

Estimando Seleção Adversa em Planos de Saúde

Sandro Leal Alves (*)

Resumo – Este trabalho procura verificar a existência do fenômeno da seleção adversa no mercado de saúde suplementar brasileiro. Através da utilização de metodologia recentemente desenvolvida não foi possível confirmar sua existência para o período pré-regulamentação. Estes resultados lançam novos desafios para a regulação atual especialmente porque alertam para o possível *trade-off* entre garantia de acesso e eficiência econômica quando ocorre o estabelecimento de coberturas mínimas.

Palavras-chave – Seleção Adversa; Regulação; Planos de Saúde

Abstract – This paper investigates the existence of adverse selection in the brazilian health plan market. By using recent developed methodology it was not possible to confirm the existence of the phenomenon in the before regulation period. These results present new challenges to the current regulation, specially because it alerts to the possible trade-off between access assurance and economic efficiency when one imposes minimum coverage of benefits.

Classificação ANPEC: ÁREA 4: Microeconomia, Economia Industrial e Mudança Tecnológica e Métodos Quantitativos
Classificação JEL: I11

* Professor da Universidade Santa Úrsula e economista da Agência Nacional de Saúde Suplementar (ANS).

1 – Introdução

Procuraremos ao longo deste trabalho contribuir para um melhor entendimento das propriedades econômicas do mercado de planos de saúde através de uma investigação empírica sobre os efeitos da assimetria de informação e, particularmente, da seleção adversa. Com este objetivo, apresentamos em seguida um resumo das principais formulações teóricas para lidar com o problema da seleção adversa. Logo após apresentamos a metodologia empírica proposta por Chiappori e Salanié (2000) para testar a existência de seleção adversa, derivada da informação assimétrica. Finalmente, aplicamos o modelo para o mercado de saúde suplementar brasileiro utilizando a base de dados da Pnad Saúde de 1998. A última seção conclui.

2 - Seleção Adversa: Aspectos Teóricos

Utilizando a abordagem da economia dos contratos tradicional, tem-se no modelo de principal-agente um interessante referencial teórico para compreender o problema da seleção adversa. Esta decorre da impossibilidade do principal identificar acuradamente os tipos ou características, dos agentes. Existe uma parte informada, que é o agente, e uma parte não informada, o principal. O agente conhece perfeitamente suas características próprias, mas o principal as desconhece. Extrapolando estes conceitos para a relação regulador-regulado, temos que as empresas reguladas conhecem seus custos e suas produtividades, mas o regulador não. No caso seguradora-segurado, o segurado conhece exatamente o seu risco mas a seguradora não.

A seleção adversa ocorre quando existe assimetria na informação entre a empresa e o consumidor no que diz respeito ao risco que este representa para a empresa. Este é um problema clássico do mercado de seguros, mas também ocorre no setor de planos odontológicos, bem como no setor de planos médicos, pela similaridade destes no que concerne ao risco.

Se uma empresa é incapaz de identificar de forma precisa os consumidores no que se refere ao risco, então esta cobra um preço médio para todos os agentes. Agindo desta maneira, os indivíduos de alto risco estão mais propensos a adquirir o plano do que os de menor risco. Para contornar este problema, as empresas procuram discriminar preços para cada tipo de risco. Este processo é conhecido como *experience rating*, onde o preço cobrado é determinado pela expectativa de gastos dos indivíduos através da análise dos seus fatores de risco. Entretanto, identificar, acuradamente, os riscos não é uma tarefa fácil. Isso explica a necessidade dos investimentos das operadoras¹ na identificação dos indivíduos e dos subseqüentes cálculos probabilísticos de avaliação de riscos.

Os segurados são heterogêneos com respeito aos seus custos esperados e possuem mais informação sobre seu risco do que a seguradora, que não é, em princípio, capaz de diferenciá-lo. Existe um custo para a seguradora na obtenção de informação sobre os tipos dos agentes. Naturalmente, o indivíduo de alto risco não tem incentivo a se “revelar” para a seguradora, no que concerne ao seu risco e, conseqüentemente, seu custo esperado.

¹ No presente contexto não cabe o tratamento diferenciado entre seguradoras e operadoras dado que ambas formas institucionais comercializam planos ou seguros de saúde. Entretanto, convém ressaltar a existência de diferenças do ponto de vista operacional e do tratamento regulatório que não estão sendo considerados neste momento.

Conforme observado por Arrow (1963) nestes mercados geralmente é observado um agrupamento de riscos (*pooling*) que denota uma tendência a equalizar ao invés de diferenciar prêmios. Isso constitui na verdade uma redistribuição de renda daqueles com baixa probabilidade de adoecimento para aqueles com alta probabilidade.

Em seguida, foi demonstrado por Akerlof (1970) que se todos os seguradores têm informação imperfeita sobre o risco individual, o mercado de seguros pode não existir, ou, se existir, pode não ser eficiente. Por isso, pessoas com mais de 65 anos têm dificuldades em contratar um plano de saúde, e os preços são maiores na medida em que condição médica média dos segurados piora, reduzindo o incentivo para que as empresas ofereçam este tipo de contrato. A contribuição destes dois autores impulsionou o desenvolvimento de um amplo conjunto de modelos a fim de explicar a seleção adversa, seu impacto na alocação dos recursos e os mecanismos para sua redução. [Dione, Doberty e Fomfaron (2000)].

Uma primeira geração de modelos desenvolveu-se no sentido de propor mecanismos de auto-seleção como alternativa para redução da ineficiência dos mercados sob seleção adversa. A idéia é que os indivíduos eram capazes de revelar suas características (risco) através da escolha dos contratos. Um indivíduo que escolhesse um contrato do tipo amplo, ou seja, com direito a um maior número de procedimentos, deveria ser um indivíduo com risco elevado. Sabendo disso, as seguradoras deveriam oferecer um menu variado de contratos, com diferentes níveis de cobertura e preços, a fim de que os indivíduos revelassem os seus riscos. Esta forma de alocação mostrou-se superior (em termos de eficiência econômica) do que aquela onde era cobrado um preço médio para todos os indivíduos. O principal trabalho nesta área pode ser atribuído a Rothschild e Stiglitz, (1976). Outras variações de modelos encontraram evidências onde a categorização de riscos, sob certas condições, melhorava a eficiência econômica assim como era possível melhorar a eficiência do mercado mediante o uso da informação relativa à experiência passada (histórico de doenças) do segurado como mecanismo de seleção.

