

# Ex-ante moral hazard: empirical evidence for private health insurance in Brazil

*Risco Moral Ex-ante: uma evidência empírica para o setor de seguro saúde privado no Brasil*

Ana Carolina Maia <sup>(1)</sup>

Mônica Viegas Andrade <sup>(2)</sup>

Flávia Chein <sup>(3)</sup>

<sup>(1)</sup> Universidade de São Paulo

<sup>(2)</sup> Universidade Federal de Minas Gerais

<sup>(3)</sup> Universidade Federal de Juiz de Fora

## Abstract

This paper explores the existence of ex-ante moral hazard in private health insurance in Brazil. Before the advent of illness, insured individuals have no incentives to seek preventive care if it is not previously contractible. The data set comprises longitudinal administrative records of health care utilization from a Brazilian employer-sponsored health insurance plan. The empirical strategy is based on an exogenous and anticipated shock in health insurance coverage not associated with health conditions. The results show an increase of up to 17% on medical visits and 22% on diagnostic tests due to the loss of health insurance. Medical visits start to increase five months before the individual leaves the health insurance pool, reaching its peak at two months prior to exit. For diagnostic tests, the increase was observed only in the last two months before the loss of health insurance coverage.

## Keywords

ex-ante moral hazard, health insurance, self-insurance, Brazilian private health insurance.

**JEL Codes** I13, I11, H50.

## Resumo

*Este artigo explora a existência de risco moral ex-ante no setor de saúde suplementar no Brasil. Antes do evento da doença, indivíduos com plano/seguro saúde não têm incentivos para procurar cuidado preventivo se esse não for contratado previamente. O conjunto de dados utilizado inclui registros administrativos longitudinais da utilização de cuidados de saúde de um plano de saúde brasileiro patrocinado pelo empregador. A estratégia empírica é baseada em uma mudança exógena e antecipada na cobertura de seguro de saúde não associada à condição de saúde. Os resultados mostram um aumento de até 17% nas consultas médicas e 22% nos exames diagnósticos. As consultas médicas começam a aumentar a partir do quinto mês antes do indivíduo deixar a carteira com um pico em dois meses. Para testes diagnósticos, o maior aumento foi observado em um e dois meses antes da perda da cobertura do plano de saúde.*

## Palavras-chave

*risco moral ex-ante, plano de saúde, auto-seguro, setor de saúde suplementar no Brasil.*

**Códigos JEL** I13, I11, H50.

## 1 Introduction

Moral hazard is a well-known fact associated with consumer behaviour in the context of health insurance coverage (Newhouse, 1993; Arrow, 1963; Pauly, 1968; Zeckhauser, 1970). Two distinct types of moral hazard are observed in individual behaviour: *ex-ante* and *ex-post* moral hazard (Zweifel & Manning, 2000). *Ex-post* moral hazard is related to the overuse of health care services observed in the presence of health insurance once an illness event has occurred. The reduction in the marginal cost perceived by insured individuals results in a higher level of health care use compared to the absence of health insurance.

In contrast, *ex-ante* moral hazard refers to situations before the advent of illness, when individuals have no incentives for preventive care. Moreover, having health insurance coverage can lead to individual underinvestment in preventive care, if it is not contractible, that can be interpreted as an *ex-ante* moral hazard (EAMH), as it usually affects the future costs of illness. Several factors may affect the individual choice of preventive care, for instance, opportunity costs, risk aversion, or the effect of ill health on the utility of such care.

Ehrlich and Becker (1972) distinguish two sources of *ex-ante* moral hazard: self-protection and self-insurance. Self-protection refers to the ability of individuals to affect the probability of illness through the avoidance of risky behaviours (e.g., smoking, drinking) and the adoption of healthy ones (healthy diet, exercise), whereas self-insurance relates to an individual's ability to affect the future costs of disease, such as regular doctor visits, dental care, immunization, and check-ups.

Some empirical evidence has already shown the presence of EAMH. Zweifel and Manning (2000) explore EAMH in contexts of health insurance, focusing on the opportunity costs of preventive care effort and individual risk aversion as the elements that can determine the amount of preventive care chosen by insured individuals. Opportunity costs are usually related to consumer time constraints, and therefore the wage rate is the shadow price of preventive care. Authors have extended the notion of opportunity costs by including the generosity of coverage of preventive care. The less one pays in out-of-pocket expenditure, the higher the consumption of preventive care. Less generous insurance is associated with decreased use of preventive medical services (secondary prevention),

which contradicts the pure intuition that less protection should cause an increase in prevention (Lilliard *et al.*, 1986).

Dave and Kastner (2009), using the Health and Retirement Study, examined the effect of health insurance on the health behaviours of the elderly population. Specifically, the authors investigated EAMH comparing changes in healthy activities among uninsured and insured individuals before and after Medicare enrolment. More precisely, the authors performed a difference in difference analysis comparing the utilization of preventive health care before and after the admission to Medicare considering two groups of individuals: those who were uninsured before age 65 and those who were already insured. The authors found evidence that Medicare enrolment increased unhealthy activities such as smoking and alcohol use for males, pointing to evidence of ex-ante moral hazard.

On the other hand, results for utilization of doctor visits and hospital stays showed that Medicare enrolment increased the probability of seeing a doctor. These results may seem to contradict the presence of EAMH related to self-insurance, but in truth, this phenomenon results from an access effect. Uninsured individuals had restricted access to health care services before Medicare enrolment. Card and Maestas (2008) documented the same findings.

