UNIVERSIDAD CARLOS III DE MADRID
working papers

Working Paper 05-59
Economics Series 28
November 2005

Departamento de Economía
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 91 624 98 75

MORAL HAZARD AND THE DEMAND FOR HEALTH SERVICES: A MATCHING ESTIMATOR APPROACH*

Pedro Pita Barros[1], Matilde P. Machado[2], Anna Sanz de Galdeano[3]

Abstract

In this paper we estimate the impact of health insurance coverage beyond National Health Insurance on the demand for several health services. Traditionally, the literature has tried to deal with the endogeneity of the private (extra) insurance decision by finding instrumental variables. It is hard to think, however, of any variable that a priori would be a good instrument and, therefore, we take a different approach. We concentrate on the most common health insurance plan in the Portuguese Health Survey, (ADSE), which is given to all civil servants and their dependants. We argue that this insurance is exogenous for most people i.e. not correlated with their health status. Under this identifying assumption we estimate the impact of having ADSE coverage on three different health services using a matching estimator technique. The measures of demand for health services are number of visits, number of blood and urine tests, and the probability of visiting a dentist. Preliminary results show large effects of ADSE for number of visits and tests among the young (18 to 30 years old) but only for tests are these effects statistically significantly different from zero. The magnitude of the effects represent 21.8 and 30 percent of the average number of visits and tests for the young. On the contrary we find no evidence of moral hazard on the probability of visiting a dentist. Finally, we argue that there is evidence of a positive cumulative effect of ADSE over the years.

Keywords: Demand for health services, Moral Hazard, Matching Estimator, Portuguese Health System

JEL Classification:.I11, C31

[1] Universidade Nova de Lisboa
[2] Corresponding autor: Departamento de Economía, Universidad Carlos III de Madrid. E.mail: matilde.machado@uc3m.es
[3] CSEF, University of Salerno

* We would like to than Marcos Vera-Hernández and Pierre-Carl Michaud for valuable comments. This article was supported by an unrestricted educational grant awarded jointly to the Universities Carlos III de Madrid and Pompeu Fabra of Barcelona by The Merck Foundation, the philanthropic arm of Merck Co. Inc., White House Station, New Jersey, USA.

1 Introduction

The widespread usage — by health insurance companies, and governments — of copayments, coinsurance, and deductibles as mechanisms to control health-care spending reflects the belief that the demand for health-care reacts to price. The literature on the field, however, has not yet produced irrefutable evidence on the magnitude of this reaction.

By decreasing the price-per-service faced by patients, health insurance increases health-care demand whenever demand is elastic. The potential effect of health insurance on demand for health services — usually denoted by moral hazard — was first identified by Arrow (1963).[1] The first study to show in an experimental setting the impact of insurance on the demand for health services was the RAND Health Insurance Experiment (RHIE) (e.g. Manning et al., 1987, and Newhouse, 1993). The RHIE suffered criticisms on its design reflecting the difficulty of implementing an experiment when health may be at stake. The greatest advantage of the RHIE relative to subsequent studies is the randomization of the insurance type across individuals. Randomization establishes the exogeneity of the insurance status and allows the identification of the increase in health services utilization with moral hazard. Since the RHIE, many researchers turned to field data to estimate the effect of health insurance on health-care demand. In non-experimental settings, however, the decision to buy (extra) insurance is not random but depends on the characteristics of the individual. In particular, the higher the individual's risk the higher is the optimal insurance coverage (Rothschild and Stiglitz, 1976). For example, within a National Health Insurance System individuals who contract private insurance are likely to be those who anticipate, based on private information, a higher than average demand for health care (e.g. Cameron et al., 1988 and Vera, 1999). Ignoring that adverse selection causes the private health insurance variable to be endogenous leads to an overestimation of its impact on the demand for health services.

The traditional way to deal with the endogeneity of the private (or extra) insurance is to find instrumental variables, which should be correlated with the decision to contract

[1]Meza (1983) criticizes attributing this effect solely to moral hazard. Vera (1999) follows the same line of argument. Moral hazard is the difference in demand of an insured individual with and without perfect information. Of course, the latter situation is not observed. A positive impact on demand from additional insurance is compatible with moral hazard but is not proof of moral hazard.

additional insurance but not correlated with the use of health services.[2] Santos Silva and Bago d'Uva (2002) argue that the only such variables are those related to an individual's risk aversion. Unfortunately, it is hard to find such variables in most health surveys. Some authors, e.g. Vera (1999) and Holly (2002), have used socioeconomic variables as instruments for the insurance decision with limited success. While Holly's work is still unfinished, Vera's coefficient estimates suffer from high standard deviations which hamper any meaningful conclusions about the true impact of the private insurance on the demand for physician services.

The presence of screening by the insurance companies and supply-induced demand by the health-care providers (Holly et al., 1998) are also potential sources of bias. Screening would bias downwards the effect from insurance on utilization while supply-induced demand would bias it upwards. Our data does not suffer from screening biases but may be subject to supply-induced demand. We think the latter is of no consequence in our context because the insurer's payments to providers are relatively low. Nonetheless, finding that extra coverage from private (or extra) insurance within a public national health system increases the demand for health-care is consistent with both moral hazard and supply-induced demand.

The aim of this paper is to estimate the impact of additional insurance coverage on the demand for health-care within a National Health System using the data from the Portuguese Health Survey (1998/1999).[3] The main contributions of this paper to the literature lie in the different approach we take to measure the impact of insurance. First, we use a dataset in which approximately ten percent of individuals are covered by an insurance plan that is unrelated to their current health status. The exogeneity of this coverage removes the need to use instrumental-variable estimation. More specifically, we focus on the most common health insurance plan in the country (ADSE), which is given to all civil servants and their dependants.[4] ADSE beneficiaries have double coverage since

[2] An exception, within studies using field data, is Chiappori et al. (1998) who use longitudinal data from a natural experiment in France.

[3] "Inquérito Nacional de Saúde 98/99". Data has been collected from October 1998 to September 1999.

[4] Civil servants in Portugal are employed for life. This is the main advantage of becoming a civil servant. Moreover, at least at the lower end of qualification the Government pays higher salaries than the private sector.

they always have access to the National Health Service just like any other citizen.

Second, we use a matching estimator technique (Abadie and Imbens, 2004), which does not impose any functional form on the impact of health insurance on the demand for health services and allows for heterogenous impacts. The treated group is composed by individuals with ADSE coverage and the control group is composed by individuals covered by the National Health Service (NHS) alone.[5]

Third, we estimate the impact of additional insurance on several health services. Traditionally, the literature has focused on "number of visits to the doctor". We believe there is measurement error in this variable: a visit is not an homogenous service, and may vary in quality and duration. More importantly, the measurement error may be correlated with insurance status. For example, those with additional (or private) insurance may have access to better quality and longer visits. In our case, it is also possible that "number of visits to the doctor" is subject to inducement since ADSE pays doctors per visit. Alternatively, we also estimate the impact of additional health insurance on the "number of blood and urine tests" as well as on "at least one visit to the dentist in 12 months." Relative to number of visits to the doctor, the former is a more objective and homogenous measure, in particular its quality is independent of the insurance status. It is possible, however, that ADSE doctors use the number of diagnostic tests to justify more visits so that similarly to number of visits, number of tests would be subject to demand-inducement. On the contrary, it is less likely that "at least one visit to the dentist in 12 months" be affected by inducement since ADSE pays dentists by type of procedure and not by visit.

Fourth, since some individuals in our treated group may have been subject to treatment for a long period of time, we split the sample into different age groups in order to control for bias that may arise in case health benefits from better treatment accumulate over time. If this were true, health differences between the treated and the untreated group should increase with age. While, the young individuals with access to the additional insurance may not have had the time to accumulate health benefits, it is plausible that among the older generation, the treated are healthier than their NHS-only counterparts. By

[5] A drawback of the matching technique is the lack of estimates for the effects of exogenous variables beyond the treatment.

reporting the results for different age groups we are also trying to separate the "immediate effect" of treatment from its cumulative effect. The presence of a dynamic effect and the effort to distinguish between short and long-term effects of treatments is, as far as we know, innovative. We think this estimation strategy may even be of some relevance in experimental settings where individuals are subject to treatment for different durations.

The more conservative average treatment effects (ATT) estimates for the overall sample are: 0.096 for number of visits, which corresponds to 6 percent of the average number of visits, and is not statistically significantly different from zero; and a value of 0.057 for number of tests, which corresponds to 15.8 percent of the average number of tests, which is statistically significantly different from zero. Results show that the ATT for number of tests is more precisely estimated than the ATT for number of visits. The ATT for number of visits and number of tests for the youngest group is always the largest, which may be evidence that ADSE beneficiaries accumulate health benefits over time. The estimated ATT represents 21.8 and 30 percent of the youngest group average number of visits and tests, respectively, but only the latter is almost statistically significantly different from zero. The effects for other age groups are not statistically significantly different from zero. Surprisingly, for at least one visits to the dentist the ATT is smaller and not statistically significantly different from zero. We interpret the latter result in light of Chiappori et al. (1998) who argued that for services where the non-monetary costs were high the demand would be more inelastic.

