *Paulo Soares Esteves, Carolina Reis*

12/06 THE IMPACT OF UNEMPLOYMENT INSURANCE GENEROSITY ON MATCH QUALITY DISTRIBUTION
— *Mário Centeno, Alvaro A. Novo*

13/06 U.S. UNEMPLOYMENT DURATION: HAS LONG BECOME LONGER OR SHORT BECOME SHORTER?
— *José A.F. Machado, Pedro Portugal e Juliana Guimarães*

14/06 EARNINGS LOSSES OF DISPLACED WORKERS: EVIDENCE FROM A MATCHED EMPLOYER-EMPLOYEE DATA SET
— *Anabela Carneiro, Pedro Portugal*

15/06 COMPUTING GENERAL EQUILIBRIUM MODELS WITH OCCUPATIONAL CHOICE AND FINANCIAL FRICTIONS
— *António Antunes, Tiago Cavalcanti, Anne Villamil*

16/06 ON THE RELEVANCE OF EXCHANGE RATE REGIMES FOR STABILIZATION POLICY
— *Bernardino Adao, Isabel Correia, Pedro Teles*

17/06 AN INPUT-OUTPUT ANALYSIS: LINKAGES VS LEAKAGES
— *Hugo Reis, António Rua*

2007

1/07 RELATIVE EXPORT STRUCTURES AND VERTICAL SPECIALIZATION: A SIMPLE CROSS-COUNTRY INDEX
— *João Amador, Sónia Cabral, José Ramos Maria*

2/07 THE FORWARD PREMIUM OF EURO INTEREST RATES
— *Sónia Costa, Ana Beatriz Galvão*

3/07 ADJUSTING TO THE EURO
— *Gabriel Fagan, Vítor Gaspar*

4/07 SPATIAL AND TEMPORAL AGGREGATION IN THE ESTIMATION OF LABOR DEMAND FUNCTIONS
— *José Varejão, Pedro Portugal*

5/07 PRICE SETTING IN THE EURO AREA: SOME STYLISED FACTS FROM INDIVIDUAL PRODUCER PRICE DATA
— *Philip Vermeulen, Daniel Dias, Maarten Dossche, Erwan Gautier, Ignacio Hernando, Roberto Sabbatini, Harald Stahl*

6/07 A STOCHASTIC FRONTIER ANALYSIS OF SECONDARY EDUCATION OUTPUT IN PORTUGAL
— *Manuel Coutinho Pereira, Sara Moreira*

7/07 CREDIT RISK DRIVERS: EVALUATING THE CONTRIBUTION OF FIRM LEVEL INFORMATION AND OF MACROECONOMIC DYNAMICS
— *Diana Bonfim*

8/07 CHARACTERISTICS OF THE PORTUGUESE ECONOMIC GROWTH: WHAT HAS BEEN MISSING?
— *João Amador, Carlos Coimbra*

9/07 TOTAL FACTOR PRODUCTIVITY GROWTH IN THE G7 COUNTRIES: DIFFERENT OR ALIKE?
— *João Amador, Carlos Coimbra*

10/07 IDENTIFYING UNEMPLOYMENT INSURANCE INCOME EFFECTS WITH A QUASI-NATURAL EXPERIMENT
— *Mário Centeno, Alvaro A. Novo*

11/07 HOW DO DIFFERENT ENTITLEMENTS TO UNEMPLOYMENT BENEFITS AFFECT THE TRANSITIONS FROM UNEMPLOYMENT INTO EMPLOYMENT
— *John T. Addison, Pedro Portugal*

12/07 INTERPRETATION OF THE EFFECTS OF FILTERING INTEGRATED TIME SERIES
— *João Valle e Azevedo*

13/07 EXACT LIMIT OF THE EXPECTED PERIODOGRAM IN THE UNIT-ROOT CASE
— *João Valle e Azevedo*

14/07 INTERNATIONAL TRADE PATTERNS OVER THE LAST FOUR DECADES: HOW DOES PORTUGAL COMPARE WITH OTHER COHESION COUNTRIES?
— *João Amador, Sónia Cabral, José Ramos Maria*

15/07 INFLATION (MIS)PERCEPTIONS IN THE EURO AREA
— *Francisco Dias, Cláudia Duarte, António Rua*

16/07 LABOR ADJUSTMENT COSTS IN A PANEL OF ESTABLISHMENTS: A STRUCTURAL APPROACH
— *João Miguel Ejarque, Pedro Portugal*

17/07 A MULTIVARIATE BAND-PASS FILTER
— *João Valle e Azevedo*

18/07 AN OPEN ECONOMY MODEL OF THE EURO AREA AND THE US
— *Nuno Alves, Sandra Gomes, João Sousa*

19/07 IS TIME RIPE FOR PRICE LEVEL PATH STABILITY?
— *Vitor Gaspar, Frank Smets , David Vestin*