A atividade do seguro têm fornecido boas condições para os testes empíricos da teoria dos contratos [Chiappori (2000)]. Segundo este autor, os dados armazenados pelas companhias seguradoras fornecem um campo vasto para testar as previsões fornecidas pela teoria pois neles estão registradas as informações referentes ao contrato, as informações disponíveis para ambas as partes, à performance do contrato e às transferências de renda entre eles.

Equilíbrio Competitivo no Mercado de Seguros

Nos termos de Rothschild e Stiglitz, (1976), podemos dizer que pelo lado da demanda, a riqueza dos indivíduos é dada por $W_1=W$ se este indivíduo não apresenta nenhuma enfermidade (sinistro) e $W_2=W_1 - d$, no caso de ocorrência da enfermidade sendo d os custos associados à assistência médica demandada para o tratamento da enfermidade. As seguradoras de saúde oferecem α_2 de indenização para o segurado em troca de α_1 de prêmio. Desta forma, a riqueza do indivíduo com seguro será $W_1 = W - \alpha_1$ e $W_2 = W - \alpha_1 + \alpha_2 - d = W + \alpha_2 - d$, onde $\alpha_2 = \alpha_2 - \alpha_1$. Se a probabilidade de ocorrência da enfermidade é dada por p , então pelo teorema da utilidade esperada podemos representar as preferências desses indivíduos por:

$$V(p, \alpha_1, \alpha_2) = (1 - p) U(W - \alpha_1) + p U(W + \alpha_2 - d)$$

Dado p , o indivíduo maximiza $V(\cdot)$ com relação a (α_1, α_2) . Os indivíduos são avessos em relação ao risco e não existe *moral hazard* no modelo, ou seja, estes não alteram a probabilidade de utilização do contrato ex-post. Pelo lado da oferta, as empresas de seguro são neutras em relação ao risco e maximizam o lucro esperado. Um contrato C_i consiste em 1 par (α_1, α_2) contendo uma quantidade específica de cobertura que um indivíduo pode comprar a um preço específico. O lucro esperado de um contrato oferecido a um indivíduo com probabilidade p é dado por:

$$\pi(p, \alpha_1, \alpha_2) = (1 - p) \alpha_1 - p(\alpha_2 - \alpha_1) = (1 - p) \alpha_1 - p\alpha_2$$

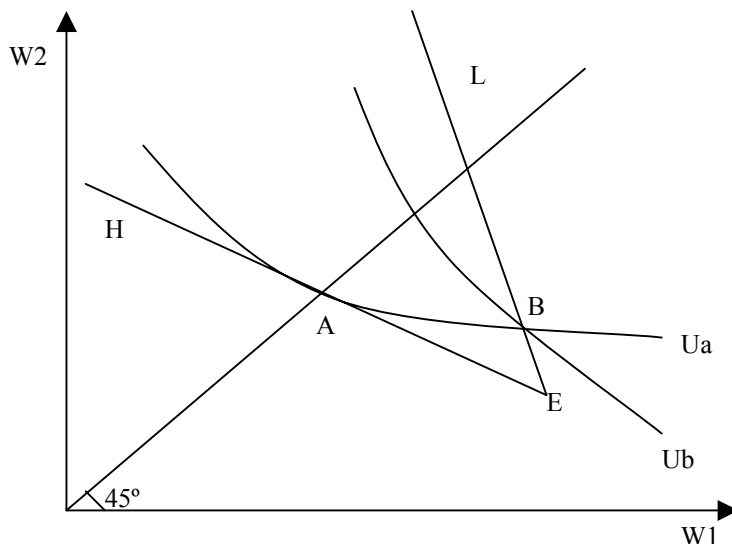
O conjunto de equilíbrio de contratos é definido como:

- consumidores maximizam a utilidade esperada
- nenhum contrato no equilíbrio pode ter lucro não negativo
- nenhum contrato fora do equilíbrio, se oferecido, produz lucro positivo.

A informação assimétrica consiste no fato de que ao adquirirem os contratos, os indivíduos conhecem suas probabilidades de utilização do contrato, ou seja, seus riscos, e as seguradoras não conhecem. Se os agentes forem idênticos, existirá um equilíbrio first-best que é equivalente ao caso com informação completa. No entanto, quando os consumidores são distintos com relação à esta probabilidade, as seguradoras vão se utilizar do comportamento desses agentes no mercado no ato da compra dos contratos para melhorar sua informação sobre essas probabilidades. Neste caso, temos os agentes de alto risco ($p = p_a$) e os agentes de baixo risco ($p = p_b$) e $p_a > p_b$. O percentual de indivíduos de alto risco é dado por λ e a probabilidade de sinistro média é dada por $p_m = \lambda p_a + (1 - \lambda) p_b$. Neste caso, dois equilíbrios são possíveis:

Equilíbrio agregador (*pooling*): Ambos os grupos compram o mesmo contrato e $(1 - p_m) \alpha_1 - p_m \alpha_2 = 0$.

Equilíbrio separados: Cada grupo distinto compra contratos distintos. Ambos os contratos devem ser tais que $(1 - p_a) \alpha_1 - p_a \alpha_2 = 0$ e $(1 - p_b) \alpha_1 - p_b \alpha_2 = 0$. Analisando o equilíbrio agregativo, Rothschild e Stiglitz, (1976) mostram que este não poderá ser um equilíbrio dado que pode ser sempre superado por um contrato que forneça lucros positivos. O único equilíbrio possível neste mercado será o equilíbrio separador, abaixo representado no espaço (W_1, W_2) .