De Preux (2011) compared changes in physical exercise, smoking, and drinking among uninsured and insured individuals before Medicare eligibility using the same database but incorporating the effect of the anticipation of Medicare benefits. The author extends the analysis of Dave and Kastner (2009), introducing the definition of anticipatory EAMH. According to de Preux (2011), as individuals anticipate the enrolment in Medicare, they tend to change their current preventive care choices just before age 65. Therefore, EAMH is related to two mechanisms. First, Medicare coverage reduces the cost of future illness care and consequently discourages healthy habits; and second, the anticipation of Medicare enrolment should reduce the future benefits of current prevention. The author found that before Medicare enrolment, uninsured individuals are less likely to exercise than are insured ones, pointing out the existence of anticipatory EAMH associated with self-protection. That is, uninsured individuals choose to reduce self-protection because they anticipate they will receive Medicare coverage.

Despite these pieces of evidence, there is scarce work related explicitly to self-insurance. This lack of evidence in part is due to difficulties in

empirically distinguishing the *coverage effect* on the use of preventive care from the *access effect* and endogeneity issues. Individuals who choose to buy health insurance may present specific attributes that may affect health care utilization.

Theoretically, the presence of EAMH relies on lower incentives for preventive care due to the presence of coverage of future illness events (coverage effect). In addition to that, for uninsured individuals, health insurance enrolment means a significant change in access to health care services (access effect). In that manner, while insured individuals tend to use fewer health care services due to the coverage effect, the uninsured tend to use more health care services due to the access effect. Therefore a comparison between uninsured and insured individuals may not be enough to distinguish the coverage effect from the access effect.

This study aims to estimate EAMH using new data from an employer-sponsored health insurance plan in Brazil provided by the Sewage Company of the State of São Paulo—Companhia de Saneamento Básico do Estado de São Paulo, Sabesp. Here, all Sabesp employees and their enrolled dependents are beneficiaries of a health insurance plan.

The data set comprises longitudinal administrative records of health care utilization during the 2004–2008 period. We take advantage of a specific form of EAMH in a patient's decision to seek preventive using an exogenous variation of health insurance coverage as our identification strategy. The expected loss of job due to retirement or dismissal is an exogenous variation of coverage status not associated with health conditions.

The anticipation of health insurance loss can affect consumer behaviour of preventive care—self-insurance. More services are chosen to compensate for the underutilization due to EAMH and consequently to reduce the future costs of medical care. In this context, the EAMH as measured in this study is the variation observed in the consumption of preventive care when individuals anticipate the loss of health insurance coverage.

In Brazil, since 1998, with the regulation of the private health insurance market (Law 9.656/1998, articles 30 and 31), in the case of dismissal or retirement of health insurance policyholders, all beneficiaries may remain insured for an additional period of a minimum of six months. In these circumstances, enrollees must formally opt to remain or exit from the health insurance pool. Because individuals know that they will lose coverage, it is possible to estimate the amount of preventive care that is unused when

they are insured. When individuals have health insurance they postpone or even avoid preventive care due to, for example, opportunity costs. Therefore, the increase in the demand for health care observed prior to their loss of health insurance is taken as the ex-ante moral hazard. Specifically, we are dealing with ex-ante moral hazard related to self-insurance, estimated through outpatient care, such as doctor visits and diagnostic tests.

## 2 Methods

### 2.1 Database

The database used in this study comprises administrative records regarding the utilization of health care services funded by the Sewage Company of Sao Paulo State (Sabesp) in Brazil. Sabesprev is a health insurance pool managed and provided by Sabesp that bears the major part of the cost of the health insurance premium for Sabesp employees, their dependents, and other relatives. Employee enrolment is mandatory, and the premium payment for employees and their dependents is a uniform wage tax of 2%. For other relatives and retired employees, Sabesp fixes the premium according to age groups. There are three types of contracts: a full plan that is mandatory for active employees; a basic plan that is available for retired employees and their relatives with only shared-room hospital accommodation; and a special plan that is also available for retired employees and their relatives with single-room accommodation. All types of contracts include the same network of providers and the same co-payments.

Originally, administrative records were organized in three separate data sets. The first set comprises monthly health care utilization records for each individual, including beneficiary ID, date of utilization, type of procedure, provider, and expenditures. The second data set contains detailed information about inpatient care, including hospitalization period, diagnosis, and provider information, which allows us to distinguish ambulatory from inpatient procedures. Finally, the last set has beneficiary information, such as birth date, gender, relationship to Sabesp employee, insurance enrolment, and exit date. This set also includes the reason for dismissal if the beneficiary is no longer enrolled.

In this study, we organized these administrative records in a monthly longitudinal data set for the 2004–2008 period with all beneficiaries of the Sabesp health insurance plan. Because beneficiaries can enrol or leave the risk pool at any time, we built an unbalanced panel. New entries in the risk pool may occur because Sabesp can hire new employees, whereas exit may occur due to firing, retirement, or death. Our data set includes all beneficiaries who were enrolled prior to January 2004 and followed for at least 13 months. Therefore, our data had 49,106 individuals, from which 46,929 remained enrolled in the health insurance pool during the entire period, and only 2,177 left the pool for any of the reasons mentioned above.

To estimate ex-ante moral hazard, we considered two groups of beneficiaries: those who remained in the risk pool during the five years and those who left the risk pool due to dismissal or retirement.

## 2.2 Empirical strategy

Our empirical strategy explores the presence of an exogenous and anticipated variation in health insurance coverage to estimate ex-ante moral hazard. We argue that the beneficiary exit is exogenous to his/her current health status and may be anticipated by enrollees. As mentioned, beneficiary exit occurs due to the dismissal or retirement of the policyholder. We assume that dismissal is not caused by a shock in health status conditions because labour law in Brazil protects unhealthy workers from being dismissed (Decree-Law 3048/99).

According to the Brazilian Social Security Laws (Giambiagi & Afonso, 2015), employees are entitled to three types of pension benefits managed by the Social Security System (*Instituto Nacional de Seguridade Social—INSS*): 1) age retirement (60 years for urban females and 65 for urban males); 2) retirement for fulfilment of minimum time employed; and 3) retirement due to disability.