20/07 IS THE EURO AREA M3 ABANDONING US?
— *Nuno Alves, Carlos Robalo Marques, João Sousa*

21/07 DO LABOR MARKET POLICIES AFFECT EMPLOYMENT COMPOSITION? LESSONS FROM EUROPEAN COUNTRIES
— *António Antunes, Mário Centeno*

2008

1/08 THE DETERMINANTS OF PORTUGUESE BANKS' CAPITAL BUFFERS
— *Miguel Boucinha*

2/08 DO RESERVATION WAGES REALLY DECLINE? SOME INTERNATIONAL EVIDENCE ON THE DETERMINANTS OF RESERVATION WAGES
— *John T. Addison, Mário Centeno, Pedro Portugal*

3/08 UNEMPLOYMENT BENEFITS AND RESERVATION WAGES: KEY ELASTICITIES FROM A STRIPPED-DOWN JOB SEARCH APPROACH
— *John T. Addison, Mário Centeno, Pedro Portugal*

4/08 THE EFFECTS OF LOW-COST COUNTRIES ON PORTUGUESE MANUFACTURING IMPORT PRICES
— *Fátima Cardoso, Paulo Soares Esteves*

5/08 WHAT IS BEHIND THE RECENT EVOLUTION OF PORTUGUESE TERMS OF TRADE?
— *Fátima Cardoso, Paulo Soares Esteves*

6/08 EVALUATING JOB SEARCH PROGRAMS FOR OLD AND YOUNG INDIVIDUALS: HETEROGENEOUS IMPACT ON UNEMPLOYMENT DURATION
— *Luis Centeno, Mário Centeno, Álvaro A. Novo*

7/08 FORECASTING USING TARGETED DIFFUSION INDEXES
— *Francisco Dias, Maximiano Pinheiro, António Rua*

8/08 STATISTICAL ARBITRAGE WITH DEFAULT AND COLLATERAL
— *José Fajardo, Ana Lacerda*

9/08 DETERMINING THE NUMBER OF FACTORS IN APPROXIMATE FACTOR MODELS WITH GLOBAL AND GROUP-SPECIFIC FACTORS
— *Francisco Dias, Maximiano Pinheiro, António Rua*

10/08 VERTICAL SPECIALIZATION ACROSS THE WORLD: A RELATIVE MEASURE
— *João Amador, Sónia Cabral*

11/08 INTERNATIONAL FRAGMENTATION OF PRODUCTION IN THE PORTUGUESE ECONOMY: WHAT DO DIFFERENT MEASURES TELL US?
— *João Amador, Sónia Cabral*

12/08 IMPACT OF THE RECENT REFORM OF THE PORTUGUESE PUBLIC EMPLOYEES' PENSION SYSTEM
— *Maria Manuel Campos, Manuel Coutinho Pereira*

13/08 EMPIRICAL EVIDENCE ON THE BEHAVIOR AND STABILIZING ROLE OF FISCAL AND MONETARY POLICIES IN THE US
— *Manuel Coutinho Pereira*

14/08 IMPACT ON WELFARE OF COUNTRY HETEROGENEITY IN A CURRENCY UNION
— *Carla Soares*

15/08 WAGE AND PRICE DYNAMICS IN PORTUGAL
— *Carlos Robalo Marques*

16/08 IMPROVING COMPETITION IN THE NON-TRADABLE GOODS AND LABOUR MARKETS: THE PORTUGUESE CASE
— *Vanda Almeida, Gabriela Castro, Ricardo Mourinho Félix*

17/08 PRODUCT AND DESTINATION MIX IN EXPORT MARKETS
— *João Amador, Luca David Opromolla*

18/08 FORECASTING INVESTMENT: A FISHING CONTEST USING SURVEY DATA
— *José Ramos Maria, Sara Serra*

19/08 APPROXIMATING AND FORECASTING MACROECONOMIC SIGNALS IN REAL-TIME
— *João Valle e Azevedo*

20/08 A THEORY OF ENTRY AND EXIT INTO EXPORTS MARKETS
— *Alfonso A. Irarrazabal, Luca David Opromolla*

21/08 ON THE UNCERTAINTY AND RISKS OF MACROECONOMIC FORECASTS: COMBINING JUDGEMENTS WITH SAMPLE AND MODEL INFORMATION
— *Maximiano Pinheiro, Paulo Soares Esteves*

22/08 ANALYSIS OF THE PREDICTORS OF DEFAULT FOR PORTUGUESE FIRMS
— *Ana I. Lacerda, Russ A. Moro*

23/08 INFLATION EXPECTATIONS IN THE EURO AREA: ARE CONSUMERS RATIONAL?
— *Francisco Dias, Cláudia Duarte, António Rua*

2009

1/09 AN ASSESSMENT OF COMPETITION IN THE PORTUGUESE BANKING SYSTEM IN THE 1991-2004 PERIOD
— *Miguel Boucinha, Nuno Ribeiro*

