HEALTH ECONOMICS
Health Econ. **21**: 883–901 (2012)
Published online 6 June 2011 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/hec.1756

ESTIMATING ADVERSE SELECTION AND MORAL HAZARD EFFECTS WITH HOSPITAL INVOICES DATA IN A GOVERNMENT-CONTROLLED HEALTHCARE SYSTEM

XIANGPING LIU[a], DANIJEL NESTIC[b] and TOMISLAV VUKINA[c,*]

[a]*Center for Environmental and Resource Economics and Policy, North Carolina State University, Raleigh, NC*
[b]*Institute for Economics, Zagreb, Croatia*
[c]*Department of Agricultural and Resource Economics, North Carolina State University, Raleigh, NC*

SUMMARY

We use invoices for hospital services from a regional hospital in Croatia to test for adverse selection and moral hazard. There are three categories of patients: with no supplemental insurance, who bought it, and who are entitled to it for free. Our identification procedure relies on the premise that the difference in the observed medical care consumption between the patients who bought the insurance and those entitled to free insurance is caused by pure selection effect, whereas the difference in healthcare consumption between the group that received the free insurance and the group that has no insurance is due to moral hazard. Results show favorable selection for patients in 20- to 30-year-old cohort and significant moral hazard for all age cohorts. The selection effect reverses its sign in older cohorts explained by the differences in risk aversion across cohorts caused by the timing of transition from socialism to market economy. Copyright © 2011 John Wiley & Sons, Ltd.

Received 17 May 2010; Revised 21 March 2011; Accepted 26 April 2011

KEY WORDS: adverse and advantageous selection; moral hazard; matching estimators

1. INTRODUCTION

The last decade, in the USA and elsewhere, has been characterized by pronounced inflation in healthcare costs. Commonly suggested explanations for this phenomenon are the distortion in the health insurance market because of asymmetric information about the latent health status of individuals that leads to adverse selection problem and runaway medical expenditures because of moral hazard. Despite being the textbook example of a market plagued with asymmetric information problems, the empirical evidence on the importance of either moral hazard or adverse selection in health insurance markets varies by country, type of health service provided, socioeconomic status of the insured, and others. For example, using Australian data, Cameron *et al.* (1998) found that more generous coverage leads to higher utilization of broad range of services because of both moral hazard and adverse selection reasons. Also for Australia, Savage and Wright (2003) found that after correcting for endogeneity, the extent of moral hazard can increase the expected length of a hospital stay by a factor of up to three. Coulson *et al.* (1995) found that supplemental insurance increases the number of prescription filled among the elderly in the USA but do not find evidence of adverse selection. In case of Spain, Vera-Hernandez (1999) found different evidence for heads of household than for their dependents. For heads of household, who presumably make the insurance decisions, there is evidence of adverse selection but no evidence of moral hazard. In contrast, for other household members, there is evidence of moral hazard. Holly *et al.* (1998), using data for Switzerland, found evidence of both moral hazard and adverse selection in hospital stays.

*Correspondence to: Department of Agricultural and Resource Economics, North Carolina State University, Raleigh, NC, 27695-8109. E-mail: tom_vukina@ncsu.edu

The main objective of this paper is to use the distinct feature of the supplemental health insurance design in Croatia to test for the presence of adverse selection and moral hazard effects and try to disentangle the two effects from each other. The healthcare system in Croatia is dominated by a single public health insurance fund: the Croatian Institute for Health Insurance (HZZO). The HZZO offers two types of insurance: the compulsory insurance and the supplemental insurance. As the name suggests, the compulsory insurance's coverage is universal, whereas the supplemental insurance can be either bought or is extended automatically free of charge to certain categories of citizens (e.g., full time students). For the group of people who bought the supplemental insurance, the asymmetric information problems consist of both moral hazard and adverse selection effects. For the group who are given the insurance free of charge, there is no selection effect, and the anticipated excess in healthcare consumption relative to the group with no supplemental insurance should be caused by the moral hazard effect alone (because it is extremely unlikely that some young people enrolled at the university for the purpose of obtaining the supplemental health insurance). Therefore, our identification procedure hinges on the premise that the difference in the observed medical care consumption between the group that bought the insurance and was given the insurance for free is due to adverse selection, whereas the difference in healthcare consumption between the group that got the insurance free of charge and the group that has no supplemental insurance at all is due to moral hazard.

Our unique data set consists of invoices for outpatient hospital services from one mid-size regional hospital in the semi-agricultural northern part of Croatia. The empirical results are obtained using matching estimators. Because the number of available conditional variables in our data is not very large, we match directly on our conditional variables (e.g., Dehejia and Wahba, 2002). The estimation approach similar to ours is Barros *et al.* (2008) who used a matching estimator technique of Abadie and Imbens (2006) to estimate the impact of extra health insurance coverage in Portugal. The major difference between the two approaches is coming from the differences in the data sets. Barros *et al.* (2008) have only two groups of patients: the control group composed of individuals covered by the national insurance alone and the treated group composed of individuals who are also covered by the government employee insurance. Their first group would roughly correspond to our group with no supplemental insurance, and the second group would correspond to our group that received the supplemental insurance for free. With these two groups of patients, they can only estimate the effect of moral hazard by matching across them, which is what we do as well. However, because we have the third group of patients that willingly bought the supplemental insurance, we can estimate the effect of adverse selection as well. Having this group of patients is essential because the members of that group selected themselves into buying the insurance based on their private information about their health status, risk aversion, and others, and comparing them against the group that has the exact same insurance but got it automatically and for free enables the estimation of the pure adverse selection effect.

Our work is also related to Fang *et al.* (2008) who provided evidence of favorable (advantageous) selection in the Medigap insurance in the USA and analyze its sources. The Croatian public healthcare system as provided by the HZZO is very similar to the Medicare system in the USA, which provides free but limited public health insurance for citizens 65 years and older. The individuals covered by Medicare can choose to purchase Medigap to cover the gaps in Medicare such as co-pays, deductibles, and uncovered expenses (e.g., prescription drugs, prolonged hospital stays, etc.). Same as Medigap, the Croatian supplemental insurance system is ideal for studying multidimensional private information problems because of the built-in restrictions on the insurance contracts that insurance companies can offer. Without these restrictions, Chiappori *et al.* (2006) showed that a suitably modified version of the positive correlation between the insurance coverage and ex-post realization of loss, characteristic of the classic equilibrium insurance model (Rothschild and Stiglitz, 1976), still holds with multidimensional private information. In Medigap, insurers can effectively price contracts only on age, gender, smoking status, and state of residence, whereas the HZZO supplemental policy, as will be discussed in detail in the next section, is priced based on the income variable alone.

Similar to Fang *et al.* (2008), we found the presence of favorable selection effect but only for the youngest patients between 20 and 30 years of age. Interestingly, the selection effect reverses its sign in the next age cohort (30–40 years) from favorable to adverse selection. The result, perhaps, can be explained by the fact that

DOI: 10.1002/hec

Croatia is a transition economy where the old socialist/communist system collapsed 20 years ago. Because the young people's entire life experience has been formed under the volatility and uncertainty of market economy, they are likely to be more risk averse than their parents and hence tend to buy more insurance. The effect wears off in older generations used to job security and free health care under the old system. We also found the presence of positive and statistically significant moral hazard effect for all age cohorts of patients. Finally, in addition to analyzing the asymmetric information effects in the general population, we also look at the behavior of farmers. Farmers are anecdotally perceived as more likely to be self-insured and also, because of arguably worse access to healthcare centers relative to urban dwellers, are more likely to be lighter users of health services. Our results show that, compared with the rest of the population, the effects of both adverse selection and moral hazard are smaller with farmers.

2. HEALTH INSURANCE IN CROATIA

The healthcare system in Croatia is still largely dominated by the institutional setup inherited from the socialist past time. The central government of Croatia plays a dual role on the market for health care. On the one hand, it acts as the purchaser and provider of health care through its influence on the public insurance fund's (HZZO) funding, and on the other hand, together with the regional and municipal governments, it is the largest owner of hospitals and public health institutions. Although the provision and funding of services are largely public, private providers and insurers also operate on the fringes of the market, and this segment of the market is steadily growing (Vukina and Nestic, 2008).