A bissetriz representa os pontos de seguro completo ($W_1 = W_2$). Pontos à direita representam seguros incompletos ($W_1 > W_2$). O ponto E é a situação inicial onde o indivíduo não possui seguro. EH é a isolucro de lucro zero para os contratos oferecidos aos agentes de risco alto e EL é a isolucro de lucro zero para os contratos oferecidos aos agentes de baixo risco. U_a é a curva de indiferença dos agentes de alto risco e U_b é a curva de indiferença para os agentes de baixo risco.

No equilíbrio separador dois tipos de contratos serão oferecidos (A e B) respectivamente para os indivíduos de alto risco e de baixo risco. Quando este vetor de contratos é oferecido, vale a condição de compatibilidade de incentivos:

$$V(p_a, \alpha_a) \geq V(p_a, \alpha_b) \text{ e } V(p_b, \alpha_b) \geq V(p_b, \alpha_a)$$

Nestes contratos os agentes de alto risco compram seguro completo e os agentes de baixo risco ficam sub-segurados configurando uma externalidade negativa dos indivíduos não saudáveis para os saudáveis. Outra característica deste equilíbrio é que sua existência está condicionada a proporcionalidade entre os agentes entre outras coisas.

As Predições Teóricas

Chiappori e Salanié (2000) propuseram um teste a fim de verificar a presença de assimetria de informação, especificamente da seleção adversa, no mercado francês de seguros para automóveis. O objetivo dos autores consistiu no desenvolvimento de um teste simples e ao mesmo tempo geral que fosse capaz de captar a presença do fenômeno.

A partir da teoria da seleção adversa, os autores identificam as seguintes predições teóricas do modelo de equilíbrio competitivo desenvolvido por Rothschild e Stiglitz (1976):

- a) Na presença de seleção adversa, os agentes igualmente observáveis estão diante de um menu de contratos dentre os quais a escolha é livre;
- b) Dentro do menu de contratos, aqueles com cobertura mais ampla possuem o maior preço unitário
- c) Os contratos com maior cobertura são escolhidos pelos agentes com maior probabilidade de utilização.

A primeira predição teórica é demasiadamente ampla pois as diferenças entre os indivíduos podem ocorrer sob várias dimensões como além do risco tais como riqueza, preferências e aversão ao risco. Portanto identificar a parcela relativa às diferenças baseadas no risco requer um modelo mais complexo. Testar a segunda predição demandaria hipóteses adicionais sobre as políticas de precificação das empresas, o que exigiria hipóteses fortes sobre a tecnologia destas empresas.

Alternativamente, a terceira predição teórica sugere um teste razoavelmente simples pois não impõe hipóteses sobre a tecnologia adotada, não depende de hipóteses sobre as preferências, bem como não requer a condição de *single crossing property*², permanece válida para o caso multidimensional e para o caso dos agentes diferirem não somente de acordo com a probabilidade de sinistro mas também de acordo com a severidade deste.

² Esta condição (também é denominada de Condição de Spence-Mirrlees) implica que as curvas de indiferença de dois agentes econômicos de riscos distintos só se cruzam uma única vez. O conteúdo econômico desta condição significa que agentes de tipos de risco elevados (maiores θ 's) estão dispostos a pagar mais por um dado aumento na qualidade do produto do que os agentes de tipo mais baixo. Ou seja, $\partial U / \partial q$ é crescente em θ .

Adicionalmente, as propriedades do teste permanecem no contexto dinâmico [Chiappori e Salanié (2000)].

A tradução empírica do teste resulta numa correlação positiva entre duas distribuições condicionais. A primeira referente à escolha do contrato e a segunda referente à ocorrência do evento. A fim de verificar a correlação positiva entre estas duas distribuições, os autores propõe o seguinte teste:

2 –O Teste de Chiappori e Salanié (2000)

O teste dedica-se a verificar a independência condicional entre a escolha dos contratos de cobertura ampla e sua utilização.

Sejam:

$i = 1, \dots, n$ os indivíduos;

X_i = vetor que representa o conjunto de variáveis exógenas para o indivíduo i ;

w_i = número de dias do ano em questão que o indivíduo i estava segurado;

Variáveis Endógenas Dicotômicas:

$y_i = 1$ se i compra o contrato com cobertura ampla

$y_i = 0$ se i compra o contrato com cobertura mínima

$z_i = 1$ se i utiliza o contrato com cobertura ampla

$z_i = 0$ se i não utiliza o contrato

O autor estima dois modelos Probits, um para a escolha da cobertura e outro para a utilização do contrato. Sendo ϵ_i e η_i os dois termos de erro aleatório iid, então:

$$y_i = X_i \beta + \epsilon_i$$

$$z_i = X_i \gamma + \eta_i$$

Após estimar as regressões, onde o peso de cada indivíduo deve ser o número de dias segurado (w_i). Em seguida, são computados os resíduos das regressões ϵ_i e η_i . Por exemplo,

$$\varepsilon_i = E(\varepsilon_i \mid y_i) = \frac{\phi(X_i \beta)}{\Phi(X_i \beta)} y_i - (1 - y_i) \frac{\phi(X_i \beta)}{\Phi(-X_i \beta)},$$

onde ϕ e Φ denotam a função de densidade e a função de distribuição acumulada da $N(0,1)$. Em seguida, seja a estatística W definida como:

$$W = \frac{\left(\sum_{i=1}^n w_i \varepsilon_i \eta_i \right)^2}{\sum_{i=1}^n w_i^2 \varepsilon_i^2 \eta_i^2}$$

Gouriéroux et al. (1987) mostram que sob a hipótese nula da independência condicional, a cov $(\varepsilon_i, \eta_i) = 0$ e W possui distribuição $\chi^2(1)$. Isto provê um teste para a seleção adversa onde a rejeição da hipótese nula de que os erros são não correlacionados indica a presença de seleção adversa.

3 – Implementando o Teste de Chiappori e Salanié para o Setor de Saúde Suplementar Brasileiro

Nesta seção procuraremos implementar o teste proposto por Chiappori e Salanié (2000) a fim de verificar a presença de seleção adversa para o mercado de planos de saúde suplementar. A estratégia está em realizar os testes sobre as consequências empíricas associadas à presença de seleção adversa tal como propõe os autores.

