Journal of Health Economics 56 (2017) 317–329

Contents lists available at ScienceDirect

Journal of Health Economics

journal homepage: www.elsevier.com/locate/econbase

Supplemental health insurance in the Colombian managed care system: Adverse or advantageous selection?

David Bardey[a,*], Giancarlo Buitrago[b]

[a] Universidad de Los Andes (Bogotá), Department of Economics, and Visiting fellow at Toulouse School of Economics, Calle 19A No 1-37, Este. Bloque W Of. 810, Bogotá, Colombia
[b] Pontificia Universidad Javeriana (Bogotá), Department of Clinical Epidemiology and Biostatistics, Cra. 7 No. 40-62, Hospital Universitario San Ignacio, Segundo Piso, Bogotá, Colombia

ARTICLE INFO

Article history:
Received 30 September 2016
Received in revised form 24 February 2017
Accepted 27 February 2017

JEL classification:
D82
I13
G22

Keywords:
Information asymmetry
Health insurance
Adverse selection
Correlation test

ABSTRACT

The aim of this article is to estimate the type of selection that exists in the supplemental health insurance market in Colombia where compulsory coverage is implemented through managed care competition. We build a panel database that combines individuals' information from the Ministry of Health and a database provided by two private health insurers. We perform the correlation test for consumption of health services frequency and supplemental coverage. Following Fang et al. (2008), we condition the estimation on health controls that are available to the econometrician but not to insurers. In both cases we obtain a positive correlation, suggesting that adverse selection predominates. In order to rule out some moral hazard effects, we estimate the correlation between previous frequency of healthcare service consumption and supplemental insurance purchase. The positive correlation obtained is robust to the inclusion of controls for diagnosis implemented by health insurers, suggesting that despite some risk selection strategies, they are not protected from adverse selection. We conclude that some subsidies to supplemental coverage purchase would lower public expenditure in Colombia.

1. Introduction

Health insurance markets are usually characterized by market failures caused by the presence of information asymmetries. More precisely, the main information asymmetries described are related to the nature of selection on the one hand, and the presence of moral hazard on the other. The first of these occurs when individuals have private information prior to the purchase of insurance, and this information is correlated both with the purchase of insurance and the probability of the occurrence of an event (Rothschild and Stiglitz, 1976; De Meza and Webb, 2001). In contrast, moral hazard is defined as an increase in the magnitude of pooled risk due to insurance (Pauly, 1968; Blomqvist, 1997..

Basically, the types of selection described are divided into two groups: adverse and advantageous.[1] In practice, insurers define risk categories using observable variables and they offer contracts calculated for the average risk within each category. Under an adverse selection scenario, (within each category) individuals with unobservable increased health risk are those who are more willing to buy health insurance coverage. As such, adverse selection leads to equilibria characterized with under-insurance (Einav et al., 2010). Conversely, advantageous selection corresponds to scenarios in which individuals that prefer insurance with greater coverage exhibit a lower health risk (De Meza and Webb, 2001; Hemenway, 1990).[2] Unlike adverse selection, advantageous selection leads to equilibria characterized with over-insurance (Einav and Finkelstein, 2011). As a result, and indeed crucially, optimal regulation in health insurance markets depends on the nature of the selection at work.

This article follows the burgeoning empirical literature on selection in health insurance markets. More precisely, we aim to estimate the type of selection at play in the Colombian health

* Corresponding author.
E-mail addresses: d.bardey@uniandes.edu.co (D. Bardey), g.buitrago@javeriana.edu.co (G. Buitrago).

[1] Another form of heterogeneity among individuals has been defined through which they can anticipate moral hazard behaviors before purchasing insurance, and this increases the likelihood of purchase. Einav et al. (2013) refer to this phenomenon as selection on moral hazard.

[2] The argument behind advantageous selection is that more risk averse individuals are usually more willing to undertake prevention activities that lower their health risk on the one hand, and more willing to buy generous insurance contracts on the other.

http://dx.doi.org/10.1016/j.jhealeco.2017.02.008
0167-6296/

system that allows for the coexistence in the market of private/supplemental insurance and compulsory insurance provided in a managed care system. This health system has created a series of incentives (subsidies) through tax cuts for individuals who want to have access to better quality health services through supplemental coverage. As we will discuss in the conclusion, the nature of the selection is an important issue to anticipate if a tax-subsidy policy is self-financing.

Using the integrated information system of the Ministry of Health and Social Protection (SISPRO), and information from two health insurers that offer compulsory and supplemental (private) coverage, a panel is constructed for 2010 and 2011 of over 400,000 individuals who are dues-paying members of the system (adult workers).[3] This contains information regarding health services consumption, type of health service, spending and related diagnostics, prevention activities, socioeconomic variables, and affiliation or not to voluntary insurance.

To begin with, we perform the test suggested in Chiappori and Salanié (2000) to determine the correlation between frequency of healthcare services consumption in 2011 and insurance coverage during the same year.[4] We condition on the consumer characteristics that determine the prices offered to each policyholder. We find that the correlation test is positive, that is, the consumption of health services correlates positively with affiliation to supplemental insurance, suggesting the presence of information asymmetries. Nevertheless, as explained by Chiappori and Salanié (2000), from this test we are unable to disentangle adverse selection from moral hazard.

Thus we use the approximation of Fang et al. (2008) and we condition the estimation on health controls – represented in trace pathologies and the frequency of healthcare services consumption during the previous year – which provide a good proxy for the health status of individuals. The frequency of healthcare services consumption during the previous year is available to the econometrician but not to insurers (or, at least, is unused by insurers to determine the premium).[5] As it is eloquently explained in Fang et al., if the sign of the correlation becomes negative, then one can conclude that advantageous selection predominates (despite moral hazard effect on board). In our estimation, in both cases, that is, with and without such controls, we obtain a positive correlation, suggesting that adverse selection may predominate.

It is worth noting that we did not have Fang et al.'s "chance" since our positive correlation may still include effects related to moral hazard behaviors. Thus, controlling by a set of individual characteristics, we take advantage of the consumption of health services in the year immediately preceding (supplemental) health insurance purchase in a different way. More precisely, we estimate the probability of enrollment to supplemental coverage in 2011 as a function of the frequency of healthcare consumption in 2010. This strategy eliminates the moral hazard effect since the correlation is estimated when individuals have not yet purchased the insurance. Our results show that the use of hospital services in the year preceding insurance purchase is positively correlated with the purchase of supplemental health insurance, corroborating that adverse selection is the predominant scenario in the private health insurance market in Colombia. This estimate is robust to the inclusion of controls for diagnosis, which correlate negatively with the purchase of insurance due to some pre-existing medical conditions applied by health insurers. It suggests that despite the strategies of risk selection used by supplemental health insurers these do not protect them from adverse selection.

Finally, the activities of primary and secondary prevention realized by policyholders before the purchase of insurance are also taken into account. More precisely, we measure the consumption of cervical Pap smear, mammogram, PSA measurement and vaccination. The first three of these are related to activities of secondary prevention, while vaccination corresponds to primary prevention activities. It is assumed that primary prevention activities are positively correlated with advantageous selection, and negatively correlated with adverse selection, while secondary prevention activities are negatively correlated with advantageous selection (and positively with adverse selection).[6] We find that there exists a positive correlation between the use of PSA and cervical Pap smear with supplemental health insurance purchase in the following year and that this correlation is robust to socio-demographic, health status and service consumption controls. Moreover, the correlation between vaccination and the purchase of voluntary insurance remains negative for all estimates. Again, all of these results are suggestive of adverse selection. Following this, we discuss the policy implications of our empirical findings in the conclusion.

This paper contributes to the empirical literature on asymmetries of information to identify the predominant selection scenario, taking into account the effects of moral hazard in the health insurance market (Finkelstein and McGarry, 2006; Cohen, 2005; Einav et al., 2010; Einav et al., 2013). Empirical tests for determining the existence and nature of the selection or moral hazard in the insurance market are difficult to develop. The difficulty is that adverse selection has similar consequences to moral hazard in the correlation test; in either of the two scenarios there is an increased use of services offered by insurance, therefore isolating the two effects is not easy and it depends on the data available. Some empirical studies (Fang et al., 2008; Resende and Zeidan, 2010) have been based on the correlation test for asymmetric information developed by Chiappori and Salanié (2000). This test estimates the correlation between the level of insurance coverage and the costs of making a claim.

In performing this correlation test, Fang et al. (2008) identify that there exists asymmetric information in the private insurance market that is linked to Medicare in the United States. They propose an empirical strategy that involves comparing the sign associated with the correlation between coverage and health expenditures, including models without controls for the health status of individuals and models with controls (again only available to the econometrician). They argue that if the positive correlation changes to a negative correlation in the health controls scenario then this is evidence of advantageous selection.