The Croatian health system has fared relatively well among the countries in the region. The system has a well-trained health workforce, a well-established system of public health programs and health delivery system, and good health outcomes in relation to countries at comparable income levels (Voncina *et al.*, 2008). However, these results have been achieved at a high cost, and the HZZO has faced growing deficits in the late 1990s. Despite the fact that the generous benefits and exemptions inherited from the old system have been politically difficult to roll back, Croatia has embarked on a number of reform initiatives that resulted in a relative decline in total spending on health, from estimated 9.2% in 1999 to 8.1% of gross domestic product estimated for 2008 (World Bank, 2008). Nevertheless, despite the reductions, the health system is still heavily burdened by arrears. According to the Croatian Ministry of Health, at the end of 2008, the accumulated debt amounted to 4.65 billion HRK ($1 = 5.19 HRK as of July 20, 2009). Solutions advocated by policy makers in Croatia heavily revolved around increasing the inflow of private funds into the system. Reforms have included enlarging the participation scheme (co-payments), reducing the number of individuals exempt from participation, the introduction of administrative fees, and some cost savings in prescription drugs expenditures. The pressure on public resources to spend more on health will intensify in the coming years and will have to be met with prudent allocation of resources and continuous efforts to improve productivity wherever possible.

The main characteristics of the system provided by the HZZO, reflective of the most recent reform launched on January 1, 2009, can be summarized as follows.

Compulsory Insurance. The compulsory health insurance, based on the principles of equality and solidarity, is mandatory for essentially all people with permanent residence in Croatia. It is funded by a combination of a 15% payroll tax (with an additional 0.5% for work-related injury insurance) and specific budgetary appropriations to cover the shortfalls. The system covers primary health care, specialists' medical care, hospital care, prescription medications, dental care, orthopedic and prosthetic devices, and the medical coverage while traveling abroad. In addition to healthcare coverage, the system also includes monetary compensation for sick leaves, for temporary disabilities, and medical care-related transportation costs. The insured people have the right to choose their primary care physician and dentist. Women also have the right to choose their gynecologist, and parents have the right to choose the pediatrician for their children. These choices are made for a minimum of 1-year period but can be easily changed. To seek specialists' medical care (inpatient or outpatient), the insured person needs a referral from his or her primary care physician (family physician,

DOI: 10.1002/hec

dentist, or gynecologist). The service will be provided at the closest facility that has an established contractual relationship with the HZZO. The referral is valid for 30 days. When it comes to prescription drugs coverage, the insured persons have access to medications from the base list and the supplemental list. The base list contains medically and economically most appropriate (mostly generic) drugs, whereas the supplemental list contains more expensive drugs. The compulsory insurance guarantees the coverage of prescription drugs up to the price of the equivalent drug from the base list.

The compulsory health insurance system distinguishes two categories of medical care services: (i) full coverage healthcare services; and (ii) co-payment-based services.

(a) The full coverage services afforded by the compulsory insurance include the following:

- full health care for children under the age of 18 years
- health care of women related to pregnancy and child birth
- preventive and curative health care related to AIDS and other infectious diseases
- mandatory vaccinations and immunizations
- hospital care for all chronic psychiatric patients
- complete treatment of all cancerous diseases
- dialysis
- organ transplants
- emergency room interventions
- house calls and at-home treatment of patients
- certain categories of ambulance transportation
- prescription drugs from the HZZO basic list

(b) The insured are required to pay 20% of the full price of medical care for the following list of services (listed co-payment amounts are valid for 2009):

- laboratory, radiological, and other diagnostics at the primary healthcare level—15.00 HRK
- specialists' visits and all outpatient services except physical therapy and rehabilitation—25.00 HRK
- specialists' diagnostics not at the primary care level—50 HRK
- orthopedic and prosthetic devices—50 HRK
- outpatient and at-home physical therapy and rehabilitation—25 HRK per day
- inpatient care—100 HRK per day
- dental care of adults 18 to 65 years of age—1000.00 HRK
- dental care of adults older than 65 years—or 500 HRK
- primary care (family, gynecologist, dentist)—15 HRK
- prescription drugs—15 HRK per subscription

The largest out-of-pocket cost-share amount that a person can pay amounts to 3000.00 HRK per one invoice.

Supplemental Insurance. This is a voluntary insurance that can be acquired by a person 18 years or older, having compulsory insurance, by signing a contract with the HZZO. Alternatively, an employer can sign a contract with the HZZO for the supplemental insurance coverage for its employees where the premium can be paid either by the company or deducted from the employees' net salary. However, the number of people who automatically received the supplemental insurance as a part of their compensation package is still likely to be very small.[1] A person having the supplemental insurance policy is entitled to full waiver of all medical expense

[1]In a private conversation with an HZZO official, we found that the number of insured for whom the supplemental insurance premiums are paid by their employers (about 2500 firms) is only about 80,000 or 6%. However, the agency cannot tell how many of these people receive this insurance coverage as a part of their compensation package and how many of them actually chose to buy the supplemental insurance themselves but pay the premiums through the payroll deduction for tax (like the FLEX accounts in the USA) or convenience purposes.

DOI: 10.1002/hec

co-payments listed in item (b) in the previous text. The premiums for the supplemental insurance are determined as follows:

- 50.00 HRK per month for a retired person with pension less than 5108.00 HRK
- 80.00 HRK per month for a retired person with pension higher than 5108.00 HRK
- 80.00 HRK per month for an active person with net income less than 5108.00 HRK
- 130 HRK per month for an active person with net income in excess of 5108.00 HRK
- 80 HRK per month for all family members and dependents.

(c) The most important feature of the HZZO supplemental insurance program is the fact that certain categories of citizens are entitled to the supplemental insurance free of charge, that is, their premiums are covered from the state budget. The list of people that have the right to free supplemental insurance and hence pay no co-payments from item (b) is quite large. Here we list only the most important ones:

- people whose per capita monthly household income does not exceed 1516.32 HRK and 1939.39 HRK for single retired individuals,
- full-time secondary school or college students,
- various categories of people with disabilities and war veterans and their family members,
- military service draftees,
- subsistence farmers 65 years and older,
- blood donors with 35 or more lifetime donations for men and 25 or more for women.[2]

The official HZZO figures for 2009 show the total number of people with supplemental insurance in Croatia is 2,735,088, of which 1,397,962 or 51% are eligible for free coverage. Among the remaining 49% who bought the insurance themselves, about 46% are active (payroll), 46% are retired, and the remaining 8% are self-employed, farmers, or unemployed.[3]

3. DATA SET

Our data consist of all invoices submitted by a midsize regional hospital to the HZZO for all outpatient services in the period between March 1 and June 30, 2009. Each observation reflects one invoice or one hospital visit. The data contain the following set of variables: a numeric code for the type of hospital service provided (there are 94 different treatments), compulsory health insurance number, supplemental insurance number (if the patient has one), period covered by the supplemental insurance, numeric code for categories entitled to supplemental insurance free of charge (exempt from paying participation fees), eligibility category for compulsory insurance (employed, self-employed, retired, farmer, etc.), cost of hospital service, part of the cost covered by compulsory insurance, part of the cost covered by supplemental insurance, part of the cost covered by participation (co-payment), and date of birth and sex of the patient.

The entire data set consists of 105,646 observations (visits). Because our objective is to compare the differences in healthcare consumption among patients with different supplemental insurance types, we initially dropped all observations pertaining to categories listed in item (a) in the previous text for which the full coverage under the compulsory insurance applies. The most important category in this group is children under the age of 18 years, but the list also includes certain illnesses and procedures that are free for the entire population regardless of age (e.g., all cancer treatments). In addition to visits from item (a), we also eliminated

[2]The share of patients eligible for free supplemental health insurance based on the blood donation criterium in our data set is only 2.7% of the total number of patients with free supplemental insurance.
[3]The percent shares by employement categories is an unofficial estimate from an HZZO source.