In the first two categories, the retirement benefits requirement is voluntary and, therefore, not usually correlated to contemporaneous changes in the health status of the policyholder. For retirement due to disabilities, it is clear that there is an association between risk pool exit and employee health status. Unfortunately, in the Sabesp data set it is not possible to identify the reason for retirement. In the case of withdrawal due to disabil-

ity, we assume that the health shock occurred before the retirement date. As the judicial process necessary to gain these benefits is very long, the enrollee has to be off work at least 24 months before the exit date. During the period that the enrollee is claiming his/her rights, he/she remains covered. In that manner, the health shock should have occurred at least 24 months before the enrollee's exit. Additionally, only 35% of the beneficiaries who left the risk pool are policyholders, and it is reasonable to assume no correlation between the retirement decision of the policyholder and the dependent's health status.

It is worth mentioning that employees and their dependents in any case of risk pool exit can anticipate the loss of health insurance coverage. Retirement is a decision taken by employees and is usually not a fast process. Dismissed workers have to be notified at least 30 days before exit and can remain in the risk pool for at least six months after the dismissal.

The moral hazard estimation considers two health care indicators: the number of doctor visits and the number of diagnostic tests. To a great extent, primary care utilization is a patient's decision, being less associated with previous medical diagnosis or referral. We included diagnostic tests because they are a deployment of medical visits and part of a routine of preventive care.

To measure moral hazard, we built six dummy variables according to time before leaving the health insurance risk pool: 1 to 6 months. Each dummy variable captures the increase in health care observed for beneficiaries who left the risk pool in the respective month compared to the ones who remained. The comparison of these *dummy* coefficients allows for analysing how this effect varies over time.

The estimated model is specified as follows:

$$y_{it} = f(a_i, t_i, x_{it}, \alpha_i, \mu_{it}) \quad (1)$$

for each individual  $i = 1, \dots, m$  and time  $t = 1, \dots, n$ ,  $y_{it}$  is the dependent variable regarding health care utilization for individual  $i$  on time  $t$ ,  $a_i$  is a vector of individual attributes that does not vary over time (for example, sex),  $b_t$  is the vector of time dummies,  $x_{it}$  is a vector of individual attributes that varies over time (for example, age),  $\alpha_i$  are the non-observable characteristics of individual  $i$  (for example, health status), and  $\mu_{it}$  is the idiosyncratic error. Chart 1 describes all independent variables included in the model.

Chart 1 Description of Independent Variables

Variable	Description
Time before leaving the pool	Number of months until leaving the pool specified by a set of six dummy variables. Each dummy variable $d_{ijt}$ assumes a value equal to one if individual $i$ during $t$ time is $j$ months away from leaving the portfolio, and zero otherwise; $j$ varies from 1 to 6 while $t$ varies from 1 to 60. These are our interest variables that measure for each indicator the average increase observed in the health care utilization $j$ months before exiting the risk pool.
Age	Age variable built by considering a set of 17 dummy variables $A_{ijt}$ regarding 5-year age groups: 0–4, 5–9, 10–14, 15–19, 20–24, 25–29, 30–34, 35–39, 40–44, 45–49, 50–54, 55–59, 60–64, 65–69, 70–74, 75–79, and over 80 years. Each dummy variable $a_{ijt}$ assumes a value equal to one if individual $i$ during $t$ time belongs to $j$ age group, and zero otherwise; $j$ varies from 1 to 17 while $t$ varies from 1 to 60. The reference category is 0–4 years. Age is a proxy variable for individual risk.
Family size	Number of beneficiaries in each family enrolled in the <i>Sabesprev</i> pool. Family size can affect opportunity costs to use health care services.
Relationship with the policyholder	Set of three dummy variables $R_{ijt}$ regarding the beneficiary's relationship with the policyholder. Three types of household membership are available: proper policyholder, aggregate, or dependent. Each dummy variable $r_{ijt}$ assumes a value equal to one if individual $i$ during time $t$ has $j$ relationship with the policyholder, and zero otherwise. Reference category: policyholder.
Sex	Dummy variable $S_{ij}$ that assumes a value equal to 1 if the beneficiary is a man, and zero otherwise. Sex is a proxy for individual risk.
Type of contract	Set of three dummy variables $C_{ijt}$ regarding each type of contract: full, basic, and special plan. The reference category is full plan.
Time	Set of 60 dummy $T_t$ variables for each year. This allows controlling for aggregate shocks that can affect health care utilization. The reference category is January 2004.

To correctly estimate model (1), it is worth taking into account two characteristics of our data. First, as clarified by Cameron and Trivedi (2013), we cannot assume a normal distribution of our interest outcome as it is a limited-range dependent variable (number of doctor visits and diagnostic tests). A commonly applied count data model for such a situation is the negative binomial that controls for data overdispersion. Second, in a longitudinal data set with counted responses correlated within subjects over time, it is crucial to deal with the violation of independence assumption. It is reasonable to assume that the utilization of health care has an observed individual pattern. As emphasized by Hilbe (2011), the generalized estimation equation (GEE), or population-averaged model, adds at least one



extra parameter to linear prediction to inform how observations within panels are to be construed.

To consider these issues, in this study we estimate a negative binomial model by the GEE method, adopting three approaches to correct the coefficient variance based on different correlation structures: AR (autoregressive), exchangeable, and unstructured.

The GEE is a method initially proposed by Liang and Zeger (1986) and constitutes an extension of the Generalized Linear Model (GLM) applied to panel data. The method evaluates the relationship between the response and predictor variables in a population context, and it is known as the marginal effect model. The essential characteristic of its specification is factoring the variance function to include a parametric correlation structure in the panel.

The GLM of  $y_{it}$  with covariates can be expressed as:

$$g\{E(y_{it})\} = x_{it}\beta, \quad y \sim F \text{ with parameters } \theta_{it} \quad (2)$$

where  $g$  is a function link and  $F$  is the distribution of the dependent variable, called family. Each pair of links and family functions defines a different model. In our case, we specify the family function as the negative binomial distribution and the link as the log function.