The approach presented herein differs from that proposed by Einav et al. (2010), and although it is close to the proposals made by Chiappori and Salanié (2000) and Fang et al. (2008), it makes a contribution by supplementing these proposals in several aspects. First, we take advantage of the consumption of services prior to health

[3] The compulsory health insurance system distinguishes enrollees on the basis of "members" (contributing) and "insured" (family members that are covered by the members' policy). Our panel database consists of member enrollee exclusively.

[4] We use the frequency of healthcare services such as the health consumption variable instead of health spending, which is related to price and affected by unknown factors.

[5] In contrast, health insurers observe trace pathologies but do not use them for pricing. Nevertheless, some of them are used to establish pre-existing medical conditions that prevent individuals from buying supplemental coverage.

[6] The intuition behind this statement is as follows: in scenarios of advantageous selection, individuals undertake more activities that reduce the risk of disease (or the probability of its occurrence). In this case, the goal of vaccination is to prevent the occurrence of disease. On the contrary, in adverse selection the individual knows their risk of disease and so the use of secondary prevention activities plays two roles: first, the individual (or medical) receives a signal about his or her risk of getting sick, which leads to a greater probability of him or her using these services; second, according to the results, individuals propensity to buy insurance may increase because of knowledge of a more likely adverse outcome. We test this assumption in section 5.3.

insurance affiliation to rule out moral hazard behaviors. Thus the positive correlation between the consumption of health services and insurance purchase at a later time corroborates our result that adverse selection predominates in the supplemental health insurance market in Colombia. Second, this paper identifies the channels or type of private information that can be related to one or another form of selection. In particular, it considers the correlation between buying supplemental insurance and the use of activities of primary and secondary prevention.

Finally, our paper is related to Olivella and Vera-Hernandez (2013) since we measure the selection effect in a private health insurance market that co-exists with a free public and (almost) universal outside option. As in the UK, private coverage in Colombia seems to be more supplementary than complementary. Even though we do not adopt the same empirical strategy due to the different institutional contexts, Olivella and Vera-Hernandez also find that adverse selection predominates in the UK private health insurance market.[7]

This article comprises five sections. The second section describes the Colombian healthcare system. The third section presents the sources of information and data. The fourth presents the identification strategy. The fifth shows the results of the empirical estimates and the final section concludes by discussing various policy implications.

2. The Colombian healthcare system

Colombia introduced mandatory social health insurance with the approval of an ambitious healthcare reform package in 1993. The aim of this reform was to improve equity through mandatory universal health insurance and increase the performance of public spending on health. To ensure this second objective, and despite the fact that this mandatory social health insurance system is financed through a combination of payroll contributions and general taxation, the 1993 reform introduced competition into both insurance and the provision of care through a managed-care model, where public and private firms intervene. This comprehensive national social insurance scheme included a contributory regime for those able to pay and a fully subsidized scheme for the poor. Historically, beneficiaries of the contributory regime had access to a wider explicit package of benefits than beneficiaries of the subsidized regime, but both packages (of health benefits) converged in 2015 after a decision taken by the Constitutional Court. Financial protection against health risk as well as spending on public health has improved dramatically during the two last decades (Escobar et al., 2010). In 2015, it was estimated that approximately 96% of the population is insured (OCDE Report, 2015). Moreover, out-of-pocket payments are pretty low and to some extent depend on individuals' income in order to reduce problems of access.

Insured individuals in both the contributory and subsidized regimes choose their health insurer, choose care providers within their insurer's network, and receive a health benefits package purchased by insurers from public and private providers through contracts. In other words, health insurers are responsible for organizing and guaranteeing the provision of health services included in the benefits package. Moreover, all participants in the contributory regime can enroll their dependents as a family unit. Insurers receive a capitation payment. The government defines this premium annually, and it is calculated according to age, gender and place of residence.[8] To give an example, in 2010 the (annual) average premium received by health insurers to cover policyholders was about US$246, and in 2011 it was about US$272.

Even though in theory the beneficiaries of both regimes have access to a wide package of treatments and medical services, in practice health insurers also work as gatekeepers and may limit access to healthcare (Vargas et al., 2010). This widespread perception, combined with the inadequate quality offered by providers that belong to health insurers' networks in the mandatory regime, has opened a window for supplemental (and private) health insurance. Thus insurers who offer supplemental health insurance do not usually supply additional coverage in the sense that they do not cover out-of-pockets for social insurance coverage, but they rather provide access to health services of better quality. In concrete terms, this better quality involves reduced waiting times, access to better hospitals and physicians, and so on.

We estimate that supplemental health insurance is bought by 4.3% and that these policyholders usually come from the contributory regime.[9] Health insurers, which supply supplemental coverage, usually participate in the contributory regime and, due to commercial policy,[10] more than the 81% of policyholders who benefit from a supplemental plan buy it from the same insurer that offers their mandatory social coverage.[11] Curiously, despite this commercial strategy applied by health insurers, it seems that they do not share information between the two levels of coverage (mandatory and supplemental). In other words, when offering supplemental coverage health insurers do not use individuals' information taken from the mandatory coverage. Moreover, despite an absence of community regulation health insurers do not adjust their premiums according to individual risk.[12] Indeed, tariffs are pre-determined according to the plan chosen by policyholders and they vary only according to age and gender. In the supplemental coverage market it seems that health insurers protect themselves against adverse selection just by excluding individuals with some pre-existing medical conditions.[13]

The fact that health insurers that supply supplemental coverage use very little information, indeed less information that which they could easily have access to, is important for our empirical strategy. At first glance, the supplemental health insurance market presents an institutional setting that differs from Medigap in the USA, due to a lack of community rating regulation. Nevertheless, the poor risk discriminatory schemes applied by those health insurers that supply supplemental coverage implies that the resulting information structure is similar to Fang et al. (2008). Thus, health status and prior healthcare utilization are not observed (or taken into account) by health insurers in the supplemental coverage market, allowing us to take advantage of such information for our empirical strategy. Finally, we include the disclaimer that we do not have an official statement to support the assertion that no information exchange takes place between the two levels of coverage. As a robustness check, we also perform a modified correlation test suggested in Finkelstein and Poterba (2014) in which it is assumed that some variables are observable by health insurers who choose not to use them in pricing. Since we observe all of the variables that come

[7] Olivella and Vera-Hernandez (2013) take advantage of collective *versus* individual private health insurance contracts to disentangle moral hazard from adverse selection.

[8] This premium is called the "*Unidad de Pago por Capitación*" (UPC).

[9] 95% Confidence Interval: 3.9–4.7. Information obtained from the 2011 Quality of Life Survey. Available at: https://www.dane.gov.co. No official statistics are available on enrollment in supplemental coverage.

[10] Some discounts are usually offered to policyholders to give them incentives to subscribe to supplemental coverage with the same insurer.

[11] This information is provided by ACEMI, the health insurers' federation.

[12] See Finkelstein and Poterba (2014) for an interesting discussion related to the rationale of observable but unused individual characteristics by health insurers.

[13] The webpage on which supplemental coverage plans are offered reveals that there exists a no-risk adjustment in premiums: https://prezi.com/pc9l3fnoolaj/otros-planes/

from the mandatory regime, we are able to tackle this alternative information structure.

3. Data

3.1. Information sources

The main source of information is the Integrated Information System of Social Protection (SISPRO), which integrates more than 10 primary sources of health-related information in a single query system.[14] SISPRO matches the citizen identification of each of the subjects included and this relates to the different sources of information. Using SISPRO we can build a database for 2010 and 2011 with information coming from three primary sources: BDUA, PILA and UPC.[15]

The BDUA corresponds to the database that contains sociodemographic information regarding those members fully identified with the different social security regimes. It is updated every month and is the main source for the transfer of the capitation payment by the government to health insurers for the mandatory coverage. From this database we retrieve all of the individuals' standard sociodemographic variables, as well as information related to membership status.

PILA is the source of information where the firms and the independent workers report their monthly income, with the objective of making contributions to the social security system in Colombia (health, retirement and other mandatory contributions). It is updated each month. From PILA we extract individuals' monthly incomes.

The UPC database contains information such as member ID and the characterization of the health service provided: place of delivery, delivery date, type of service provided, related diagnostic, health institution that provided the service, and the value paid by the health insurer. Every year, this information is supplied by all health insurers to the government and it is used to calculate the UPC. From UPC database we have information related to the consumption of healthcare services during a calendar year. All of these databases correspond to administrative data and are administered under high-quality standards.

Finally, the information panel was paired with information from two health insurers that offer the compulsory insurance as well as the supplemental coverage. Information for the two insurers is used to identify whether individuals benefited from supplemental coverage or not.[16] This information was available for the years 2010 and 2011. In short, we build a panel database that contains information on an individual level for each of the health expenses paid by the compulsory coverage along with related diagnoses, demographic information of the individual and whether they benefited from supplemental coverage.

