

UNIVERSIDADE DE SÃO PAULO  
FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE  
DEPARTAMENTO DE ECONOMIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

UTILIZAÇÃO DE MODELOS DE CONTAGEM NA ESTIMAÇÃO DA DEMANDA POR  
CONSULTAS MÉDICAS

Autor: Paloma Vaissman Uribe

Orientador: Prof. Dr. Denisard Cneio de Oliveira Alves

SÃO PAULO

2008

Prof. Dra. Suely Vilela  
Reitor da Universidade de São Paulo

Profa. Dr. Carlos Roberto Azzoni  
Diretora da Faculdade de Economia, Administração e Contabilidade

Prof. Dr. Joaquim José Martins Guilhoto  
Chefe do Departamento de Economia

Prof. Dr. Dante Mendes Aldrighi  
Coordenador do Programa de Pós-Graduação em Economia

**PALOMA VAISSMAN URIBE**

**UTILIZAÇÃO DE MODELOS DE CONTAGEM NA ESTIMAÇÃO DA DEMANDA  
POR CONSULTAS MÉDICAS**

Dissertação apresentada ao Departamento de  
Economia da Faculdade de Economia,  
Administração e Contabilidade da  
Universidade de São Paulo como requisito pra  
obtenção do título de Mestre em Economia

**Orientador: Prof. Dr. Denisard Cneio de Oliveira Alves**

**SÃO PAULO**

**2008**

## **FICHA CATALOGRÁFICA**

Elaborada pela Seção de Processamento Técnico do SBD/FEA/USP

Uribe, Paloma Vaissman

Utilização de modelos de contagem na estimação da  
demanda por consultas médicas / Paloma Vaissman

Uribe. – São Paulo, 2008.

106 p.

Dissertação (Mestrado) – Universidade de São Paulo, 2008

Bibliografia

1. Economia da saúde 2. Econometria 3. Sistema unificado  
de saúde I. Universidade de São Paulo. Faculdade de Economia,  
Administração e Contabilidade. II. Título.

CDD – 338.433621

## AGRADECIMENTOS

Em primeiro lugar gostaria de agradecer aos meus familiares, que sempre me ajudaram de todas as formas possíveis, seja no plano financeiro, seja no plano pessoal. Estarei sempre grata a vocês pelo apoio e incentivo ao estudo e conhecimento, atributos altamente valorizados na nossa família. Em especial gostaria de prestar uma homenagem aos meus avós, Fani e Isaac, duas pessoas fantásticas e lutadoras, que construíram uma vida pautada em carinho, respeito e valorização do outro.

Agradeço imensamente à minha mãe, Magda, pelos conselhos, pela alegria e estímulo que ela me traz toda vez que a encontro, e também pelas palavras mais assertivas e impositivas que me fazem não perder o rumo e auxiliam a me restaurar diante de um obstáculo. Também sou grata à Carol, minha irmã, que mesmo de longe, acredita muito em mim e me incentiva, sempre positivamente. Ao meu pai Javier, agradeço pelas valiosas discussões sobre o sistema de saúde, a economia e a política e pelo estímulo à pesquisa. Agradeço ainda a meu irmão, Alexandre e à Beth, que sempre participaram da minha vida e me proporcionaram muitos momentos de alegria. À Natália, figura única e especial na minha vida, com seu jeito doce, teimoso e peculiar, que me ajudou a crescer e que sempre procurou me proteger com suas orações.

Gostaria de agradecer muito também ao Prof. Denisard, que acreditou em mim desde o início, mesmo que não me conhecendo há tempos. Serei eternamente grata aos seus ensinamentos de econometria e de vida e às contribuições valiosas para o andamento deste trabalho. Com seu carisma e jeito próprio, foi o professor com que mais me identifiquei no mestrado, e que mesmo com suas intervenções mais duras, foi peça fundamental na conclusão desta dissertação. Agradeço também à Dora, pela sua paciência, simpatia e auxílio nas tarefas burocráticas.

Aos meus amigos do mestrado, Marina, Paula, Jaqueline, Priscila, Pedro, Everton, Érica, Raone, Sérgio: muito obrigada pelo auxílio nas intermináveis listas de exercício e pelas discussões enriquecedoras. Muito obrigada também pelos momentos inesquecíveis de lazer, e pela amizade com que me receberam.