2/09 FINITE SAMPLE PERFORMANCE OF FREQUENCY AND TIME DOMAIN TESTS FOR SEASONAL FRACTIONAL INTEGRATION
— *Paulo M. M. Rodrigues, Antonio Rubia, João Valle e Azevedo*

3/09 THE MONETARY TRANSMISSION MECHANISM FOR A SMALL OPEN ECONOMY IN A MONETARY UNION
— *Bernardino Adão*

4/09 INTERNATIONAL COMOVEMENT OF STOCK MARKET RETURNS: A WAVELET ANALYSIS
— *António Rua, Luís C. Nunes*

5/09 THE INTEREST RATE PASS-THROUGH OF THE PORTUGUESE BANKING SYSTEM: CHARACTERIZATION AND DETERMINANTS
— *Paula Antão*

6/09 ELUSIVE COUNTER-CYCLICALITY AND DELIBERATE OPPORTUNISM? FISCAL POLICY FROM PLANS TO FINAL OUTCOMES
— *Álvaro M. Pina*

7/09 LOCAL IDENTIFICATION IN DSGE MODELS
— *Nikolay Iskrev*

8/09 CREDIT RISK AND CAPITAL REQUIREMENTS FOR THE PORTUGUESE BANKING SYSTEM
— *Paula Antão, Ana Lacerda*

9/09 A SIMPLE FEASIBLE ALTERNATIVE PROCEDURE TO ESTIMATE MODELS WITH HIGH-DIMENSIONAL FIXED EFFECTS
— *Paulo Guimarães, Pedro Portugal*

10/09 REAL WAGES AND THE BUSINESS CYCLE: ACCOUNTING FOR WORKER AND FIRM HETEROGENEITY
— *Anabela Carneiro, Paulo Guimarães, Pedro Portugal*

11/09 DOUBLE COVERAGE AND DEMAND FOR HEALTH CARE: EVIDENCE FROM QUANTILE REGRESSION
— *Sara Moreira, Pedro Pita Barros*

12/09 THE NUMBER OF BANK RELATIONSHIPS, BORROWING COSTS AND BANK COMPETITION
— *Diana Bonfim, Qinglei Dai, Francesco Franco*

13/09 DYNAMIC FACTOR MODELS WITH JAGGED EDGE PANEL DATA: TAKING ON BOARD THE DYNAMICS OF THE IDIOSYNCRATIC COMPONENTS
— *Maximiano Pinheiro, António Rua, Francisco Dias*

14/09 BAYESIAN ESTIMATION OF A DSGE MODEL FOR THE PORTUGUESE ECONOMY
— *Vanda Almeida*

15/09 THE DYNAMIC EFFECTS OF SHOCKS TO WAGES AND PRICES IN THE UNITED STATES AND THE EURO AREA
— *Rita Duarte, Carlos Robalo Marques*

16/09 MONEY IS AN EXPERIENCE GOOD: COMPETITION AND TRUST IN THE PRIVATE PROVISION OF MONEY
— *Ramon Marimon, Juan Pablo Nicolini, Pedro Teles*

17/09 MONETARY POLICY AND THE FINANCING OF FIRMS
— *Fiorella De Fiore, Pedro Teles, Oreste Tristani*

18/09 HOW ARE FIRMS' WAGES AND PRICES LINKED: SURVEY EVIDENCE IN EUROPE
— *Martine Druant, Silvia Fabiani, Gabor Kezdi, Ana Lamo, Fernando Martins, Roberto Sabbatini*

19/09 THE FLEXIBLE FOURIER FORM AND LOCAL GLS DE-TRENDED UNIT ROOT TESTS
— *Paulo M. M. Rodrigues, A. M. Robert Taylor*

20/09 ON LM-TYPE TESTS FOR SEASONAL UNIT ROOTS IN THE PRESENCE OF A BREAK IN TREND
— *Luis C. Nunes, Paulo M. M. Rodrigues*

21/09 A NEW MEASURE OF FISCAL SHOCKS BASED ON BUDGET FORECASTS AND ITS IMPLICATIONS
— *Manuel Coutinho Pereira*

22/09 AN ASSESSMENT OF PORTUGUESE BANKS' COSTS AND EFFICIENCY
— *Miguel Boucinha, Nuno Ribeiro, Thomas Weyman-Jones*

23/09 ADDING VALUE TO BANK BRANCH PERFORMANCE EVALUATION USING COGNITIVE MAPS AND MCDA: A CASE STUDY
— *Fernando A. F. Ferreira, Sérgio P. Santos, Paulo M. M. Rodrigues*

24/09 THE CROSS SECTIONAL DYNAMICS OF HETEROGENOUS TRADE MODELS
— *Alfonso Irarrazabal, Luca David Opromolla*

25/09 ARE ATM/POS DATA RELEVANT WHEN NOWCASTING PRIVATE CONSUMPTION?
— *Paulo Soares Esteves*