The total for matched individuals between the base obtained from SISPRO and the contents of the membership status for supplemental health insurance was 484,005 individuals (6.4% of contributors in Colombia in 2011), which corresponds to the total of dues-paying members of the two health insurers included. From the UPC database we identify the cost of services used, the frequency of use between inpatient and outpatient services, the diagnosis associated with the services used and whether the individual is affiliated with supplemental coverage or not. In addition, the frequency of use of primary and secondary prevention services (cervical Pap smear, mammogram, PSA measurement and vaccination) are characterized by the codes of the Unified Classification of Health Procedures (CUPS). We use the International Statistical Classification of Diseases and Related Health Problems 10th Revision (ICD-10) to identify the diagnostics included. We use the codes for high blood pressure (hypertension), chronic kidney disease (CKD), diabetes mellitus (DM), cancer, pneumonia, urinary tract infection (UTI), preeclampsia, gastrointestinal disease, upper respiratory tract infection (URTI), chronic obstructive pulmonary disease (COPD), asthma and mental illness.

3.2. Descriptive statistics

As shown in Table 1, 3.17% of contributing members bought supplemental coverage. In general, the average age of the population was 37 years, being higher in affiliates of supplemental coverage (43 years on average). Likewise, there are more women than men in the study population. 40% of individuals in the database are female but this proportion is higher for individuals with supplemental coverage (54%). The average income is close to the legal minimum wage but is higher (17%) for individuals who benefit from supplemental coverage. In the Appendix A, we give more information about these variables.

It is also observed that small differences exist in the proportion of individuals with diseases between the groups with and without supplemental coverage, although there are proportionally fewer people with chronic diseases that are affiliated with supplemental coverage. This finding is due to the restrictive policy of supplemental coverage insurance companies for the enrollment of people with a pre-existing condition. Additionally, both the average consumption of health services (both ambulatory and hospital) and the average cost of services are higher in individuals affiliated with supplemental coverage (frequency of ambulatory services is 3.82 higher; frequency of hospital services is 1.2 higher; and average cost is 1.23 higher).

Finally, the proportion of individuals who have a Pap smear, mammography or PSA test is greater in the sample with supplemental coverage. Specifically, 1.2% of individuals with supplemental coverage have a Pap smear, 7.5% mammography and 8.4% PSA, while 0.3% have a Pap smear, 3.8% mammography and 3.9% PSA in the sample without supplemental coverage. Regarding vaccination, individuals can use different classes of vaccines. The use of vaccines depends on individual conditions, like pregnancy, previous exposure to a vaccine or infectious agent, work related exposure, chronic diseases, and so on. The descriptive information shows that 8.6% (n = 195,558) of the women use vaccines compared to 1.07% (n = 288,447) of the men. This difference is explained by the fact that pregnant women need to receive certain vaccines.[17] In particular, the average of frequency of vaccination use by individuals with supplemental coverage is 0.011 (SD 0.142), while this average is 0.030 (SD 0.254) in the sample of individuals without supplemental coverage (Table 1). Additionally, 63.3% of the women (who use vaccines) have received the Tetanus and Diphtheria vaccine, 23.8% the Flu vaccine (Influenza) and 12.9% other vaccines. In men, 75.8% have received the Flu vaccine, 8.4% the Yellow Fever vaccine, 3.9% Pneumococcal vaccines and 11.9% other vaccines.

[14] SISPRO handles quality standards and information validation, allowing for reliability with respect to those reported in each of the databases that SISPRO uses. More information: http://www.sispro.gov.co/

[15] BDUA refers to the Base de Datos Única de Afiliados; PILA to the Planilla Integrada de Pago de Aportes; and UPC to the Unidad de Pago por Capitación.

[16] The match between the two databases was performed for affiliated contributors; that is to say, beneficiaries were not included in the final database. The members of the contributory system are divided into contributors (the workers who are deducted a percentage of their salary for health) and beneficiaries (those who are economically dependent on the contributor, usually a spouse and/or children).

[17] See "Guidelines of the Colombian Health Ministry". Available at: http://gpc.minsalud.gov.co/gpc_sites/Repositorio/Conv_500/GPC_embarazo/gpc_embarazo.aspx

Table 1
Descriptive statistics of individuals.

	Total		Supplemental health insurance				
			Yes		No		
	Average	SD	Average	SD	Average	SD	P-value
Age	37,14	11,88	43,07	12,22	36,94	11,82	0.000
Female	0,40	0,49	0,54	0,50	0,40	0,49	0.000
Average monthly income 2010	589.296,90	692.114,90	684.314,70	989.596,60	585.999,70	679.221,00	0.000
HT	0,10	0,31	0,09	0,29	0,10	0,31	0.000
CKD	0,01	0,08	0,00	0,05	0,01	0,08	0.000
DM	0,02	0,15	0,03	0,16	0,02	0,14	0.000
Cancer	0,02	0,13	0,01	0,12	0,02	0,13	0.000
Pneumonia	0,01	0,08	0,01	0,07	0,01	0,08	0.107
UTI	0,06	0,24	0,05	0,21	0,06	0,24	0.000
Preeclampsia	0,00	0,05	0,00	0,02	0,00	0,05	0.000
Gastrointestinal disease	0,12	0,32	0,14	0,35	0,12	0,32	0.000
HRTI	0,06	0,24	0,04	0,20	0,06	0,25	0.000
COPD	0,01	0,10	0,01	0,11	0,01	0,10	0.023
Asthma	0,01	0,12	0,01	0,10	0,01	0,12	0.000
Mental disorder	0,00	0,05	0,00	0,05	0,00	0,04	0.035
Use of hospital services 2011	0,48	8,96	1,70	16,49	0,44	8,60	0.000
Use of ambulatory services 2011	8,21	12,25	9,87	14,67	8,16	12,16	0.000
Use of hospital services 2010	0,42	9,37	1,78	33,79	0,38	7,29	0.000
Use of ambulatory services 2010	7,43	11,19	9,09	13,93	7,37	11,08	0.000
Spending 2010	486.469,30	4.399.673,00	597.543,00	3.730.686,00	482.646,30	4.420.851,00	0.002
Spending 2011	557.202,50	4.123.044,00	672.578,60	2.949.511,00	553.231,30	4.157.489,00	0.001
Pap smear	0.003	0.063	0.012	0.115	0.003	0.060	0.000
Mammography	0.039	0.201	0.075	0.270	0.038	0.198	0.000
PSA	0.041	0.238	0.084	0.366	0.039	0.232	0.000
Vaccination	0.029	0.251	0.011	0.142	0.030	0.254	0.000
Observations	484.005		15.376		468.629		

SD: Standard Deviation; HT: Hypertension; CKD: Chronic Kidney Disease, DM; Diabetes Mellitus; UTI: Urinary Tract Infection; URTI: Upper Respiratory Tract. Infection; COPD: Chronic Obstructive Pulmonary Disease.
F-statistic for joint test of regression = 214.10 (p = 0.000).
The Appendix A has the description for all variables.

4. Empirical strategy

This paper aims to identify the nature of the selection that dominates in the voluntary health insurance market in Colombia. To this end, we first apply the same empirical strategy as Fang et al. (2008). Nevertheless, due to the fact that health spending is linked to the price that the insurer pays each provider, and that this price is affected by many factors that cannot be observed, the previous consumption of health services is considered. We assume that the frequency of healthcare consumption in 2011 is a function of having bought supplemental coverage, socio-demographic characteristics, health status and previous health consumption of the individual. Using ordinary least squares, we estimate the following equation:

$$Y_{i,t} = \beta_0 + \beta_1 MP_{i,t} + \beta_2 X_i + \beta_3 Dx_i + \beta_4 Y_{i,t-1} + \mu_i, \quad (1)$$

where $Y_{i,t}$ is the consumption of health services for individual i for 2011, which is measured in frequency use, $MP_{i,t}$ is a dichotomous variable that indicates whether the individual is affiliated or not in 2011 to a supplemental coverage and $X_i$is a vector of the individual's characteristics. In particular, $X_i$includes age and gender, which constitute the individual characteristics used by health insurers for pricing.[18] Dx_i measures the state of health of the individual through the diagnosis of certain trace pathologies. More specifically, Dx_i includes chronic pathologies sometimes used by insurers to avoid the enrollment of individuals characterized by some pre-existing medical conditions, and other acute pathologies. $Y_{i,t-1}$is the frequency of healthcare consumption in the previous year. It is worth noting that $Y_{i,t-1}$ is available to the econometrician but is not observed, or at least not used, by health insurers for pricing.[19] In this estimation, our parameter of interest is β_1. Its sign determines the correlation between having bought supplemental coverage and the consumption of health services.