DOI: 10.1002/hec

Table I. Basic statistics by insurance type

	'NO' supplemental insurance	'FREE' supplemental insurance	'BOUGHT' supplemental insurance	Total/ Average
No. of visits (invoices)	4927	24,266	41,666	70,859
No. of patients	2485	7932	12,540	22,957
Visits per patient	1.98	3.06	3.32	3.09
Total cost per visit (HRK)	140.25	160.80	142.34	148.51
Co-payment per visit (HRK)	41.27	–	–	3.05
Average age (years)	40.53	54.83	53.77	52.70
Share of female patients (%)	46.8	59.0	58.4	57.4
Composition of insured persons by category (%):	100.0	100.0	100.0	100.0
Payroll	75.0	17.6	42.6	37.5
Unemployed	6.3	7.7	4.3	5.7
Retired	9.1	47.7	41.8	40.3
Farmer	4.6	22.5	8.0	12.7
Welfare recipient	0.8	3.3	1.3	1.9
Self-employed	4.2	1.1	2.0	1.9

the invoices for patients where we could not clearly discern the insurance type because of the coding ambiguities (9655 such invoices). The remaining 70,859 invoices are used for matching estimation purposes. In addition, we will also use the invoices covered in full by the compulsory insurance based on diagnosis to test for the differences in latent health status among those eligible for free coverage and the rest of the patients' population. However, only portion of these invoices (3711) are actually usable because for others, we cannot discern their insurance coverage type.[4]

The breakdown of the data set by the type of insurance is presented in Table I. There are 22,957 patients in the data set, 57.4% of them are women, they are on average 52.7 years old, and a large share of them are retired —40.3%, followed by active (payroll)—37.5%, and farmers—12.7%. The fact that we are dealing with an older population is caused by the fact that patients under 18 years of age were dropped from the sample for reasons explained before. There are only 2485 patients (10.8%) with no supplemental insurance, the rest have supplemental insurance either because they qualified for free access (34.6%) or they bought the policy (54.6%).[5] When it comes to farmers, only 3.9% of them have no supplemental insurance compared with the general population of 10.8%. Also, more of them are eligible for free supplemental insurance—61.5%, relative to the general population where only 34.6% qualify.

By simply eyeballing the data, one can clearly see that the number of visits per patient is larger in the group of patients with supplemental insurance (more than three visits) than in the group of patients with no supplemental insurance (two visits). Also, total cost per visit is higher for the group who have the supplemental insurance, but the difference is quite negligible when it comes to comparison of people that bought the supplemental insurance and those who do not have it. The difference is substantially larger (about 20 HRK per visit) when one compares the group with free access to supplemental insurance with the group who does not have the supplemental insurance at all.

Next, we look at some specific services provided by the hospital. The most frequently used hospital services by the total number of visit are listed in Table II. It is interesting to note that only eight categories of services

[4]These are the patients who, in addition to seeking medical attention for fully covered diagnoses under compulsory coverage, also sought medical attention for services subject to co-payments, and hence, for the most part, their insurance status has been deciphered. The problem is that once the hospital knows that the case is covered in full based on the diagnosis, it frequently does not care asking about the supplemental insurance coverage, so people who never showed up for any other treatment typically do not have their supplemetal insurance status recorded.

[5]The share of those eligible for free supplemental insurance who are actually enrolled is 100%. This is because our data include only users of hospital services, hence everybody eligible is automatically enrolled because the hospital can check his or her eligibility status.

DOI: 10.1002/hec

Table II. Most frequently provided hospital services

	No. of visits	Proportion of total visits (%)
Outpatient physical therapy	13,634	19.2
Medical biochemistry	5832	8.2
Radiology—classical and contrast exams	4003	5.6
Ophthalmology	3832	5.4
Dermatology and STDs	3399	4.8
ORL	2907	4.1
Ultrasound	2529	3.6
Physical medicine and rehabilitation	2408	3.4
Total (selected services)	38,544	54.3
Total visits (overall sample)	70,859	100.0

account for well over half of the total visits during the period covered by the data. The largest number of visits falls into the category of physical therapy (19.2% of visits). The percentage would be even larger if combined with a similar type of service (physical medicine and rehabilitation—3.4%). Another interesting feature of the most frequently provided hospital services is that they all seemed to be largely non-life-threatening diagnoses, which made them desirable for the purposes of this study. Because we are interested to estimate the impact of health insurance on the level of consumption of medical care, preferably we would like to deal with illnesses and conditions where the decision regarding whether to see a doctor or not is not a matter of life and death. The basic statistics for the subsample of the top eight hospital services are presented in Table III.

The comparison of raw statistics in Tables I and III reveals several results that are worth mentioning. First, the demographic characteristics of the patients (age, female representation, occupation) are very similar in the whole sample and in the top eight non-life-threatening diagnoses. However, the total number of visits and the number of visits per patient within the category of patients with NO supplemental insurance is smaller for the top eight diagnoses than in the entire sample. In the top eight services, people with NO supplemental insurance made 6.5% of the total number of visits, whereas in the entire sample, this number is 7%. The number of visits per patient in the top eight group is 1.85, and for all services combined, this number is 1.98. These results make perfect sense indicating that when it comes to non-life-threatening conditions, people are more cost conscious, and those who are not fully covered by the insurance may economize with doctors' visits. Finally, the services provided under the top eight category are on average cheaper than other services. The total cost per visit in the entire sample is 140.25 HRK, whereas in the top eight category, it is only 111.46 HRK.

Table III. Basic statistics by insurance type for the top eight hospital services

	'NO' supplemental insurance	'FREE' supplemental insurance	'BOUGHT' supplemental insurance	Total/ Average
No. of visits (invoices)	2488	12,258	23,798	38,544
No. of patients	1343	4626	7938	13,907
Visits per patient	1.85	2.65	3.00	2.77
Total cost per visit(HRK)	111.46	121.44	120.57	119.98
Participation per visit (HRK)	34.95	–	–	3.05
Average age (years)	41.32	56.25	54.91	54.04
Share of female patients (%)	53.31	59.30	58.60	58.32
Composition of insured persons by category (%):	100.00	100.00	100.00	100.00
Payroll	77.36	16.93	42.29	37.24
Unemployed	3.20	21.03	6.64	11.10
Retired	9.01	50.97	43.78	42.81
Farmer	5.73	7.22	4.02	5.25
Welfare recipient	0.74	2.79	1.02	1.58
Self-employed	3.95	1.06	2.25	2.02

DOI: 10.1002/hec

4. IDENTIFICATION: CONCEPTUAL FRAMEWORK

Imagine a hypothetical experiment where we can randomly divide the population into two groups: group X and group Y, such that there are no systematic differences between them. We award insurance to group X free of charge and offer people in group Y the chance to purchase the insurance if they choose to do so. Depending on their health status, longevity expectations, risk attitudes, cognitive ability, and others, some people will buy the insurance (call it Y1) and some people will not (call it Y2). The people with free insurance (group X) and the people who bought the insurance (group Y1) enjoy precisely the same coverage in all respects. Obviously, people in group X are not affected by adverse selection because they obtained this coverage automatically and not by choice. Their behavior is only impacted by moral hazard (having the insurance changed their behavior relative to when they were uninsured). Contrary to this, the behavior of people in the group who actually bought the insurance (Y1) has been impacted by both adverse selection effect (they bought the insurance because of some private information) and moral hazard. In this kind of experiment, we can identify the 'pure' adverse selection effect but only the 'jammed' moral hazard effect. The clean adverse selection is obtained by comparing the consumption of healthcare services in group Y1 (who were given the choice to buy the insurance and took it) with their identical counterparts in group X who were assigned the insurance for free. The jammed moral hazard effect is obtained by comparing the consumption of healthcare services by group X with their observably identical counterparts in group Y2 who chose not to buy insurance.

To illustrate, we can write a person's medical expenditure (or consumption) M_i as a function of his or her insurance status I_i, unobserved fixed effects as a measure of poor health status h_i, and some noise ε_i with mean zero:

$$M_i = \alpha I_i + h_i + \varepsilon_i.$$

An individual can obtain the insurance in which case $I_i = 1$ either because he or she is eligible to receive it for free or chooses to buy it ($B_i = 1$). Obviously, in this specification, α captures the effect of moral hazard.[6] Now, the average medical expenditure difference between groups Y1 who chose to buy the insurance and Y2 who chose not to buy can be expressed as follows:

$$E[M|B_i = 1] - E[M|B_i = 0] = \alpha + \{E[h_i|B_i = 1] - E[h_i|B_i = 0]\}. \quad (1)$$

The average medical expenditure difference between group X who is entitled to free insurance and Y2 who chose not to buy the insurance is as follows:

$$E[M|\mathbf{free}] - E[M|B_i = 0] = \alpha + \{E[h_i] - E[h_i|B_i = 0]\}. \quad (2)$$

This is the measure of asymmetric information that we call the 'jammed' moral hazard because it is equal to pure moral hazard α plus the difference between the mean fixed effects for the free insurance group and the group who choose not to purchase the insurance. Finally, the mean medical expenditure difference between group Y1 and group X is as follows:

$$E[M|B_i = 1] - E[M|\mathbf{free}] = E[h_i|B_i = 1] - E[h_i] \quad (3)$$

which represents the pure adverse selection effect.

