

Moral hazard and the demand for health services: a matching estimator approach*

Pedro Pita Barros

Matilde P. Machado

Universidade Nova de Lisboa

Universidad Carlos III de Madrid

Anna Sanz de Galdeano

Universitat de Girona and IZA

May 23, 2007

Abstract

We estimate the impact of extra health insurance coverage beyond a National Health System 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. Since a priori instrumental variables are hard to find we take a different approach. We focus on the most common health insurance plan in Portugal, ADSE, which is given to all civil servants and their dependents. We argue that this insurance is exogenous, i.e., not correlated with beneficiaries' health status. This identifying assumption allows us to estimate the impact of having ADSE coverage on the demand for three different health services using a matching estimator technique. The health services used are number of visits, number of blood and urine tests, and the probability of visiting a dentist. Results show large positive effects of ADSE coverage for number of visits and tests among the young (18 to 30 years old) but only the latter is statistically significantly different

*We would like to thank Miguel Gouveia, Pierre-Carl Michaud, João Santos Silva, Marcos Vera-Hernández, and an anonymous referee for valuable comments on previous versions of this paper. This article was supported by grants SEJ2004-00670 and SEJ2004-03276 from the Spanish Ministry of Education and Science and an unrestricted educational grant awarded jointly to the Universities Carlos III de Madrid and Pompeu Fabra de Barcelona by The Merck Foundation, the philanthropic arm of Merck Co. Inc., White House Station, New Jersey, USA.

from zero. 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.

1 Introduction

The widespread usage — by health insurance companies, and governments — of co-payments, coinsurance, and deductibles as mechanisms to control health-care spending reflects the belief that the demand for health-care reacts to price. The literature, 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 relative to price. 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 the impact of insurance on the demand for health services in an experimental setting was the RAND Health Insurance Experiment (RHIE) (e.g., Manning et al., 1987, and Newhouse, 1993). The RHIE received criticisms for 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 have 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 (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

¹Meza (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 but is not proof of moral hazard.

additional insurance but not correlated with the use of health services.² Bago d’Uva and Santos Silva (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 et al. (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 that hamper any meaningful conclusions about the true impact of the private insurance on the demand for physician services.

Screening by insurance companies and supply-induced demand by health-care providers (Holly et al., 1998) are other potential sources of bias. Screening would bias the effect from insurance on utilization downwards, while supply-induced demand would bias it upwards. Our data do 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 (NHS) 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 data set in which approximately 10 percent of individuals are covered by an extra 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 dependents.⁴ ADSE beneficiaries have double coverage since they also have access to the NHS just like any other citizen.

²An 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 have been collected from October 1998 to September 1999.

⁴Civil 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.

Second, we use a matching estimator technique (Abadie and Imbens, 2006a) that does not impose any functional form on the impact of health insurance on the demand for health services and allows for heterogeneous impacts. The control group is composed of individuals covered by the NHS alone, and the treated group is composed of individuals who are also covered by ADSE.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 homogeneous 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 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 homogeneous measure; in particular its quality is independent of insurance status. It is possible, however, that ADSE doctors use the number of diagnostic tests to justify more visits, so this measure could be subject to demand-inducement. In contrast, it is less likely that “at least one visit to the dentist in 12 months” would 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 to control for bias that may arise if the effects from better treatment accumulate over time. If, for example, ADSE beneficiaries have better (unobserved) health because they have had access to better treatment for long periods of time, we would expect health differences between the treated and control groups to increase with age. While the young individuals with access to additional insurance may not have had the time to accumulate more health benefits, it is plausible that among the older generation the treated are healthier than their NHS-only counterparts. By reporting the results for different age groups we are also trying to

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

separate the “immediate effect” of treatment from its cumulative effect. The presence of a dynamic effect and the effort to distinguish between short-term and long-term effects of treatments is, as far as we know, innovative in this literature. We believe this estimation strategy may also 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 but is not statistically significantly different from zero; and 0.057 for number of tests, which corresponds to 15.8 percent of the average number of tests, and 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 ATTs for both number of visits and number of tests for the youngest group are large, representing 21.8 and 30 percent of the youngest group average number of visits and tests, respectively. However, only the latter is close to being statistically significantly different from zero. The effects for other age groups are smaller and not statistically significantly different from zero, which may constitute evidence that ADSE beneficiaries accumulate health benefits over time. For at least one visit 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 nonmonetary 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 double coverage 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 primarily reflect moral hazard (and/or supply-induced demand) and not adverse selection because it is unlikely that these individuals 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-care system; Section 3 describes the data set; Section 4 makes an exploratory analysis of the data; Section 5 describes the matching estimator methodology; Section 6 describes the main results; Section 7 discusses

the quality of the matching and the plausibility of exogeneity of the ADSE status; and Section 8 concludes. The Appendix contains tables with results.

2 Background

2.1 Review of the literature

In most settings, ignoring the role of adverse selection in the decision to obtain private or additional insurance 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 used to estimate the impact of insurance on health-care utilization varies substantially in the literature, depending both on 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 associated more 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, Bago d’Uva and Santos Silva (2002) argue that the appropriate instruments should be (unobserved) variables related to risk aversion which are absent from most data sets. Because appropriate instrumental variables are hard to find, Chiappori et al. (1998) avoid using them by relying on a natural experiment in France. 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

Two other potential sources of endogeneity of insurance status have received little attention in the literature, perhaps because their perceived effects are small. The first

⁶We recently became aware of simultaneous work by Trujillo et al. (2005), in which 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.

one is screening by private insurance companies. Screening consists of denying insurance coverage to the highest risk individuals. 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 a fee-for-service to doctors or hospitals. If insurance choice is affected only 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, most likely, be biased upwards.⁷ The estimates of price elasticities in a context with high incentives to induce demand in Van de Voorde et al. (2001) 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 because of both 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 than for non-heads-of-household. For heads-of-household, presumably those who make the insurance decision, there is evidence of adverse selection but no evidence of moral hazard. In contrast, for other household members, there is evidence of moral hazard. More recently, Olivella and Vera (2005) and Gardiol et al. (2005) show significant evidence of adverse selection using data from the British household panel survey and Swiss health insurance claims, respectively. 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

⁷Presumably doctors/hospitals induce more demand from those patients who 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.

for Switzerland, find evidence of both adverse selection and moral hazard in hospital stays. Deb and Trivedi (2002) use data from the RHIE where the insurance choice is exogenous. They find that, everything else held constant, an increase in the coinsurance rate decreases utilization, although this effect is only statistically significantly different from zero at the 10 percent level. Finally, Chiappori et al. (1998) find evidence of moral hazard for general practitioner (GP) home visits but not for office visits to a GP or to a specialist. The authors argue that the presence of high nonmonetary costs associated with office visits translates a small change in price into a negligible change in the total cost borne by the patient. On the contrary, for GP home visits, the nonmonetary costs are virtually inexistent, making a small change in price more noticeable.