Next, we estimate the following equation:

$$Ln\left(\frac{Pr(MP_{i,t}=1)}{1-Pr(MP_{i,t}=1)}\right) = \alpha_0 + \alpha_1 Dx_{i,t-1} + \alpha_2 P_{i,t-1} + \alpha_3 Y_{i,t-1} + \alpha_4 X_i + \varepsilon_i, \quad (2)$$

where $Pr(MP_{i,t}=1)$ is the probability that the individual is affiliated to supplemental coverage and X_i is the same variable as in (1). $Dx_{i,t-1}$ and $Y_{i,t-1}$ also denote the same variables as in (1), but taken in *t-1*. $P_{i,t-1}$, which is not observed by health insurers, measure the activities of primary and secondary prevention that the individual performs during the previous year.[20] Thus Eq. (2) is twofold: i) it allows us to identify the nature of the existing selection ruling out the presence of moral hazard; ii) it describes the characteristics of individuals who purchase supplemental insurance. The estimation for (2) is performed using a logit model.

As pointed out in Finkelstein and Mc Garry (2006), "*conditional on the information set used by the insurance company, the existence of any individual characteristic that is observed by the econometrician, but not by the insurer and that is correlated with both insurance coverage and risk occurrence indicates the presence of asymmetric information. This result is true regardless of the sign of the correla-*

[18] Additionally, X_i includes variables observed but not used by the insurer, like municipality of residence and monthly income.

[19] In section 5.4, we perform the modified positive correlation test suggested in Finkelstein and Poterba (2014) in order to take into account the possibility that health insurers observe but does not use this variable.

[20] Through information regarding services consumed, four types of service related to activities of primary and secondary prevention were identified: vaccination services, PSA, mammography and Pap smear.

tion." Beyond the fact that Eq. (2) allows us to rule out moral hazard effects, we take advantage of some of the information related to prevention activities undertaken by individuals in order to identify various channels that explain the selection nature at work in the supplemental health insurance market.

5. Results

This section presents the results of the empirical estimation and is divided into three sub-sections. The first shows the results related to the estimation of Eq. (1), which represents the positive correlation test (Chiappori and Salanié, 2000). The second sub-section describes the effects of the relationship between health status and consumption in the previous year on membership to supplemental coverage (Eq. 2). The third sub-section presents the characteristics of individuals who buy supplemental coverage, especially those related to the activities of primary and secondary prevention. The fourth sub-section provides a series of robustness checks in order to strengthen our interpretation.

5.1. Correlation test

Following Fang et al. (2008), Table 2 presents the estimation of Eq. (1) using ordinary least squares where the dependent variable is the frequency of healthcare services consumption. The coefficient associated with having bought supplemental coverage in 2011 remains positive despite the inclusion of individuals' characteristics used for pricing (coefficient = 0.719, $p < 0.01$) and other variables perhaps observed, but not used by health insurers. Regarding the other coefficients, these are generally consistent between different estimates (coefficient = 0571, $p < 0.01$); the presence of diseases correlates with increased consumption frequency of health services, and the same occurs for age and being female. At first glance, these results suggest the presence of adverse selection.

However, as eloquently pointed out in Chiappori and Salanié (2000), one has to be cautious before concluding that adverse selection exists.[21] In particular, the aforementioned authors argue that this positive correlation may be caused by nonlinearities in either the pricing schedule or the accident probabilities. Such nonlinearities imply that the linear functional forms are a valid test only in a homogeneous population.

Another usual difficulty where a positive correlation is obtained for health insurance contracts is that some moral hazard may also be present. Even though this difficulty remains true in general, the magnitude of *ex post* moral hazard found in several empirical works dealing with health insurance requires that this concern be tackled.[22]

In the subsequent sections we perform several tests in order to corroborate adverse selection. First, we take advantage of our panel dimension to disentangle moral hazard behaviors from the adverse selection phenomenon (Section 5.2). Second, thanks to various data related to prevention decisions undertaken by policyholders, we attempt to determine the channels that may underlie this adverse selection result (Section 5.3). Third, we subject our results to additional robustness checks: i) we resume our linear estimation that we restrict to more homogeneous groups, and ii) we take advantage of unused observable information to perform the modified correlation test suggested in Finkelstein and Poterba (2014) (Section 5.4).

[21] "*In other words, we believe that a negative conclusion, like the one we get, is probably much more robust than any positive finding could be.*" Chiappori and Salanié (2000).

[22] See for instance Vera-Hernandez (2003).

5.2. Nature of the selection

Ex post moral hazard corresponds to the increase of health expenditure and the increase in frequency of use of healthcare caused by health insurance coverage (Pauly, 1968). Because we observe individuals before they purchase supplemental insurance, we can rule out moral hazard effects. Thus the sign associated with $Y_{i,t-1}$ in Eq. (2) determines which type of selection is predominant in the supplemental health insurance market.

Table 3 presents the estimation of Eq. (2) to show the correlation between the frequency of use of healthcare services in the period preceding membership year (in 2010) and the probability of enrollment to supplemental coverage during 2011. This estimate allows us to eliminate the effect of moral hazard (increased consumption of services frequency after insuring), so that the whole sample is used (panel A – left) and the sample excludes individuals who had supplemental health insurance in 2010 (panel B – right).

We observe that the frequency of use of hospital services in 2010 is positively correlated with the purchase of supplemental coverage in 2011 for all models. More precisely, the increase in one unit of hospital services consumption frequency in 2010 from the mean increases the probability of enrollment to supplemental coverage in 2011 by 0.006 percentage points (pp) ($p = 0.000$) for the full sample of individuals (marginal effect (ME) of parameter in column 4), while this effect is equal to 0.008 pp ($p = 0.000$) for individuals who did not have supplemental coverage in 2010 (ME in column 8). Again, it is worth noting that the former coefficient may include some moral hazard effect due to individuals that benefit from supplemental coverage in 2010, while the latter is free from moral hazard and only indicates the selection nature.

Additionally, in the models that include all control variables (columns 4 and 8) the frequency of use of ambulatory services is positively correlated with the purchase of supplemental coverage in 2011.[23] In particular, the increase in one unit of ambulatory services consumption frequency in 2010 from the mean increases the probability of enrollment to supplemental coverage in 2010 by 0.004 pp ($p = 0.02$) (ME in column 4). Finally, pathologies are negatively related to the purchase of supplemental insurance. As mentioned above, this is explained by the fact that health insurers use some chronic pathologies as pre-existing medical conditions to protect themselves against adverse selection.

All in all, **this finding suggests the presence of adverse selection**, where individuals have private information that insurers cannot observe and which increases the likelihood of insurance purchase.[24] This result is consistent despite the presence of pre-existing medical conditions applied by the health insurers. Thus from this result we may conclude that risk selection strategies implemented by health insurers through pre-existing medical conditions are not sufficient by themselves in preventing adverse selection.

[23] The coefficient in column 8 is significant at $p < 0.2$.

[24] An alternative interpretation would be a selection on moral hazard, as seen in Einav et al. (2013). For hospital services, it would be unlikely that individuals consume more medical services in 2010 simply due to supplemental coverage in 2011, since individuals do not really have decision-making power over the frequency of hospital services consumed. In contrast, the selection on moral hazard is a more plausible interpretation for ambulatory services consumption. Nevertheless, even though we find that the frequency of use of ambulatory services in 2010 is positively correlated with the purchase of supplemental coverage in 2011 on the one hand, and that the frequencies of use of ambulatory services in 2010 and 2011 are positively correlated on the other hand (Table 2, column 8, panel B), we cannot establish that we have moral hazard in ambulatory healthcare consumption in 2011. Therefore, we prefer to maintain our more conservative interpretation of adverse selection.

Table 2
Effect of Supplemental Health Insurance (SHI) on the frequency of use of healthcare services (Eq. (1)).

	Panel A. Hospital services in 2011				Panel B. Ambulatory services in 2011			
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SHI 2011	**1.259*** **	**0.875*** **	**0.719*** **	**0.571*** **	**1.718*** **	**0.810*** **	**-0.386*** **	**0.216** **
	(0.134)	(0.141)	(0.150)	(0.145)	(0.120)	(0.121)	(0.128)	(0.103)
SHI 2010		1.239***	1.037***	0.812***		2.929***	0.152	0.581***
		(0.147)	(0.147)	(0.139)		(0.127)	(0.121)	(0.0953)
Age			0.0228***	0.0104***			0.228***	0.0418***
			(0.00200)	(0.00207)			(0.00200)	(0.00143)
Female			0.324***	0.137***			5.258***	2.139***
			(0.0317)	(0.0317)			(0.0404)	(0.0346)
AMI 2010			-0.163***	-0.0792*			-1.481***	-0.453***
			(0.0475)	(0.0461)			(0.0649)	(0.0517)
Hospital services 2010				0.190***				0.0193***
				(0.0209)				(0.00480)
Ambulatory services 2010				0.0211***				0.531***
				(0.00319)				(0.00405)
HT				-0.265***				3.782***
				(0.0700)				(0.0932)
CKD				0.167				6.427***
				(0.489)				(0.321)
DM				1.215***				3.314***
				(0.210)				(0.147)
Cancer				0.870***				3.580***
				(0.216)				(0.153)
Pneumonia				2.457***				2.385***
				(0.595)				(0.265)
ITU				0.342***				2.454***
				(0.100)				(0.0807)
Preeclampsia				1.620***				0.298
				(0.575)				(0.652)
Gastrointestinal disease				0.139**				2.818***
				(0.0557)				(0.0754)
URTI				-0.396***				1.998***
				(0.0534)				(0.0662)
COPD				0.984***				3.895***
				(0.302)				(0.229)
Asthma				0.127				2.181***
				(0.206)				(0.158)
Mental disorder				1.811***				4.683***
				(0.701)				(0.471)
Constant	0.443***	0.414***	1.632***	0.741	8.155***	8.086***	17.60***	6.461***
	(0.0126)	(0.0123)	(0.620)	(0.600)	(0.0178)	(0.0178)	(0.856)	(0.684)
Municipality control	NO	NO	YES	YES	NO	NO	YES	YES
Observations	484,005	484,005	409,905	409,905	484,005	484,005	409,905	409,905
R-squared	0.001	0.001	0.002	0.049	0.001	0.002	0.088	0.404