Finally, we provide some evidence of the exogeneity of the treatment by taking a sample of individuals, the unemployed, who have extra insurance through a family member and compare them with the unemployed from the control group. The effect of treatment on the unemployed's demand for health services should only reflect moral hazard (and/or supply induced demand) and not adverse selection because it is unlikely that they decided the job status of their family members.

The paper is organized as follows: Section 2 summarizes the recent literature on the subject and briefly describes the Portuguese health system; Section 3 describes the dataset; Section 4 makes a exploratory analyses of the data; Section 5 describes the matching estimator methodology; Section 6 describes the main results; Section 7 discusses the plausibility of exogeneity of the ADSE status; Finally, Section 8 concludes. The Appendix contains tables with results.

2 Background

2.1 Short Review of the Literature

The onset of the literature started with Arrow's prediction (Arrow, 1963) that by decreasing the effective price of health services faced by patients, health insurance triggers an increase in the demand for health services. The Rand Health Insurance Experiment, RHIE, (e.g. Manning et al., 1987, and Newhouse, 1993) was the first study to offer empirical evidence of Arrow's theory using data from a controlled experiment. As pointed out in the Introduction, however, in non-experimental settings the decision to buy (extra) insurance is not random but suffers from adverse selection, i.e. the higher the individual's risk the higher is the optimal insurance coverage (Rothschild and Stiglitz, 1976).

The two other potential sources of endogeneity of insurance status have been almost ignored by the received literature, perhaps because their perceived effects are small. The first one is screening from private insurance companies. Screening consists in denying insurance coverage to the highest risks. The second source is supply-induced demand. A doctor/hospital may induce more demand from those patients holding more generous insurance since these patients pay a smaller fraction of the fees. The inducement of demand is likely to be stronger when insurance companies pay fee-for-service to doctors or hospitals.

If insurance choice is affected by adverse selection, ignoring it will lead to an overestimation of the moral hazard effect in the demand for health care services. The traditional way of controlling for this potential bias is to use Instrumental Variables (e.g. Cameron et al.,1988, Coulson et al., 1995, Holly et al., 1998, Vera, 1999, Savage and Wright, 2003). The econometric model for health services usage varies substantially in the literature, depending on both the characteristic of the dependent variable, e.g. count or binary, and the convenience of the model. For example, Cameron et al. (1988) start with a negbin specification but switch to a linear specification when they instrument for the insurance variable. The typical instruments used are socioeconomic variables that tend to be more associated with the insurance decision than with health care utilization. It is hard, however, to justify that any of these variables is an appropriate instrument. While Vera (1999) justifies the use of these variables as instrumental variables by deriving a structural model of demand for health care and insurance, Santos Silva and Bago d'Uva (2001) argue that the appropriate instruments should be (unobserved) variables related to risk aversion which are absent from most datasets. We agree that appropriate instrumental variables are hard to find. Chiappori et al. (1998) avoid the use of instrumental variables by relying on a natural experiment in France. Yet, natural experiments are a rare phenomenon. Our paper also avoids the use of instrumental variables and instead argues that the civil servant insurance scheme in Portugal (ADSE) is exogenous to the beneficiaries' health.[6]

If insurance choice is only affected by screening from insurance companies then ignoring this effect may lead to an underestimation of the moral hazard effect (Coulson et al, 1995). The literature has not paid attention to this potential bias.

Finally, when supply-induced demand is correlated with insurance status the estimated effect of moral hazard will be, most likely, biased upwards.[7] Van de Voorde's et al. (2001) estimates of price elasticities in a context with high incentives to induce demand are, in general, similar to those obtained in the RHIE. The authors conclude that, at least in the short-run, the level of demand inducement is low.

Most of the empirical literature has shown evidence of both moral hazard and adverse selection in the health care market. The literature has also shown that the level of moral hazard differs across health services. Cameron et al. (1988) using data for Australia find that for a broad range of services more generous coverage leads to higher utilization both because of moral hazard and adverse selection. Savage and Wright (2003), also for Australia, find that private insurance increases hospital length of stay. Vera (1999) using data for a Spanish region finds different evidence for heads-of-household and non-

[6]We were recently aware of simultaneous work by Trujillo et al. (2005) where the authors rely on a propensity score matching approach, similar to our matching estimator, to evaluate the impact of a social subsidy on the demand for health care from the poor in Colombia.

[7]Presumably doctors/hospitals induce more demand on those that have more generous coverage. It is possible, however, that the level of inducement is directly related to the amount the insurer pays to the doctor/hospital for the visit, in which case the bias may be reversed.

heads-of-household. For heads of household, presumably those that take the insurance decision, there is evidence of adverse selection but no evidence of moral hazard. On the contrary, for non-heads of household, there is evidence of moral hazard. More recently Olivella and Vera (2005) show evidence of adverse selection using data from the British household panel survey. Coulson et al. (1995) find that supplemental insurance increases the number of prescriptions filled among the elderly in the United States, i.e. moral hazard, but do not find evidence of adverse selection. Holly et al. (1998) using data for Switzerland find both evidence of adverse selection and moral hazard in hospital stays. Deb and Trivedi (2002) use data from the RAND Health Insurance Experiment where the insurance choice is exogenous. They find that given everything else constant, an increase in the coinsurance rate decreases the utilization although this effect is only significant at the 10% level. Finally, Chiappori et al. (1998) find evidence of moral hazard for GP home visits but not for office visits to a GP or to a specialist. The authors argue that the presence of high non-monetary costs associated with office visits translates a small change in price into a negligible change in the total cost bared by the patient. On the contrary, for GP home visits, the non-monetary costs are virtually inexistent, making a small change in price more noticeable.

Most studies control for individual's subjective assessment of health status because these are strong predictors of health services utilization (e.g. Cameron et al., 1988, Coulson et al., 1995, Vera, 1999). These variables are, however, likely to be endogenous, i.e. correlated with the unobserved variables in the health demand equation (Windjmeyer and Santos Silva, 1997). Windjmeyer and Santos Silva (1997) suggest to use long-term determinants of health, such as smoking and drinking, as instruments for the subjective health measures. In this paper, we present results with and without matching on subjective health measures but give more credibility to the results without subjective measures. In Appendix A we provide a theoretical motivation for not including subjective health variables in the regressions.

2.2 The Portuguese Health Care System

This Section will borrow heavily on Bentes et al. (2004) report, which describes at length the Portuguese health system from its beginning until 2002. The Portuguese National Health Service (NHS) was born in 1979 when legislation was passed establishing the right of all citizens to health protection, a guaranteed right to universal free health-care through the NHS, and access to the NHS for all citizens regardless of economic and social status.

Before 1979, the state only covered the costs of health-care for civil servants, and provided limited preventive care, maternal and child health-care, mental health, as well as some control of infectious diseases. The evolution of the health system in Portugal implies relatively better access to health-care by the elder cohorts of ADSE beneficiaries in our sample in comparison with their NHS-only counterparts. Our results by age group show that the impact of ADSE is higher for the youngest generation while smaller and not statistically significantly different from zero for the older groups. This result is compatible with the elderly ADSE beneficiaries being (unobservably) healthier than their NHS-only matches.

After 1979, some aspects of the pre-NHS period remained such as the health subsystems (from the Portuguese *subsistemas*). These are health insurance schemes for which membership is based on professional or occupational category. Bentes et al. (2004) state that these schemes were kept because trade unions were not willing to give up the good service and easy access to a wide range of providers. As a consequence, by 2004 about 25% of the population had, in addition to the NHS, coverage from some of the health subsystems (for a brief description of these health subsystems see Bentes et al., 2004 pages 21-24). The subsystems either use their infrastructure to provide treatment or contract with private or public providers. The largest subsystem and the focus of our paper, ADSE, covers 15% of the population including all employees of the NHS. This is a compulsory scheme for civil servants for which they pay 1% of their salary. ADSE beneficiaries have access to three types of assistance: First, ADSE has agreements with providers where the beneficiaries enjoy reduced copayments; Second, the beneficiaries may decide to choose a provider outside the ADSE agreement network in which case they pay a higher copayment. Finally, ADSE beneficiaries may also benefit from all services offered by the NHS network subject to the same small copayments and exemptions as any user of the NHS. According to Bentes et al. (2004), generally the subsystems offer larger benefits than the NHS. The NHS is predominantly funded through general taxation. The health subsystems are funded through employer and employee contributions although for example the ADSE needs to be topped up with money from the Government budget.

After 1992, hospitals and health centers started applying small copayments to NHS users. There is, however, a substantial fraction of the NHS users who are exempt.[8] Copayments and spending in private doctor visits are deductible from income tax, which means a de facto subsidy by the state. Copayments in the NHS are homogenous across specialties and only vary with the nature of the visit — emergency visits are more expensive— and with the type of health care center —visits at central and larger hospitals are more expensive. Copayments in the NHS range between 2 Euros for GP visits at the local health care center and 5 Euros for emergency visits to general hospitals. Diagnostic tests are also subject to copayments.[9]

Although the NHS is expected to offer all services demanded, in practice there is very limited provision of certain services such as dental care for adults. Adults demanding dental care usually consult a private dentist. Bentes et al., 2004 refer that according to 1995/1996 Health Survey 92% of all dentist visits are private. For the ADSE beneficiaries, the coverage for private dentists with no ADSE agreement is 80% up to a limit amount and the copayment for dentists with ADSE agreements depend on the specific treatment but vary between 0.95 Euros to 9.18 Euros with no limit. Legally, NHS-only beneficiaries may get partial reimbursements for their dentist visits and treatments but these are so low that few people take the trouble of filling the paperwork. In short, for dental care ADSE beneficiaries face a much better coverage than their NHS-only counterparts who are, in practice, left without coverage.