Most studies control for an 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, however, are likely to be endogenous, i.e., correlated with the unobserved variables in the health demand equation. Windmeijer and Santos Silva (1997) suggest using 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 those 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

The Portuguese National Health Service (NHS) was instituted in 1979 when legislation established 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.⁸

Before 1979, the state only covered the costs of health care for civil servants, and provided limited preventive care, maternal and child health care, and mental health treatment, as well as some control of infectious diseases. Therefore, the evolution of the health system in Portugal implies that the elder cohorts of ADSE beneficiaries received relatively

⁸This Section will borrow heavily from Bentes et al. (2004), which describes at length the Portuguese health system from its beginning until 2002.

better access to health care 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 pp. 21-24). 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 through which the beneficiaries enjoy reduced copayments. Second, the beneficiaries may decide to choose a provider outside the ADSE agreement network in which case they would 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 better 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 supplemented with money from the government budget.

After 1992, hospitals and health centers started requiring small copayments from NHS users. However, a substantial fraction of NHS users are exempt.⁹ Copayments and spending on private doctor visits are partially deductible from income taxes (at a 30% rate), which means they are a de facto subsidy by the state. Copayments in the NHS are ho-

⁹NHS users exempt from copayments are people with low incomes, the unemployed, the chronically ill, pregnant women, children up to 12 years old, drug abusers under treatment, and the mentally ill. Current estimates put their numbers somewhere in the interval 40-50% of the population.

mogeneous 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.¹⁰

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) report that, according to the 1995-1996 Health Survey, 92 percent of all dentist visits are private. For ADSE beneficiaries, the coverage for private dentists with no ADSE agreement is 80 percent up to a limit, and the copayment for dentists with ADSE agreements depends on the specific treatment but varies between 0.95 and 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 to fill out 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 number of services provided through the private sector has been growing steadily and in 1998-99 accounted for about 32% of the specialist visits, while the public sector accounted for a wide majority of the GP visits (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 data set

The data set covers 48,606 individuals belonging to 18,186 families and 17,491 households. The sample is selected following the multistage design of the Portuguese census: geographical areas are drawn randomly and in several stages with probability proportional to their population. Within each of the smallest geographical area levels, containing

¹⁰Source: http://www.arsc.online.pt/scripts/cv.dll?sec=sns&pass=guia_utente#15. Law published in 1992 (Portaria n^o338/92, 11 April). The more recent legal document is Portaria n.º 395-A/ 2007, 30 March 2007.

around 300 households, random draws select those households to be interviewed.

The data set includes information on sociodemographic 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 the health insurance status of each individual. Most individuals are only beneficiaries of the public NHS (84%), followed by civil servants and their dependents who are beneficiaries of ADSE (9.99%), those with private insurance (1.7%), and those who are linked to the military and the police who have up to five different insurance schemes.¹¹ 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; and (3) whether the individual has visited the dentist in the last year. Importantly, a serious drawback of this data set is that there is no indication of how many of the visits or tests were done in the private sector and how many were done in the NHS.

We obtain our working sample after dropping a few observations from the data. We drop 7 observations corresponding to situations where the insurance status of the individual was not known. We drop 191 observations of pregnant women whose visits to the doctor were related to the pregnancy. We deleted observations with inconsistent answers including 40 individuals who declare not to have visited the doctor in the previous 3 months but who also report going to a private doctor, health center, or hospital. Survey information may be provided by a different person in the household; therefore, to avoid potential measurement errors or missing values, we restrict our sample to observations where the individual answers for himself, which excludes minors under the age of 15 from the sample as well as another 165 observations provided by nonfamily members. 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

¹¹The data also contain information regarding the risks covered by the insurance plan for those individuals who declare having insurance. Only 4.77% of the respondents declare having insurance, which reflects a lack of understanding of the health system.

dependents (0.29%), banking employees and their dependents (1.38%), as well as people with private insurance (1.7%). We decided to delete all these insurance types from the control group because the insurance coverage and copayments may differ from the rest of the control group, which may affect the beneficiaries' behavior, and because there may be non-observable variables correlated with the insurance type that may also affect individuals' behavior.

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 21,151 to 21,908 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 that an individual visits a GP before visiting a specialist, which 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 the number of blood and urine tests. Statistics for the number of tests are shown in Table 2. The unconditional mean 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 offers 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 greater if one looks at certain predetermined variables (see Table 4). The NHS-only group is relatively poorer and older than the ADSE group. In the ADSE group there are fewer married or widowed individuals and more single and divorced people than in the NHS-only group. The regional distribution of the two groups is also unequal in the full sample. Civil servants are more concentrated in the capital and less concentrated in the other areas of the country with the exception of Alentejo, which is the least densely populated region in the country.¹² 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 fewer 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 himself/herself is statistically the same in both insurance groups.

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 among the NHS-only group with the exception of allergies, which are more common among the ADSE beneficiaries. The ADSE group, perhaps because of the younger age, practices more exercise, drinks less wine and beer and has fewer weight problems than the NHS-only group. The percentage of smokers is, however, statistically the same in both groups. Finally, ADSE beneficiaries practice better dental hygiene than the NHS-only group.

¹²The five regional variables correspond to the five regional health administrations.

5 Methodology

Our treatment group is composed of ADSE beneficiaries and represents roughly 10 percent of our sample. The matching estimator proposed by Abadie and Imbens (2006a) is used to estimate the “average treatment effect on the treated” (ATT) i.e., the average increase in the demand for health services among ADSE beneficiaries due to their double coverage. We argue it is very unlikely that individuals want to become civil servants just 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 offered to banking sector employees and their dependents, offer much better coverage than ADSE. If our argument is true, then we may rule out the existence of adverse selection in ADSE. 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 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, and more importantly, the candidate’s expected health status in the future is not taken into consideration (e.g., genetic diseases).¹³ In short, we believe there are good arguments to discard selection bias in our ATT estimates.