Aos meus colegas de trabalho, principalmente à Maria Gabriela e ao Edgar, que me deram o suporte necessário para que eu pudesse tirar períodos de folga para escrever a dissertação. À Maria Gabriela, pela enorme simpatia e compreensão, e pelos valiosos conselhos de quem também já passou por esta fase.

Finalmente, agradeço ao Renato, pela paciência, carinho e atenção do dia-a-dia e por ter se disposto a enfrentar a difícil tarefa de revisão do texto, a fim de tornar a tarefa do leitor mais fácil.

## RESUMO

Este trabalho tem por objetivo identificar e quantificar o efeito dos determinantes para a demanda de consultas médicas no Brasil. Utilizando a base de dados do suplemento sobre saúde da PNAD de 2003, foram estimados modelos de contagem e comparados segundo critérios estatísticos. O modelo escolhido foi o *Hurdle* Binomial Negativo, que guarda relação com a teoria do agente principal aplicada à Economia da Saúde. Na especificação foram consideradas variáveis socioeconômicas como renda, gênero, idade, escolaridade, raça e região, e variáveis de saúde, como informação sobre morbidades, auto-avaliação da saúde, tipo de provimento (público ou privado) e cobertura de seguro-saúde. Além disso, estimou-se um preço sombra para o setor público de forma a incorporar a variável preço dos serviços de saúde na estimação. De forma geral, concluiu-se que existem iniquidades no acesso à saúde da população favorecendo os indivíduos de renda elevada.

***ABSTRACT***

The present work intends to identify and to measure the effect of some drivers for the healthcare demand. Using the database of the health supplement of the national household survey of the year 2003, several count data models were compared. According to some statistical criteria, the Hurdle Negative Binomial model was considered the best and was founded to be related to the agency theory applied to Health Economics. The specification was composed by social and economic variables such as income, gender, age, education, race and region and by health related variables, like morbidity, self assessed health, health system and insurance status. Moreover, it was estimated a shadow price for public sector in order to incorporate the price of health services in the healthcare frequency model. In general, it was founded iniquities in the access to the healthcare services.



## SUMÁRIO

<b>LISTA DE TABELAS.....</b>	<b>2</b>
<b>LISTA DE FIGURAS .....</b>	<b>3</b>
<b>INTRODUÇÃO .....</b>	<b>5</b>
<b>1 POLÍTICAS PÚBLICAS E O SISTEMA DE SAÚDE BRASILEIRO.....</b>	<b>9</b>
1.1 Sistema de saúde público e falhas de mercado .....	12
1.2 Breve discussão sobre equidade.....	15
1.3 Características gerais do sistema de saúde brasileiro.....	17
<b>2 ANÁLISE DESCRITIVA DOS DADOS DA PNAD.....</b>	<b>19</b>
2.1 Metodologia da PNAD.....	19
2.2 Perfil socioeconômico e de saúde da população brasileira segundo a PNAD .....	20
2.2.1 Características socioeconômicas da população brasileira.....	20
2.2.2 Características de saúde dos indivíduos segundo a PNAD (2003) .....	24
<b>3 MODELOS DE CONTAGEM.....</b>	<b>35</b>
3.1 A distribuição Poisson.....	36
3.1.1 Regressão Poisson – modelo de contagem padrão.....	37
3.2 Distribuição e modelo de regressão Binomial Negativo.....	37
3.3 Modelos <i>Hurdle</i> .....	39
3.4 Modelos <i>Finite Mixture</i> .....	46
3.4.1 Modelos <i>Zero-Inflated</i> .....	47
3.4.2 Modelos <i>Finite Mixture</i> genéricos .....	48
<b>4 MODELO ECONOMÉTRICO .....</b>	<b>53</b>
4.1 Estimação do preço dos serviços de saúde.....	53
4.1.1 Gastos com consultas médicas segundo a POF 2002/03 .....	60
4.2 Estimação do uso de serviços de saúde: consultas.....	62
4.2.1 Modelos Poisson e Binomial Negativo .....	64
4.2.2 Modelos <i>Hurdle</i> .....	71
4.2.3 Modelos <i>Finite Mixture</i> .....	73
4.2.3.1 Modelos <i>Zero-inflated</i> .....	73
4.2.3.2 Modelo <i>Finite Mixture</i> Binomial Negativo.....	81
4.3 Comparação entre modelos .....	85
4.4 Resultados .....	86
<b>5 CONSIDERAÇÕES FINAIS .....</b>	<b>89</b>
<b>APÊNDICE .....</b>	<b>94</b>
<b>REFERÊNCIAS .....</b>	<b>95</b>