The private sector has been growing steadily and in 1998/99 accounted for about 32% of the specialists visits while the wide majority of the GP visits were carried out in the

[8] NHS users exempt from copayments are: the low income, the unemployed, the chronically-ill, pregnant women, children up to 12 years old, drug abusers under treatment, and the mentally-ill.

[9] source: http://www.arsc.online.pt/scripts/cv.dll?sec=sns&pass=guia_utente#15. Law published in 1992 (Portaria n^{o}338/92, 11 April). The more recent regulation is Portaria 103/2004, 23 January.

public sector (Bentes et al., 2004). While the percentage of private hospitals beds is only about 23% of total stock, the presence of the private sector in dental services, blood tests, X-rays, dialysis, and physiotherapy is considerably higher (Bentes et al., 2004).

3 The dataset

The dataset covers 48606 individuals belonging to 18186 families and 17491 households. The sample is selected following the multi-stage design of the census: geographical areas are drawn randomly and in several stages with probability proportional to their population. Within each of the smallest geographical area level, containing around 300 households, random draws select which households will be interviewed.

The dataset has information on socio-demographic characteristics, income levels, doctor and hospital visits, medical procedures, expenditures on physician services, objective and subjective measures of health status (e.g. obesity, and "feels in very good health"), and consumption habits that may affect health (e.g. tobacco and alcohol consumption). Finally, there is also information about health insurance status of each individual. Most individuals are only beneficiaries of the public National Health System (84%), followed by civil servants and their dependants who are beneficiaries of ADSE (9.99%), third those with private insurance (1.7%) and lastly there are up to 5 different insurance schemes linked to the military and the police.[10] We use three different variables to capture the demand for health services: 1) the number of physician visits in the previous 3 months; 2) the number of blood and urine tests in the previous 3 months; 3) whether the individual has visited the dentist in the last year. Importantly, a serious drawback of this dataset is not to know how many of the visits or tests were done in the private sector and how many were done in the NHS.

Our working sample was obtained after dropping a few observations from the data. First, we dropped 7 observations corresponding to situations where the insurance status of the individual was not known. Second, we drop 191 observations of pregnant women whose visits to the doctor were related to the pregnancy. Third, we deleted observations

[10]The data also contains information regarding the risks covered by the insurance plan for those individuals who declare having an insurance. Only 4.77% of the respondents declare having an insurance, which reflects a lack of understanding of the health system.

with inconsistent answers: 14 individuals declare not to have visited the doctor in the previous 3 months but yet declare to have visited a private doctor; 19 individuals who also declare not to have gone to the doctor in 3 months but declare to have gone to their nearest health center; a total of 7 people declare not to have visited the doctor but have gone either to a public or private hospital. Fourth, the information about a given individual in the dataset may be provided by the individual himself, as long as older than 15 years of age, or by a different person in the household. We dropped 165 observations corresponding to information provided by non-family members due to potential measurement error or missing values in important variables. We also dropped 2854 observations corresponding to people with special insurance status different from ADSE and the National Health System. These observations correspond to the military (1.09%), the police forces (1.45%), the judicial system employees and their dependants (0.29%), banking employees and their dependants (1.38%), as well as people with private insurance (1.7%). We decided to delete all these insurance types from the control group for two reasons: first, because the insurance coverage and copayments may differ, which may affect the beneficiaries behavior. Second, there may be non-observable variables correlated with the insurance type that may also affect individual's behavior.

Finally, we restrict the sample to observations where the individual answers for himself. This excludes minors from the sample. We also delete observations with missing values on the exogenous variables used. In the case we wish to include a variable with too many missing values such as the daily wine consumption (in litres) we create a dummy variable (winens) that equals one whenever the variable wine is missing, and zero otherwise. Altogether, our working samples range from 21151 to 21908 observations.

4 Preliminary and Exploratory analysis of the data

In this Section we perform a very preliminary analysis of the data. Table 1 shows some statistics relative to the first measure of health care services, number of visits to a doctor in the previous 3 months. The second column of Table 1 shows that, regardless of their insurance status, 53 percent of the people had at least one visit to the doctor in the previous 3 months. The p-value of the t-test indicates that the unconditional probability of at least one visit does not significantly differ across insurance types. However, as the first column of Table 1 also shows the average number of visits to the doctor is significantly smaller for the beneficiaries of ADSE than for those with the NHS-only. The higher number of visits for the beneficiaries of the NHS-only may reflect the requirement to visit a GP before visiting a specialist, this is not a requirement for the ADSE beneficiaries. This requirement is one of the main reasons why we consider visits to the doctor to be a less than perfect variable to study the prevalence of moral hazard in this context.

In order to circumvent the potential measurement error in number of visits, we also analyze blood and urine tests. Statistics for number of tests are shown in Table 2. The unconditional mean for number of tests for the ADSE group is significantly higher than the mean number of tests for the NHS-only group.

Another measure of interest is dental services. ADSE has a generous coverage for dental services while the NHS-only beneficiaries, in practice, visit private dentists and pay the full fee. As expected, the unconditional probability of having at least one visit to the dentist in the previous 12 months is significantly higher for the ADSE group than for the NHS-only group as table 3 shows.

The differences between the two types of insurance are wider if one looks at certain pre-determined variables (see Table 4). The NHS-only group is relatively poorer and older than the ADSE group. In the ADSE group there are less married or widows than in the NHS-only group and more single and divorced people. The regional distribution of the two groups is also unequal in the full sample. Most civil servants are concentrated in the capital and less in the other areas of the country with the exception of Alentejo, which is the least dense populated region in the country.[11] The ADSE group is more educated than the NHS-only group. The difference is even greater when the comparison involves the no-student population of the two groups since the ADSE group contains a larger proportion of students. The ADSE group also contains less people who are out of the labor force, on sick leave for more than 3 months, not working for some other reason, or unemployed. The percentage of people answering the questionnaire about him/her self is statistically the same in both insurance groups.

[11] The five regions correspond to the five regional health administrations.

In the ADSE group people feel healthier than in the NHS-only group as one can observe by the statistics on subjective health on Table 5. This makes sense since the presence of physical limitations and chronic diseases is more prevalent amongst the NHS-only group with the exception of allergies, which are more common amongst the ADSE beneficiaries. The ADSE group, perhaps because of the younger age, practices more exercise, drinks less wine and beer and have less weight problems than the NHS-only group. The percentage of smokers is, however, statistically the same in both groups. Finally, the ADSE beneficiaries practice a better dental hygiene than the NHS-only group.

5 Methodology

Our treatment group is composed of ADSE beneficiaries and represents roughly 10 percent of our sample. ADSE is the most common health insurance plan in the country, which is given to all civil servants and their dependants. The matching estimator proposed by Abadie and Imbens (2004) is used to estimate the "average treatment effect on the treated" (ATT) i.e. the average increase in the demand for health services amongst ADSE beneficiaries due to their double coverage. We argue it is very unlikely that individuals want to become civil servants just in order to benefit from ADSE health insurance since the NHS offers practically universal service. Moreover, if health coverage were the main objective, other health subsystems, such as the one of the banking sector, offer much better coverage than ADSE. If our argument is true then we may rule out the typical adverse selection in additional insurance. It is also implausible for the state to choose individuals on the basis of health variables unobserved to us. The hiring process for civil servants is highly regulated and the starts with a public call for applicants who need to fulfill certain criteria. Beyond being physically capable for the job, health status is not included among the criteria.[12] In short, we believe there are good arguments to discard selection bias in our ATT estimates.

Our ATT estimates may be underestimated if ADSE beneficiaries enjoy more or better treatment than NHS-only beneficiaries. Over the individuals' life, this better treatment would translate into a significant accumulation of health advantages by the ADSE benefi-

[12] The law is written in Decreto-Lei n^o. 204/98, de 11.07, article 29.

ciaries relative to their NHS-only counterparts. In this case, the impact of ADSE should be larger for the young beneficiaries who did not yet had the time to accumulate health advantages, and smaller for the old beneficiaries sho would be healthier than their NHS-only matches and, therefore, demand less health services. In order to identify this effect we estimate the ATT for each age group separately.

Finally, the ATT may overestimate moral hazard if there is supply-induced demand for ADSE beneficiaries. Supply inducement is more likely to occur in the number of visits to the doctor and in the number of tests since ADSE pays per visit except to dentists who are paid by procedure.[13] However, the ADSE payments to doctors are low so we think the magnitude of this effect, if positive, is relatively small.

Various methods of semiparametric estimation of average treatment effects under exogeneity have recently been proposed in the econometric literature (see Imbens, 2004, for a review and references). In this paper, we apply matching estimators in order to estimate the impact of having additional health insurance coverage on the demand for health services. Matching estimators have not been applied before in this context and have the advantage of avoiding the imposition of functional form restrictions.[14] In particular, we apply the matching estimators proposed by Abadie and Imbens (2004), "AI" in what follows, who derive the large sample properties of matching estimators of average treatment effects and consistent estimators for their variance.[15]

We are interested in estimating the average effect of a binary treatment (having ADSE health insurance coverage) on some outcomes (several measures of health care utilization: physician visits, tests and dentist visits within different time periods). Each individual i, with $i = 1, ..., N$, is characterized by a pair of potential outcomes, $Y_i(0)$ for the outcome

[13] Doctors may request more tests in order to justify more visits. Inducement by dentists would be reflected in a misreport of the type of services provided to the patient and not in the number of visits to the dentist.