Our ATT estimates may, however, be a biased estimate of moral hazard. On one hand, we may underestimate moral hazard if ADSE beneficiaries enjoy more or better treatment than NHS-only beneficiaries. For example, suppose civil servants enjoy regular check-ups at their place of work and NHS-only beneficiaries do not.¹⁴ Over the individual’s life, this better treatment would translate into a significant accumulation of health advantages by the ADSE beneficiaries relative to their NHS-only counterparts. In this case, the impact of ADSE should be larger for the younger beneficiaries who have not yet had the time to accumulate health advantages, and smaller for the older beneficiaries who would be healthier than their NHS-only matches and, therefore, demand fewer health services. In

¹³The law is written in Decreto-Lei n°. 204/98, de 11.07, article 29. It requires candidates to perform a simple health test such as a simple chest X-ray to dismiss those with tuberculosis. Tuberculosis, however, is extremely rare nowadays. Other jobs have similar or stringent requirements.

¹⁴Notice that if the availability of regular check-ups were the difference in quality then civil servant dependents, who also are ADSE beneficiaries, would not be entitled to this benefit.

order to identify this effect we estimate the ATT for each age group separately.

On the other hand, we 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 doctors, except dentists, per visit.¹⁵ However, the ADSE payments to doctors are low, so we believe the magnitude of this effect, if positive, would be relatively small.

Finally, the positive impact of ADSE on the demand for health services may not be due to moral hazard but may be the consequence of undertreatment in the NHS, for example due to the existence of capacity constraints. We discard this interpretation since the Portuguese NHS has (and had already at the time of the survey) agreements with private providers to deliver those services which the NHS is incapable of delivering.

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, to our knowledge, and have the advantage of avoiding the imposition of functional form restrictions.¹⁶ In particular, we apply the matching estimator proposed by Abadie and Imbens (2006a) (hereafter “AI”). While AI focus on “covariate matching,” many recent studies have relied on propensity score matching. Baser (2006) offers a particularly useful guide in any empirical application to choose amongst the different propensity score matching estimators. The main reason for not using propensity score matching in our paper is the lack of a variance estimator when the propensity score is estimated rather than known, which explains why researchers very often use bootstrapping methods. Abadie and Imbens (2006b), however, show that the standard bootstrap is generally invalid and provide, for the time being, the only consistent and formally justified variance estimator. Further-

¹⁵Doctors may request more tests in order to justify more visits. ADSE pays dentists by type of procedure and not per visit. Inducement by dentists would then be reflected in a misreport of the type of services provided to the patient and not in the number of visits to the dentist.

¹⁶Both 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.

more, in order to validate our approach and guarantee some robustness to our exercise we report the following: 1) different matching estimators, in particular we show results based on simple matching and biased adjusted matching for one and four matches, $M=1$ and 4, respectively; 2) results for two different specifications: the first one includes a large set of controls, some of which may not be exogenous, and a second one which restricts the controls to a subset, we believe, is exogenous. The estimated treatment effect on the treated is larger with the first specification. In our conclusions, however, we put more emphasis on the more conservative results obtained with the second specification; 3) We provide statistics of the matching quality and of the overlapping distribution of the propensity score between treated and control individuals, both of which reassure us about our results.

The rest of this section follows closely AI, including their notation, assumptions and some terminology. We would like to estimate the average effect of treatment, i.e. having ADSE health insurance coverage, on several measures of health-care utilization: physician visits, number of tests and dentist visits within different time periods. Denote by $Y_i(0)$ the outcome obtained by individual i , $i = 1, \dots, N$, if under the control group (i.e., NHS-only), and $Y_i(1)$ the outcome individual i would obtain if under the treatment group (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, $W_i \in \{0, 1\}$ reflects whether individual i received treatment or not, and Y_i denotes the realized outcome:

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

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 observed only for treated units, and the untreated outcome, $Y(0)$, is observed only for comparison units. Hence, only one of the potential outcomes is observed for each individual and the other is unobserved.

We are interested in what AI denote the “population average treatment effect on the treated” (ATT) $(\tau^{p,t})$ and in the “sample average treatment effect on the treated” $(\tau^{s,t})$:

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

where $N_1 = \sum_{i=1}^N W_i$ stands for the number individuals in the treated group. 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 .*

As stated in AI, assumption 1 is referred to as strong ignorability (Rosenbaum and Rubin, 1983). The overlap assumption 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. If (ii) does not hold, all individuals with a given covariate pattern would receive the same treatment and there would be no similar individuals of the opposite treatment group. Our estimated propensity score is actually bounded away from zero and one when considering the whole sample and also when we split the sample by age groups. Figure 1 in Appendix B shows the estimated propensity score for all treated and control units used in the matches. Immediately, we see that the last bin for controls is empty, which may indicate a problem. Note, however, that this is no longer the case when the propensity score is estimated separately by age groups, as shown in Figures 2, 3, and 4, also in Appendix B. These figures suggest that, when carrying out our analyses separately by age groups, the degree of overlap is enough to proceed with the estimation.

The unconfoundedness assumption, also known as “selection on observables,” implies:

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

This assumption is crucial because it allows to use the realized outcome of individuals with the same covariates values from the opposite group. 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 .

We believe that the unconfoundedness assumption is reasonable in our context for the youngest age group but may be violated for the eldest group due to unobserved health benefits that may have accumulated over time for the treatment group. This difference motivates the separate estimation of the ATTs by age-group. In this respect, however, it is worth remarking that the NHS is available to everyone almost for free and is of relatively high quality. Furthermore, as argued in the beginning of the Section, it is unlikely that individuals anticipating a high health-care demand choose to become civil servants in order to obtain ADSE coverage, just as it is unlikely that the state selects employees on the basis of their current or future health status.

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, in terms of covariates, to individual i based on the distance measure by the norm $\|\cdot\|$, among the individuals in the opposite treatment group. Following AI, $j_m(i)$ is defined as the index j that solves:

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

where $1 \{ \cdot \}$ is the indicator function.¹⁷ We will do matching with replacement, i.e. we allow each individual in the control group to be used in more than one match since this technique produces matches of higher quality than matching without replacement by increasing the set of possible matches.

The simple matching estimator for the ATT estimates the missing potential outcomes $Y(0)$ when $W_i = 1$ as the average of the outcomes of the nearest neighbors of control group:

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

¹⁷These definitions can easily be generalized to allow for the presence of ties.

where $I_M(i)$ denotes the set of indices for the first M matches for individual i . Hence, the simple matching estimator for the average treatment effect for the treated discussed in AI is:

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

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. They propose to combine the matching process with a regression adjustment in order to adjust the differences within the matches for the differences in their covariate values. The adjustment is based on an estimate of the regression function $\mu_w(x) \equiv E[Y(w)|X = x]$ for the control group $W = 0$ since we are interested in estimating the ATT.¹⁸ Given the estimated regression function for the controls, the missing potential outcomes are predicted as:

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

The bias-corrected matching estimator of the ATT is then written as:

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

This bias adjustment makes matching estimators $N^{1/2}$ -consistent.¹⁹ Our application suggests that the bias obtained with the simple matching estimator may be large.