Notice that even in this true natural experiment, we cannot precisely disentangle the moral hazard effect from the adverse selection effect. This is due to the fact that we essentially observe three means:

$$E[M|B_i = 1] = \alpha + E[h_i|B_i = 1]$$

$$E[M|B_i = 0] = E[h_i|B_i = 0]$$

[6]In this simple model, we ignore observable health components as they would only increase notational clutter. In models that allow for multidimensional private information, in addition to health status, h_i can also include things like risk aversion, cognitive ability, longevity expectations, and others.

DOI: 10.1002/hec

$$E[M|\mathbf{free}] = \alpha + E[h_i]$$

but on the right-hand side of these equations, we have four unknowns, and there is no way to recover the four unknowns from the three observed variables on the left-hand side. Given the previously mentioned specification of the medical expenditure function, in the randomly assigned groups $E[h_i] \geq E[h_i|B_i = 0]$ and the jammed moral hazard measure in expression (2) would systematically overestimate the pure moral hazard. It is also straightforward to see that the sum of expressions (2) and (3) yields expression (1), which means that the difference in average medical expenditure between the group that bought the insurance (Y1) and the group that refused to buy it (Y2) is the sum of jammed moral hazard and pure adverse selection effects. Consequently, in randomly assigned groups, expression (1) would systematically overestimate the combined effect of pure moral hazard and adverse selection.

The institutional design of the health insurance in Croatia mimics the previously mentioned experimental setting but with an important difference. The patients who are entitled to supplemental insurance for free correspond to group X in the previously mentioned hypothetical experiment, and the remainder of the population who had to choose between buying and not buying the supplemental insurance corresponds to group Y. However, the composition of our group X has been decided based on some predetermined set of criteria, and therefore it is not random, and hence the leftover segment Y is nonrandom as well. The nonrandom group assignments could complicate the estimation of both moral hazard and adverse selection effects.

The analytical bias introduced by the term $E[h_i] - E[h_i|B_i = 0]$ into the measurement of the true moral hazard effect in expression (2) depends on the difference between the average fixed effect (health status) of the general patient population and the patients who opted not to buy the insurance. In our concrete situation, the first group is not the general patient population but the patients with preferential insurance status, that is, those who were entitled to receive the supplemental insurance free of charge (FREE), and in the leftover group (1-FREE), we have patients who decided not to buy the supplemental insurance (NO). If our FREE group consists of patients with better than average health (lower h) relative to the general population and more similar to the patients in the other group who chose not to buy the insurance (NO), it is actually possible that our nonrandom assignment design could attenuate or totally eliminate the positive analytical bias (jam). Of course, if the opposite were true, the problem of overestimation of the moral hazard would be exacerbated.

The nonrandom assignment of groups could also complicate the estimation of the adverse selection effect. Even without the analytical bias, here the problem is caused by the fact that $E[h_i]$ is not representative of the entire patient population but of the FREE group, and $E[h_i|B_i = 1]$ represents the average health status of the portion of the leftover (1-FREE) group who decided to purchase the supplemental insurance (BOUGHT). If our FREE group consists of people with better than average health (lower h) relative to the general population, then the residual (1-FREE) group will necessarily have worse average health (higher h) than the random group and, among them, people who decided to buy the insurance will have worse health than people who would decide to buy the insurance in a randomly selected group. In such a case, our data would overestimate the actual adverse selection effect. If the opposite were true, then our data would underestimate the actual adverse selection effect.

5. ESTIMATION APPROACH

The empirical estimation of moral hazard and adverse selection effects is carried out based on expressions (2) and (3) by the use of matching estimators. The matching method compares the outcomes of program participants with those of matched nonparticipants where matches are chosen based on the similarity in observed characteristics. The framework assumes that there are two potential outcomes W_0 and W_1 that represent control (no treatment) and treatment states and two groups of individuals, $D = 1$ indicating persons who participated in the program and $D = 0$ for persons who did not. An individual can only be in one state at a time, hence only one of the outcomes is observed. The outcome that is not observed is called a counterfactual

DOI: 10.1002/hec

outcome. Assessing the results of a treatment requires making inferences about the outcomes, which would have been observed in the control state. In our case, we run two matching estimations. First, we estimate the effect of moral hazard based on expression (2) where we designate our FREE group to be a control group and the NO group to be our treated group. Second, the effect of adverse selection is estimated based on expression (3) where, again, we designate our FREE group to be a control group and the BOUGHT group to be the treated group. In the first case, the counterfactuals reflect how much medical services would have been consumed by a patient with NO supplemental insurance had that person have FREE insurance. Thus, the problem in estimating expression (2) is to construct the unobservable counterfactuals for the patients who have no supplemental insurance using the patients who are entitled to free supplemental insurance. In the second case, we match the FREE patients with the BOUGHT patients to estimate the effect of adverse selection as the average difference in health consumption for the BOUGHT patients and the constructed counterfactuals from the FREE patients.

To evaluate the effects of moral hazard and adverse selection, we use an average treatment effect on the treated estimator, $\mathrm{ATT} = E[(W_1 - W_0)|D = 1]$, which gives the mean impact of the program on program participants. Matching estimators typically rely on the conditional independence assumption. This assumption requires that there exist a set of observed characteristics Z such that potential outcomes (W_0, W_1) are independent of the participation status D conditional on Z. However, if the parameter of interest is ATT, a weaker version of the conditional independence assumption, known as the conditional mean independence (CMI) assumption, which can be expressed as $E(W_0|Z, D = 1) = E(W_0|Z, D = 0) = E(W_0|Z)$, is sufficient (Heckman *et al.*, 1998). The CMI assumption asserts that the outcome W_0 for the participants had they not participated in the program is the same as that of nonparticipants, conditional on Z. However, the method allows the individuals who expect to experience different levels of W_1 to select into the program on the basis of that information. For estimating the ATT parameter, matching methods allow selection into the treated group to be based on possibly unobserved components of the anticipated program impact, but only to the extent that the program participation decisions are based on the unobservable determinants of W_1 and not those of W_0 (Todd, 2008).

Based on the CMI assumption, we define our FREE group as a control group and our BOUGHT and NO groups as treated groups. The reason is that the BOUGHT and NO groups made decisions about buying or not buying the insurance based on their unobserved health status (and perhaps some other private information), whereas the selection into the FREE group was not determined based on the health status but rather some exogenous set of rules (poor people, veterans, students, etc.). This designation of the control and treatment groups allows us to estimate the negative of the jammed moral hazard from the ATT in expression (2) and the pure adverse selection effect from the ATT in expression (3).

The differences between the hypothetical experimental design described before and the actual institutional design of the health insurance in Croatia that generated our data could cause some estimation problems. Specifically, as a result of nonrandom group composition, the distribution of patients' latent health status could be systematically different between the FREE and 1-FREE groups. In an unlikely extreme case, the latent health status might not overlap at all such that one cannot find any patients from the FREE group who has the same health status in the BOUGHT or NO groups. In more likely cases, we could possibly encounter biases (overestimation or underestimation) of asymmetric information effects discussed in the previous section. Because the latent health status is unobservable, we need to find variables that are good proxies for it. Among many different genetic, socioeconomic, and environmental determinants of health, age is probably the most important and also the most predictable one in the sense that younger people are on average healthier than older people. Another quite important and predictable determinant of health is probably wealth or income. Wealthier people are likely to lead healthier lifestyles and have better access to medical care. These effects can accumulate and have more pronounced positive effects on the health status of elderly people than on young people. The same can be probably argued about the positive health effect of education.

Because our invoices data do not contain patients' income (wealth) and education variables, we are unable to explicitly match patients on these variables. This problem can be potentially mitigated by the fact that we match patients on the eligibility category for compulsory insurance (payroll, unemployed, self-employed, retired,

DOI: 10.1002/hec

farmer, welfare recipient), which are under Croatian circumstances fairly good proxies for both. After controlling for these factors, to the extent that there could be some leftover uncontrolled unobservables influencing the patients' latent health status, we argue that their impact should be heterogeneous across different age groups.[7] In particular, young patients in the 20- to 40-year-old cohort are generally quite healthy, and any potential differentiated impacts of wealth or education on their individual health status did not yet have a chance to materialize. Hence, young adults who are entitled to supplemental insurance for free (e.g., because of their full-time student status) should be equally healthy regardless of whether they attend college or not and as such similar to their counterparts in the 1-FREE group. Obviously the same cannot be argued for older cohorts. For them, any potential differentiated impacts of wealth or education on their individual health status could have materialized, and because the FREE group automatically includes poor people, it is reasonable to anticipate that in older age cohorts, the unobserved average heath status in the FREE group is likely to be worse than in the 1-FREE group. To identify the differential age effect on adverse selection and moral hazard, we examine differences in the number of visits and the total cost per patient for 20–30, 30–40, 40–50, 50–60, and over 60-year age groups separately.