A base de dados que suporta a análise é a Pesquisa Nacional por Amostra de Domicílio (PNAD) desenvolvida pelo IBGE, no ano de 1998. Neste ano, o IBGE incluiu na pesquisa o suplemento de saúde, que possibilitou a análise do setor de planos de saúde.

a) Variáveis Endógenas

a1) Construção da Variável de Escolha (E)

A variável de escolha é definida da seguinte forma:

$$E = 1, \text{ se o indivíduo é detentor de um contrato do tipo amplo}$$

$$E = 0, \text{ se o indivíduo é detentor de um contrato do tipo mínimo}$$

Para o contrato do tipo mínimo (CM) foi considerado aquele que oferece cobertura para o grande risco, pelo menos. Entendemos a cobertura para internações como sendo a cobertura mínima necessária para garantir proteção contra o grande risco. O contrato que cobre pelo menos esses eventos são considerados contratos mínimos. No entanto, isso não significa que os contratos mínimos cobrem apenas as internações hospitalares. Tais contratos podem oferecer coberturas adicionais mas sempre combinadas com as internações.

Para o contrato amplo (CA) foi considerado o contrato que oferece além da cobertura para internações, cobertura para consultas, para exames complementares e para procedimentos odontológicos. Portanto, definimos estas variáveis da forma a seguir:

$$CA = 1, \text{ se o indivíduo tem direito a internações, consultas médicas, exames complementares e procedimentos odontológicos;}$$

$$CA = 0, \text{ caso contrário; e}$$

$$CM = 1, \text{ se o indivíduo tem cobertura para internações pelo menos;}$$

$$CM = 0, \text{ caso contrário.}$$

Obviamente, o indivíduo que possui um contrato amplo possui necessariamente um contrato do tipo mínimo mas o inverso não é necessariamente verdadeiro.

A2) Construção da Variável de Utilização

Chiappori e Salanié (2000), construíram a segunda variável endógena, utilização, a partir da simples observação do evento associado á utilização ou não do contrato. No caso de saúde, esta forma de computação da variável não nos parece ser adequada pois a utilização do contrato deve estar relacionada ao risco de sinistro. A aplicação direta do conceito dos autores seria a definição de uma variável dicotômica, do tipo 0-1, onde a não utilização do contrato assumiria valor zero e a utilização assumiria valor um. No entanto, a sua simples aplicação teria o inconveniente de incluir na variável utilização todos os procedimentos utilizados para fins de prevenção. Entendemos que o comportamento de prevenção não deve estar relacionado com o risco de ocorrência de um evento. A ação preventiva não pode estar associada ao uso do contrato.

Nosso problema está em criar uma variável de utilização, que agregue informações de naturezas distintas. Para isso, precisamos desenvolver um pouco mais o conceito de utilização de serviços de saúde suplementar a fim de criar um índice que nos permita discriminar os indivíduos que utilizaram mais os seus contratos daqueles que utilizaram menos. Inicialmente procuramos criar uma variável que indique o grau desta utilização do contrato. Esta variável foi construída a partir da ponderação pela participação dos gastos referentes a cada tipo de cobertura sobre o total de gastos ocorridos em consultas, exames complementares e internações. O índice de utilização proposto assume a seguinte forma:

$$Iu_i = \beta_I I + \beta_C C + \beta_E E + \beta_O O$$

Onde:

$$\beta_I = \frac{\sum_{i=1}^n G_I}{\sum_{i=1}^n G_{\cdot T}}; \quad \beta_C = \frac{\sum_{i=1}^n G_C}{\sum_{i=1}^n G_{\cdot T}}; \quad \beta_E = \frac{\sum_{i=1}^n G_E}{\sum_{i=1}^n G_{\cdot T}}; \quad \beta_O = \frac{\sum_{i=1}^n G_O}{\sum_{i=1}^n G_{\cdot T}}$$

- i = indivíduo detentor de plano de saúde na amostra onde $i = 1, 2, 3, \dots, n$.
- I = número de internações do indivíduo i no período;
- C = número de consultas médicas realizadas pelo indivíduo i no período;
- E = número de exames complementares realizados pelo indivíduo i no período;
- O = utilização de procedimentos odontológicos;
- β_I = peso das internações no total de gastos com saúde;
- β_C = peso das consultas médicas no total de gastos com saúde;
- β_E = peso dos exames no total de gastos com saúde;
- β_O = peso dos gastos com odontologia no total de gastos com saúde;
- U_i = Índice de utilização do indivíduo i .
- G_I = total de gastos com internações;
- G_C = total de gastos com consultas;
- G_E = total de gastos com exames complementares;
- G_O = total de gastos com procedimentos odontológicos;

- G_T = total de gastos com consultas, internações e exames complementares.
- C = número de consultas médicas realizadas nos últimos 12 meses;
- I = número de internações do indivíduo nos últimos 12 meses;
- E = número de exames complementares realizados pelo indivíduo nas duas semanas anteriores à pesquisa;
- O = número de procedimentos odontológicos realizados nas últimas duas semanas;

Calculamos os β_{i_s} com base na amostra total, ou seja, para os 344.975 indivíduos e não somente para aqueles detentores de planos de saúde. Os ponderadores são apresentados em seguida.

$$Iu_i = 0.1(E) + 0.2 (I) + 0.53 (O) + 0.17 (C)$$

Assim, construímos a variável Iu_i para o conjunto de indivíduos presentes em nossa amostra (5129). Agora, a dificuldade que se apresenta é transformar esta variável, que assume valores no intervalo (0-17,63), em uma variável dicotômica 0-1, para que a implementação do teste seja possível.

Neste caso, é possível estabelecer um ponto de corte a fim de separar os eventos. Para indivíduos com valores acima deste limite são considerados como se tivessem utilizado seus contratos e os indivíduos que se encontrassem abaixo sendo considerados como se não os tivessem utilizado. A variável utilização é definida como:

- $U_i = 0$, se $Iu_i \leq$ ponto de corte;
- $U_i = 1$, se $Iu_i >$ ponto de corte.