6 Results

In this Section we show estimates for the ATT using different specifications.²⁰

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

¹⁸AI use nonparametric estimation to impute the value for the untreated.

¹⁹For details on the properties of the matching estimators and their variance, see AI.

²⁰Results were obtained from running different versions of the Abadie and Imbens Matlab programs provided on their web pages.

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 11 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 diseases such as asthma, diabetes, bronchitis, allergies, high blood pressure, and back pain, living habits such as exercise, smoking, and 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 believe are exogenous in this context.²¹ 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 literature typically includes them as controls for unobserved conditions.²² First, perceived health may be endogenous (Windmeijer and Santos Silva, 1997). A shock affecting the number of visits, for example, 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 than the NHS. 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 intermediate outcomes may bias the impact of ADSE upwards, as can be verified from the comparison of results in tables 6 and 7.

The first segments of Tables 6 and 7 show the results obtained with the whole sample ($N = 21,151$, $N_1 = 2,251$ and $N = 21,908$, $N_1 = 2,326$, respectively where N_1 refers

²¹We excluded diabetes from matching for the eldest cohort because it may be the outcome of bad eating habits and, therefore, related to the quality of health services received. In contrast, diabetes is more likely to be a genetic condition for the youngest cohorts. In previous versions of the paper, we estimated the ATT leaving out only subjective health measures, which led to higher values than the ones obtained in Table 7. We may regard the ATT estimates in table 7, therefore, as the most conservative estimates.

²²One of the problems with perceived measures of health identified in the literature is their sensitivity to the order of questions 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.

to treated individuals). 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 increasing the number of matches 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 an OLS regression.²³ The three bottom segments in Tables 6 and 7 show the estimated ATT obtained after splitting the sample into three age groups.²⁴

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 is the case for 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 a GP before the first visit to a specialist, artificially increases the number of visits for the NHS-only group; second, higher-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. To isolate the latter bias

²³Reported 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.

²⁴The 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 $\|z - x\|$ as 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.

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 highly 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 represents only 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 covariates 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 errors. 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 are 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 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 re-estimated the ATTs from Table 7 by 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.²⁵

Finally, we also estimate the ATT for the variable at least one visit to the dentist in the 12 months prior to the interview. Since ADSE beneficiaries have a much higher coverage than the NHS-only group for dental services, 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 for the age group subsamples controlling for the smallest set of regressors. Results show a positive effect for all the samples but the middle-age group, but none of the estimates are statistically significantly different from zero.²⁶ Chiappori

²⁵Due 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 youngest 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 when all covariates are used for matching are they statistically significantly different from zero for the middle-age cohort. Similar to the results in Tables 6 and 7, the estimated ATT is larger for the youngest cohort.

²⁶Relative to the visits to the doctor, an anonymous referee wondered if we were calling moral hazard to an effect which in part could be the result of “undertreatment”. Although dentist visits are not prohibitively costly in themselves, some treatments may be, in which case the patient may decide not to undergo dental treatment and, therefore, stop visiting the dentist. There are two reasons why we think there is not an undertreatment effect for the variable chosen. 1) If the NHS-only individuals were subject to undertreatment in dental services due to financial constraints, then our estimated ATT effect should overestimate the moral hazard effect. Yet, given the small and statistically insignificant ATTs we obtain for dental services, it seems unlikely that there is undertreatment for the NHS-only individuals. 2) The variable chosen is a dummy variable for at least one visit to the dentist in the last year. One annual visit to the dentist is perfectly affordable (specially given that we are matching individuals within the same education, and profession, and income group (and there are 10 income groups and a no-reply group). The undertreatment argument would be more plausible if we were using *number* of dentist visits.

et al. (1998) argue that when nonmonetary costs are large the demand is more inelastic and this may well be the explanation for the small or inexistent effect of ADSE in dental care. Alternatively, 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 Robustness checks

In this Section we intend to look deeper at two important issues. First, given the high standard deviations of some of our estimates, it is important to check the quality of the matching, i.e., whether individuals in the treatment and control group are really alike. Second, we would like to present some evidence of the exogeneity of ADSE which is our main identifying assumption.

7.1 The quality of matching

To establish the quality of matched pairs used in our estimation we follow the same strategy as Abadie and Imbens (2006a) in their Table 3. For brevity, we present in Table 10 evidence of the quality of matching only for the variables used in the more conservative estimations from our Table 7.

First, all the covariates were normalized to have mean zero and variance equal to one. The first two columns in Table 10 show the average for the ADSE group and the NHS-only group before matching. The difference between the first and the second column is reported in the third column to facilitate the reading of the table. The fourth and fifth columns represent the average of the covariates for the ADSE and NHS-only groups, respectively, computed with those observations used in the single matching case ($M = 1$). The sixth and seventh columns represent the average difference within the matched pairs for each covariate and its standard deviation.

The matching is not perfect. However, for most variables the difference in averages between the treatment and the control group is much smaller after the matching than before the matching (compare columns 3 and 6). In fact, only for three of the covariates (“age,” “female,” and “centro”) is the difference between treated and control averages higher after the matching. The large size of the difference in averages after matching

for the age variable indicates that matching for the overall sample is not of high quality and is consistent with the improvement on the matching quality found when splitting the sample by age group. For a few of the indicator variables the matching is even exact (i.e. “student,” “clerks,” “skilled agricultural worker,” “income C,” “income D”) while for the remaining covariates the average difference within the matched pairs is close to zero and never statistically different from zero.²⁷

7.2 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 unconfoundedness 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. If this holds, then those individuals who have been ADSE beneficiaries for a longer period of time would visit the doctor less and demand fewer 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

²⁷We have also produced the same table for the matching by age group. An expected improvement relative to Table 10 is that the difference in age after the matching is always much smaller than before the matching. Results are available from the author upon request.

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 unconfoundedness 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 run OLS regressions of risky lifestyle 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 lifestyle behaviors were identical in both groups. These regressions do not offer clear evidence that ADSE beneficiaries are more risk-averse than NHS-only beneficiaries.