To better understand the GEE algorithm proposed by Liang and Zeger (1986), equation (3) specifies the variance function of the general model, estimated by GLM:

$$V(\mu_{it}) = \left[ D(V(\mu_{it}))^{1/2} R_{n_i X_{n_i}} D(V(\mu_{it}))^{1/2} \right]_{n_i X_{n_i}} \quad (3)$$

where  $V(\mu_{it})$  is the GLM variance function defined in terms of the mean. In the negative binomial, the variance function is  $\mu + \alpha\mu^2$ .  $D$  is a diagonal matrix of the variance functions of  $y_{it}$ , and  $R$  is the specified correlation matrix. If we assume a zero correlation between subsequent measures of a subject within panels, we have  $R_{n_i X_{n_i}} = I_{n_i X_{n_i}}$ . In this case, the correlation structure is independent.

In the exchangeable structure, it is assumed that the correlation between subsequent measures of a subject within panels is always the same. On the other hand, in the unstructured approach, all correlations are considered different and are estimated from the data. In the AR correlation

structure, one might expect that there is a decline in the value of the correlation coefficient as it increases within panel time intervals.

Despite its contribution, the GEE model does not take into account the individual-specific unobserved effect. Since the seminal article by Hausman *et al.* (1984), there has been some application in the literature of fixed models for the negative binomial. However, as Allison and Waterman (2002) discuss, this model does not meet the usual properties required by the fixed effect method as a control of all stable covariates. According to the authors, this occurs due to the decomposition of the overdispersion parameter instead of the usual average decomposition (Allison & Waterman, 2002). Although they present alternatives such as the conditioning of the negative multinomial, these are not secure (Hilbe, 2011).

In this study, as an additional step, we opted to estimate a random effect model—a negative binomial with random beta-distributed effect, following Hausman *et al.* (1984). In this model, the conditional expected value and the variance are:

$$E(y_{it} | x_{it}, \alpha_i) = \alpha_i \mu_{it} \text{ e } V(y_{it} | x_{it}, \alpha_i) = \alpha_i \mu_{it}^* (1 + \alpha_i)^{-1} \tag{4}$$

where  $\hat{\mu}_{it} = \exp(x_{it}'\beta)$  with  $x_{it}$  has exogenous covariates in time  $t$ , and  $(1 + \alpha_i)^{-1}$  is a random beta-distributed variable.

In this case, the estimated coefficients are consistent only if the random effect is not correlated with exogenous variables. The random individual-specific effect can be interpreted as different attitudes regarding health protection, and the previous assumption can be violated. However, the variation of covariates among individuals is more relevant than the variation of individuals; in this scenario, the fixed effect can generate inconsistent estimators (Chamberlain, 1984).

### 3 Results

Table 1 presents the descriptive statistics of observable individual attributes considering the status of health insurance coverage. The first group regards individuals who remained in the risk pool during the 2004–2008 period, and the second one those who lost health insurance coverage in 2005 due to dismissal or retirement of the policyholder. As some attributes

vary over time, it is only possible to compare individual characteristics between the two groups for each year separately.

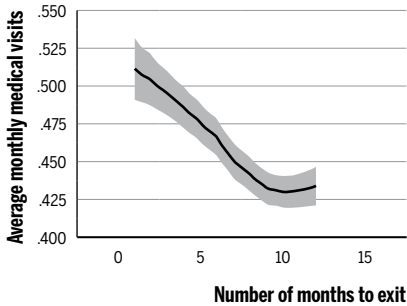
In 2005, both groups were quite similar regarding age, gender, relationship to policyholder, and type of contract: 50% of beneficiaries were men, on average they were 35 years old, almost 60% were dependent, and around 90% had full coverage. The only difference between active beneficiaries and the ones who exited the risk pool concerns the family size and type of health insurance contract. It is expected that type of contract differs between active and non-active beneficiaries because the full plan is available only for Sabesp policyholders who are still working. Family size is slightly larger among beneficiaries who remained in the risk pool during the entire period.

Table 1 Beneficiaries' Attributes by Status of Health Insurance Coverage

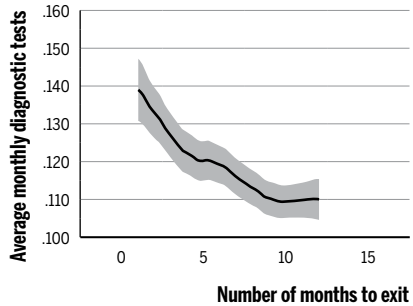
Active beneficiaries during the whole period		Beneficiaries who lost health insurance coverage in 2005	Difference
	(1)	(2)	(2) – (1)
Age	34.30 (0.089)	35.13 (0.823)	-0.82 (0.814)
Sex (1 if male)	0.49 (0.002)	0.50 (0.020)	0.00 (0.021)
<b>Relationship to policyholder</b>			
Policyholder	0.36 (0.002)	0.33 (0.019)	0.03 (0.020)
Aggregate	0.03 (0.001)	0.04 (0.008)	0.009 (0.008)
Dependent	0.58 (0.002)	0.63 (0.020)	-0.04** (0.020)
<b>Type of health insurance plan</b>			
Full	0.89 (0.001)	0.92 (0.011)	-0.03** (0.013)
Basic	0.019 (0.001)	0.012 (0.005)	0.01 (0.006)
Special	0.089 (0.001)	0.066 (0.010)	0.022* (0.011)
Family size	3.71 (0.006)	3.52 (0.068)	0.18*** (0.058)
Observations	46,927	573.00	

Graphs 1 to 4 present the average number of doctor visits and diagnostic tests received by individuals according to the number of months before exiting the risk pool. For each figure, the X-axis shows the number of months before the departure, and zero means the exit date. The shaded area corresponds to the 95% confidence interval for health care indicators.