[14] Both matching and traditional regression analysis assume that selection into treatment is "on the observables". Contrary to regression analysis, matching does not make the linear functional form assumption. Moreover, if selection is actually on the observables but linearity does not hold, then matching estimates are consistent while regression estimates are not.

[15] While AI focus on "covariate matching", many recent studies have relied on propensity score matching. No variance estimator has been proposed for the case when the propensity score is estimated rather than known and researchers often use bootstrapping methods. However, Abadie and Imbens (2005) show that the standard bootstrap is generally invalid and AI provide, for the time being, the only analytic variance estimators that are formally justified.

under the control treatment i.e. the NHS-only, and $Y_i(1)$ for the outcome under the active treatment i.e. ADSE. For each individual i we observe the triple (W_i, X_i, Y_i), where X_i is a vector of covariates, W_i, $W_i \in \{0, 1\}$ indicates the treatment received and Y_i, denotes the realized outcome:

$$Y_i \equiv Y_i(W_i) = \begin{cases} Y_i(0) & \text{if } W_i = 0 \\ Y_i(1) & \text{if } W_i = 1. \end{cases} \tag{1}$$

The realized outcome is equal to the $Y_i(0)$ if the individual is not an ADSE beneficiary and equals $Y_i(1)$ otherwise. Note that the treated outcome, $Y(1)$, is only observed for treated units while the untreated outcome, $Y(0)$, is only observed for comparison units. Hence, only one of the potential outcomes is observed for each individual and the other is unobserved.

Our main focus is on the population average treatment effect on the treated ($\tau^{p,t}$) and its sample counterpart ($\tau^{s,t}$):

$$\tau^{p,t} = E[Y_i(1) - Y_i(0)|W_i = 1] \text{ and } \tau^{s,t} = \frac{1}{N_1} \sum_{i:W_i=1} (Y_i(1) - Y_i(0)),$$

where $N_1 = \sum_{i=1}^{N} W_i$ is the number of treated individuals. The main idea behind matching estimators is that, if assignment to treatment is independent of the potential outcomes for individuals with similar values of the covariates, the unobserved potential outcomes can be imputed by using only the outcomes of similar individuals of the opposite treatment group. The following key assumptions about the treatment assignment are made:

Assumption 1: *For all x in the support of X,*

(i) (unconfoundedness) W is independent of $(Y(0), Y(1))$ conditional on $X = x$;

(ii) (overlap) $c < Pr(W = 1|X = x) < 1 - c$, for some c.

The combination of unconfoundedness and overlap is referred to as strong ignorability (Rosenbaum and Rubin, 1983). The overlap assumption restricts the joint distribution of observables and it requires that the conditional probability of receiving treatment, also known as the propensity score (Rosenbaum and Rubin, 1983), is bounded away from zero and one. Without overlap in the covariates distributions, all individuals with a given covariate pattern would receive the same treatment and there would be no similar individuals of the opposite treatment group. The unconfoundedness assumption, also known as "selection on observables", validates comparisons of individuals with the same covariates values, i.e. the following is true:

$$E[Y(w)|X = x] = E[Y(w)|W = w, X = x] = E[Y|W = w, X = x]. \tag{2}$$

That is, systematic differences in outcomes between treated and control units with the same values for the covariates are attributable to the treatment. Thus, the average treatment effect can be recovered by averaging $E[Y|W = 1, X = x] - E[Y|W = 0, X = x]$ over the distribution of X.

Given the richness of the Portuguese Health Survey, we believe that the unconfoundedness assumption is not unreasonable in our context. In this respect, it is worth remarking that the NHS is available to everyone almost for free and of relatively high quality. As argued in the beginning of the Section, it is unlikely that individuals anticipating a high healthcare demand choose to become civil servants in order to obtain ADSE coverage.

In many studies, the number of exogenous variables is large and an exact match may be impossible. Therefore, matching is based on the observations that are "close" in terms of their covariate values. More precisely, let $j_m(i)$ be the index of the individual that is the m-th closest match to individual i in terms of the distance measure based on the norm $\|\cdot\|$, among the individuals with the treatment opposite to that of individual i. Formally, $j_m(i)$ is the index j that solves $W_j = 1 - W_i$ and

$$\sum_{l:W_l=1-W_i}^{N} 1\{\| X_l - X_i \| \leq \| X_j - X_i \|\} = m, \tag{3}$$

where $1\{\cdot\}$ is the indicator function, equal to one if the expression in brackets is true and zero otherwise.[16] Matching is carried out with replacement, allowing each individual to be used as a match more than once, which produces matches of higher quality than matching without replacement by increasing the set of possible matches.

The simple matching estimator for the average treatment effect for the treated that AI propose uses the following estimates for the imputed potential outcomes for control individuals:

[16] These definitions can easily be generalized to allow for the presence of ties.

$$\hat{Y}_i(0) = \begin{cases} Y_i & \text{if } W_i = 0, \\ \frac{1}{M}\sum_{j \in I_M(i)} Y_j & \text{if } W_i = 1, \end{cases} \tag{4}$$

where $I_M(i)$ denotes the set of indices for the first M matches for individual i. In words, the missing potential outcomes are estimated by averaging the outcomes of the nearest neighbors of the opposite treatment group. Hence, the simple matching estimator for the average treatment effect for the treated discussed in AI is:

$$\hat{\tau}^{sm,t} = \frac{1}{N_1} \sum_{i:W_i=1} (Y_i - \hat{Y}_i(0)) \tag{5}$$

where N_1 denotes the number of treated individuals in the sample.

AI show that due to matching discrepancies this estimator has a bias of order $O(N^{-1/K})$, where K is the number of continuous covariates. To remove this bias, they propose a bias-corrected matching estimator that adjusts the differences within the matches for the differences in their covariate values by combining the matching process with a regression adjustment. The adjustment is based on an estimate of the regression function $\mu_w(x) \equiv E[Y(w)|X = x]$ for $W = 0$ since we are interested in estimating the average treatment effect for the treated.[17] Given the estimated regression function for the controls, the missing potential outcomes are predicted as:

$$\tilde{Y}_i(0) = \begin{cases} Y_i & \text{if } W_i = 0, \\ \frac{1}{M}\sum_{j \in I_M(i)} (Y_j + \hat{\mu}_0(X_i) - \hat{\mu}_0(X_j)) & \text{if } W_i = 1, \end{cases} \tag{6}$$

with corresponding estimator for the ATT:

$$\hat{\tau}^{bcm,t} = \frac{1}{N_1} \sum_{i:W_i=1} (Y_i - \tilde{Y}_i(0)) \tag{7}$$

This bias adjustment makes matching estimators $N^{1/2}$-consistent. [18] Our application suggests that the bias obtained with the simple matching estimator may be large.

[17] AI use non-parametric estimation to impute the value for the untreated.

[18] For details on the properties of the matching estimators and their variance see AI.

6 Results

In this Section we show estimates for the ATT using different specifications.[19]

Table 6 shows results for number of visits to the doctor and number of tests when matching is done on a wide set of covariates. We classify the set of covariates into two groups. Group 1 consists of individual characteristics such as age, female dummy, marital status, number of family members, region dummies, dummies for the month of the interview, years of schooling, employment status, occupational dummies, and up to eleven income level dummies; Group 2 consists of variables that are more related to the individual's health status or habits such as on-sick-leave for more than 3 months, on-sick-leave for less than 3 months, other reasons for not working, daily average consumption of wine (in litres), underweight, overweight, obesity, restricted activity, and other diseases such as asthma, diabetes, bronchitis, allergies, high blood pressure, back pain, living habits such as exercise, smoking, intake of sleeping pills, and subjective measures of health.

Table 7 shows the estimated ATT when we restrict the variables used for matching to the variables in group 1, and a few variables in group 2 such as asthma, allergies, and diabetes, which we think are exogenous in this context.[20] The other variables in group 2 are potentially problematic. For example, there are at least two problems with the subjective measures of health although the received literature typically includes them as controls for unobserved conditions.[21] First, perceived health may be endogenous (Windjmeyer and Santos Silva, 1997). A shock affecting the number of visits may also impact the individual's perceived health. Second, perceived health may be an intermediary output. For example, ADSE beneficiaries would have a higher perceived health if ADSE offered higher quality services. The intermediary output classification also applies to the excluded health status and health habits variables in group 2. In appendix A, we present a simple regression model showing that variables such as subjective health measures or

[19] Results were obtained from running different versions of the Abadie & Imbens Matlab programs provided in their web pages.

[20] We excluded diabetes from matching for the eldest cohort because it may be the outcome of bad eating habits and, therefore, related to quality of health services received. On the contrary, it is more likely to be a genetic condition for the youngest cohorts.

[21] One of the problems with perceived measures of health identified in the literature is their sensitivity to the order in the survey. It has been shown that if perceived health is asked at the outset, individuals tend to tell the truth. This is the case in the Portuguese Health Survey.

intermediate outcomes may bias the impact of ADSE upwards, as it can be verified from the comparison of results in tables 6 and 7.