Some people, however, may argue that consumption of alcohol and cigarettes do not reflect risk aversion but additive behavior. For that reason, we run the same regressions on two additional measures that we believe reflect risk aversion. The first additional measure reflects preventive behavior and is drawn from the reason for the last visit to the doctor. We say an individual exerts prevention ($prev=1$) if the reason for his last visit to the doctor was one of the following: 1) routine visit; 2) visit related to occupational health often also known as occupational safety and health; 3) blood pressure check not related to any illness. Any other reason for the last visit to the doctor is not classified as prevention ($prev=0$). Moreover, if the individual did not visit any doctor in the previous 3 months he is classified as not exerting prevention ($prev=0$). The second additional measure is whether the individual brushes his/her teeth (conditional on having teeth). Being a beneficiary of ADSE had not a coefficient statistically significantly different from zero in any of the two additional regressions.²⁸

To support the first and the second identification assumptions, we focus on a subsample of ADSE beneficiaries who 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 data set does not allow us to identify whether

²⁸Results are available from the authors upon request.

individuals are covered by ADSE in their own name or through a family member. Hence, we 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 his sample between heads-of-households (in principle, those who make the decision of contracting private insurance) and non-heads-of-households (beneficiaries of private insurance who do not make the contract decisions) in order to test for adverse selection in the contract of private insurance.

Table 11 shows the estimated ATT for number of visits to the doctor, number of tests, and at least one visit to the dentist in the previous 12 months 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, average number of tests, and probability of a dentist visit respectively, although none of these estimates are 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 copayments in NHS but are subject to copayments under ADSE.

In conclusion, due to the low number of observations, it is hard to perform a conclusive test on the reasonability of our identifying assumptions. Despite that, the evidence found does not contradict them, and the size of the point estimates seems 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 at least one visit to 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 data set where 10 percent of the sample enjoys

an exogenous double health insurance coverage denoted by ADSE; Second, by using a matching estimator technique (Abadie and Imbens, 2006a), which does not impose any functional form on the impact of health insurance on the demand for health services and allows for heterogeneous impacts. Third, by estimating the impact of the additional insurance on several services, particularly on blood and urine diagnostic tests. And fourth, we allow for a dynamic impact of the additional coverage by splitting the sample into different 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% of the average number of visits, 15.8% of the average number of tests, and 7% of the average probability of visiting a dentist at least once in 12 months. The effects of ADSE are particularly large for the youngest cohort, 18 – 30 years old, where they reach 21.8, 30 and 11.6% of the average number of visits, tests, and probability of visiting the dentist, respectively, for that age group. Due to the imprecision of estimates we cannot conclude that there is moral hazard in the number of visits since the average treatment effect on the treated (ATT) is not statistically different from zero. For number of tests, we do find evidence of moral hazard for the overall sample and for the youngest cohort. For the probability of at least one visit to the dentist, 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 nonmonetary costs for this type of service.

Our data do not allow us to differentiate between the effects of moral hazard and supply-induced demand. The presence of supply-induced demand would lead to an over-estimation of the moral hazard effect. Out of the three variables used to proxy demand for health services, we expect at least one visit to the dentist to be less sensitive to supply inducement since ADSE pays dentists per procedure and not per visit, and indeed we obtained the lowest ATT for this variable. However, our results also show a larger ATT for the younger cohort (18 – 30 years old) when inducement should be larger for the eldest cohort who, especially the retired, have a lower opportunity cost of time and, therefore, a more inelastic demand (e.g., van de Voorde et al., 2001).

The ATT variation by age group is also consistent with long-term positive effects

from ADSE. If ADSE’s double coverage has a positive impact on individuals health, for example because treatment is of better quality or because it allows individuals to exert more prevention²⁹, then, older generations, who have been subject to double coverage for longer periods of time, may accumulate health benefits and enjoy better unobserved health than their NHS-only counterparts. The better health of the elder ADSE beneficiaries would reduce their demand for health care relative to their NHS-only counterparts. In this case the estimated ATT for the younger cohort should be larger than the estimated ATT for the older cohorts, just as we find in our results.

References

- [1] Abadie, Alberto and Imbens, Guido (2006a): “Large Sample Properties of Matching Estimators for Average Treatment Effects,” *Econometrica*, 74 (1), pp 235-267..
- [2] Abadie, Alberto and Imbens, Guido (2006b): “On the Failure of the Bootstrap for Matching Estimators,” mimeo, Harvard University.
- [3] Arrow, Kenneth J. (1963): “Uncertainty and the Welfare Economics of Medical Care,” *American Economic Review*, 53(5), pp 485-973.
- [4] Bago d’Uva, Teresa and João M. C Santos Silva (2002): “Asymmetric Information in the Portuguese Health Insurance Market,” ISEG, Universidade Técnica de Lisboa, mimeo.
- [5] Baser, Onur (2006): “Too Much Ado about Propensity Score Models? Comparing Methods of Propensity Score Matching,” *Value in Health*, Vol 9, N° 6, pp 377-385.
- [6] 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.
- [7] Cameron, A.C. and Trivedi, P.K. (1998): “Regression Analysis of Count Data,” Cambridge University Press.
- [8] Cameron, A.C., Trivedi, P.K., Milne, Frank, and Piggott, J. (1988): “A Microeconomic Model of the Demand for Health Care and Health Insurance in Australia,” *Review of Economic Studies*, pp 85-106.
- [9] 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.
- [10] Coulson, N. Edward, Terza, Joseph V., Neslusan, Cheryl A., and Bruce Stuart (1995): “Estimating the Moral-Hazard Effect of Supplemental Medical Insurance in the Demand for Prescription Drugs by the Elderly,” *The American Economic Review*, Vol. 85, No. 2 Papers and Proceedings of the Hundredth and Seventh Annual Meeting of the American Economic Association, pp 122-126.

²⁹In fact, as we comment in Section 7, we do not find evidence supportive of more preventive behavior among ADSE beneficiaries.