Robust standard errors *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.
SHI: Supplemental Health Insurance; AMI: Average Monthly Income; HT: Hypertension; CKD: Chronic Kidney Disease, DM; Diabetes Mellitus; UTI: Urinary Tract Infection; URTI: Upper Respiratory Tract Infection; COPD: Chronic Obstructive Pulmonary Disease.

5.3. Activities of prevention

Now that our adverse selection results seem to be confirmed by estimations presented in Section 5.2, we attempt to identify some potential channels. Table 4 shows the association of the activities of primary and secondary prevention of individuals in the previous year with the chance to buy supplemental health insurance in 2011. This analysis is performed for the sample of men and women separately due to prevention activities being related exclusively to gender. It is noted that, controlling for socio-demographic and health characteristics, in 2010 practicing one cervical Pap smear or having the PSA test (secondary prevention) are correlated with an increased likelihood of buying supplemental health insurance in 2011 (columns 1–8). More precisely, the use of the Pap smear in 2010 increases the probability of buying supplemental health insurance in 2011 in 2.10 pp ($p = 0.000$) (ME in column 6), while PSA increases the probability in 2.76 pp ($p = 0.000$) (ME in column 8). This finding contrasts with that of vaccination (primary prevention), where vaccination is correlated with a decrease in the probability of buying supplemental health insurance in the previous year.

Regarding the relation between secondary prevention activities and the decision to buy supplemental health insurance, it must be said that, at this stage, two interpretations remain plausible. On the one hand, secondary prevention activities may indicate individuals' risk aversion. In other words, individuals characterized by

Table 3
Characteristics of individuals who had Supplemental Health Insurance (SHI) in 2011 (Equation (2)).

	Panel A. Full sample				Panel B. Sample without individuals who had supplemental health insurance in 2010			
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Hospital services 2010	**0.00458*** **	**0.00317*** **	**0.00322*** **	**0.00325*** **	**0.00422*** **	**0.00422*** **	**0.00429*** **	**0.00436*** **
	(0.000578)	(0.000589)	(0.000609)	(0.000633)	(0.000656)	(0.000656)	(0.000681)	(0.000618)
Ambulatory services 2010	0.0103***	0.00447***	-0.00416***	0.00194**	0.00336***	0.00336***	-0.00677***	0.00162
	(0.000597)	(0.000678)	(0.000836)	(0.000871)	(0.000873)	(0.000873)	(0.00112)	(0.00113)
SHI 2010		3.009***	2.809***	2.762***				
		(0.0197)	(0.0212)	(0.0214)				
Age			0.0241***	0.0279***			0.0226***	0.0271***
			(0.000723)	(0.000776)			(0.000847)	(0.000912)
Female			0.376***	0.386***			0.549***	0.559***
			(0.0192)	(0.0194)			(0.0228)	(0.0230)
AMI 2010			0.817***	0.791***			1.021***	0.993***
			(0.0271)	(0.0271)			(0.0282)	(0.0281)
HT				-0.509***				-0.630***
				(0.0345)				(0.0443)
CKD				-0.987***				-1.402***
				(0.181)				(0.288)
DM				0.00667				0.125*
				(0.0613)				(0.0756)
Cancer				-0.404***				-0.520***
				(0.0746)				(0.0975)
Pneumonia				-0.179				-0.100
				(0.131)				(0.151)
ITU				-0.334***				-0.414***
				(0.0424)				(0.0511)
Preeclampsia				-1.072***				-1.066***
				(0.369)				(0.381)
Gastrointestinal disease				0.0253				-0.0146
				(0.0273)				(0.0332)
URTI				-0.400***				-0.462***
				(0.0444)				(0.0533)
COPD				-0.0282				-0.0857
				(0.0897)				(0.118)
Asthma				-0.241**				-0.202*
				(0.0942)				(0.106)
Mental disorder				-0.0370				-0.0392
				(0.182)				(0.236)
Constant	-3.504***	-3.835***	-15.65***	-15.41***	-3.827***	-3.827***	-18.36***	-18.12***
	(0.00975)	(0.0114)	(0.363)	(0.362)	(0.0121)	(0.0121)	(0.379)	(0.377)
Municipality control	NO	NO	YES	YES	NO	NO	YES	YES
Observations	484,005	484,005	409,905	409,905	467,777	467,777	393,677	393,677

Robust standard errors *** p < 0.01, ** p < 0.05, * p < 0.1.
SHI: Supplemental health insurance; AMI: Average Monthly Income; HT: Hypertension; CKD: Chronic Kidney Disease, DM; Diabetes Mellitus; UTI: Urinary Tract Infection; URTI: Upper Respiratory Tract Infection; COPD: Chronic Obstructive Pulmonary Disease.

higher risk aversion levels may have a higher willingness to buy supplemental health insurance and, for the same reason, they may be more engaged in secondary prevention activities. On the other hand, it is likely that individuals who have received a signal that indicates that they may face a higher health risk are more willing to opt for secondary prevention activities. While in the first case secondary prevention activities can be viewed as a proxy for risk aversion; the second case constitutes a channel by which to explain our adverse selection outcome.

We thus explore the relation between secondary prevention activities and health risk occurrence. To do this, we estimate the correlation between secondary prevention activities consumed during 2010 and individuals' health risk occurrence in 2011. To approximate individuals' health risk occurrence we use two variables: the probability of being hospitalized in the intensive care unit (ICU) and the number of days spent in hospital. For the first variable we apply a logit model, while the second dependent variable is treated through a standard OLS. Moreover, we perform these two estimations for individuals who did not present chronic disease during 2010 (*i.e.* for individuals who were not characterized by a severe health risk while they were consuming secondary prevention activities), and for individuals without supplemental coverage.

Table 5 presents the estimation of the logit and OLS models, respectively. In panel A (left), we find that individuals who consume secondary prevention activities during 2010 (Pap smear for women and PSA for men) are more likely to be hospitalized in ICU services in 2011. More precisely, the use of the Pap smear in 2010 increases the probability of admission to ICU services in 2011 in 0.29 pp ($p = 0.000$) (ME in column 3) and the use of PSA increases the probability in 0.04 pp ($p = 0.001$) (ME in column 4).[25] The results presented in Panel B (right) go into the same direction: a positive correlation between secondary prevention activities (PSA and

[25] As expected, the sign of the correlation is negative for primary prevention (we restricted the estimation for women due to the low variation of vaccination services for men). The higher the frequency of vaccination during 2010, the lower is the probability of being hospitalized in emergency services in 2011.

Table 4
Activities of primary and secondary prevention in 2010 and relation to supplemental health insurance purchase in 2011.

	Panel A. FULL SAMPLE				Panel B. Sample without individuals who had supplemental health insurance in 2010			
	FEMALES		MALES		FEMALES		MALES	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Pap smear 2010	**0.824*** **	**0.783*** **			**0.908*** **	**0.885*** **		
	(0.0737)	(0.0747)			(0.0865)	(0.0876)		
Mammography 2010	−0.0529	−0.0397			−0.0870*	−0.0558		
	(0.0350)	(0.0357)			(0.0487)	(0.0497)		
PSA 2010			**0.255*** **	**0.286*** **			**0.231*** **	**0.276*** **
			(0.0269)	(0.0282)			(0.0390)	(0.0406)
Vaccination 2010	**−0.946*** **	**−0.857*** **	**−0.313** **	−0.0278	**−1.002*** **	**−0.919*** **	−0.106	0.119
	(0.106)	(0.102)	(0.157)	(0.145)	(0.125)	(0.122)	(0.190)	(0.184)
Age	0.0300***	0.0362***	0.0410***	0.0493***	0.0132***	0.0190***	0.0276***	0.0349***
	(0.00103)	(0.00116)	(0.00125)	(0.00135)	(0.00133)	(0.00150)	(0.00159)	(0.00171)
AMI 2010	0.985***	0.933***	0.868***	0.809***	1.265***	1.214***	1.064***	1.011***
	(0.0399)	(0.0399)	(0.0424)	(0.0427)	(0.0458)	(0.0458)	(0.0496)	(0.0499)
Hospital services 2010		0.00608***		0.00333***		0.00528***		0.00341***
		(0.000869)		(0.000742)		(0.000977)		(0.000799)
Ambulatory services 2010		0.00389***		0.00292*		0.00165		0.000742
		(0.00110)		(0.00162)		(0.00158)		(0.00217)
Constant	−17.26***	−16.69***	−16.74***	−16.18***	−20.72***	−20.11***	−19.20***	−18.68***
	(0.538)	(0.537)	(0.566)	(0.570)	(0.619)	(0.619)	(0.662)	(0.665)
Municipality control	YES	YES	YES	YES	YES	YES	YES	YES
Diagnosis control	NO	YES	NO	YES	NO	YES	NO	YES
Observations	140,476	140,476	170,561	170,561	132,879	132,879	165,248	165,248

Robust standard errors *** p < 0.01, ** p < 0.05, * p < 0.1.
AMI: Average Monthly Income; PSA: Prostate-Specific Antigen.