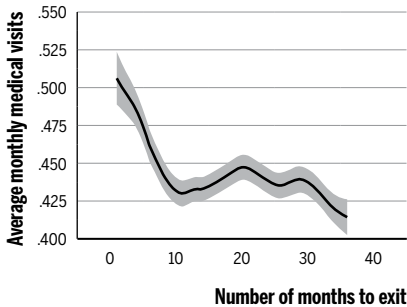
**Graph 1 Number of medical visits along 12 months before the risk pool exit**



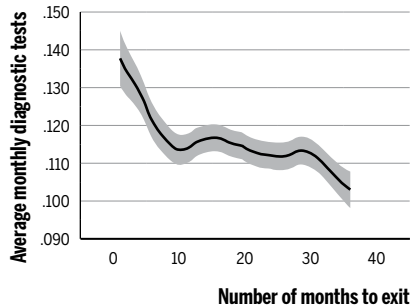
**Graph 2 Number of diagnostic test along 12 months before the risk pool exit**



**Graph 3 Number of medical visits along 36 months before the risk pool exit**



**Graph 4 Number of diagnostic test along 36 months before the risk pool exit**



■ 95%CI — Ipoly smooth: medical visits or diagnostic tests

Source: Sabesprev data.

Note: Graphics estimated using a kernel-weighted local polynomial regression for each indicator by number of months before the risk pool exit.

Graphs 1 and 2 show the kernel function of health care utilization considering only 12 months before the exit, whereas, in Graphs 3 and 4, this

period extends to 36 months. It is noteworthy that for both health care indicators, the number of procedures grew as the risk pool exit became closer. These results suggest that individuals tend to intensify preventive health care utilization as the loss of health insurance coverage approaches. Additionally, this increase starts 10 months before the risk pool exit.

Table 2 presents the results for regression models estimated for medical visits. The first three columns display the population effect model using the different correlation structures of the variance matrix (AR1, exchangeable, and unstructured), whereas the fourth corresponds to the random effect model. Our interest outcomes are the dummy variables regarding the number of months before the risk pool exit (time dummies). To better understand time dummy effects, semi-elasticities for each health care indicator are reported in Table 2. The semi-elasticity coefficient is the percentage increase observed in health care utilization for enrollees who lost health insurance coverage in comparison to beneficiaries who remain in the risk pool.

The main findings confirm the presence of ex-ante moral hazard for both types of health care utilization. The analysis of time dummies shows that these effects start at different months before the risk pool exit, depending on the kind of care. These results are found independently of the estimation method, pointing to the robustness of the analysis.

The increase in medical visits starts four months before the risk pool exit when the dummy coefficient turns statistically significant. The highest effect is observed two months before the departure, and the magnitude of ex-ante moral hazard varies from 14% to 17% conditioned on the model estimated. It is worth noting that one month before losing health insurance coverage, the number of medical visits proves similar among both groups (Table 2).

Compared to medical visits, the utilization of diagnostic tests increases closer to the exit date (Table 3). For these procedures, the effect is observed only in the last two months before the risk pool exit. The semi-elasticity coefficients are higher than those estimated for medical visits, varying from 17% to 22%. As we shall see, a time lapse is observed between the increase in the frequency of medical appointments and diagnostic tests. This behaviour reflects the need for a physician referral to perform diagnostic tests.

Table 2 Semi-elasticity Coefficients Estimated for Medical Visits

Method	Generalized estimation equations			Random effect model
	Correlation structure			
	ARI	Exchangeable	Non-structured	
last month on the portfolio (1 month to go)	0.072* (0.037)	0.038 (0.037)	0.036 (0.037)	0.029 (0.033)
2 months to go	0.173*** (0.034)	0.140*** (0.034)	0.141*** (0.033)	0.142*** (0.031)
3 months to go	0.102*** (0.034)	0.084** (0.033)	0.083** (0.033)	0.073** (0.030)
4 months to go	0.116*** (0.031)	0.103*** (0.030)	0.108*** (0.029)	0.097*** (0.027)
5 months to go	0.056* (0.029)	0.044 (0.028)	0.048* (0.028)	0.046* (0.027)
6 months to go	0.045 (0.029)	0.042 (0.028)	0.046 (0.028)	0.035 (0.026)
5 to 9 years	-0.435*** (0.014)	-0.321*** (0.011)	-0.298*** (0.011)	-0.333*** (0.008)
10 to 14 years	-0.761*** (0.016)	-0.522*** (0.013)	-0.510*** (0.013)	-0.579*** (0.010)
15 to 19 years	-0.634*** (0.016)	-0.417*** (0.014)	-0.424*** (0.014)	-0.493*** (0.012)
20 to 24 years	-0.414*** (0.019)	-0.307*** (0.018)	-0.298*** (0.017)	-0.384*** (0.014)
25 to 29 years	-0.347*** (0.018)	-0.248*** (0.017)	-0.221*** (0.016)	-0.319*** (0.013)
30 to 34 years	-0.343*** (0.016)	-0.233*** (0.015)	-0.206*** (0.015)	-0.297*** (0.013)
35 to 39 years	-0.348*** (0.016)	-0.212*** (0.015)	-0.186*** (0.014)	-0.272*** (0.012)
40 to 44 years	-0.295*** (0.016)	-0.164*** (0.014)	-0.142*** (0.014)	-0.220*** (0.012)
45 to 49 years	-0.229*** (0.016)	-0.110*** (0.014)	-0.087*** (0.014)	-0.165*** (0.012)
50 to 54 years	-0.164*** (0.016)	-0.041*** (0.014)	-0.023 (0.014)	-0.101*** (0.013)
55 to 59 years	-0.125*** (0.018)	0.023 (0.016)	0.030* (0.015)	-0.036*** (0.013)

(continues on next page)

Table 2 (continued)