The first segment on Tables 6 and 7 shows the results obtained with the whole sample ($N = 21,151$ and $N = 21,908$, respectively). The first two rows correspond to the estimated ATT obtained with a single match ($M = 1$) while rows three and four correspond to the estimated ATT obtained with four matches ($M = 4$). Note that by increasing the number of matches, one increases the precision of the estimates at the cost of greater bias. Row five shows the unconditional difference between the ADSE and the NHS-only group. Row six shows the coefficient of the ADSE dummy in a OLS regression.[22] The three bottom segments in Tables 6 and 7 show the estimated ATT obtained after splitting the sample into three age groups.[23]

Results show that, especially for number of visits to the doctor, the simple-matching estimator of ATT produces at times very different results from the bias-adjusted matching estimator. As discussed in the previous Section, the simple-matching estimator is biased when there are continuous covariates, as it is the case of the age variable. Hence, we regard the bias-adjusted estimates as more reliable.

Table 6 shows that for the overall sample, the simple-matching produces a strong and statistically significantly different from zero ATT for number of visits to the doctor whereas the bias-adjusted matching shows a small and not statistically significantly different from zero effect. For number of tests, the difference between the simple-matching and the bias-adjusted ATT is smaller and both estimates are positive and statistically significantly different from zero. The statistically insignificant impact of ADSE on number of visits and the larger effect on number of tests may have several explanations: First, the heterogeneity of visits, e.g. the requirement to visit the GP before the first visit to a specialist artificially increases the number of visits for the NHS-only group; Second, higher

[22]Reported standard deviations for the unconditional mean difference and the OLS regression are corrected for clustering. This correction may be important since more than one family member may be present in the sample implying error terms are not independent. Correcting for clustering, however, hardly affected the standard deviations.

[23]The variance for the ATT has been calculated using one match and allowing for heteroskedasticity. Regarding the metric used to measure the distance between covariates, let $\| x \|_V = (x'Vx)^{1/2}$ be the vector norm with positive definite weight matrix V and define $\| z - x \|$ to be the distance between the vectors x and z. We follow Abadie and Imbens and define V as the diagonal matrix of the inverse of the covariate variances.

quality services provided to the ADSE group may reduce the number of visits needed to treat the same condition; Third, the ADSE group is unobservably healthier, which would occur if the ADSE group enjoyed better services. In order to isolated the latter bias we present the estimated ATT by age group with the conviction that if it exists, the bias should be larger for the older generation, which has had more time to accumulate health benefits relative to their NHS-only counterparts, and smaller or inexistent for the younger generation. The estimation by age group improved the quality of the matching thereby reducing the disparity between the simple-matching and the bias-adjusted ATT estimates.

The youngest cohort in Table 6 has the largest estimated ATT for number of visits. The ATT of 0.533 (bias-adjusted and $M = 1$) represents 48% of the average number of visits for that cohort and is very statistically significantly different from zero. As argued, it is likely that the youngest cohort has not accumulated health benefits so this large effect is solely due to moral hazard. The second largest estimated ATT (0.342) is obtained for the oldest cohort but here, although statistically significantly different from zero, it only represents 18% of the average number of visits for this cohort. The estimated ATT for the middle-age cohort is very small (0.04) and not statistically significant. The smaller effects of ADSE on the middle-age and eldest cohorts may be due to the accumulated health benefits derived from better health care over the years.

The estimated ATT for number of tests in Table 6 is typically smaller than for number of visits but statistically significantly different from zero for all cohorts but the eldest. The estimated ATT is largest for younger groups, possibly also reflecting accumulated health benefits from ADSE coverage. The estimated ATT for the youngest cohort represents 65% of the average number of tests for this age group and 22% of the average for the middle-age groups. These effects are quite substantial.

When comparing results from Table 6 with those in Table 7 notice that, at least for number of visits, the differences between the simple-matching and the bias-adjusted estimates are smallest in the latter. Also, the exclusion of most of the variables in group 2 leads, in general, to a drop in the ATT estimates and, except for the youngest cohort, to a rise in the standard deviations. The estimated ATT for number of visits in Table 7 are still large for the youngest and oldest cohorts, representing 21.8 and 12.6 percent of the average number of visits respectively, but so imprecisely estimated that none of these estimates are statistically significantly different from zero. In contrast, the ATT for number of tests is statistically significant for the overall sample and almost statistically significantly different from zero for the youngest cohort, representing 15.8 and 30 percent of the average number of tests, respectively.

The samples in the previous Tables 6 and 7 may contain more than one member per family. In these situations the error terms are not independent, which causes a bias in the standard deviations. In table 8 we reestimated the ATTs from Table 7 restricting the sample to one member per family in order to obtain correct standard deviations. Most results are similar to the ones in Table 7 except the ATT for number of tests for the youngest cohort, which is now considerably larger and strongly statistically significantly different from zero.[24]

Finally, we also estimate the ATT for the variable at least one visit to the dentist in the 12 months prior to the interview. This variable is very different from the previous ones since the public coverage of dental services is scant and presents long waiting lists. Most NHS-only beneficiaries, therefore, consult private dentists for which they have no insurance coverage. Hence, for dental services, ADSE beneficiaries have a much higher coverage than the NHS-only group. Theoretically, we expect a positive impact of ADSE in the probability of visiting a dentist. Table 9 shows the ATT estimates for the whole sample and the age group subsamples controlling for the smallest set of regressors. Results show a positive effect for all the samples but the middle-age, but none of the estimates is statistically significantly different from zero. Chiappori et al. (1998) argued that when non-monetary costs are large the demand is more inelastic and this may well be the explanation for the small or inexistence effect of ADSE in dental care. Alternatively,

[24]Due to the high number of zeros and ones in number of visits and number of tests we decided to estimate the ATTs also on the binary versions of these variables i.e. *at least one visit* and *at least one test* both in the previous 3 months. The comparison of the estimated ATTs share similar features to the previous tables. For example, when excluding most covariates from group 2, the simple-matching ATT estimates became similar to the bias-adjusted estimates, the estimated ATTs dropped, and the standard deviations increased for all but the yongest cohort. However, the ATT estimates were, in general, imprecisely estimated so none of the bias-adjusted ATTs for number of visits are statistically significantly different from zero. For number of tests, again the ATTs are statistically significant for the overall and the youngest cohort but only statistically significantly different from zero for the middle-age cohort when all covariates are used for matching. Similarly to the results in Tables 6 and 7, the estimated ATT is larger for the youngest cohort.

this result may be the consequence of a coarse dependent variable and if instead we had number of visits to the dentist we may have found an effect.

7 On the Exogeneity of the Insurance plan

Our identification strategy relies on several assumptions. First, we believe that those individuals who expect to use more health services do not select themselves to become civil servants in order to benefit from ADSE coverage. As argued above the greatest benefit of becoming a civil servant is to hold a job for life and, for some job categories, the wage offered by the Government is higher than the wage offered by the private sector. Moreover, other subsystems such as the one associated with the banking sector offer better health insurance coverage than ADSE. Still, there is the possibility that those who expect to use more health services, because they are sicker or because they are more risk averse, would more likely become civil servants. Second, it must be true that the state does not select its employees on the basis of unobservable (to us) health variables. As we argue in the previous Section, the Government must make a public call for applicants and the process is highly regulated and objective. Applicants must fulfill certain criteria and apart from being physically able for the job, health status is not part of the criteria. Third, for the unconfoundness assumption to hold, it must be the case that ADSE beneficiaries are not unobservably healthier, for example because they have enjoyed more years of better treatment. As discussed above if the latter holds then those individuals who have been ADSE beneficiaries for a longer period of time would visit the doctor less and demand less tests than their NHS-only matches. This would imply a smaller impact of ADSE on the old than on the young because the latter have not yet had the time to accumulate these health benefits. The results by age group discussed in the previous Section show that the impact of ADSE is larger for the young cohort and, at least for the case with all covariates, strongly statistically positive. This suggests that the unconfoundness assumption may not hold for the middle-age and eldest cohorts or, in other words, that there is a long-term effect from ADSE.

In order to support the first identification assumption, i.e. whether ADSE beneficiaries are more risk averse than their NHS-only counterparts, we ran OLS regressions of risky life-style habits such as smoking and drinking against the restricted set of covariates and an ADSE dummy. We find that ADSE beneficiaries consume statistically significantly more wine but less whisky. All other life-style behaviors were identical in both groups. These regressions do not offer clear evidence that ADSE beneficiaries are more risk adverse than NHS-only beneficiaries.

To support the first and the second identification assumptions, we focus on a subsample of ADSE beneficiaries that obtain ADSE coverage through a family member rather than in their own name. In principle, people who enjoy ADSE through someone else would only be subject to moral hazard and, therefore, the first and second identification assumptions should hold by default. Unfortunately, our dataset does not allow us to identify whether individuals are covered by ADSE in their own name or through another family member. Hence, we take a conservative approach and restrict our sample of ADSE beneficiaries to individuals who are unemployed and, therefore, cannot be civil servants. This approach is similar to the one followed by Vera (1999) who splits the sample between heads of households (in principal, those that take the decision of contracting private insurance) and non-head of households (beneficiaries of the private insurance who do not take the decision) in order to test for adverse selection in the contract of private insurance.