Table 5
Activities of primary and secondary prevention and relation to probability of admission to ICU and days of hospital stay in 2011.

	Panel A. ICU ADMISSION IN 2011				Panel B. HOSPITAL STAY IN 2011			
	FULL SAMPLE		Sample without individuals who had supplemental health insurance		FULL SAMPLE		Sample without individuals who had supplemental health insurance	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Pap smear 2010	0.948***		1.013***		0.121***		0.127***	
	(0.204)		(0.227)		(0.0418)		(0.0444)	
Mammography 2010	−0.213		−0.0279		0.0120		0.0271	
	(0.196)		(0.205)		(0.0261)		(0.0289)	
PSA 2010		0.288***		0.314***		0.0657***		0.0556***
		(0.0773)		(0.0940)		(0.0157)		(0.0168)
Vaccination 2010	−0.585**		−0.654**		0.000829	−0.0236***	−0.0310***	−0.0210***
	(0.244)		(0.285)		(0.0335)	(0.00392)	(0.00449)	(0.00398)
Age	−0.0144**	0.0479***	−0.0187***	0.0451***	0.00170**	0.00222***	0.00121**	0.00167***
	(0.00589)	(0.00489)	(0.00660)	(0.00522)	(0.000713)	(0.000288)	(0.000614)	(0.000286)
AMI 2010	0.0670	−0.338	0.0112	−0.268	−0.00734	0.124	−0.0202	0.140
	(0.194)	(0.222)	(0.237)	(0.241)	(0.0175)	(0.0979)	(0.0134)	(0.109)
Constant	−6.090**	−3.949	−5.298*	−4.791	0.126	−1.686	0.304*	−1.882
	(2.560)	(2.949)	(3.131)	(3.196)	(0.236)	(1.297)	(0.176)	(1.441)
Municipality control	YES	YES	YES	YES	YES	YES	YES	YES
Diagnosis control	YES	YES	YES	YES	YES	YES	YES	YES
Observations	127,573	202,778	116,657	192,309	127,629	203,550	116,708	193,990

Robust standard errors in parentheses
*** p < 0.01, ** p < 0.05, * p < 0.1
AMI: Average Monthly Income; PSA: Prostate-Specific Antigen.
Panel A. Logit model for probability to admission to ICU in 2011.
Panel B. OLS model for days of hospital stay in 2011.
Columns 1, 3, 5 and 7: sample of women.
Columns 2, 4, 6 and 8: sample of men.

cytology) and the number of days spent in hospital on the one hand, and a negative correlation with primary prevention activities on the other hand.

Again, these findings corroborate the result of adverse selection: health insurers do not observe the frequency of secondary prevention consumption before the supplemental health insurance purchase, that is to say, this information is only known by the individual before affiliation to supplemental insurance. Since individuals with some predispositions for some diseases are more willing to undertake secondary prevention actions,[26] the positive correlation between the use of secondary prevention services and the purchase of supplemental insurance indicates the presence

[26] For example, consider a case where the individual has a family history of prostate cancer or uterine cervical cancer. This signal makes the individual more prone to the use of secondary activities for prevention to aid in early detection, thus allowing for more effective treatment.

of adverse selection. Furthermore, in the case of positive results, knowledge of the results of these tests can also make the individuals more prone to affiliation with broader coverage.[27]

5.4. Additional robustness checks

In this section we provide some robustness checks in order to check our adverse selection results – both applied to the estimation of Eq. (2). First, we restrict this estimation to more homogeneous groups. Second, we take advantage of some observable but unused variables to perform the modified positive correlation test suggested in Finkelstein and Poterba (2014).

5.4.1. Homogeneous groups

According to Chiappori and Salanié (2000), when one obtains a positive correlation result, it is prudent to perform this test on homogeneous groups before concluding that asymmetric information is present. Thus we run the same estimation for more homogenous groups according to age and gender. Regarding the age variable, we consider two groups: group-1 for individuals between 18 and 50 years, and group-2 for individuals older than 50 years.[28]

Table 6 presents the results of the same estimation as in Table 3 (Eq. (2)), but this time applied to group-1 (panel A – left) and group-2 (panel B – right). Again, the parameter of interest is associated with the frequency of use of hospital services during 2010. In all models, this estimator remains positive and statistically significant with a stronger effect for women. More precisely, the marginal effect of the frequency of use of hospital services in 2010 on the probability of being affiliated to supplemental insurance is 0.016 pp (p = 0.000) (column 2) for women in the full sample, while this marginal effect is equal to 0.015 pp (p = 0.000) for women without supplemental insurance in 2010 (column 5). For men, these two marginal effects are equal to 0.004 pp (p = 0.042) (column 3) and 0.006 pp (p = 0.000) (column 6), respectively. Additionally, we observe that these marginal effects are stronger for group-1 (panel A) than for group-2 (panel B).

These results corroborate our findings obtained in Section 5.2.: ruling out moral hazard effects and considering more homogeneous groups, individuals characterized by a higher probability of buying supplemental coverage in 2011 tend to consume with higher frequency hospital services during 2010. Again, these results are in line with an adverse selection interpretation.

Table 7 replicates the estimation of Table 4 for individuals without a previous chronic disease during 2010 and exhibits these estimations for the individuals of group-1 (panel A) and group-2 (panel B). Our parameters of interest are those associated with secondary prevention (Pap smear and PSA). As in Table 4, these parameters are positive and statistically significant. More precisely, we observe that the marginal effect of use of Pap smear in 2010 on the probability of buying a supplemental coverage in 2011 is lower for group-1 than for group-2: 2.1 pp *versus* 4.4 pp, respectively (column 3 *versus* column 7). Nevertheless, for men, this difference is much smaller (0.62 pp *versus* 0.68 pp, respectively).

Again, these findings corroborate the channel depicted in Section 5.3 to sustain our adverse selection interpretation; that is, for

[27] Moreover, when primary prevention activities are positively correlated with the purchase of insurance, this indicates the existence of advantageous selection (lower-risk individuals are those who prefer higher coverage), on the other hand negative correlation indicates adverse selection. We find that the correlation between vaccination in the previous year and supplemental health insurance remains negative across all estimates.

[28] We apply 50 years as the threshold due to the fact that guidelines recommended by the Colombian Health Ministry use it for PSA and cytology. See (in Spanish): http://gpc.minsalud.gov.co/SitePages/default.aspx

Table 6
Hospital services use in 2010 and relation to supplemental health insurance purchase in 2011 in sample of individuals according to age.