Method	Generalized estimation equations			Random effect model
	Correlation structure			
	AR1	Exchangeable	Non-structured	
60 to 64 years	-0.103*** (0.021)	0.074*** (0.018)	0.078*** (0.017)	0.025* (0.015)
65 to 69 years	-0.022 (0.024)	0.135*** (0.020)	0.137*** (0.020)	0.090*** (0.016)
70 to 74 years	0.025 (0.025)	0.163*** (0.021)	0.171*** (0.021)	0.122*** (0.018)
75 to 79 years	0.004 (0.027)	0.160*** (0.024)	0.154*** (0.023)	0.122*** (0.020)
over 80 years	-0.182*** (0.033)	0.076*** (0.027)	0.047* (0.027)	0.034 (0.023)
male	-0.345*** (0.007)	-0.352*** (0.008)	-0.350*** (0.008)	-0.393*** (0.007)
aggregate	-0.026 (0.022)	0.002*** (0.020)	-0.004 (0.019)	-0.005 (0.016)
dependent	-0.101*** (0.009)	-0.122*** (0.010)	-0.116*** (0.009)	-0.104*** (0.009)
basic plan	0.043*** (0.023)	-0.093*** (0.027)	-0.044* (0.023)	-0.105*** (0.017)
special plan	-0.080*** (0.016)	-0.133*** (0.016)	-0.111*** (0.015)	-0.125*** (0.012)
family size	-0.031*** (0.002)	-0.031*** (0.002)	-0.032*** (0.002)	-0.034*** (0.002)
time dummies	yes	yes	yes	yes
number of observations	2,888,257	2,889,284	2,889,284	2,889,284
number of individuals	49,087	49,106	49,106	49,106
number of obs. per individual	min.	13	13	13
	avg.	580.8	580.8	580.8
	max.	60	60	60
link function	log	log	log	
family	binomial (k = 1.3043)	binomial (k = 1.3043)	binomial (k = 1.3043)	

Source: Sabesprev data.

Note: Reference categories: first month (Jan. 2004), from zero to four years of age, women, holder of the health care plan, and full plan standard deviation in parentheses.

\*significant at 10%, \*\* significant at 5%, \*\*\* significant at 1%.

Table 3 Semi-elasticity Coefficients for Diagnostic Tests

Method	Generalized estimation equations			Random effect model
	Correlation structure			
	AR1	Permutable	Non-structured	
last month on the portfolio (1 month to go)	0.148** (0.072)	0.172** (0.069)	0.171** (0.068)	0.110** (0.044)
2 months to go	0.190*** (0.063)	0.203*** (0.062)	0.208*** (0.062)	0.227*** (0.041)
3 months to go	-0.013 (0.065)	0.013 (0.063)	0.011 (0.063)	0.069* (0.041)
4 months to go	0.046 (0.063)	0.085 (0.061)	0.094 (0.061)	0.017 (0.038)
5 months to go	-0.075 (0.060)	-0.034 (0.057)	-0.031 (0.057)	-0.025 (0.038)
6 months to go	0.041 (0.057)	0.074 (0.055)	0.072 (0.055)	-0.005 (0.036)
5 to 9 years	-0.153*** (0.021)	-0.123*** (0.019)	-0.121*** (0.019)	-0.261*** (0.012)
10 to 14 years	-0.245*** (0.024)	-0.159*** (0.022)	-0.163*** (0.021)	-0.422*** (0.012)
15 to 19 years	-0.011*** (0.025)	0.058*** (0.023)	0.047** (0.023)	-0.281*** (0.012)
20 to 24 years	0.357*** (0.028)	0.371*** (0.027)	0.369*** (0.027)	0.014 (0.014)
25 to 29 years	0.526*** (0.026)	0.544*** (0.024)	0.545*** (0.024)	0.157*** (0.013)
30 to 34 years	0.639*** (0.023)	0.664*** (0.022)	0.663*** (0.021)	0.198*** (0.012)
35 to 39 years	0.706*** (0.023)	0.754*** (0.021)	0.751*** (0.021)	0.223*** (0.011)
40 to 44 years	0.850*** (0.022)	0.890*** (0.020)	0.887*** (0.020)	0.298*** (0.011)
45 to 49 years	0.977*** (0.022)	1.006*** (0.020)	1.003*** (0.020)	0.386*** (0.011)
50 to 54 years	1.098*** (0.022)	1.111*** (0.020)	1.107*** (0.020)	0.466*** (0.011)
55 to 59 years	1.177*** (0.023)	1.200*** (0.022)	1.192*** (0.021)	0.534*** (0.012)

(continues on next page)



Table 3 (continued)

Method	Generalized estimation equations			Random effect model
	Correlation structure			
	AR1	Permutable	Non-structured	
60 to 64 years	1.213*** (0.027)	1.271*** (0.024)	1.261*** (0.024)	0.589*** (0.013)
65 to 69 years	1.311*** (0.029)	1.354*** (0.027)	1.344*** (0.026)	0.695*** (0.014)
70 to 74 years	1.378*** (0.031)	1.420*** (0.029)	1.412*** (0.028)	0.749*** (0.014)
75 to 79 years	1.368*** (0.033)	1.407*** (0.029)	1.397*** (0.029)	0.733*** (0.015)
over 80 years	1.230*** (0.036)	1.323*** (0.032)	1.309*** (0.032)	0.617*** (0.016)
male	-0.367*** (0.009)	-0.374*** (0.010)	-0.372*** (0.010)	-0.344*** (0.004)
aggregate	-0.129*** (0.026)	-0.122*** (0.025)	-0.119*** (0.025)	-0.096*** (0.010)
dependent	-0.221*** (0.011)	-0.234*** (0.011)	-0.231*** (0.011)	-0.172*** (0.005)
basic plan	0.087*** (0.030)	0.044 (0.030)	0.048 (0.030)	-0.024** (0.010)
special plan	0.152*** (0.018)	0.103*** (0.018)	0.111*** (0.018)	-0.040*** (0.007)
family size	-0.033*** (0.002)	-0.035*** (0.003)	-0.036*** (0.003)	-0.022*** (0.001)
time dummies	yes	yes	yes	yes
number of observations	2,888,257	2,889,284	2,889,284	2,889,284
number of individuals	49,087	49,106	49,106	49,106
number of obs. per individual	min.	13	13	13
	avg.	580.8	580.8	580.8
	max.	60	60	60
link function	log	log	log	
family	binomial (k = 0.0921)	binomial (k = 0.0921)	binomial (k = 0.0921)	