## LISTA DE TABELAS

Tabela 1 – proporção de idosos na população por região e Brasil (%).....	5
Tabela 2 – dados sobre longevidade da população por país .....	6
Tabela 3 – proporção de gastos públicos em saúde (%) .....	12
Tabela 4 – proporção de gastos privados em saúde (%) .....	12
Tabela 5 - perfil de rendimentos de idosos e população total (%).....	21
Tabela 6 - anos de estudo da população (%).....	22
Tabela 7 – rendimento domiciliar e escolaridade da população .....	23
Tabela 8 - distribuição dos indivíduos por região geográfica (%).....	24
Tabela 9 - auto-avaliação da saúde: população e idosos (%).....	25
Tabela 10 - morbidades (% de respostas positivas) .....	26
Tabela 11 - pessoas de referência na família e titularidade dos planos (%).....	27
Tabela 12 - pessoa responsável pelo pagamento da mensalidade do plano (%).....	27
Tabela 13 - População coberta por plano de saúde (%) .....	28
Tabela 14 - faixas de mensalidade vs. faixas etárias (%).....	29
Tabela 15 - local de atendimento que costuma frequentar (%).....	29
Tabela 16 - média e desvio padrão do número de consultas por faixa etária.....	30
Tabela 17 - média de consultas no ano segundo gênero e cobertura por plano .....	32
Tabela 18 - motivo do principal atendimento realizado nas últimas duas semanas (%).....	32
Tabela 19 – tipo do principal atendimento (%).....	33
Tabela 20 - exemplo custo/benefício de um plano de saúde.....	45
Tabela 21 - resultados do procedimento de Heckman para estimação do preço.....	57
Tabela 22 - médias regionais - mensalidade paga a um plano de saúde privado.....	59
Tabela 23 - média regional do valor pago em uma consulta segundo a POF 2002/03 .....	62
Tabela 24 - descrição das variáveis explicativas utilizadas .....	63
Tabela 25 - resultados do modelo de regressão Poisson .....	67
Tabela 26 - resultados do modelo de regressão Binomial Negativo .....	68
Tabela 27 - probabilidades estimadas pelos modelos Poisson e Binomial Negativo .....	69
Tabela 28 - teste de sobredispersão Poisson e Binomial Negativo .....	71
Tabela 29 - critérios de informação Poisson e Binomial Negativo.....	71
Tabela 30 - resultados dos modelos <i>Hurdle</i> Poisson e Binomial Negativo .....	72
Tabela 31 - teste de sobredispersão <i>Hurdle</i> Poisson e <i>Hurdle</i> .....	73
Tabela 32 - critérios de informação modelos <i>Hurdle</i> .....	73
Tabela 33 - resultados da estimação do “Always-zero group” .....	75
Tabela 34 - resultados do modelo <i>Zero-inflated</i> Poisson.....	77
Tabela 35 - resultados do modelo <i>Zero-inflated</i> Binomial Negativo.....	78
Tabela 36 - probabilidades estimadas pelos modelos <i>Zero-inflated</i> .....	79
Tabela 37 - teste de sobredispersão modelos <i>Zero-inflated</i> .....	80
Tabela 38 - critérios de informação modelos <i>Zero-inflated</i> .....	80
Tabela 39 - resultados do modelo <i>Finite Mixture</i> : “low-users” .....	83
Tabela 40 - resultados do modelo <i>Finite Mixture</i> : “heavy-users” .....	84
Tabela 41 - resumo critérios de informação.....	85
Tabela 42 - valores previstos para o número de consultas por modelo .....	86
Tabela 43 - efeitos marginais dos estágios do <i>Hurdle</i> .....	94

**LISTA DE FIGURAS**

Figura 1 – gastos em saúde por país como proporção do PIB (%) .....	9
Figura 2 – gastos em saúde <i>per capita</i> por país em USS internacional .....	10
Figura 3 - histograma do número de consultas .....	30
Figura 4 - Proporção de pessoas que não realizaram consulta médica no último ano por faixa etária .....	31
Figura 5 – média de consultas realizadas no ano e faixas de rendimento domiciliar .....	31
Figura 6 - relação entre renda <i>per capita</i> e valor gasto com consultas médicas .....	61
Figura 7 - ajuste gráfico das probabilidades estimadas pelo modelo Poisson .....	70
Figura 8 - ajuste gráfico das probabilidades estimadas pelo modelo Binomial Negativo .....	70
Figura 9 - ajuste gráfico das probabilidades estimadas pelo modelo <i>Zero-Inflated</i> Poisson ...	79
Figura 10 - ajuste gráfico das probabilidades estimadas pelo modelo <i>Zero-inflated</i> Binomial Negativo .....	80