We directly match the observationally equivalent patients among different groups using observed demographics. For each patient in the NO (BOUGHT) group, we find their observationally equivalent patients in the FREE group. Specifically, we match the patients if they have the same age, sex, and the eligibility category for compulsory insurance. We allow 2-year age difference between matched patients because patients born at the end of a year are not much older than the ones born early the following year. We use a matching with replacement rule, meaning that, if a patient in the FREE group has the same sex, eligibility category, and similar age to multiple patients in the NO (BOUGHT) group, the patient is used multiple times. Next, we calculate the counterfactual for each patient in the NO (BOUGHT) group by taking the mean outcomes (number of visit and total cost per patient) of the matched patients from the FREE group. By doing so, we constructed counterfactual outcomes for all patients in the NO (BOUGHT) group for whom matches in the FREE group have been found. Finally, we conduct a standard *t*-test to compare the mean of the outcomes between the patients in the NO (BOUGHT) group and their constructed counterfactuals.

6. RESULTS

We start the empirical analysis by an attempt to assess the gravity of the potential bias introduced by the nonrandom composition of the FREE and 1-FREE groups. To test for the differences in latent health status among those eligible for free coverage and the rest of the patients' population, we use invoices covered in full by the compulsory insurance based on the diagnosis. Recall, these are the invoices pertaining to chronic psychiatric diseases, cancers, dialysis, organ transplants, and others. Because these illnesses are covered in full for everybody, we can treat them as exogenous and hence not impacted by peoples' choices that are likely to be influenced by having or not having access to insurance. Therefore, this indicator could be a good proxy for the unobservable health status. We ran a simple probit where the dependent variable takes the value of 1 if the patient made a fully insured visit to the hospital and 0 if not, and the independent variable takes the value of 1 if the patient is in the FREE group and 0 if the patient is in the 1-FREE group. We ran this regression with the entire data set and then for the different age cohorts separately.[8]

[7]Recall from Section 2, item (c), that the three most important categories comprising the FREE group are the following: (i) poor people; (ii) secondary school and college students, and military draftees; and (iii) various categories of people with disabilities, war veterans, and their families. When it comes to age distribution, the poor people are most likely predominantly older, whereas the second two groups are largely composed of younger people, most of them ages 18–40 years. The war veterans are the participants of the Croatian independence war of 1990–1995 and hence still relatively young. Some of them are amputees (or with some other physical impairment), but most of them suffer from various forms and degrees of post-traumatic stress disorders but are otherwise probably of similar health as the rest of the population in the same age group.

[8]There are only 1564 patients (3437 invoices) with a visit to the hospital for a fully covered diagnosis. The reason for 274 fewer invoices is because we deleted invoices related to pregnancies and childbirths, which are obviously not indicators of inferior health.

DOI: 10.1002/hec

Table IV. Testing the relationship between fully covered hospital visits and being in the FREE supplemental insurance group

Probit	Number of observations	Marginal effects	z-statistic	$P>\|z\|$
All ages	22957	−0.0624	−1.79	0.074
20–30	2815	−0.0323	−3.79	0.000
30–40	2661	0.0027	0.25	0.803
40–50	4007	−0.0003	−0.03	0.977
50–60	5075	−0.0006	−0.08	0.934
60+	8399	−0.0050	−0.91	0.363

Left-hand side variable equals 1 if the number of visits ≥1, else 0; right-hand side variable equals 1 if patient is in the FREE group, else 0. Constants suppressed for brevity.

As seen from Table IV, the marginal effect of the FREE indicator is negative but insignificant (at the standard 5% level), indicating that the probability of having a visit to the hospital for one of the fully covered diagnosis is not significantly impacted by the membership in the FREE group. Therefore, to the extent that this variable correctly measures the unobservable health status, the two groups appear to be similarly healthy or unhealthy. The separate results for different age cohorts also show insignificant FREE effects, except for the 20- to 30-year-old cohort. This marginal effect is negative and significant indicating that the probability of having a visit to the hospital in regard to a fully covered diagnosis is 3.2% lower for the FREE group relative to the 1-FREE group. This means that the youngest cohort in the data set in the FREE group is on average healthier than their counterparts in the rest of the population.[9]

The matching estimation results are presented in Table V. We show one set of results for all age groups combined and also disaggregated results for five different age groups. There are two outcome variables: the number of visits per patient and the total cost per patient expressed in the local currency—Croatian kuna (HRK). Each table has two panels of results, labeled as A (moral hazard) and B (adverse selection), corresponding to the comparisons of observed and counterfactual variables outlined in the estimation section. In each of the two panels, the first column of numbers always represents the observed values (NO, BOUGHT), and the second column presents the counterfactual values (FREE, FREE). The third column represents its difference (ATT), and the remaining two columns are the t-statistics and the number of patients in the NO and BOUGHT groups for whom matches are found in the FREE group. As mentioned before, matching is carried out directly on age, sex, and the eligibility category.[10]

The results for all age cohorts combined show that the patients with supplemental insurance, purchased or given for free, have more hospital visits and higher total cost than the ones with compulsory insurance only. As seen from panel A, all ATT numbers are negative and significant indicating that the healthcare consumption of patients with NO insurance is smaller than that of patients with FREE insurance, hence there is positive moral hazard. For patients of all ages, the numbers in panel B are positive and significant indicating that the healthcare consumption of those who BOUGHT the insurance is larger than of those with FREE insurance, hence there is adverse selection. For example, during the period under consideration (4 months), patients with NO supplemental insurance have on average 0.95 (34%) less hospital visits than people with FREE supplemental insurance, and patients who BOUGHT the insurance have on average 0.23 (8%) hospital visits more than patients with FREE insurance.[11] There is also the difference in costs. Patients with NO insurance

[9]Almost identical results are obtained by testing the differences in the mean number of vists (unconditional and conditional on age) for fully covered diagnoses between FREE and (1-FREE) groups.

[10]We have also tried to match on the number of visits for the fully (compulsory insurance) covered diagnosis as a proxy for the underlying unobserved health status, but the procedure failed because of the insufficient number of patients with such invoices.

[11]The percent changes in parentheses are always calculated as the effect of the treatment relative to that of no treatment, that is, (treatment—control)/control. The combined effect of moral hazard and adverse selection is obtained as the sum of two percentages.

DOI: 10.1002/hec

Table V. Average impact of supplemental insurance on the consumption of medical care for the top eight hospital services for all patients

	A. Moral hazard					B. Adverse selection				
	NO	FREE	ATT	*t*-value	Matched	BOUGHT	FREE	ATT	*t*-value	Matched
Number of visits										
All ages	1.86	2.80	−0.95	−14.41	1332	3.00	2.77	0.23	5.39	7901
	(0.06)	(0.02)	(0.07)			(0.04)	(0.01)	(0.04)		
20–30	1.48	2.69	−1.21	−12.45	309	2.20	2.61	−0.41	−2.89	500
	(0.08)	(0.07)	(0.10)			(0.14)	(0.05)	(0.14)		
30–40	2.04	2.74	−0.69	−4.65	325	2.88	2.69	0.19	1.35	763
	(0.14)	(0.04)	(0.15)			(0.14)	(0.03)	(0.14)		
40–50	1.95	2.81	−0.86	−6.56	310	3.33	2.81	0.51	4.46	1398
	(0.13)	(0.04)	(0.13)			(0.11)	(0.01)	(0.12)		
50–60	2.10	3.13	−1.03	−5.39	248	3.38	3.20	0.18	1.88	1987
	(0.19)	(0.04)	(0.19)			(0.09)	(0.02)	(0.10)		
60+	1.66	2.70	−1.05	−6.20	128	2.81	2.55	0.26	4.21	3175
	(0.16)	(0.08)	(0.17)			(0.06)	(0.01)	(0.06)		
Cost per patient										
All ages	174	255	−81	−16.91	1332	295	273	22	5.91	7901
	(4.60)	(1.81)	(4.81)			(3.66)	(0.69)	(3.67)		
20–30	131	208	−77	−9.62	309	184	204	−20	−1.90	500
	(7.46)	(3.58)	(8.00)			(10.36)	(2.79)	(10.52)		
30–40	187	254	−66	−6.20	325	268	252	15	1.38	763
	(10.07)	(2.85)	(10.67)			(10.93)	(1.71)	(11.05)		
40–50	185	272	−87	−9.44	310	313	273	39	4.21	1398
	(9.18)	(2.60)	(9.25)			(9.31)	(1.14)	(9.35)		
50–60	199	292	−92	−6.90	248	330	311	19	2.40	1987
	(12.52)	(4.02)	(13.37)			(7.93)	(1.41)	(8.05)		
60+	172	276	−103	−7.17	128	292	268	24	4.31	3175
	(13.26)	(6.61)	(14.41)			(5.57)	(0.91)	(5.58)		