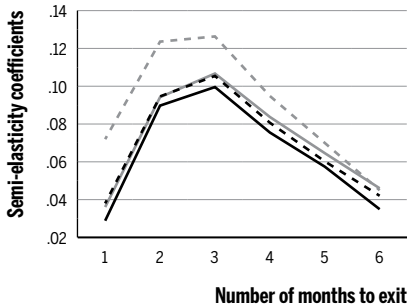
Source: Sabesprev data.

Note: Reference categories: first month (Jan. 2004), from zero to four years of age, women, holder of the health care plan, and full plan standard deviation in parentheses.

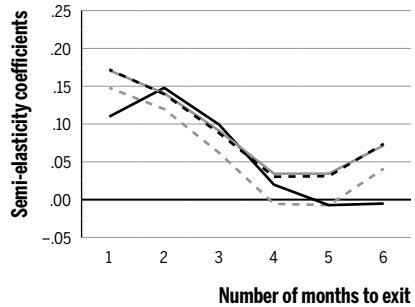
\*significant at 10%, \*\* significant at 5%, \*\*\* significant at 1%.

To compare the time coefficients pattern among the several models estimated, smoothed curves are shown in Graphs 5 and 6. It is noteworthy that the same pattern is observed independently of the estimation method for both medical visits and diagnostic tests.

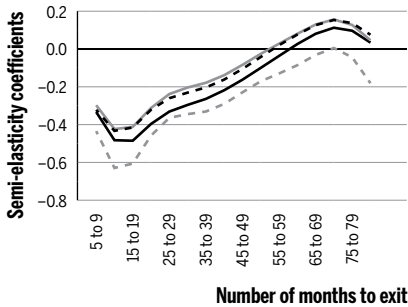
Graph 5 Time to exit effects on utilization of Medical Visits



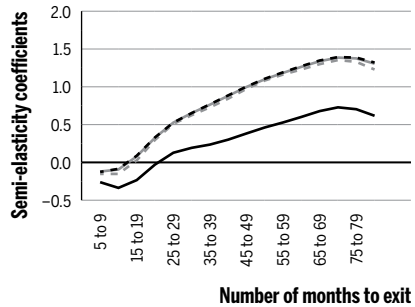
Graph 6 Time to exit effects on utilization of Diagnostic Tests



Graph 5 Age effects on utilization of Medical Visits



Graph 6 Age effects on utilization of Diagnostic Tests



----- ARI    - - - - Exchangeable    ——— Non Structured    ——— Random Effects

Age and gender dummies present the typical pattern observed along the life cycle (Yamamoto, 2013). The utilization of medical visits is higher for the first five years, decreasing until age 15, and then increasing smoothly until age 75. This pattern is observed for all estimation methods (Graph 7). Male enrollees receive fewer doctor visits and diagnostic tests than do women, around 35% for both health care indicators.

The use of diagnostic tests increases smoothly with age, presenting a lower decrement only for individuals over 80 years old (Graph 8). The same behaviour is observed for all methods.

## 4 Discussion

To the best of our knowledge, this is the first study to explore the existence of ex-ante moral hazard as a result of a change in the health insurance coverage status in Brazil. Previous studies usually analyse EAMH in the context of obtaining health insurance coverage through Medicare enrolment (de Preux, 2011; Card & Maestas, 2008; Dave & Kastner, 2009).

This study implements an empirical strategy similar to de Preux (2011), with at least three main advantages. First, its database contains a population with different age groups, rather than only the elderly, as is the case with other studies. The change of health insurance coverage status occurs for all beneficiaries and not just for the policyholder as in the Medicare experiment.

Because Sabesp is an employment-based health benefit, all employees and their dependents are automatically included in the risk pool when they are hired. Therefore, our data avoids the self-selection bias of health insurance adherence. Also, the longitudinal design allows robust estimation of EAMH because it takes into account unobserved heterogeneity and is a more powerful strategy to identify causality.

The second advantage refers to the mechanism that identifies EAMH. In both exercises, the assumption is that individuals anticipate the change in health insurance coverage. Differing from de Preux (2011), this study refers to a loss of health insurance coverage, and its time window occurs just before the change of insurance coverage. Therefore, it is possible to guarantee that all beneficiaries have equal access to health care services and providers during the entire period, which allows for controlling the access effect.

Finally and yet importantly, in this study, the anticipatory EAMH relates to self-insurance, not self-protection. As there is no effective change in the health insurance coverage and consequently no access effect, it is possible to measure EAMH based on health care utilization instead of lifestyle.

The main findings showed that individuals tend to use fewer preventive health services when they are insured by Sabesp. These results are in

accordance with Ehrlich and Becker (1972), who showed that when prevention is not contractible, health insurance coverage reduces self-insurance due to indirect costs such as displacement and time costs. It applies directly to the case of Sabesp because there is no enforcement for the use of preventive care by the beneficiaries.

The results revealed an increase of up to 17% on medical visits and 22% on diagnostic exams. Medical visits begin to increase from the fifth month before leaving the portfolio, with a peak at two months before exit. For diagnostic tests, the higher increase was observed at one and two months before the loss of health insurance coverage. This result probably occurs because the use of diagnostic tests requires a referral by providers and is usually not a patient's decision.

These results are relevant because the presence of EAMH may affect future costs of illness. In that manner, one policy recommendation is for the introduction of regulatory mechanisms to induce mandatory preventive care. There is substantial evidence for the conclusion that preventive care saves future costs of illness (Cohen, Neumann, & Weinstein, 2008).