Objetivando reduzir o grau de arbitrariedade no estabelecimento do ponto de corte, criamos cinco variáveis de utilização, observando diferentes pontos de corte. Utilizamos a média, a mediana, a moda, o quarto percentil e o sexto percentil como ponto de corte. Desta forma podemos observar a sensibilidade do teste com relação à definição da variável de utilização.

Tabela 1 - Estatísticas Descritivas da Variável Utilização

Média	0,75
Moda	0,00
Mediana	0,51
Quarto Percentil	0,34
Sexto Percentil	0,61
Mínimo	0,00
Máximo	17,63
Desvio-padrão	1,03

Fonte: Elaboração Própria

b) Variáveis Exógenas de Controle

- Auto-avaliação do Estado de Saúde (AES);

A variável auto-avaliação do estado de saúde (AES) assume os seguintes valores:

AES = 0, se o indivíduo se auto-avalia como ruim ou muito ruim;

AES = 1, se o indivíduo se auto-avalia como regular;
AES = 2, se o indivíduo se auto-avalia como bom ou muito bom;

- Qualidade do saneamento básico (SAN);

A variável escolhida para qualidade do saneamento básico refere-se ao escoadouro sanitário, que assume os seguintes valores:

SAN=0, no caso de fossa rudimentar, vala ou esgoto jogado direto no rio, lago ou mar;
SAN=1, no caso de fossa séptica não ligada à rede coletora;
SAN=2, no caso de fossa séptica ligada à rede coletora e existência de rede coletora de esgoto;

- Nível de Educação (EDUC);

A variável nível de educação assume os seguintes valores:

EDUC = 0, para indivíduos sem instrução ou com menos de 1 ano de estudo;
EDUC = 1, para indivíduos entre 1 e 3 anos de estudo;
EDUC = 2, para indivíduos entre 4 e 10 anos de estudo;
EDUC = 3, para indivíduos com 11 ou mais anos de estudo;

- Sexo (S);

A variável sexo assume os seguintes valores:

S=0, no caso de sexo feminino;
S=1, no caso de sexo masculino.

- Nível de Renda (R);

- Idade (ID);

As variáveis idade e número de dependentes no plano assumem valores contínuos assim como a variável nível de renda, definida como sendo a renda mensal familiar em reais (R\$).

- Presença de Coparticipação (Cop)³;

É uma variável binária onde 1 indica que o indivíduo tem algum mecanismo de coparticipação e 0 indica a não existência de tal mecanismo.

- Preço (P)

Esta variável assume os seguintes valores:

³ Coparticipação é um instrumento utilizado pelas companhias seguradoras ou pelas operadoras de planos para reduzir a utilização dos serviços. Uma coparticipação de 10% significa que o indivíduo assume 10% do custo do evento e a empresa assume os outros 90 %.

P=0 se a mensalidade encontra-se entre R\$30,00 e R\$50,00;
P=1 se a mensalidade está entre R\$50,00 e R\$100,00;
P=2 se a mensalidade está entre R\$100,00 e R\$200,00;
P=3 se a mensalidade está entre R\$200,00 e R\$300,00;
P=4 se a mensalidade está entre R\$300,00 e R\$500,00;
P=5 se a mensalidade é superior a R\$500,00.

Tratamento dos Dados

O arquivo original da PNAD possuía 344.975 observações. Realizamos o procedimento de filtragem a fim de selecionarmos apenas os indivíduos detentores de planos de saúde do tipo individual ou familiar. A escolha deste tipo de contrato se justifica pois é neste caso onde o consumidor se depara com um menu de contratos para exercer sua escolha. Nos planos empresariais, a escolha não é baseada no risco de adoecimento do indivíduo que compra o plano. A empresa contratante define a compra do plano de saúde e suas coberturas.

O filtro foi realizado na variável que define quem paga pelo plano de saúde⁴ onde selecionamos apenas os titulares de planos de saúde que efetuam o pagamento diretamente para a operadora de planos, sem a intermediação da empresa onde trabalha. Após este procedimento trabalhamos com 10.460 observações.

Em seguida selecionamos para compor nossa amostra apenas os indivíduos que têm planos e cujas informações referem-se ao próprio indivíduo. Eliminamos da amostra as observações onde o informante não reside no domicílio ou é outra pessoa moradora do domicílio que não a detentora do plano de saúde individual (um dependente por exemplo). A nova amostra conta com 5.436 observações.

Após eliminarmos as observações com variáveis faltando ou não identificadas, nossa amostra final ficou com 5.129 indivíduos.

4 - Resultados do Modelo

Estimamos as seguintes regressões Probits, independentemente:

$$E_i = X_i \beta + \epsilon_i$$

Onde:

E é a escolha entre um contrato amplo e um contrato mínimo;
X_i são as variáveis exógenas definidas anteriormente;
ε_i são os resíduos da regressão.

$$U_i = X_i \gamma + \eta_i$$

Onde:

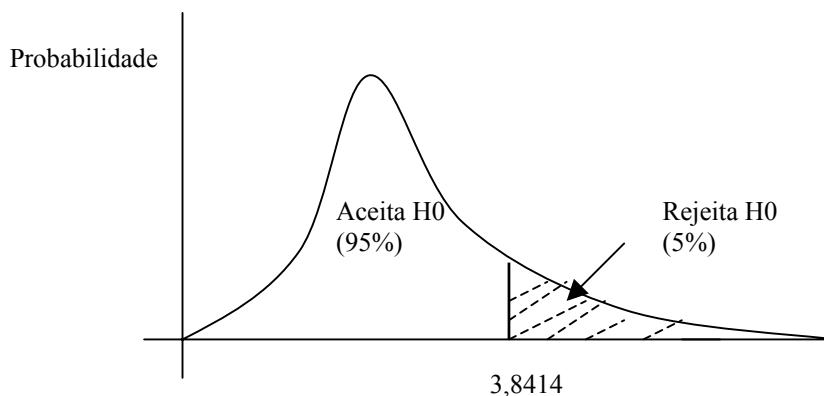
U é a variável de utilização;
X_i são as variáveis exógenas definidas anteriormente;
η_i são os resíduos da regressão.