Standard errors are in parentheses.
Matched A. Number of patients with NO that found their match in the FREE category.
Matched B. Number of patients with BOUGHT that found their match in the FREE category.

spent 81 HRK (32%) per person less than patients with FREE insurance, and patients who BOUGHT the insurance spent 22 HRK more (8%) than patients with FREE insurance. Therefore, as indicated by expression (1), the average combined effect of moral hazard and adverse selection amounts to 1.18 hospital visits or 103 HRK per patient. This means that a patient who BOUGHT the insurance, on average, outspent a patient who had the opportunity but chose not to buy the insurance (NO) by 42% in terms of the number of visits to the hospital or 40% in terms of costs.

Next, we look at the results for different age groups. When it comes to estimating the effect of moral hazard (panel A), the results indicate the presence of positive and statistically significant moral hazard effect for all age cohorts. For the youngest cohort of 20–30 years of age, the jammed moral hazard effect amounts to 45% increase in the number of visits and 37% increase in total cost per patient. To the extent that the members of the FREE group in the 20- to 30-year-old cohort are healthier than the members of the 1-FREE group (as indicated by the results in Table IV), the analytical bias $E[h_i] - E[h_i|B_i = 0]$ attenuates the jam and contributes to a more precise estimation of the pure moral hazard effect. When it comes to estimating the effect of adverse selection, the results obtained on the aggregate data level do not hold for all age groups. As seen from panel B, we found the presence of favorable selection effect (negative adverse selection) for the group of patients in the 20- to 30-year-old cohort and the presence of adverse selection effect for all other age groups. If the 20- to 30-year-old cohort in the FREE group consist of people with better than average health, the same cohort in the residual 1-FREE group will necessarily have patients with worse than average health, and among them, people who decided to buy the insurance will have worse health then people who would decide to buy the insurance in a

DOI: 10.1002/hec

random group. As a result, our formula $E[h_i|B_i = 1] - E[h_i]$ will tend to overestimate the actual adverse selection effect (underestimate the favorable selection effect) in the youngest cohort.

The statistically significant favorable selection effect for the youngest age cohort of patients (20–30 years) is robust with respect to the outcome variable. Patients who bought the supplemental insurance are visiting the hospital less frequently and spend less money than patients who obtained the supplemental insurance free of charge. Although somewhat unexpected, the results are not that unusual and can be easily rationalized.[12] The 20- to 30-year-old cohort who bought the insurance is dominated by people who completed education entirely after the collapse of socialism and are likely to be more employable than older cohorts. At the same time, these young people are used to the lack of job security and the uncertainties of free enterprise markets and are likely to be more risk averse than their parents and hence more likely to purchase the insurance.[13] The reversal of the selection result from favorable selection to adverse selection in the older cohorts is interesting, yet quite plausible as well. An individual that is currently 40 years old was 20 when socialist Yugoslavia collapsed. These and especially older people were used to socialized medicine where healthcare access was completely free for all, and therefore they are very sensitive to out-of-pocket payments for access to health care. Consequently, those with somewhat or significantly impaired health will typically choose to buy the supplemental insurance.

The obtained results for farmers turn out to be qualitatively similar to the results for the rest of the population in that we see the presence of positive moral hazard and adverse selection effects as well. However, the magnitudes of the estimated effects are always smaller compared with the entire population, and the ATTs measuring the adverse selection effects are generally insignificant. In the case of farmers, the procedure for controlling potentially confounding influences of wealth and education on the unobserved health status by looking at matching results for different age cohorts failed for some cohorts because of the small sample size. However, because farmers are generally more homogeneous when it comes to income and education than the rest of the population, the measured effects at the aggregate level should be more precise than the corresponding set of results for the entire population.[14]

7. ROBUSTNESS CHECKS

To investigate how important are the possible systematic differences in unobservables between FREE and 1-FREE groups, we experiment with several additional specifications of the basic matching model. Recall that among the people eligible for free supplemental insurance are various categories of handicapped citizens. Because only the FREE group contains the handicapped patients, their appropriate matches may not be found in the NO and BOUGHT groups. This is the rationale for excluding them from the comparison with the other two groups. The people with disabilities entitled to free supplemental insurance are either war veterans or civilians. As mentioned briefly before, virtually all Croatian war veterans are the participants of the Croatian independence war of 1990–1995 and hence still relatively young. Many of them acquired the status of invalids (people with disabilities) based on the various forms and degrees of post-traumatic stress disorders (which is typically difficult to accurately diagnose) but are otherwise probably equally healthy as the rest of the population in the same age group. Therefore, one can argue that they should not be excluded from the analysis. Because we do not have the information on the nature and the gravity of the disablements, we decided to exclude all people with disabilities (but not their family members) regardless of whether their handicap is war related or not and

[12]See also Finkelstein and McGarry (2006), Fang *et al.* (2008), and Zheng and Zimmer (2008) in the case of US farmers.

[13]Note that heterogeneity in risk preferences is not neccessary to obtain the advantageous selection result. In fact, as shown by Fang *et al.* (2008) who were able to analyze various sources of advantageous selection in the Medigap insurance market, risk preferences do not appear to be its main source. Instead, they suggest that cognitive ability plays a prominent role. Our data does not allow us to empirically investigate the importance of various sources of favorable selection, but our story could have been equally convincingly told by resorting to superior cognitive ability of the younger generation rather than their higher risk aversion.

[14]Separate estimation results for farmers are not presented to economize with space but are available from the authors upon request.

DOI: 10.1002/hec

regardless of the degree of disablement. Among 4626 patients in our data set for the top eight hospital services (Table III) that are entitled to FREE supplemental insurance, these exclusions eliminate 674 patients.

The results are presented in Table VIa. Same as before, we present the results for all ages groups combined and for five age cohorts separately. Several results are worth mentioning. First, the presence of statistically significant favorable selection effect for the youngest age cohort of patients (20–30 years) was confirmed for both the number of visits and the total cost per patient. Second, the effect of moral hazard is smaller with the truncated FREE group than with the entire FREE group, which makes sense because of the elimination of some of the supposedly heavy users from the FREE group. Finally, the effect of decreased moral hazard has been approximately offset by the increase in the adverse selection.