- [11] Deb, Partha and Pravin K. Trivedi (2002): “The Structure of Demand for Health Care: Latent Class versus Two-Part Models,” *Journal of Health Economics*, 21, pp 601-625.
- [12] Gardiol, Lucien, Geoffard, Pierre-Yves, and Chantal Grandchamp (2005): “Separating Selection and Incentive Effects in Health Insurance” CEPR DP 5380.
- [13] 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.
- [14] 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.
- [15] Imbens, Guido (2004): “Nonparametric Estimation of Average Treatment Effects under Ecogeneity”, *Review of Economics and Statistics* 86(1), 4-29.
- [16] 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
- [17] Meza, D. (1983): “Health Insurance and the Demand for Health Care,” *Journal of Health Economics*, 2, 47-54.
- [18] Newhouse, Joseph and the Insurance Experiment Group (1993): “Free for All? Lessons from the RAND Health Insurance Experiment,” Harvard University Press, Cambridge.
- [19] Olivella, Pau and Vera-Hernandéz, Marcos (2005): “Testing for Adverse Selection in the National Health Service”, mimeo.
- [20] 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.
- [21] Rothschild, M and Stiglitz, J. (1976): “Equilibrium in Competitive Insurance Markets,” *Quarterly Journal of Economics*, 90, pp 629-649.
- [22] Santos Silva, João M. C. and Frank Windmeijer (2001): “Two-part Multiple Spell Model for Health Care Demand,” *Journal of Econometrics*, vol104 (1), pp 67-89.
- [23] 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.
- [24] 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.
- [25] 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.
- [26] Vera-Hernandéz, Marcos (1999): “Duplicate Coverage and Demand for Health Care. The case of Catalonia,” *Health Economics* vol 8 pp 579-598.

- [27] Windmeijer, Frank A. G. and João 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.

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:

$$y = \alpha + \beta ADSE + \gamma H + \varepsilon \quad (8)$$

$$H = a_0 + a_1 ADSE + u. \quad (9)$$

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 omitted variable bias and we would underestimate β , the partial effect.

Now what if the error ε and u are (negatively) 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 if we do not include H ,

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

and we would estimate the total effect of ADSE without bias.

10 Appendix B

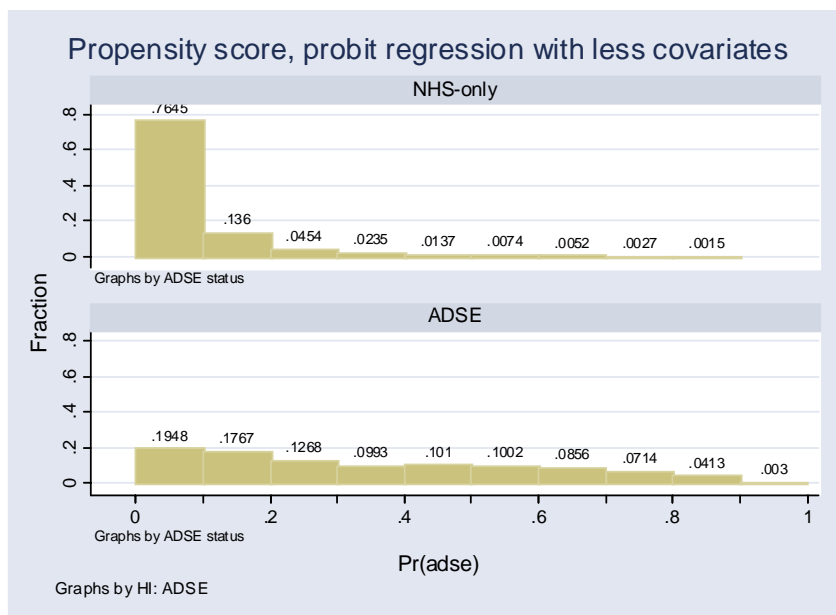


Figure 1: Estimated propensity scores from probit regressions.

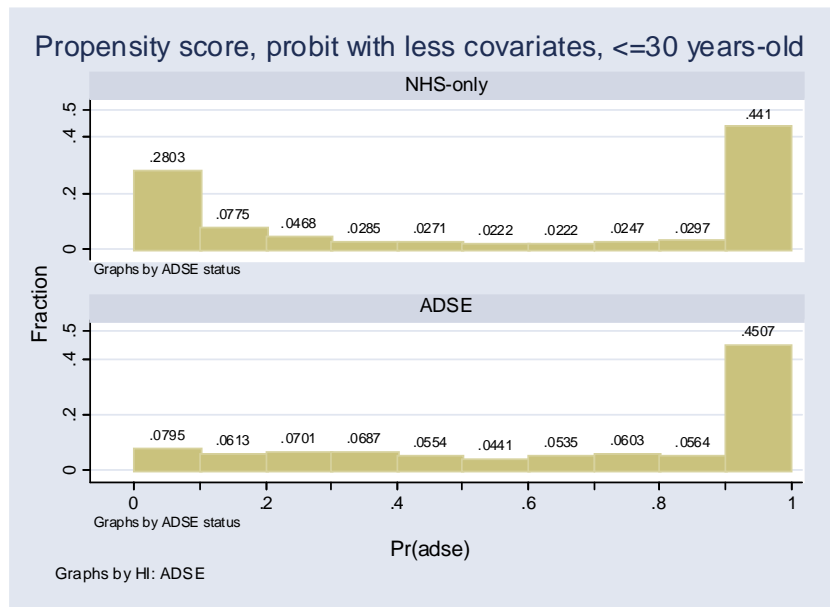


Figure 2: Estimated propensity scores from probit regressions for the young cohort

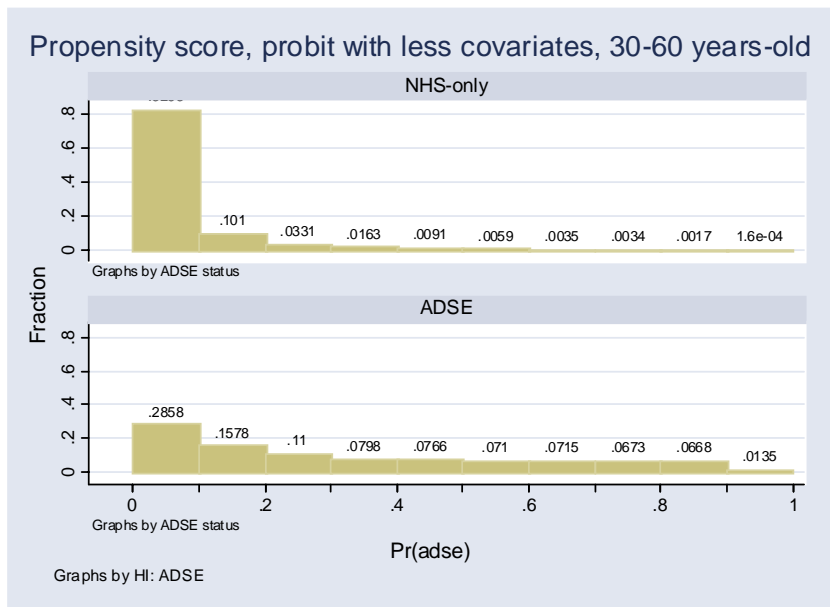


Figure 3: Estimated propensity scores from probit regressions for the middle-age cohort