⁴ V1332 na nomenclatura da PNAD.

Após estimarmos as duas probits independentes para escolha e utilização, implementamos o teste W, supondo que todos os indivíduos tivessem o mesmo tempo de seguro, ou seja, os pesos (w_i) são iguais para todos eles. Na hipótese de independência condicional [$cov(\varepsilon_i, \eta_i)=0$], a estatística W calculada apresenta uma distribuição qui-quadrada com 1 grau de liberdade ($\chi^2(1)$). Isto nos permite testar a presença de seleção adversa mediante o seguinte teste de hipótese:

H0: $cov(\varepsilon_i, \eta_i)=0$;

H1: $cov(\varepsilon_i, \eta_i)\neq 0$;



Logo, aceitar a hipótese nula significa aceitar a ausência de covariância entre os erros aleatórios das duas Probits, que é exatamente aceitar a ausência de seleção adversa no modelo. Rejeitar esta hipótese significa que não se pode descartar a existência de covariância entre os erros e portanto é possível a existência de seleção adversa.

Uma crítica que poderia surgir com relação ao teste proposto diz respeito ao seu grau de arbitrariedade com que propomos transformar o índice de utilização em uma variável dicotômica para então podermos aplicar o teste proposto por Chiapporri e Salanié. Primeiramente, escolhemos a mediana como sendo a medida de posição com que realizamos a separação entre aqueles indivíduos que utilizaram o contrato daqueles que não o utilizaram. Com o intuito de verificar a sensibilidade do teste à definição da variável utilização, fizemos flutuar o ponto de corte que separa a variável utilização para o quarto (40%) e o sexto percentis (60%), ou seja, deslocamos este ponto para cima e para baixo da mediana. Fizemos ainda o exercício utilizando a média e a moda como ponto de corte. Assim, abaixo seguem os resultados encontrados.

Tabela - Resultados Encontrados

Equação Probit	Equação Probit	Valor de W	Teste de hipótese	Presença de Seleção Adversa
Escolha	Utilização (ponto de corte na mediana)	0,0529	Aceita H0	Não é possível Aceitar
Escolha	Utilização (ponto de corte na média)	0,0555	Aceita H0	Não é possível Aceitar
Escolha	Utilização (ponto de corte na moda)	0,0856	Aceita H0	Não é possível Aceitar
Escolha	Utilização (ponto de corte no quarto percentil)	0,0809	Aceita H0	Não é possível Aceitar
Escolha	Utilização (ponto de corte no sexto percentil)	0,0589	Aceita H0	Não é possível Aceitar

Fonte: Elaboração Própria

Como podemos observar, os valores encontrados para a estatística W situam-se na região de aceitação da hipótese nula, indicando a ausência de seleção adversa no modelo. O resultado apresenta-se pouco sensível às variações nos pontos de corte da variável de utilização pois situa-se sempre na região de aceitação.

Embora o objetivo principal encontre-se na análise da estatística W, os modelos estimados, cujas regressões estão apresentadas no Anexo F, fornecem informações importantes sobre o comportamento dos indivíduos detentores de planos de saúde que merecem alguns breves comentários.

A variável saneamento não foi estatisticamente significativa na maioria das regressões estimadas (exceto na regressão da utilização quando o ponto de corte escolhido foi a moda). Isto sugere que o grau de saneamento básico das moradias não explica nem a escolha do plano de saúde nem a sua utilização.

A variável auto-avaliação do estado de saúde (AES) não é estatisticamente significativa na determinação da escolha do plano mas é altamente significativa na determinação da utilização dos serviços. O sinal negativo com que aparece significa que quanto pior for a auto-avaliação do estado de saúde do indivíduo, maior será a sua utilização.

A variável coparticipação é altamente significativa na regressão da escolha e nas regressões de utilização (embora só o seja a 10% de significância no caso da moda como ponto de corte). Seu sinal negativo indica que a presença deste mecanismo de contenção do risco moral de fato inibe a utilização dos serviços assim como a escolha do contrato.

A variável educação é estatisticamente significativa em todas as regressões e possui sinal positivo indicando que o indivíduo mais educado tende a comprar planos mais amplos assim como tende a utilizá-lo mais.

A variável idade não foi estatisticamente significativa na decisão da escolha do plano de saúde embora nas equações de utilização o tenha sido. Quanto maior a idade, maior a utilização do indivíduo, evidenciado pelo sinal positivo deste coeficiente.

A renda se apresentou como sendo estatisticamente significativa em apenas duas das

regressões estimadas (escolha e utilização tendo como ponto de corte a moda), tendo sinais positivos em ambos os casos.

O nível de preço do plano se apresentou como sendo altamente significativo em todas as regressões exceto a regressão da utilização com a moda como ponto de corte.

Finalmente, o sexo foi altamente significativo em todas as regressões e apresentou sinal negativo indicando que as mulheres tendem a comprar seguros mais amplos e usá-los mais intensamente que os homens.

Vale ressaltar que também procuramos encontrar evidências da presença de seleção adversa por um método relativamente diferente do idealizado pelos autores. No entanto, optamos por investir mais esforço no método anteriormente apresentado por este estar mais fortemente documentado na literatura. No entanto, vamos fazer um breve relato do teste proposto.

Estimamos duas equações. A primeira equação foi a Probit da escolha tal como apresentada anteriormente ($E_i = X_i \beta + \epsilon_i$) A segunda equação estimada foi a do índice de utilização ($IU_i = X_i \gamma + \eta_i$), que não é uma variável dicotômica. Estimamos pelo método tradicional de regressão múltipla contra as mesmas variáveis de controle utilizadas na equação de escolha. Os resultados foram erros gerados por uma probit e erros gerados por uma regressão linear. Fizemos uma terceira regressão entre os dois resíduos ($\epsilon_i = \alpha + \beta \eta_i + \xi$) e testamos a significância do seu coeficiente através do teste t-student. O coeficiente β estimado não é estatisticamente diferente de zero a um nível de 10% de significância indicando, portanto, a ausência de correlação entre os erros das duas equações. Este resultado também está no sentido de não aceitar a presença da seleção adversa. Este resultado, embora confirme todos os outros resultados apresentados anteriormente, ainda necessita de um esforço teórico a fim de garantir sua validade. Por certo, o caso apresentado em Chiappori e Salanié (2000) mostra-se como sendo um caso mais geral que até poderia admitir a solução proposta. Entretanto, a confirmação da validade vai além do escopo do presente trabalho e, por hora, optamos pelos testes dos referidos autores já conhecidos na literatura.