The second problem associated with the design of the FREE group is the income census criterion for obtaining the free supplemental insurance. There are two categories of people eligible for free supplemental insurance under this criterion: (i) people whose per capita household monthly income is below 1516.32 HRK; and (ii) single retired individuals with income below 1939.39 HRK. By the nature of the program design, these poor people cannot be found in any of the other two groups (BOUGHT or NO), so it would have been impossible to find their matches based on income if we had the income variable in our dataset. So, to the extent that income matters, either for the unobserved health status or for the actual consumption of the health services, it is reasonable to exclude these people from the analysis. However, the problem is that these established eligibility thresholds do not represent some meaningful breaks in the income distribution. Based on the fact that for 80% of the population in Croatia the per capita income falls within 1 standard deviation around the

Table VIa. Average impact of supplemental insurance on the consumption of medical care for the top eight hospital services with truncated FREE group (people with disabilities)

	A. Moral hazard					B. Adverse selection				
	NO	FREE	ATT	*t*-value	Matched	BOUGHT	FREE	ATT	*t*-value	Matched
Number of visits										
All ages	1.86	2.70	−0.85	−12.97	1328	3.00	2.65	0.35	8.03	7889
	(0.06)	(0.02)	(0.07)			(0.04)	(0.01)	(0.04)		
20–30	1.48	2.58	−1.10	−12.03	308	2.20	2.51	−0.31	−2.20	499
	(0.07)	(0.05)	(0.09)			(0.14)	(0.04)	(0.14)		
30–40	2.05	2.71	−0.67	−4.49	325	2.89	2.66	0.23	1.63	759
	(0.14)	(0.04)	(0.15)			(0.14)	(0.03)	(0.14)		
40–50	1.95	2.83	−0.88	−6.66	310	3.33	2.82	0.50	4.38	1398
	(0.13)	(0.03)	(0.13)			(0.11)	(0.02)	(0.12)		
50–60	2.10	2.81	−0.71	−3.72	248	3.38	2.93	0.45	4.66	1987
	(0.19)	(0.04)	(0.19)			(0.09)	(0.02)	(0.10)		
60+	1.63	2.55	−0.92	−5.55	125	2.80	2.44	0.36	5.78	3168
	(0.16)	(0.07)	(0.17)			(0.06)	(0.01)	(0.06)		
Cost per patient										
All ages	174	247	−73	−15.33	1328	295	263	32	8.63	7889
	(4.59)	(1.72)	(4.79)			(3.67)	(0.68)	(3.69)		
20–30	131	208	−77	−9.66	308	184	203	−19	−1.82	499
	(7.49)	(3.50)	(7.99)			(10.38)	(2.67)	(10.54)		
30–40	187	243	−56	−5.27	325	268	243	25	2.22	759
	(10.07)	(2.58)	(10.56)			(10.98)	(1.59)	(11.06)		
40–50	185	272	−87	−9.35	310	313	273	40	4.26	1398
	(9.18)	(2.80)	(9.31)			(9.31)	(1.22)	(9.36)		
50–60	199	272	−73	−5.43	248	330	293	37	4.57	1987
	(12.52)	(4.19)	(13.44)			(7.93)	(1.65)	(8.13)		
60+	169	255	−86	−6.25	125	292	257	36	6.35	3168
	(12.94)	(5.72)	(13.75)			(5.57)	(0.83)	(5.59)		

Standard errors are in parentheses.
Matched A. Number of patients with NO that found their match in the FREE category.
Matched B. Number of patients with BOUGHT that found their match in the FREE category.

DOI: 10.1002/hec

Table VIb. Average impact of supplemental insurance on the consumption of medical care for the top eight hospital services with the truncated FREE group (people with disabilities and income census self-employed)

	A. Moral hazard					B. Adverse selection				
	NO	FREE	ATT	*t*-value	Matched	BOUGHT	FREE	ATT	*t*-value	Matched
Number of visits										
All ages	1.86	2.75	−0.89	−13.28	1301	3.00	2.68	0.32	7.27	7803
	(0.06)	(0.02)	(0.07)			(0.04)	(0.01)	(0.04)		
20–30	1.48	2.58	−1.10	−12.03	308	2.20	2.51	−0.31	−2.20	499
	(0.07)	(0.05)	(0.09)			(0.14)	(0.04)	(0.14)		
30–40	2.04	2.74	−0.71	−4.74	320	2.89	2.67	0.21	1.53	752
	(0.14)	(0.04)	(0.15)			(0.14)	(0.03)	(0.14)		
40–50	1.97	2.94	−0.97	−6.93	300	3.32	2.87	0.48	3.88	1384
	(0.13)	(0.04)	(0.14)			(0.11)	(0.02)	(0.12)		
50–60	2.10	2.88	−0.78	−3.82	237	3.39	3.00	0.39	3.82	1924
	(0.19)	(0.07)	(0.20)			(0.09)	(0.03)	(0.10)		
60+	1.64	2.56	−0.93	−5.56	124	2.80	2.45	0.36	5.74	3166
	(0.16)	(0.07)	(0.17)			(0.06)	(0.01)	(0.06)		
Cost per patient										
All ages	173	249	−76	−15.65	1301	294	264	30	8.11	7803
	(4.63)	(1.83)	(4.86)			(3.67)	(0.75)	(3.70)		
20–30	131	208	−77	−9.66	308	184	203	−19	−1.82	499
	(7.49)	(3.50)	(7.99)			(10.38)	(2.67)	(10.54)		
30–40	186	243	−57	−5.51	320	268	244	24	2.16	752
	(9.99)	(2.58)	(10.42)			(10.99)	(1.55)	(11.06)		
40–50	187	278	−91	−9.46	300	312	274	37	4.03	1384
	(9.39)	(3.22)	(9.64)			(9.19)	(1.33)	(9.24)		
50–60	198	277	−79	−5.60	237	330	297	33	3.94	1924
	(12.96)	(4.87)	(14.13)			(8.07)	(2.03)	(8.37)		
60+	169	256	−87	−6.34	124	292	257	35	6.27	3166
	(13.02)	(5.40)	(13.80)			(5.56)	(0.82)	(5.59)		

Standard errors are in parentheses.
Matched A. Number of patients with NO that found their match in the FREE category.
Matched B. Number of patients with BOUGHT that found their match in the FREE category.

mean, it is very likely that there are a lot of people in the BOUGHT and NO groups whose income is only negligibly higher than in the FREE group.[15] For all practical purposes, the purchasing powers, consumption patterns, lifestyles, and others for the people on either side of this income threshold are likely to be indistinguishable.

To somehow address this problem, we kept all single retired people and all multi-person household members in the FREE group except for self-employed individuals. This approach is predicated on a conjecture that if there are any meaningful differences between poor people and the rest of the population, they have to be captured in the category of self-employed individuals where the new class of private entrepreneurs and successful businessmen should be found. Hence, this particular truncation of the FREE group includes all patients with disabilities as in Table VIa plus all self-employed individuals whose per capita household monthly income is below 1516.32 HRK. The matching estimators for all patients with such truncated FREE group are presented in Table VIb. The obtained results are very similar to the results presented in Table VIa. The favorable selection effect for the youngest cohort is still there, the aggregate adverse selection effect is somewhat smaller, and the moral hazard effect is somewhat larger than in Table VIa.

[15]Our estimates based on the 2008 Croatian Household Budget Survey point to a relatively tight income distribution. The ratio of top 10 to bottom 10 percentiles of total income per capita distribution for the entire population is 3.9. The top 10 percentile is about two times higher than the median, and the bottom 10 percentile is about a half of the median income. Similar findings for income distribution in 2002 and 2004 are presented in World Bank (2006).

DOI: 10.1002/hec

Finally, to investigate whether the obtained results are sensitive with respect to the type of the outpatient services used, we re-estimated our model with the invoices data for all hospital services (see Table I for the basic statistics). The results are quite similar to the hereto obtained results with the subsample of the top eight most widely used services.[16] First, the presence of statistically significant favorable selection effect for the youngest age cohort of patients (20–30 years) is reaffirmed for both outcome variables. Second, the magnitude of the aggregate (all ages) moral hazard effect as measured by the number of visits per patient is about the same as for the top eight services (about 30% reduction in the NO group relative to the FREE group) but larger when measured by the cost per patient (35% for all services relative to 30% for the top eight services). This makes sense as those other services now included in the estimation are somewhat pricier. Finally, the magnitude of the adverse selection effect appears to be smaller. When estimated with all hospital services, the effect amounts to 11% increase (BOUGHT over FREE) in the number of hospital visits (4% in the cost) per patient relative to 13% increase in visits (12% in the cost) per patient when estimated with the top eight services presented in Table VIa.

8. CONCLUSIONS

The institutional setting under which the system of health insurance in Croatia is organized is likely to be plagued by asymmetric information problems, both of moral hazard and adverse selection types. First, the problem of moral hazard is present in both compulsory and supplemental insurance because having health insurance is likely to change peoples' behavior in response to the fact that healthcare services are now available at zero or significantly reduced price. On the other hand, the problem of adverse selection is only relevant when pricing the supplemental insurance policies because the problem of adverse selection completely disappears under conditions of universal coverage and a single-payer system, both of which are satisfied in the compulsory insurance program. Because our primary objective was to disentangle the effect of moral hazard from adverse selection, we focus on the supplemental insurance only and rely on the distinct feature of the program design to accomplish this goal.