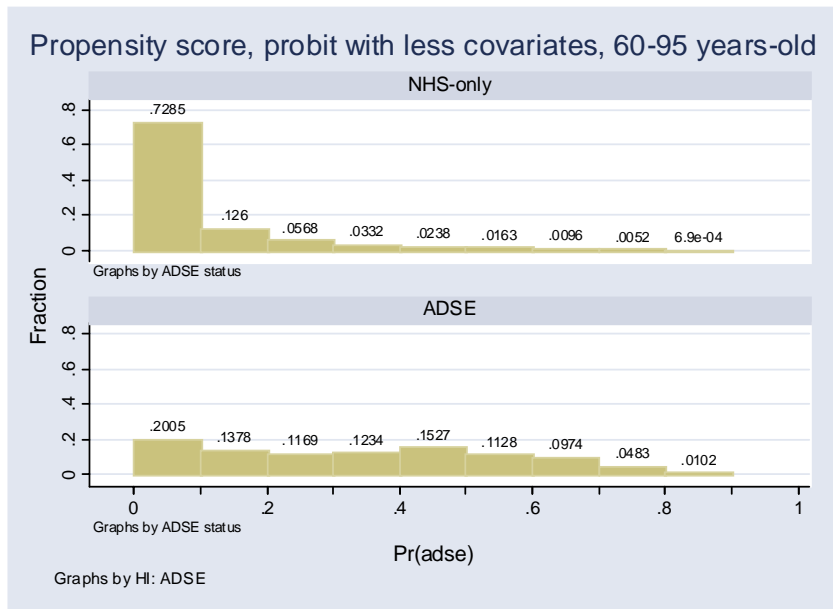


Figure 4: Estimated propensity scores from probit regressions for the eldest cohort

Table 1: Doctor visits by health insurance status. Full sample

	# 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-only	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

	# 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-only	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-only	0.35	0.48	0	1	33331
t-value	21.32				
p-value	0.0000				

Table 4: Socioeconomic characteristics by health insurance status. Full sample

Variable	ADSE			NHS			H ₀ :equal means	
	N	Mean	S.D.	N	Mean	S.D.	t-value	p-value
Age	4814	37.53	21.09	40535	42.39	23.20	14.97	0.000
Female	4814	0.558	0.497	40535	0.519	0.500	5.12	0.000
HH size	4814	3.321	1.187	40535	3.293	1.412	1.54	0.124
Respondent \neq Individual	4814	0.515	0.500	40535	0.515	0.500	0.03	0.972
Married	4814	0.511	0.500	40535	0.544	0.498	4.33	0.000
Single	4814	0.415	0.493	40535	0.350	0.477	8.72	0.000
Widow	4814	0.042	0.201	40535	0.085	0.279	13.36	0.000
Divorced/Separated	4814	0.032	0.175	40535	0.021	0.144	4.08	0.000
Norte (North)	4814	0.212	0.409	40535	0.322	0.467	17.38	0.000
Centro (Center)	4814	0.187	0.390	40535	0.202	0.401	2.49	0.013
Lisboa (Lisbon)	4814	0.312	0.463	40535	0.244	0.429	9.66	0.000
Alentejo (Alentejo)	4814	0.170	0.376	40535	0.115	0.319	9.84	0.000
Algarve (Algarve)	4814	0.119	0.324	40535	0.117	0.322	0.35	0.729
Years of schooling	4814	8.730	5.483	40499	4.858	3.935	47.57	0.000
Managers	4803	0.009	0.093	40488	0.028	0.164	12.10	0.000
Professionals	4803	0.158	0.364	40488	0.012	0.108	27.59	0.000
Technicians	4803	0.084	0.278	40488	0.024	0.155	14.62	0.000
Clerks	4803	0.077	0.267	40488	0.035	0.183	10.75	0.000
Service Workers	4803	0.038	0.191	40488	0.060	0.237	7.35	0.000
Skilled Agr. Workers	4803	0.010	0.099	40488	0.056	0.230	25.06	0.000
Craft Workers	4803	0.029	0.167	40488	0.114	0.318	29.64	0.000
Machine Operators	4803	0.026	0.159	40488	0.045	0.207	7.46	0.000
Elementary Occupations	4803	0.063	0.243	40488	0.058	0.235	1.19	0.234
Unemployed	4814	0.006	0.076	40533	0.038	0.192	22.37	0.000
Student	4814	0.252	0.434	40533	0.145	0.352	16.42	0.000
Not Working	4814	0.246	0.431	40533	0.377	0.485	19.65	0.000
IncomeA	4814	0.007	0.085	40535	0.084	0.277	41.47	0.000
IncomeB	4814	0.016	0.127	40535	0.117	0.321	41.25	0.000
IncomeC	4814	0.042	0.200	40535	0.133	0.339	27.29	0.000
IncomeD	4814	0.072	0.259	40535	0.128	0.334	13.78	0.000
IncomeE	4814	0.082	0.275	40535	0.124	0.329	9.66	0.000
IncomeF	4814	0.088	0.284	40535	0.109	0.312	4.68	0.000
IncomeG	4814	0.103	0.304	40535	0.085	0.278	4.02	0.000
IncomeH	4814	0.154	0.361	40535	0.079	0.269	14.06	0.000
IncomeI	4814	0.151	0.358	40535	0.044	0.205	20.32	0.000
IncomeJ	4814	0.234	0.423	40535	0.035	0.184	32.20	0.000
IncomeNSNR	4814	0.049	0.217	40535	0.063	0.244	4.14	0.000