	Panel A. Individuals between 18 and 50 years old						Panel B. Individuals over 50 years					
	Full sample			Sample without individuals who had supplemental health insurance in 2010			Full sample			Sample without individuals who had supplemental health insurance in 2010		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Hospital services 2010	0.00456***	0.00592***	0.00340**	0.00805***	0.00648***	0.00513***	0.00229***	0.00502***	0.000516	0.00276***	0.00471***	0.00105
	(0.000974)	(0.000908)	(0.00167)	(0.00125)	(0.000973)	(0.00102)	(0.000798)	(0.00127)	(0.000572)	(0.00107)	(0.00155)	(0.00170)
Ambulatory services 2010	−0.000822	−0.000311	−0.00210	−0.00282	−0.00216	−0.00374	0.00575***	0.00415***	0.00818***	0.00897***	0.00790***	0.0108***
	(0.00135)	(0.00154)	(0.00261)	(0.00175)	(0.00193)	(0.00326)	(0.00115)	(0.00148)	(0.00177)	(0.00144)	(0.00191)	(0.00218)
SHI 2010	2.661***	2.348***	3.078***				2.959***	2.946***	2.969***			
	(0.0270)	(0.0355)	(0.0397)				(0.0365)	(0.0514)	(0.0520)			
Age	0.0380***	0.0326***	0.0438***	0.0365***	0.0301***	0.0465***	0.00388	−0.00533	0.0160***	−0.00424	−0.0186***	0.0122**
	(0.00138)	(0.00185)	(0.00205)	(0.00162)	(0.00206)	(0.00228)	(0.00248)	(0.00339)	(0.00371)	(0.00379)	(0.00567)	(0.00524)
Female	0.491***			0.687***			0.147***			0.205***		
	(0.0232)			(0.0278)			(0.0363)			(0.0474)		
AMI 2010	0.770***	0.844***	0.684***	1.025***	1.049***	0.809***	0.762***	0.939***	0.670***	1.060***	1.469***	0.860***
	(0.0315)	(0.0455)	(0.0437)	(0.0371)	(0.0465)	(0.0443)	(0.0538)	(0.0949)	(0.0647)	(0.0568)	(0.0895)	(0.0658)
Constant	−15.50***	−15.71***	−14.68***	−18.88***	−18.31***	−16.41***	−13.69***	−15.37***	−13.18***	−17.19***	−21.64***	−15.45***
	(0.419)	(0.608)	(0.579)	(0.494)	(0.623)	(0.585)	(0.742)	(1.298)	(0.902)	(0.804)	(1.279)	(0.944)
Municipality control	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Diagnosis control	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	334,729	134,392	200,337	293,580	128,132	195,830	75,176	33,115	42,061	69,298	30,014	39,284

Robust standard errors in parentheses
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
AMI: Average Monthly Income; SHI: Supplemental Health Insurance.
Panel A. Logit model for probability to SHI purchase in 2011 in individuals between 18 and 50 years old.
Panel B. Logit model for probability to SHI purchase in 2011 in individuals over 50 years.
Columns 1, 4, 7 and 10: full sample (women and men together).
Columns 2, 5, 8 and 11: sample of women.
Columns 3, 6, 9 and 12: sample of men.

Table 7
Activities of primary and secondary prevention in 2010 and relation to supplemental health insurance purchase in 2011 in individuals who did not present chronic disease during 2010 according to age.

	Panel A. Individuals between 18 and 50 years old				Panel B. Individuals over 50 years			
	Full sample		Sample without individuals who had supplemental health insurance in 2010		Full sample		Sample without individuals who had supplemental health insurance in 2010	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Pap smear 2010	0.783***		0.819***		0.999***		1.281***	
	(0.0938)		(0.110)		(0.206)		(0.246)	
Mammography 2010	0.0213		−0.0126		0.0146		−0.108	
	(0.0655)		(0.0824)		(0.0573)		(0.0868)	
PSA 2010		0.562***		0.468***		0.266***		0.240***
		(0.0578)		(0.0743)		(0.0385)		(0.0569)
Vaccination 2010	−0.829***	0.0676*	−0.863***	0.0838**	−1.557***	−0.729*	−1.876*	−1.170*
	(0.118)	(0.0398)	(0.141)	(0.0415)	(0.559)	(0.442)	(0.977)	(0.650)
Age	0.0382***	0.0535***	0.0268***	0.0436***	0.0242***	0.0364***	−0.0185**	0.00679
	(0.00193)	(0.00206)	(0.00227)	(0.00243)	(0.00402)	(0.00407)	(0.00718)	(0.00639)
AMI 2010	0.815***	0.654***	1.030***	0.776***	1.016***	0.610***	1.483***	0.818***
	(0.0443)	(0.0401)	(0.0493)	(0.0447)	(0.0801)	(0.0570)	(0.103)	(0.0681)
Hospital services 2010	0.00909***	0.00506**	0.00809***	0.0112***	0.00456**	0.00223*	0.00463*	0.00396
	(0.00180)	(0.00246)	(0.00181)	(0.00404)	(0.00206)	(0.00115)	(0.00255)	(0.00275)
Ambulatory services 2010	0.00211	−0.00716*	0.00174	−0.00890**	0.00777***	0.0129***	0.0100***	0.0144***
	(0.00186)	(0.00366)	(0.00232)	(0.00418)	(0.00179)	(0.00225)	(0.00252)	(0.00293)
Constant	−15.19***	−14.27***	−17.95***	−15.87***	−17.25***	−12.95***	−21.80***	−14.69***
	(0.592)	(0.530)	(0.661)	(0.589)	(1.118)	(0.805)	(1.490)	(0.999)
Municipality control	YES	YES	YES	YES	YES	YES	YES	YES
Diagnosis control	YES	YES	YES	YES	YES	YES	YES	YES
Observations	108,218	175,363	103,218	171,343	19,275	28,005	17,213	25,902

Robust standard errors in parentheses
*** p < 0.01, ** p < 0.05, * p < 0.1
AMI: Average Monthly Income; PSA: Prostate-Specific Antigen.
Panel A. Logit model for probability to SHI purchase in 2011 in individuals between 18 and 50 years old.
Panel B. Logit model for probability to SHI purchase in 2011 in individuals over 50 years.
Columns 1, 3, 5 and 7: sample of women.
Columns 2, 4, 6 and 8: sample of men.

individuals without chronic disease, the probability of buying supplemental coverage in 2011 increases in the secondary prevention actions undertaken during 2010.

5.4.2. *Testing for asymmetric information using unused observables (Finkelstein and Poterba, 2014)*

As we pointed out in Section 2, where we describe the Colombian healthcare system, it is reasonable to think that some variables are observable by health insurers who choose not to use them in pricing. Thus we perform the modified positive correlation test suggested in Finkelstein and Poterba (2014) in order to take into account this slightly modified information structure. According to these authors, this modified positive correlation test consists in using variables that are both demand-related and correlated with individuals' health risk.

It is plausible that the frequency of use of hospital services during 2010 satisfies this condition. Indeed, since individuals usually buy their mandatory and supplemental coverage from the same insurer, it is likely that this information is available and costless for health insurers. This modified positive correlation test involves two estimations.[29] The first estimates the correlation between frequency of use of hospital services in 2010 (observed but unused variable) and the probability of buying supplemental coverage in 2011. The results of this first estimation are already provided in Table 3, which exhibits positive and statistically significant coefficients (Columns 1–8).

The second equation estimates the correlation between an unobserved variable and individual health risk. In Table 8 we present the results obtained using an OLS regression that estimates the correlation between the frequency of use of hospital services in 2010 and the same variable in 2011 (dependent variable). Again, under this different specification we obtain positive and statistically significant coefficients (columns 1–8) that strengthen our adverse selection interpretation.

6. Conclusions

Our results suggest that adverse selection is predominant in the Colombian supplemental health insurance market. Our empirical strategy allows us to rule out moral hazard effects. Our results are corroborated by findings related to the behavior of individuals through the consumption of secondary prevention. Risk selection strategies used by insurers based on pre-existing medical conditions do not prevent adverse selection when the risk adjustment use for pricing does not incorporate some health status variables.

The nature of selection in the health insurance market (that is, what kind of individuals are more likely to purchase insurance) is of great importance for both the design of public policies related to insurance and the optimal design of policies in the private sector. In particular, we believe that our results have two strong policy implications. First, the adverse selection outcome implies that the consumption of supplemental coverage is sub-optimal, and that potential government intervention that encourages supplemental coverage consumption would be a step in the right direction. However, though our insights indicate the direction for a welfare-improving government intervention, we lack sufficient information

[29] Following Finkelstein and Poterba (2014), the unused observables test can be formalized using the following equations: (1) $C_i = X_i\beta + W_i\alpha + \varepsilon_i$; and (2) $L_i = X_i\gamma + W_i\delta + \mu_i$; where X_i denotes the attributes that are used to assign a potential insurance buyer to a risk class, C_i denotes the demand for insurance coverage, L_i denotes risk of loss and W_i denotes an unused observable variable. Rejecting $\alpha = 0, \delta = 0$ is equivalent to rejecting the null hypothesis of symmetric information.

Table 8
Positive correlation test with observable but unused variables.