Tabela 2 – Estudo das Variáveis Independentes

Variável	Sinais	Coefficiente Estimado	Comentários
SAN	Nd	Não significativa	Não explicou a decisão nem a utilização
AES	-	Significante, exceto na reg. escolha	Quanto pior a auto-avaliação, maior será a utilização. Mas não explica a escolha.
EDUC	+	Significante em todas as reg.	Quanto maior o nível de educação, maior a probabilidade de comprar um contrato amplo e maior a utilização do contrato.
IDAD	+	Significante, exceto na reg. escolha	Quanto maior a idade, maior a utilização
RENDA	+	Significante em apenas duas reg. (escolha e util com moda)	Quanto maior a renda, maior a probabilidade de se escolher um contrato amplo
PREÇO	+	significante	Quanto maior o preço maior a probabilidade de escolha de um contrato amplo e maior a utilização
SEXO	-	significante	Mulheres escolhem contratos mais amplos e utilizam mais
COP	-	significante	A coparticipação reduz a utilização

Fonte: Elaboração Própria

5 – Conclusões

Como resultado da aplicação da metodologia desenvolvida por Chiappori e Salanié (2000) para mercado brasileiro de saúde suplementar não obtivemos a evidência empírica necessária para suportar a hipótese da existência do fenômeno da seleção adversa. Este resultado pode ser explicado pela não incorporação no teste de um modelo multidimensional que incorpore outros elementos de assimetria de informação além do próprio risco individual tais como o grau de aversão ao risco e a probabilidade de sinistro. Esta explicação, no entanto, demanda um esforço de pesquisa adicional para que possa ser corroborada.

Outra explicação possível está no próprio reconhecimento de que os agentes econômicos procuram reduzir a assimetria de informação antes do estabelecimento dos contratos. Neste caso, as próprias empresas investem na identificação dos seus riscos assumidos através de entrevistas qualificadas que buscam conhecer o estado de saúde do segurado e, conseqüentemente, estimar o prêmio de risco. Adicionalmente, a relativa liberdade de oferecimento de contratos antes da regulamentação do setor em 1998, permitia à empresa o estabelecimento de menus de contratos diferenciados onde os agentes se auto-revelavam no ato da compra.

O reconhecimento da plausibilidade da segunda explicação lança luz sobre o possível *trade-off* entre acesso ao mercado e eficiência econômica. A regulamentação do setor procurou proteger os consumidores de planos de saúde através da obrigatoriedade de oferecimento de contratos com coberturas mínimas. Este procedimento, como mostram alguns autores pode não levar a economia para uma alocação *second-best* além de estar relacionada à redução de oferta e perdas de bem-estar, como mostram Neudeck and Podczeck (1996) e Finkelstein (2002).

6 – Bibliografia

AKERLOF, G. (1970) “The Market for Lemons: Quality Uncertainty and the Market Mechanism” *Quarterly Journal of Economics* n. 74, p.488-500.

ANDRADE, M.V. (2000) “Ensaio em Economia da Saúde”. Tese de Doutorado. EPGE/FGV.

ARROW, K (1963) “Uncertainty and the Welfare Economics of Medical Care”. In: *The American Economic Review*, v.LIII, n.5, dec.

CHIAPPORI, P. A (2000) “Econometric Models of Insurance under Asymmetric Information”. In: *Handbook of Insurance*, p.365-393.

_____ and SALANIÉ B. (2000). “Testing for Asymmetric Information in Insurance Markets”. *Journal of Political Economy*, 108, 56-78.

DIONE, DOBERTY E FOMFARON (2000) “Adverse Selection in Insurance Market” In: *Handbook of Insurance*, p.185-243.

FINKELSTEIN, A (2002) “Minimum Standard and Insurance Regulation: Evidence from the Medigap Market” NBER Working Paper Series 8917.

NEUDECK W. and PODCZECK K. (1996). “Adverse Selection and Regulation in Health Insurance Markets”. In: *Journal of Health Economics*, v.15, p.387-408.

PINDYCK, R.D, and RUBINFELD, D.L. (1998) “Econometric Models and Economic Forecast” MC Graw Hill 4th-ed.

ROTHSCHILD, M., and STIGLITZ, J., (1976) “Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information”, *Quarterly Journal of Economics*; 80:629-649.

SALANIÉ, B. (1997). *The Economics of Contracts – A Primer*. The MIT Press.

VARIAN, H. (1992) “*Microeconomic Analysis*” . W.W. Norton and Company. 3 ed.

APÊNDICE - Resultados das Regressões Estimadas

1 - Regressão Probit da Equação de Escolha

Dependent Variable: ESCOLHA

Method: ML - Binary Probit

Sample: 2 5130

Included observations: 5129

Convergence achieved after 4 iterations

Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
AES	0.069943	0.048570	1.440053	0.1499
COP	-0.388665	0.058383	-6.657207	0.0000
EDUC	0.073244	0.037800	1.937707	0.0527
LOGIDADE	0.151287	0.101813	1.485922	0.1373
LOGRENDA	0.161323	0.068697	2.348324	0.0189
PRECO	0.414035	0.029685	13.94744	0.0000
SAN	-0.002460	0.037805	-0.065063	0.9481
SEX	-0.211915	0.052058	-4.070771	0.0000
Mean dependent var	0.894521	S.D. dependent var	0.307199	
S.E. of regression	0.287532	Akaike info criterion	0.582174	
Sum squared resid	423.3758	Schwarz criterion	0.592379	
Log likelihood	-1484.986	Hannan-Quinn criter.	0.585746	
Avg. log likelihood	-0.289527			
Obs with Dep=0	541	Total obs	5129	
Obs with Dep=1	4588			