Table 10 shows the estimated ATT for number of visits to the doctor, number of tests, and probability of visiting the dentist for unemployed ADSE beneficiaries when excluding most variables from group 2. The first thing to notice is the imprecision of most ATT estimates which is due, most likely, to the small sample size (878 observations). We restrict our description to the bias-adjusted ATT for one match ($M = 1$) because the quality of the matches for $M = 4$ is not good with such a small number of observations. The magnitude of the ATT for number of visits, number of tests, and for at-least-one dentist visit are much higher than the ATTs for the overall sample (in Tables 7 and 9), representing 20 %, 114 % and 17% of the average number of visits, tests, and probability of dentist visit respectively, although none of these estimates is precisely estimated. For number of visits and tests we expected a smaller impact of ADSE in this case given that unemployed people do not have to pay co-payments in NHS but are subject to copayments under ADSE.

In conclusion, due to the low number of observations it is hard to find support for our identifying assumptions although the size of the point estimates seem to indicate that the ADSE effect is mostly due to moral hazard.

8 Conclusion

This paper estimates the impact of additional coverage on the demand of visits to the doctor, diagnostic tests and the probability of visiting the dentist within the Portuguese National Health System. Our paper's contribution to the large literature on moral hazard is fourfold: First, by using a dataset where 10 percent of the sample enjoys an exogenous double health insurance coverage; Second, by using a matching estimator technique (Abadie and Imbens, 2004), which does not impose any functional form on the impact of health insurance on the demand for health services and allows for heterogenous impacts. Third, by estimating the impact of the additional insurance on several services, particularly on the diagnostic blood and urine tests. And fourth, by allowing for a dynamic impact of the additional coverage splitting the sample by age groups.

In general we find that the impact of ADSE is positive and large. For the whole sample the ADSE effect corresponds to 6 percent of the average number of visits, 15.8 percent of the average number of tests, and 7 percent of the average probability of visiting a dentist. The effects of ADSE are particularly large for the youngest cohort $18-30$ years old where they reach 21.8, 30 and 11.6 percent of the average number of visits, tests and probability of visiting the dentist, respectively for that age group. Due to the imprecision of estimates, however, the Average Treatment Effect on the Treated (ATT) for number of visits is not statistically different from zero. For number of tests, we do find evidence of moral hazard for the overall sample and the youngest cohort. On the contrary, for dental services, where we expected to find the greatest ATT, we do not find evidence of moral hazard. We argue following Chiappori et al. (1998) that the inexistence of moral hazard for dental visits is the consequence of large non-monetary costs for this type of services.

Supply inducement if it existed would be less likely to occur in the at-least-one-visit to the dentist since ADSE pays all doctors but dentists per visit. Our data does not allow us to differentiate between the effects of moral hazard and supply-induced demand but the presence of supply-induced demand would lead to an overestimation of the moral hazard effect. However, there is no reason why demand inducement should be larger for the younger cohort (18-30 years old). Quite the contrary, we should expect inducement to be larger for the older cohort since these, specially the retirees, have a lower opportunity cost of time and, therefore, a more inelastic demand (e.g. van de Voorde et al., 2001).

Our results are also consistent with long-term positive effects from ADSE since the estimated ATT is lower for older generations who may have accumulated health benefits from better treatment and enjoy better unobserved health than their NHS-only counterparts.

References

[1] Abadie, Alberto and Imbens, Guido (2004): "Large Sample Properties of Matching Estimators for Average Treatment Effects," mimeo, Harvard University.

[2] Abadie, Alberto and Imbens, Guido (2005): "On the Failure of the Bootstrap for Matching Estimators," mimeo, Harvard University.

[3] Álvarez, Begoña (1999): "Especificación y validación de modelos de demanda de asistencia sanitaria , absentismo y actitudes de los desempleados: aplicación al caso Español," PhD dissertation, Universidad Carlos III de Madrid.

[4] Arrow, Kenneth J. (1963): "Uncertainty and the Welfare Economics of Medical Care," American Economic Review, 53(5), pp 485-973.

[5] Bentes M, Dias CM, Sakellarides C, Bankauskaite V (2004). Health care systems in transition: Portugal. Copenhagen, WHO Regional Office for Europe on behalf of the European Observatory on Health Systems and Policies.

[6] Cameron, A.C. and Trivedi, P.K. (1998): "Regression Analysis of Count Data," Cambridge University Press.

[7] Cameron, A.C., Trivedi, P.K., Milne, Frank, and Piggott, J. (1988): "A Microeconometric Model of the Demand for Health Care and Health Insurance in Australia," Review of Economic Studies, pp 85-106.

[8] Chiappori, Pierre-André, Durand, Frank, Geoffard, Pierre-Yves (1998): "Moral Hazard and the Demand for physician services: First Lessons from a French Natural Experiment," European Economic Review, 42, pp 499-511.

[9] Deb, Partha and Pravin K. Trivedi (2002): "The Struture of Demand for Health Care: Latent Class versus Two-Part Models," Journal of Health Economics, 21, pp 601-625.

[10] Holly, Alberto Lucien Gardiol and Jacques Huguenin (2002): "Hospital services utilization in Switzerland: The role of supplementary insurance," Institute of Health Economics and Management, University of Lausanne, manuscript.

[11] Holly, Alberto Lucien Gardiol, Gianfranco Domenighetti and Brigitte Bisig (1998): "An econometric model of health care utilization and health insurance in Switzerland," *European Economic Review* 42, pp 513-522.

[12] Imbens, Guido (2004): "Nonparametric Estimation of Average Treatment Effects under Ecogeneity", *Review of Economics and Statistics* 86(1), 4-29.

[13] Jimenez-Martin, Sergi and José M. Labeaga and.Maite Martinez-Granado (2001): "Latent class versus two-part models in the demand for physician services across the European Union," Health Economics.

[14] Manning, Willard G., et al. 1987. Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment. Santa Monica: RAND Corporation (Pub. no. R-3476-HHS). Abridged version published in American Economic Review 77:251-277

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

[16] Newhouse, Joseph and the Insurance Experiment Group (1993): "Free for All? Lessons from the RAND Health Insurance Experiment," Harvard University Press, Cambridge.

[17] Olivella, Pau and Vera-Hernandéz, Marcos (2005): "Testing for Adverse Selection in the National Health Service", mimeo.

[18] Savage, Elizabeth and Donald J. Wright (2003): "Moral Hazard and Adverse Selection in Australian Private Hospitals: 1989-1990," Journal of Health Economics, 22, pp 331-359.

[19] Trujillo, Antonio J., Portillo, Jorge E., Vernon, John A. (2005): "The Impact of Subsidized Health Insurance for the Poor: Evaluating the Colombian Experience Using Propensity Score Matching," *International Journal of Health Care Financing and Economics*, 5, 211-239.

[20] Vera-Hernandéz, Marcos (1999): "Duplicate Coverage and Demand for Health Care. The case of Catalonia," Health Economics vol 8 pp 579-598.

[21] Windmeijer, Frank A. G. and Joao M. C. Santos-Silva (1997): "Endogeneity in Count Data Models: An application to demand for health care," Journal of Applied Econometrics, vol 12 281-294.

[22] Rosenbaum, P. R. and D.B. Rubin (1983): "The Central Role of the Propensity Score in Observational Studies for Causal Effects," Biometrika 70(1), 41-55.

[23] Rothschild, M and Stiglitz, J. (1976): "Equilibrium in Competitive Insurance Markets," Quarterly Journal of Economics, 90, pp 629-649.

[24] Santos-Silva, Joao M. C. and Frank Windmeijer (2001): "Two-part Multiple Spell Model for Health Care Demand," Journal of Econometrics, vol104 (1), pp 67-89.

[25] Santos-Silva, Joao M. C and Teresa Bago d'Uva (2002?): "Asymmetric Information in the Portuguese Health Insurance Market," ISEG, Universidade Técnica de Lisboa, mimeo.

[26] Savage, Elizabeth and Donald J. Wright (2003): "Moral Hazard and Adverse Selection in Australian Private Hospitals," *Journal of Health Economics*, 22 pp 331-359.

[27] Van de Voorde, Carine, Van Doorslaer, Eddy and Erik Schokkaert (2001): "Effects of Cost Sharing on Physician Utilization under Favourable Conditions for Supplier-Induced Demand," *Health Economics*, 10, pp 457-471.

[28] Wagstaff, Adam and Menno Pradhan (2003): "Evaluating the impacts of health insurance: looking beyond the negative," mimeo.

9 Appendix A

Here we present an argument for excluding subjective measures from some of our regressions. Denote by y the number of visits to the doctor and H the subjective measure of health or any intermediate output (for simplicity take H to be a continuous variable). For notation simplicity assume all other characteristics X are constant. then the demand for visits could be modelled as:

$$\begin{aligned} y &= \alpha + \beta ADSE + \gamma H + \varepsilon && (8) \\ H &= a_0 + a_1 ADSE + u && (9) \end{aligned}$$

Suppose for the moment that ε and u are not correlated. We check the bias under correlation below. The partial derivative (obtained when including subjective health measures) is

$$\frac{\partial y}{\partial ADSE} = \beta \quad (10)$$

but the total effect (total derivative) is in reality:

$$\frac{dy}{dADSE} = \beta + \gamma a_1 < \beta \text{ if } \gamma < 0 \quad (11)$$

The total effect is more interesting. Of course if we do not include H then we would have an ommitted variable bias and we would underestimate β, the partial effect.