## References

- ALLISON, P. D.; WATERMAN, R. P. Fixed-effects negative binomial regression models. *Sociological Methodology*, v. 32, n. 1, p. 247–266, 2002 <https://doi.org/10.1111/1467-9531.00117>.
- ARROW, K. J. (1963). Uncertainty and the welfare economics of medical care. *The American Economic Review*, v. 53, n. 5, p. 941–973, 1963. <https://doi.org/S0042-96862004000200013> [pii].
- BRASIL. Lei nº 9.656, de 3 de junho de 1998. Dispõe sobre os planos e seguros privados de assistência à saúde. Brasília (DF): Senado; 1998.
- CAMERON, C. A.; TRIVEDI, P. 2013. *Regression analysis of count data*, 2<sup>nd</sup> edition. [NEEDS THE PUBLISHER AND CITY].
- CARD, D.; MAESTAS, N. Care utilization : Evidence from Medicare. *American Economic Review*, v. 98, n. 5, p. 2242–2258, 2008. <https://doi.org/10.1257/aer.98.5.2242>.The.
- CHAMBERLAIN G. Panel Data," Chapter 22 in *Handbook of Econometrics*, Vol. by Z. Griliches and MD Intrilligator. Amsterdam: Elsevier Science Publishers BV.
- COHEN JT, NEUMANN PJ, WEINSTEIN MC. Does preventive care save money? Health economics and the presidential candidates. *New England Journal of Medicine*. 2008 Feb 14; 358 (7): 661–3.
- DAVE, D.; KAESTNER, R. Health insurance and ex ante moral hazard: Evidence from Medicare. *International Journal of Health Care Finance and Economics*, v. 9, n. 4, p. 367–390, 2009.

<https://doi.org/10.1007/s10754-009-9056-4>.

- DE PREUX, L. B. Anticipatory ex ante moral hazard and the effect of Medicare on prevention. *Health Economics*, v. 20, n. 9, p. 1056–72, 2011. <https://doi.org/10.1002/hec.1778>.
- DECRETO No 3.048, DE 6 DE MAIO DE 1999 (1999). Brazil. Retrieved from [http://www.planalto.gov.br/ccivil\\_03/decreto/d3048.htm](http://www.planalto.gov.br/ccivil_03/decreto/d3048.htm).
- EHRlich, I.; BECKER, G. S. Market insurance, self-insurance, and self-protection. *The Journal of Political Economy*, v. 80, n. 4, p. 623–648, 1972. <https://doi.org/10.2307/1829358>.
- GIAMBIAGI, F.; AFONSO, L. E. (2015). Previdência do setor público e INSS: a fotografia e o filme. In: TAFNER, P.; BOTELHO, C.; ERBISTI, R. (Eds.), *Reforma da Previdência – a visita da velha senhora*. Brasília: Editora Gestão Pública. 2015. 1a, p. 107–132.
- HAUSMAN, J.; HALL, B. H.; GRILICHES, Z. Econometric model for count data with an application to the patents-R&D relationship. *Econometrica*, v. 52, n. 4, p. 909–938, 1984.
- HILBE, J. M. *Negative binomial regression*. Cambridge: Cambridge University Press, 2011. <https://doi.org/10.1111/j.1540-6210.2010.02207.x>.
- LIANG, K.-Y.; ZEGER, S. L. Longitudinal data analysis using generalized linear models. *Biometrika*, v. 73, n. 1, p. 13–22, 1986. <https://doi.org/10.1093/biomet/73.1.13>.
- LILLARD, L. A., MANNING, W. G., PETERSON, C. E., LURIE, N., GOLDBERG, G. A., & PHELPS, C. E. *Preventive medical care: Standards usage and efficacy*. RAND Publication No. R-3266-HCFA. Santa Monica, CA: RAND, US Department of Health and Human Services, 1986.
- NEWHOUSE, J. and the Insurance Experiment Group. *Free for all? Lessons from the RAND health insurance experiment*. Cambridge, MA: Harvard University Press, 1993.
- PAULY, M. V. The economics of moral hazard: Comment. *American Economic Review*, v. 58, n. 3, p. 531–537, 1968. <https://doi.org/10.2307/1813785>.
- YAMAMOTO, D. H. and Society of Actuaries. Health Care Costs – From Birth to Death. Health Care Cost Institute’s Independent Report Series, Report 2013-1, 2013.
- ZECKHAUSER, R. Medical insurance: A case study of the tradeoff between risk spreading and appropriate incentives. *Journal of Economic Theory*, v. 2, n. 1, p. 10–26, 1970. [https://doi.org/10.1016/0022-0531\(70\)90010-4](https://doi.org/10.1016/0022-0531(70)90010-4).
- ZWEIFEL, P.; MANNING, W. G. Chapter 8 moral hazard and consumer incentives in health care. In: *Handbook of health economics*, Elsevier, 2000. vol. 1. p. 409–459.

### About the authors

Ana Carolina Maia – [anacmaia@usp.br](mailto:anacmaia@usp.br)

Faculdade de Economia, Administração e Atuária da Universidade de São Paulo, São Paulo, SP, Brasil.

ORCID: <https://orcid.org/0000-0002-4956-5948>.

Mônica Viegas Andrade – [mviegas@cedeplar.ufmg.br](mailto:mviegas@cedeplar.ufmg.br)

Centro de Desenvolvimento e Planejamento Regional, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brasil.

ORCID: <https://orcid.org/0000-0003-4606-2795>.

Flávia Chein – [flavia.chein@ufff.edu.br](mailto:flavia.chein@ufff.edu.br)

Departamento de Economia, Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil.

ORCID: <https://orcid.org/0000-0003-4002-2522>.

Authors are grateful to CAPES, Cnpq, and FAPEMIG for financial support.

### **About the article**

Submission received on July 22, 2018. Approved for publication on March 29, 2019.