The presence and the magnitude of the asymmetric information problems have important consequences for the pricing strategies of insurance companies and their long-run profitability and viability. In case of the HZZO, the pricing strategies for supplemental insurance are definitely not designed with the primary objective of solving the moral hazard problem but are predominantly ruled by the social welfare considerations. As a matter of fact, the HZZO appears to be fully cognizant of the moral hazard problem and wants to mitigate it by introducing and constantly modifying the system of co-payment-based services. However, most of the really expensive services (pregnancies and births, cancer treatments, organ transplants, etc.) are still covered in full through the compulsory insurance policy, whereas the system of co-payments only applies to a wide list of less expensive care, and the total maximum out-of-pocket expense is limited. In addition, a person who obtained the supplemental insurance policy (either by purchase or by the government grant) is entitled to full waiver of all co-payments and can consume essentially all medical services at zero prices.

When it comes to assessing the importance of adverse selection problems in pricing the supplemental insurance policies, there are two possibilities. First, it is possible that the HZZO cannot precisely ascertain each individual's health status at a reasonably low cost. The theoretically correct solution to this problem is to offer a menu of contracts where people would self-select themselves into purchasing a contract that is reflective of their type (health status). This menu would involve various combinations of benefits and premiums (prices) that people can choose from. On the other hand, it is also possible that there is no fundamentally important asymmetry of information in the sense that using readily observable individual socioeconomic and other

[16]The results for all hospital services with excluded patients with disbilities from the FREE group and the results with excluded patients with disabilities plus all self-employed individuals whose per capita household monthly income is below 1516.32 HRK from the FREE group are available from the authors upon request.

DOI: 10.1002/hec

characteristics, the HZZO can correctly predict the probability of a person having a loss (in this context, going to the hospital and consuming a health service) and can sell insurance policies to different categories of people at different prices. Because the HZZO is neither offering a menu of contracts nor meaningfully price discriminating when selling their supplemental insurance policies, its pricing strategy is not revealing of the true nature and the magnitude of the adverse selection problem. It is possible that it is selling all policies at essentially the same price because the asymmetric information problems are substantial, but offering a menu of contracts may be administratively too costly to apply. Alternatively, the asymmetric information problems could be small or trivial, and the reasons for limited or no price discrimination can be explained by equity and solidarity considerations.[17]

Using the unique data set of hospital invoices from a regional hospital in Croatia and relying on the methodology of matching estimators, we were able to separately estimate the effect of adverse selection and moral hazard in the consumption of medical services related to non-life-threatening diagnoses. Our best estimates show that, on average for the population of patients 18 years and older, the combined distortion because of adverse selection and moral hazard problems amounts to an increase in total number of visits per patient during a 4-month period by about 44%. There is also a statistically significant increase in the total cost per patient by about 42%. The pure effect of adverse selection amounts to 12% more visits and about 11% higher total cost per patient. The (jammed) moral hazard effect amounts to 32% increase in hospital visits and about 31% increase in total cost per patient. To fine-tune the estimates, we also looked at the asymmetric information effects across age cohorts. When it comes to estimating the effect of moral hazard, the results indicate the presence of positive and statistically significant moral hazard effect for all age cohorts. However, for the youngest age cohort of patients (20–30 years), we found the presence of statistically significant favorable selection effect. Interestingly enough, the selection effect reverses its sign in the next age cohort (30–40 years) from favorable to adverse selection. Although somewhat unusual, the obtained results can be explained by the differences in risk aversion across cohorts caused by the timing of transition from socialism to market economy.

The obtained results could have an important policy implication for the future reform of the Croatian state-administered supplemental health insurance plan. Given that adverse selection appears to be generally small, the improvements in the healthcare system should focus on curbing the moral hazard side of the problem by introducing a universal system of co-payments. This proposal has two prongs: (i) the current system of wide exceptions from paying co-payments needs to be abolished or significantly trimmed; and (ii) the co-payment rates have to be widely applied to the majority of health services and need to be increased to the point where they begin to effectively 'bite'. Estimating the elasticity of demand for various health services with respect to co-payments could be a topic of an interesting future research.

Finally, the results could also have much wider implications. The results showing the presence of favorable selection in one age cohort and the presence of adverse selection in other cohorts are particularly interesting for at least two reasons. First, they highlight the importance of future theoretical and empirical work on identification and estimation of insurance models with multidimensional screening. Second, they also contribute to the growing empirical evidence that the gravity of asymmetric information problems because of adverse selection in the market provision of insurance is not that large. However, it is also worth mentioning that because of the use of invoices data, any generalization of our results would be tenuous because the inclusion of nonuser contract holders could influence the findings. For example, in the measurement of adverse selection, it is not known how many nonusers hold contracts in the FREE group relative to nonusers holding contracts in the BOUGHT group. Addressing this problem is beyond the scope of this paper as it would require an altogether different type of data set.

[17]As mentioned in Section 2, the HZZO is selling supplemntal insurance policies at three different monthly premiums (50, 80, and 130 HRK, which is the equivalent of $10, $16, and $26). Obviously, the HZZO pricing policy is not governed by the actuarial considerations because the supplemental insurance to retirees, who are the most likely to have a loss (to use it), is sold at a discount.

DOI: 10.1002/hec

ACKNOWLEDGEMENTS

The support for this project came through a grant to the University of Delaware from the State Department's Title VIII Program. No potential conflict of interest exists. The manuscript contains original unpublished work and is not being submitted for publication elsewhere.

REFERENCES

Abadie A, Imbens G. 2006. Large sample properties of matching estimators for average treatment effects. *Econometrica* **74**(1): 235–267.

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

Cameron, AC, Trivedi PK, Milne F, Piggott J. 1998. A microeconometric model of the demand for health care and health insurance in Australia. *Review of Economic Studies* **1**: 85–106.

Chiappori P-A, Jullien B, Salanie B, Salanie F. 2006. Asymmetric information in insurance: general testable implications. *The Rand Journal of Economics* **37**(Winter): 783–798.

Coulson NE, Terza JV, Neslusan CA, Bruce S. 1995. Estimating the moral-hazard effect of supplemental medical insurance in the demand for prescription drugs by the elderly. *The American Economic Review* **85**(2): 122–126.

Dehejia R, Wahba S. 2002. propensity score-matching methods for non-experimental causal studies. *The Review of Economics and Statistics* **84**(1): 151–161.

Fang H, Keane MP, Silverman D. 2008. Sources of advantageous selection: evidence from the Medigap insurance market. *Journal of Political Economy* **116**(2): 303–350.

Finkelstein A, McGarry K. 2006. Multiple dimensions of private information: evidence from the long-term care insurance market. *The American Economic Review* **96**: 938–58.

Heckman J, Ichimura H, Todd P. 1998. Matching as an econometric evaluation estimator. *The Review of Economic Studies* **65**(2): 261–294.

Holly A, Lucien G, Gianfranco D, Brigitte B. 1998. An econometric model of health care utilization and health insurance in Switzerland. *European Economic Review* **42**: 513–522.

Rothschild M, Stiglitz J. 1976. Equilibrium in competitive insurance markets: an essay on the economics of imperfect information. *Quarterly Journal of Economics* **90**(November): 629–649.

Savage E, Wright DJ. 2003. Moral hazard and adverse selection in Australian private hospitals: 1989–1990. *Journal of Health Economics* **22**: 331–359.

Todd PE. 2008. In Matching Estimators. Durlauf SN, Blume LE (eds.), *The New Palgrave Dictionary of Economics*, Palgrave Macmillan, The New Palgrave Dictionary of Economics Online. Palgrave Macmillan. 13 December 2010, doi:10.1057/9780230226203.1059

Vera-Hernandez AM. 1999. Duplicate coverage and demand for health care. The case of Catalonia. *Health Economics* **8**: 579–598.

Voncina L, Jemiai N, Merkur S, Golna C, Maeda A, Chao S, Dzakula A. 2008. Croatia: health system review. *Health Systems in Transition* **8**(7): 1–108.

Vukina T, Nestic D. 2008. Asymmetric information in health insurance: some preliminary evidence from the Croatian state-administered supplemental plan. *Privredna kretanja i ekonomska politika* **115**: 25–47.

World Bank. 2006. Croatia living standard assessment: promoting social inclusion and regional equity. *Report No. 37992*, The World Bank, Washington D.C.

World Bank. 2008. Croatia: restructuring public finance to sustain growth and improve public services. *Report No. 37321-HR*, The World Bank, Washington D.C.

Zheng X, Zimmer DM. 2008. Farmers' health insurance and access to health care. *American Journal of Agricultural Economics* **90**(1): 267–279.