Now what if the error ε and u are (negative) correlated? It is very likely that the shocks that affect your perceived health affect your decision to go to the doctor conditional on H. That means that all estimates in 8 would be inconsistent and in particular γ. So what happens if we do not include H?

$$y = \alpha + \gamma a_0 + (\beta + \gamma a_1) ADSE + \gamma u + \varepsilon \quad (12)$$

we would estimate the total effect of ADSE without bias.

10 Appendix B

Table 1: Doctor Visits by Health Insurance Status. Full sample

	Nr. of Doctor Visits				**At Least 1 Visit (%)**				**N**
	Mean	S.D.	Min.	Max.	Mean	S.D.	Min.	Max.	
ADSE	1.16	1.89	0	30	0.53	0.50	0	1	4808
NHS	1.27	2.04	0	30	0.53	0.50	0	1	40484
|t-value|	3.56				0.05				
p-value	0.0004				0.9576				

Table 2: Blood Tests by Health Insurance Status. Full sample

	Nr. of Tests				**At Least 1 Test (%)**				**N**
	Mean	S.D.	Min.	Max.	Mean	S.D.	Min.	Max.	
ADSE	0.30	0.73	0	10	0.23	0.42	0	1	4811
NHS	0.27	0.67	0	10	0.22	0.41	0	1	40481
|t-value|	2.42				2.15				
p-value	0.0155				0.0312				

Table 3: Dentist Visits by Health Insurance Status. Full sample

	At Least 1 Dentist Visit (%)				**N**
	Mean	S.D.	Min.	Max.	
ADSE	0.52	0.50	0	1	4251
NHS	0.35	0.48	0	1	33331
|t-value|	21.32				
p-value	0.0000				

Table 4: Socio-Economic Characteristics by Health Insurance Status. Full sample

	ADSE			NHS			H0: Equal Means	
Variable	N	Mean	S.D.	N	Mean	S.D.	|t-value|	p-value
Age	4814	37.528	21.086	40535	42.394	23.196	14.97	0.0000
Female	4814	0.558	0.497	40535	0.519	0.500	5.12	0.0000
HH size	4814	3.321	1.187	40535	3.293	1.412	1.54	0.1243
Respondent $\neq$ Individual	4814	0.515	0.500	40535	0.515	0.500	0.03	0.9722
Married	4814	0.511	0.500	40535	0.544	0.498	4.33	0.0000
Single	4814	0.415	0.493	40535	0.350	0.477	8.72	0.0000
Widow	4814	0.042	0.201	40535	0.085	0.279	13.36	0.0000
Divorced/Separated	4814	0.032	0.175	40535	0.021	0.144	4.08	0.0000
Norte	4814	0.212	0.409	40535	0.322	0.467	17.38	0.0000
Centro	4814	0.187	0.390	40535	0.202	0.401	2.49	0.0128
Lisboa	4814	0.312	0.463	40535	0.244	0.429	9.66	0.0000
Alentejo	4814	0.170	0.376	40535	0.115	0.319	9.84	0.0000
Algarve	4814	0.119	0.324	40535	0.117	0.322	0.35	0.7285
Years of schooling	4814	8.730	5.483	40499	4.858	3.935	47.57	0.0000
Currently working	4814	0.494	0.500	40533	0.433	0.495	8.10	0.0000
Unemployed	4814	0.006	0.076	40533	0.038	0.192	22.37	0.0000
On sick leave (< 3 months)	4814	0.005	0.067	40533	0.006	0.078	1.48	0.1390
On sick leave (> 3 months)	4814	0.006	0.076	40533	0.010	0.098	3.27	0.0011
Student	4814	0.252	0.434	40533	0.145	0.352	16.42	0.0000
Not in labour force	4814	0.484	0.500	40533	0.508	0.500	3.13	0.0018
Not working (other reason)	4814	0.009	0.092	40533	0.005	0.070	2.60	0.0094
IncomeA	4814	0.007	0.085	40535	0.084	0.277	41.47	0.0000
IncomeB	4814	0.016	0.127	40535	0.117	0.321	41.25	0.0000
IncomeC	4814	0.042	0.200	40535	0.133	0.339	27.29	0.0000
IncomeD	4814	0.072	0.259	40535	0.128	0.334	13.78	0.0000
IncomeE	4814	0.082	0.275	40535	0.124	0.329	9.66	0.0000
IncomeF	4814	0.088	0.284	40535	0.109	0.312	4.68	0.0000
IncomeG	4814	0.103	0.304	40535	0.085	0.278	4.02	0.0001
IncomeH	4814	0.154	0.361	40535	0.079	0.269	14.06	0.0000
IncomeI	4814	0.151	0.358	40535	0.044	0.205	20.32	0.0000
IncomeJ	4814	0.234	0.423	40535	0.035	0.184	32.20	0.0000
IncomeNS	4814	0.041	0.199	40535	0.058	0.233	5.40	0.0000
IncomeNR	4814	0.008	0.091	40535	0.006	0.074	2.07	0.0389

Note:

Table 5: Health-related Indicators by Health Insurance Status. Full sample

	ADSE			NHS			H0: Equal Means	
Variable	N	Mean	S.D.	N	Mean	S.D.	|t-value|	p-value
Very good health	3207	0.081	0.273	25400	0.037	0.188	9.00	0.0000
Good health	3207	0.488	0.500	25400	0.355	0.479	14.21	0.0000
Normal health	3207	0.339	0.473	25400	0.380	0.485	4.69	0.0000
Bad health	3207	0.077	0.267	25400	0.180	0.385	19.45	0.0000
Very bad health	3207	0.015	0.121	25400	0.047	0.213	12.84	0.0000
Walking limitations (age>10)	4326	0.009	0.093	37051	0.027	0.162	10.97	0.0000
Diabetes	4814	0.036	0.186	40493	0.056	0.230	6.85	0.0000
Asthma	4811	0.046	0.210	40501	0.062	0.241	4.75	0.0000
Chronic bronchitis	4811	0.020	0.140	40496	0.030	0.170	4.55	0.0000
Allergy	4814	0.162	0.369	40506	0.142	0.350	3.50	0.0005
High blood pressure	4806	0.124	0.330	40434	0.178	0.383	10.57	0.0000
Back pain	4814	0.301	0.459	40504	0.415	0.493	16.10	0.0000
Sleeping pills (age>14)	3942	0.124	0.329	34772	0.130	0.336	1.16	0.2469
Smoker	4809	0.174	0.379	40512	0.176	0.380	0.23	0.8195
Physical exercise (age>14)	3942	0.166	0.372	34784	0.083	0.276	13.49	0.0000
No brushing teeth	4734	0.016	0.126	39905	0.059	0.236	19.79	0.0000
No teeth	4734	0.006	0.079	39905	0.024	0.152	12.47	0.0000
Obese	3630	0.102	0.303	32434	0.132	0.339	5.63	0.0000
Overweight	3630	0.333	0.471	32434	0.374	0.484	5.04	0.0000
Normal weight	3630	0.534	0.499	32434	0.472	0.499	7.12	0.0000
Underweight	3630	0.032	0.175	32434	0.022	0.147	3.21	0.0013
Wine (lt.)	2039	0.211	0.332	17518	0.264	0.331	6.78	0.0000
Beer (lt.)	1483	0.201	0.381	11634	0.261	0.404	5.61	0.0000
Bagaco (lt.)	388	0.016	0.062	3872	0.018	0.048	0.68	0.4998
Whisky (lt.)	843	0.018	0.044	5608	0.018	0.051	0.29	0.7727

Note:

Table 6: Matching and regression estimates of the impact of ADSE on physician visits and blood and urine tests. Standard deviations for regressions are corrected for clustering. Dependent variables are counts, specification with all covariates

Age Group	**M**	**Estimator**	**Nr. Doctor Visits**		**Nr. of Tests**	
			ATT	**(St. Error)**	**ATT**	**(St. Error)**
All, N=21,151						
	1	Simple Matching	0.215	(0.063)	0.116	(0.026)
		Bias-adjusted	0.060	(0.062)	0.096	(0.026)
	4	Simple matching	0.158	(0.057)	0.089	(0.023)
		Bias-adjusted	0.017	(0.056)	0.042	(0.023)
		Mean difference	-0.124	(0.048)	0.064	(0.020)
		Regression	0.030	(0.053)	0.038	(0.023)
18-30, N=2,741						
	1	Simple Matching	0.623	(0.175)	0.147	(0.049)
		Bias-adjusted	0.533	(0.174)	0.137	(0.049)
	4	Simple matching	0.462	(0.154)	0.103	(0.042)
		Bias-adjusted	0.340	(0.153)	0.105	(0.042)
		Mean difference	0.177	(0.158)	0.053	(0.038)
		Regression	0.318	(0.168)	0.047	(0.041)
30-60, N=10,422						
	1	Simple Matching	0.189	(0.072)	0.082	(0.032)
		Bias-adjusted	0.040	(0.071)	0.081	(0.032)
	4	Simple matching	0.089	(0.067)	0.033	(0.030)
		Bias-adjusted	-0.039	(0.066)	0.015	(0.030)
		Mean difference	-0.124	(0.056)	0.036	(0.023)
		Regression	-0.036	(0.062)	0.019	(0.029)
60-95, N=7,988						
	1	Simple Matching	0.206	(0.165)	0.108	(0.070)
		Bias-adjusted	0.342	(0.163)	0.053	(0.070)
	4	Simple matching	0.298	(0.123)	0.126	(0.057)
		Bias-adjusted	0.247	(0.121)	0.051	(0.056)
		Mean difference	0.001	(0.110)	0.194	(0.048)
		Regression	0.093	(0.118)	0.044	(0.051)