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

Variable	ADSE			NHS			H ₀ : equal means	
	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	0.000
Good health	3207	0.488	0.500	25400	0.355	0.479	14.21	0.000
Normal health	3207	0.339	0.473	25400	0.380	0.485	4.69	0.000
Bad health	3207	0.077	0.267	25400	0.180	0.385	19.45	0.000
Very bad health	3207	0.015	0.121	25400	0.047	0.213	12.84	0.000
Walking diff. (age>10)	4326	0.009	0.093	37051	0.027	0.162	10.97	0.000
Diabetes	4814	0.036	0.186	40493	0.056	0.230	6.85	0.000
Asthma	4811	0.046	0.210	40501	0.062	0.241	4.75	0.000
Chronic bronchitis	4811	0.020	0.140	40496	0.030	0.170	4.55	0.000
Allergy	4814	0.162	0.369	40506	0.142	0.350	3.50	0.001
High blood pressure	4806	0.124	0.330	40434	0.178	0.383	10.57	0.000
Back pain	4814	0.301	0.459	40504	0.415	0.493	16.10	0.000
Sleeping pills (age>14)	3942	0.124	0.329	34772	0.130	0.336	1.16	0.247
Smoker	4809	0.174	0.379	40512	0.176	0.380	0.23	0.820
Exercise (age>14)	3942	0.166	0.372	34784	0.083	0.276	13.49	0.000
No brushing teeth	4734	0.016	0.126	39905	0.059	0.236	19.79	0.000
No teeth	4734	0.006	0.079	39905	0.024	0.152	12.47	0.000
Obese	3630	0.102	0.303	32434	0.132	0.339	5.63	0.000
Overweight	3630	0.333	0.471	32434	0.374	0.484	5.04	0.000
Normal weight	3630	0.534	0.499	32434	0.472	0.499	7.12	0.000
Underweight	3630	0.032	0.175	32434	0.022	0.147	3.21	0.001
Wine (lt.)	2039	0.211	0.332	17518	0.264	0.331	6.78	0.000
Beer (lt.)	1483	0.201	0.381	11634	0.261	0.404	5.61	0.000
Bagaço (lt.)	388	0.016	0.062	3872	0.018	0.048	0.68	0.500
Whisky (lt.)	843	0.018	0.044	5608	0.018	0.051	0.29	0.773

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	# Doctor Visits		# of Tests	
			ATT	(St. Error)	ATT	(St. Error)
All, N=21,151						
N ₁ =2,251	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						
N ₁ =252	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						
N ₁ =1,445	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						
N ₁ =554	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).

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 fewer covariates

Age Group	M	Estimator	# Doctor Visits		# of Tests	
			ATT	(St. Error)	ATT	(St. Error)
All, N=21,908						
N ₁ =2,326	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						
N ₁ =318	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						
N ₁ =1,449	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						
N ₁ =559	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).

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 fewer covariates. Sample contains only one observation per family

Age Group	M	Estimator	# Doctor Visits		# of Tests	
			ATT	(St. Error)	ATT	(St. Error)
All, N=16,452						
N ₁ =2,001	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						
N ₁ =235	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						
N ₁ =1,275	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						
N ₁ =491	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 fewer covariates

Age Group	M	Estimator	At Least 1 Dentist Visit	
			ATT	(St. Error)
All, N=19,979 N ₁ =2,232	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 N ₁ =300	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 N ₁ =1,399	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 N ₁ =533	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: Mean covariate differences in matched pairs

Variable	Before matching			After Matching (M=1)			
	ADSE Mean	NHS-only Mean	Diff. Mean	ADSE Mean	NHS-only Mean	Diff. Mean	S.D.
Age	-0.212	0.025	-0.237	-0.181	1.273	1.454	2.026
Female	-0.006	0.001	-0.006	0.145	0.131	-0.014	3.012
HH size	0.051	-0.006	0.058	-0.097	-0.147	-0.050	0.460
Single	0.127	-0.015	0.142	-0.211	-0.344	-0.133	0.621
Widow	-0.169	0.020	-0.190	-0.241	-0.363	-0.122	0.597
Divorced/Separated	0.114	-0.014	0.127	-0.006	-0.050	-0.045	1.038
Norte	-0.198	0.023	-0.221	0.127	0.021	-0.106	0.767
Centro	-0.063	0.008	-0.071	-0.169	-0.251	-0.082	0.507
Lisboa	0.205	-0.024	0.229	0.113	0.092	-0.022	0.349
Alentejo	0.089	-0.011	0.099	-0.198	-0.126	0.073	0.840
Years of schooling	1.098	-0.130	1.228	-0.031	-0.041	-0.010	0.249
Unemployed	-0.171	0.020	-0.192	-0.217	-0.225	-0.008	0.194
Student	0.190	-0.023	0.213	-0.065	-0.065	0	0
Managers	-0.111	0.013	-0.124	0.016	-0.013	-0.029	0.393
Professionals	0.940	-0.112	1.052	1.096	0.862	-0.234	0.719
Technicians	0.509	-0.060	0.570	1.161	0.891	-0.271	0.921
Clerks	0.266	-0.032	0.298	-0.111	-0.111	0	0
Service workers	-0.053	0.006	-0.059	0.938	0.931	-0.007	0.201
Skilled agricultural	-0.228	0.027	-0.255	0.508	0.508	0	0
Craft workers	-0.217	0.026	-0.242	0.265	0.263	-0.0021	0.176
Machine operators	-0.064	0.008	-0.072	-0.053	-0.055	-0.002	0.147
Elementary occup.	0.058	-0.007	0.065	-0.228	-0.233	-0.005	0.145
Income B	-0.326	0.039	-0.365	0.059	0.047	-0.012	0.219
Income C	-0.260	0.031	-0.291	-0.172	-0.172	0	0
Income D	-0.129	0.015	-0.145	0.192	0.192	0	0
Income E	-0.044	0.005	-0.049	-0.326	-0.329	-0.003	0.194
Income F	-0.008	0.001	-0.009	-0.259	-0.265	-0.006	0.238
Income G	0.106	-0.013	0.119	-0.129	-0.121	0.008	0.271
Income H	0.271	-0.032	0.303	-0.044	-0.065	-0.021	0.274
Income I	0.464	-0.055	0.519	-0.009	-0.012	-0.003	0.279
Income J	0.742	-0.088	0.830	0.107	0.089	-0.018	0.323
Income NS-NR	-0.026	0.003	-0.029	0.269	0.263	-0.006	0.247
Diabetes	-0.097	0.012	-0.108	0.462	0.452	-0.010	0.212
Asthma	-0.087	0.010	-0.097	0.740	0.732	-0.008	0.190
Allergies	0.035	-0.004	0.039	-0.027	-0.031	-0.004	0.134

Note: All variables have been normalized to have mean zero and variance 1. The first two columns represent the averages of the variables before matching for the treatment and control, respectively. The third column is the difference between the first and the second columns. The fourth and fifth columns represent the average of the variables for the matched units only when M=1. The sixth and seventh columns represent the average difference within the matched pairs and its standard deviation, also for M=1. We decided not to show the statistics for month of the interview because it would increase the length of the table considerably. The mean difference for months of the interview (corresponding to column 5) is between -0.1887 for March and 0.1140 for February but none is statistically significant.

Table 11: 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 fewer covariates. Unemployed sample

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

Note: Robust standard errors in parentheses. N=878 and N₁=15 for doctor and tests regressions. N=788 and N₁=15 for dentist visit 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).