2 - Probit Utilização (Ponto de Corte: Media)

Dependent Variable: UTILIZME

Method: ML - Binary Probit

Sample: 2 5130

Included observations: 5129

Convergence achieved after 3 iterations

Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
AES	-0.664000	0.037847	-17.54452	0.0000
COP	-0.173112	0.051789	-3.342663	0.0008
EDUC	0.093154	0.029910	3.114516	0.0018
LOGIDADE	0.218027	0.076164	2.862603	0.0042
LOGRENDA	0.055141	0.051109	1.078877	0.2806
PRECO	0.073686	0.017454	4.221718	0.0000
SAN	0.029394	0.030889	0.951603	0.3413
SEX	-0.564056	0.039659	-14.22263	0.0000
Mean dependent var	0.322675	S.D. dependent var	0.467545	
S.E. of regression	0.439912	Akaike info criterion	1.142965	
Sum squared resid	991.0297	Schwarz criterion	1.153170	
Log likelihood	-2923.134	Hannan-Quinn criter.	1.146537	
Avg. log likelihood	-0.569923			
Obs with Dep=0	3474	Total obs	5129	
Obs with Dep=1	1655			

3 - Probit Utilização (Ponto de Corte: Mediana)

Dependent Variable: UTILIZMD

Method: ML - Binary Probit

Sample: 2 5130

Included observations: 5129

Convergence achieved after 4 iterations

Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
AES	-0.684956	0.038299	-17.88465	0.0000
COP	-0.182180	0.049351	-3.691548	0.0002
EDUC	0.123318	0.029145	4.231139	0.0000
LOGIDADE	0.496344	0.074413	6.670088	0.0000
LOGRENDA	0.008046	0.049581	0.162278	0.8711
PRECO	0.058746	0.016952	3.465381	0.0005
SAN	0.007555	0.029712	0.254268	0.7993
SEX	-0.549261	0.037934	-14.47928	0.0000
Mean dependent var	0.422694	S.D. dependent var	0.494036	
S.E. of regression	0.463873	Akaike info criterion	1.239509	
Sum squared resid	1101.929	Schwarz criterion	1.249714	
Log likelihood	-3170.722	Hannan-Quinn criter.	1.243082	
Avg. log likelihood	-0.618195			
Obs with Dep=0	2961	Total obs	5129	
Obs with Dep=1	2168			

4 - Probit Utilização (Ponto de Corte: Moda)

Dependent Variable: UTILIZMO

Method: ML - Binary Probit

Sample: 2 5130

Included observations: 5129

Convergence achieved after 3 iterations

Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
AES	-0.637997	0.051981	-12.27370	0.0000
COP	-0.089047	0.054454	-1.635258	0.1020
EDUC	0.147878	0.033938	4.357271	0.0000
LOGIDADE	0.855932	0.088714	9.648266	0.0000
LOGRENDA	0.136504	0.057222	2.385529	0.0171
PRECO	0.030756	0.019707	1.560648	0.1186
SAN	0.055136	0.033174	1.662028	0.0965
SEX	-0.622179	0.043323	-14.36153	0.0000
Mean dependent var	0.806785	S.D. dependent var	0.394859	
S.E. of regression	0.376812	Akaike info criterion	0.896449	
Sum squared resid	727.1152	Schwarz criterion	0.906654	
Log likelihood	-2290.943	Hannan-Quinn criter.	0.900021	
Avg. log likelihood	-0.446665			
Obs with Dep=0	991	Total obs	5129	
Obs with Dep=1	4138			

5 - Probit Utilização (Ponto de Corte: Sexto Percentil)

Dependent Variable: UTILIZ60

Method: ML - Binary Probit

Sample: 2 5130

Included observations: 5129

Convergence achieved after 3 iterations

Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
AES	-0.687929	0.038109	-18.05161	0.0000
COP	-0.183302	0.049845	-3.677415	0.0002
EDUC	0.108793	0.029285	3.714990	0.0002
LOGIDADE	0.434911	0.074700	5.822075	0.0000
LOGRENDA	0.025184	0.049876	0.504932	0.6136
PRECO	0.062233	0.017054	3.649288	0.0003
SAN	0.020290	0.029958	0.677286	0.4982
SEX	-0.566526	0.038279	-14.79991	0.0000
Mean dependent var	0.398908	S.D. dependent var		0.489722
S.E. of regression	0.459044	Akaike info criterion		1.219898
Sum squared resid	1079.102	Schwarz criterion		1.230103
Log likelihood	-3120.427	Hannan-Quinn criter.		1.223470
Avg. log likelihood	-0.608389			
Obs with Dep=0	3083	Total obs		5129
Obs with Dep=1	2046			

6 - Probit Utilização (Ponto de Corte: Quarto Percentil)

Dependent Variable: UTILIZ40

Method: ML - Binary Probit

Sample: 2 5130

Included observations: 5129

Convergence achieved after 4 iterations

Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
AES	-0.664009	0.039966	-16.61414	0.0000
COP	-0.197472	0.048444	-4.076283	0.0000
EDUC	0.097177	0.029183	3.329890	0.0009
LOGIDADE	0.656893	0.074771	8.785445	0.0000
LOGRENDA	0.026851	0.049435	0.543155	0.5870
PRECO	0.062246	0.016927	3.677226	0.0002
SAN	0.025155	0.029556	0.851109	0.3947
SEX	-0.577123	0.037488	-15.39475	0.0000
Mean dependent var	0.536362	S.D. dependent var		0.498725
S.E. of regression	0.468599	Akaike info criterion		1.257259
Sum squared resid	1124.495	Schwarz criterion		1.267464
Log likelihood	-3216.242	Hannan-Quinn criter.		1.260831
Avg. log likelihood	-0.627070			
Obs with Dep=0	2378	Total obs		5129
Obs with Dep=1	2751			