Note: Robust variance. Means of the number of physician visits are: 1.53 (whole sample), 1.12 (18-30), 1.37. (30-60) and 1.87. (60-95). Means of the *number of tests* are: 0.36 (whole sample), 0.21 (18-30), 0.36 (30-60) and 0.44 (60-95).ALL

Table 7: Matching and regression estimates of the impact of ADSE on physician visits and blood and urine tests. Standard deviations for regressions are corrected for clustering. Dependent variables are counts, specification with less covariates

Age Group	**M**	**Estimator**	**Nr. Doctor Visits**		**Nr. of Tests**	
			ATT	**(St. Error)**	**ATT**	**(St. Error)**
All, N=21,908						
	1	Simple Matching	0.132	(0.100)	0.069	(0.028)
		Bias-adjusted	0.096	(0.099)	0.057	(0.028)
	4	Simple matching	0.054	(0.074)	0.061	(0.024)
		Bias-adjusted	0.015	(0.073)	0.044	(0.024)
		Mean difference	-0.129	(0.047)	0.059	(0.019)
		Regression	0.002	(0.055)	0.029	(0.022)
18-30, N=3,203						
	1	Simple Matching	0.279	(0.164)	0.072	(0.041)
		Bias-adjusted	0.235	(0.165)	0.060	(0.041)
	4	Simple matching	0.225	(0.150)	0.069	(0.034)
		Bias-adjusted	0.168	(0.150)	0.060	(0.034)
		Mean difference	0.112	(0.133)	0.029	(0.031)
		Regression	0.191	(0.148)	0.025	(0.033)
30-60, N=10,501						
	1	Simple Matching	0.021	(0.117)	0.035	(0.041)
		Bias-adjusted	-0.043	(0.117)	0.036	(0.041)
	4	Simple matching	-0.081	(0.096)	-0.003	(0.034)
		Bias-adjusted	-0.127	(0.096)	0.014	(0.034)
		Mean difference	-0.127	(0.056)	0.035	(0.023)
		Regression	-0.059	(0.068)	0.015	(0.030)
60-95, N=8,204						
	1	Simple Matching	0.184	(0.175)	0.072	(0.075)
		Bias-adjusted	0.236	(0.175)	0.009	(0.075)
	4	Simple matching	0.103	(0.134)	0.092	(0.059)
		Bias-adjusted	0.114	(0.135)	0.035	(0.059)
		Mean difference	0.011	(0.110)	0.195	(0.048)
		Regression	0.094	(0.124)	0.045	(0.052)

Note: Robust variance. Means of the number of physician visits are: 1.52 (whole sample), 1.08 (18-30), 1.38. (30-60) and 1.87. (60-95). Means of the *number of tests* are: 0.36 (whole sample), 0.20 (18-30), 0.33 (30-60) and 0.44 (60-95).NONE

Table 8: Matching and regression estimates of the impact of ADSE on physician visits and blood and urine tests. Standard deviations for regressions are corrected for clustering. Dependent variables are counts, specification with less covariates. Sample contains only one observation per family

Age Group	M	Estimator	Nr. Doctor Visits		Nr. of Tests	
			ATT	**(St. Error)**	**ATT**	**(St. Error)**
All, N=16,452						
	1	Simple Matching	0.091	(0.106)	0.075	(0.034)
		Bias-adjusted	0.123	(0.106)	0.065	(0.034)
	4	Simple matching	0.038	(0.080)	0.060	(0.032)
		Bias-adjusted	0.016	(0.079)	0.049	(0.032)
		Mean difference	-0.123 (5)	(0.052)	0.063	(0.020)
		Regression	0.026	(0.061)	0.042 (10)	(0.024)
18-30, N=2,115						
	1	Simple Matching	0.277	(0.220)	0.132	(0.048)
		Bias-adjusted	0.256	(0.219)	0.144	(0.049)
	4	Simple matching	0.192	(0.196)	0.104	(0.042)
		Bias-adjusted	0.183	(0.196)	0.104	(0.043)
		Mean difference	0.148	(0.169)	0.051	(0.039)
		Regression	0.205	(0.193)	0.043	(0.042)
30-60, N=8,032						
	1	Simple Matching	-0.016	(0.149)	0.036	(0.044)
		Bias-adjusted	-0.066	(0.150)	0.032	(0.044)
	4	Simple matching	-0.066	(0.108)	0.016	(0.039)
		Bias-adjusted	-0.121	(0.108)	0.030	(0.039)
		Mean difference	-0.129 (5)	(0.059)	0.040	(0.025)
		Regression	-0.033	(0.072)	0.032	(0.033)
60-95, N=6,305						
	1	Simple Matching	0.112	(0.183)	0.103	(0.074)
		Bias-adjusted	0.087	(0.181)	0.083	(0.074)
	4	Simple matching	0.195	(0.139)	0.080	(0.059)
		Bias-adjusted	0.185	(0.141)	0.062	(0.060)
		Mean difference	0.038	(0.121)	0.184	(0.048)
		Regression	0.136	(0.137)	0.055	(0.051)

Note: Robust variance. Means of the number of physician visits are: 1.53 (whole sample), 1.15 (18-30), 1.37 (30-60) and 1.85 (60-95). Means of the number of tests are 0.36 (whole sample), 0.21 (18-30), 0.33 (30-60) and 0.44 (60-95).

Table 9: Matching and Regression Estimates of the Impact of ADSE on the Demand for Dental Care. Dependent variable is dichotomic, Specification with less covariates

Age Group	**M**	**Estimator**	**At Least 1 Dentist Visit**	
			ATT	**(St. Error)**
All, N=19,979	1	Simple Matching	0.044	(0.019)
		Bias-adjusted	0.024	(0.019)
	4	Simple matching	0.046	(0.015)
		Bias-adjusted	0.019	(0.015)
		Mean difference	0.181	(0.011)
		Regression	0.035	(0.013)
		Probit	0.038	(0.014)
18-30, N=2,875	1	Simple Matching	0.077	(0.048)
		Bias-adjusted	0.058	(0.049)
	4	Simple matching	0.063	(0.040)
		Bias-adjusted	0.034	(0.041)
		Mean difference	0.099	(0.032)
		Regression	0.035	(0.035)
		Probit	0.036	(0.035)
30-60, N=9,715	1	Simple Matching	0.024	(0.024)
		Bias-adjusted	-0.019	(0.024)
	4	Simple matching	0.035	(0.020)
		Bias-adjusted	-0.022	(0.020)
		Mean difference	0.164	(0.014)
		Regression	0.029	(0.017)
		Probit	0.030	(0.019)
60-95, N=7,390	1	Simple Matching	0.063	(0.036)
		Bias-adjusted	0.043	(0.036)
	4	Simple matching	0.071	(0.028)
		Bias-adjusted	0.033	(0.028)
		Mean difference	0.176	(0.022)
		Regression	0.032	(0.024)
		Probit	0.035	(0.026)

Note: Robust variance. Means of the dentist visit indicator are: 0.34 (whole sample), 0.50 (18-30), 0.40 (30-60) and 0.22 (55-95).

Table 10: Matching and regression estimates of the impact of ADSE on physician and dentist visits and blood and urine tests. Standard deviations for regressions are corrected for clustering. Specification with less covariates. Unemployed sample

M	**Estimator**	**Nr. visits**	**At least 1 visit**	**Nr. tests**	**At Least 1 Test**	**At Least 1 Dentist visit**
1	Simple Matching	-0.067	-0.200	-0.200	-0.067	0.156
		(1.113)	(0.193)	(0.434)	(0.148)	(0.184)
	Bias-adjusted	0.262	-0.079	0.352	0.065	0.067
		(1.093)	(0.185)	(0.613)	(0.185)	(0.175)
4	Simple matching	-0.100	-0.200	-0.133	-0.067	0.250
		(0.975)	(0.145)	(0.367)	(0.121)	(0.140)
	Bias-adjusted	0.025	-0.028	-0.159	-0.084	0.064
		(1.001)	(0.186)	(0.483)	(0.156)	(0.164)
	Mean difference	0.232	-0.193	0.095	-0.051	0.281
		(0.835)	(0.123)	(0.264)	(0.104)	(0.123)
	Regression	0.464	-0.143	0.130	0.020	0.187
		(0.849)	(0.132)	(0.278)	(0.107)	(0.132)
	Probit	-	-0.149	-	0.014	0.202
			(0.133)		(0.103)	(0.133)

Note: Robust standard errors in round brackets. N. Obs.=878 for doctor and tests regressions and N. Obs=788 for dentist visits regressions. Sample means of the dependent variables: number of physician visits (1.30), at least one physician visit (0.52), number of tests (0.31), at least one test (0.25), at least one dentist visit (0.39).