	Panel A. Full sample				Panel B. Sample without individuals who had supplemental health insurance in 2010			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Hospital services in 2010	**0.204*** **	**0.203*** **	**0.192*** **	**0.190*** **	**0.214*** **	**0.214*** **	**0.194*** **	**0.191*** **
	(0.0240)	(0.0240)	(0.0210)	(0.0209)	(0.0383)	(0.0383)	(0.0335)	(0.0334)
Ambulatory services in 2010	0.0312***	0.0303***	0.0281***	0.0211***	0.0265***	0.0265***	0.0245***	0.0184***
	(0.00279)	(0.00278)	(0.00281)	(0.00319)	(0.00297)	(0.00297)	(0.00279)	(0.00317)
SHI 2010		1.043***	0.966***	0.974***				
		(0.129)	(0.128)	(0.128)				
Age			0.0123***	0.0109***			0.0105***	0.00952***
			(0.00196)	(0.00207)			(0.00196)	(0.00205)
Female			0.148***	0.143***			0.186***	0.182***
			(0.0313)	(0.0317)			(0.0308)	(0.0315)
AMI 2010			−0.0731	−0.0575			−0.0303	−0.0160
			(0.0477)	(0.0475)			(0.0484)	(0.0482)
Constant	0.165***	0.136***	0.605	0.447	0.160***	0.160***	0.115	−0.0348
	(0.0169)	(0.0171)	(0.623)	(0.620)	(0.0161)	(0.0161)	(0.632)	(0.630)
Municipality control	NO	YES	YES	YES	NO	YES	YES	YES
Diagnosis control	NO	NO	YES	YES	NO	NO	YES	YES
Observations	484,005	484,005	409,905	409,905	467,777	467,777	393,677	393,677
R-squared	0.049	0.049	0.048	0.049	0.034	0.034	0.030	0.032

Robust standard errors in parentheses
*** p < 0.01, ** p < 0.05, * p < 0.1
SHI: Supplemental Health Insurance.

to infer what would be the optimal supplemental health insurance purchase.[30]

Second, and beyond the deadweight loss, supplemental health insurance purchase, subsidies for a more pragmatic budgetary concern is also present. As pointed out in Lopez-Nicolás and Vera-Hernandez (2008), it is important to determine if tax-subsidies for supplemental health insurance are self-financing. The presence of adverse selection into supplemental coverage implies that subsidies to supplemental health insurance coverage reduce public expenses, that is, they are self-financing (Buitrago and Bardey, 2015). Thus, our adverse selection findings reinforce the idea that the Colombian government should subsidize the purchase of supplemental health insurance coverage.

Finally, the presence of adverse selection in the supplemental health insurance market is a warning for the risk adjustment mechanism at work in the mandatory regime. Indeed, the set of variables that is used to calculate the capitation payment in the mandatory regime is very close to the set of variables applied by health insurers (that provide supplemental coverage) for pricing.[31] The adverse selection is not really an issue under a mandatory regime, since individuals are unable to escape from the risk pooling and, on top of this, the package benefit is standardized. Nevertheless, as shown in Castaño and Zambrano (2006), the risk concentration is relatively high in the mandatory regime. In light of the adverse selection outcome found in the supplemental coverage, we recommend enriching the set of variables used in the risk adjustment mechanism applied to the mandatory regime. A richer risk adjustment should lower the risk concenion in the mandatory regime, and thereby reduce adverse selection in the supplemental coverage.

Acknowledgements

We thank two anonymous referees and Thomas McGuire the guest editor of this special issue, for their valuable comments and suggestions. We also thank Adriana Camacho, Sandra Garcia, Maria Fernanda Rosales, Diego Amador and participants to seminars held at the University of Los Andes for their suggestions.

Appendix A. Definitions of variables.

Supplemental Health Insurance (SHI): 1 if the individual has SHI, 0 otherwise.

Age: individual age in years.

Female: 1 if the individual is a women, 0 otherwise.

Average Monthly Income: Monthly income (in 2010 Colombian pesos) reported to PILA.

Hypertension (HT): 1 if the individual has hypertension reported in database of UPC, 0 otherwise.

Chronic Kidney Disease (CKD): 1 if the individual has CKD reported in database of UPC, 0 otherwise.

Cancer: 1 if the individual has Cancer reported in database of UPC, 0 otherwise.

Pneumonia: 1 if the individual has Pneumonia reported in database of UPC, 0 otherwise.

Urinary Tract Infection (UTI): 1 if the individual has UTI reported in database of UPC, 0 otherwise.

Preeclampsia: 1 if the individual has Preeclampsia reported in database of UPC, 0 otherwise.

Gastrointestinal disease: 1 if the individual has gastrointestinal disease reported in database of UPC, 0 otherwise.

High Respiratory Tract Infection (HRTI): 1 if the individual has HRTI reported in database of UPC, 0 otherwise.

Chronic Obstructive Pulmonary Disease (COPD): 1 if the individual has COPD reported in database of UPC, 0 otherwise.

Asthma: 1 if the individual has asthma reported in database of UPC, 0 otherwise.

Mental disorder: 1 if the individual has mental disorder reported in database of UPC, 0 otherwise.

Use of hospital services: number of hospital services used (frequency). The hospital services include drugs, diagnostic tests, days of hospital stay, and others hospital procedures (like surgeries, invasive procedures, etc.).

Use of ambulatory services: number of ambulatory services used (frequency). The ambulatory services include drugs, diagnostic tests, consultations, and others ambulatory procedures (like

[30] Some price variations caused by fiscal reform could be exploited in 2013 to replicate the analysis undertaken in Bundorf et al. (2012). We have not obtained these data yet, but it does comprise part of our research agenda.

[31] Additionally to age and gender, the capitation payment includes the geographical location.

ambulatory surgeries, invasive procedures, etc.). Does not include emergency room services.

Spending: total cost of health services used (in 2010 Colombian pesos).

Pap smear: 1 if the individual use Pap smear, 0 otherwise.

Mammography: 1 if the individual use mammography, 0 otherwise.

Prostate-specific antigen (PSA): 1 if the individual use PSA, 0 otherwise.

Vaccination: number of vaccines used.

Intensive Care Unit (ICU) admission: 1 if the individual is admitted to ICU, 0 otherwise.

Hospital Stay: numbers of days of hospital stay.

Bibliography

Blomqvist, A., 1997. Optimal non-linear health insurance. J. Health Econ. 16, 303–321.

Buitrago G., Bardey D., 2015. Voluntary Health Plan Subsidies and Public Expenditure Working Paper No. 605 (October) TSE working paper.

Bundorf, M.K., Levin, J., Mahoney, N., 2012. Pricing and welfare in health plan choice. Am. Econ. Rev. 102 (7), 3214–3248.

Castaño, R., Zambrano, A., 2006. Biased selection within the social health insurance market in Colombia. Health Policy 3, 313–324.

Chiappori, P.A., Salanié, B., 2000. Testing for asymmetric information in insurance markets. J. Polit. Econ. 108, 56–78.

Cohen, Alma, 2005. Asymmetric information and learning: evidence from the automobile insurance market. Rev. Econ. Stat. 87, 197–207.

De Meza, D., Webb, D., 2001. 'Advantageous selection in insurance markets'. advantageous selection in insurance markets. Rand J. Econ. 32, 249–262.

Einav, L., Finkelstein, A., 2011. Selection in insurance markets: theory and empirics in pictures. J. Econ. Perspect. 25 (1), 115–138.

Einav, L., Finkelstein, A., Cullen, M., 2010. Estimating welfare in insurance markets using variation in prices. Q. J. Econ. 125 (3), 877–921.

Einav, L., Finkelstein, A., Ryan, S., Schrimpf, P., Cullen, M.R., 2013. Selection on moral hazard in health insurance. Am. Econ. Rev. 103 (1 (February)), 178–219.

Escobar, M.L., Giedeon, U., Giuffrida, A., Glassman, A., 2010. Colombia: after a decade of health system reform. In: Glassman, A., Escobar, M.L., Giuffrida, A., Giedeon, U. (Eds.), From Few to Many: Ten Years of Health Insurance Expansion in Colombia. Inter-American Development Bank and Brookings Institution, Washington, DC.

Fang, H., Keane, M., 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. Am. Econ. Rev. 96 (4), 938–958.

Finkelstein, A., Poterba, J., 2014. Testing for asymmetric information using Unused observables in markets: evidence from the U. K. annuity market. Insurance 81 (4), 709–734.

Hemenway, D., 1990. Propitious selection. Q. J. Econ. 105 (4), 1063–1069.

Vera-Hernandez, A., Lopez-Nicolás, M., 2008. Are tax subsidies for private medical insurance self-financing? evidence from a microsimulation model. J. Health Econ. 27, 1285–1298.

OCDE, 2016. OCDE Reviews of Health Systems: Colombia. OCEDE Publishing, Paris, pp. 2015.

Olivella, P., Vera-Hernandez, M., 2013. Testing for asymmetric information in private health insurance. The Economic Journal 123 (567), 96–130.

Pauly, M.V., 1968. 'The economics of moral hazard: comment'. Am. Econ. Rev. 58 (13), 531–537.

Resende, M., Zeidan, R., 2010. Adverse selection in the health insurance market: some empirical evidence. Eur. J. Health Econ. 11 (4), 413–418.

Rothschild, M., Stiglitz, J., 1976. Equilibrium in competitive insurance markets? Q. J. Econ. 90 (4), 629–649.

Vargas, I., Vasquez, M.-L., Mogollón-Peréz, A.-S., Unger, J.-P., 2010. 'Barriers of access to care in a managed competition model: lessons from Colombia'. BMC Health Serv. Res. 10, 297.

Vera-Hernandez, M., 2003. Structural estimation of a principal-agent model: moral hazard in medical insurance. Rand J. Econ. 34 (4), 670–693.