## INTRODUÇÃO

Uma questão importante para a economia e para a saúde pública é o envelhecimento populacional. Nos países desenvolvidos, este processo iniciou-se há algumas décadas, porém, o fato de que a faixa etária acima dos 60 anos<sup>1</sup> está crescendo mais rapidamente do que as demais faixas tem se tornado uma realidade para países em desenvolvimento, como o Brasil. Em 1991, os idosos representavam aproximadamente 7,3% da população brasileira. Em 2005, este número passou a ser de 9,2% da população, com destaque para a região sudeste, cujo percentual neste mesmo ano era de 10%. O estado do Rio de Janeiro é aquele que possui a maior proporção da população representada por pessoas idosas (11,5% em 2005).

**Tabela 1 – proporção de idosos na população por região e Brasil (%)**

	1991	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
Brasil	7.3	7.9	8.6	8.8	9.0	8.6	8.7	8.8	8.9	9.1	9.2
Região Norte	4.6	5.0	5.8	6.0	6.0	5.5	5.5	5.6	5.7	5.8	5.8
Região Nordeste	7.3	7.8	8.4	8.5	8.8	8.4	8.5	8.6	8.7	8.9	9.0
Região Sudeste	7.9	8.6	9.5	9.6	10.0	9.3	9.4	9.5	9.6	9.9	10.0
Região Sul	7.7	8.4	8.6	9.0	8.9	9.2	9.3	9.4	9.6	9.8	9.9
Região Centro-Oeste	5.2	5.9	6.1	6.5	7.0	6.6	6.7	6.8	6.9	7.2	7.3

Fonte: DATASUS

Segundo Camarano (2002), as conseqüências do envelhecimento têm sido vistas, em geral, com preocupações por acarretarem pressões para transferência de recursos na sociedade, colocando desafios para o Estado, os setores produtivos e as famílias, especialmente no que diz respeito aos sistemas de previdência e assistência social.

De acordo com Barros *et al* (1999), a dificuldade de inserção dos idosos no mercado de trabalho faz com que estes sejam potencialmente vulneráveis ao estado de pobreza. A partir de certo momento, os indivíduos passam a depender cada vez mais de fontes de renda diferentes do salário tais como aposentadorias e rendimentos dos demais moradores do domicílio. Por outro lado, a estrutura dos gastos tende a ser mais volátil que a maioria da

<sup>1</sup> Embora sob o ponto de vista demográfico sejam classificadas como idosos as pessoas com 65 anos ou mais, a Organização Mundial da Saúde (OMS) propõe que para países em desenvolvimento o limite etário deve ser 60 anos.

população, havendo maior probabilidade de gastos elevados e inesperados, decorrentes da maior debilidade física e mental.

Além dos idosos, as crianças são fortes demandantes de consultas médicas. No Brasil, a taxa de mortalidade infantil ainda é alta, mesmo se comparada com países latino-americanos, conforme a tabela seguinte, a qual mostra a expectativa de vida e a mortalidade adulta para ambos os sexos, além da taxa de mortalidade infantil.

**Tabela 2 – dados sobre longevidade da população por país**

<i>País</i>	<b>Expectativa de vida</b>		<b>Mortalidade adulta</b>		<b>Mortalidade</b>
	<i>Homens</i>	<i>Mulheres</i>	<i>Homens</i>	<i>Mulheres</i>	<b>infantil</b>
<b>Canadá</b>	78	83	90	56	5
<b>França</b>	77	84	128	58	4
<b>Reino Unido</b>	77	81	101	62	5
<b>Alemanha</b>	76	82	110	57	4
<b>EUA</b>	75	80	137	81	7
<b>Chile</b>	74	81	128	64	8
<b>Argentina</b>	72	78	162	86	14
<b>México</b>	72	77	162	94	22
<b>China</b>	71	74	155	98	23
<b>Brasil</b>	68	75	225	118	28

Fonte: WHO, 2007 (dados de 2005)

Pela tabela acima, pode-se perceber que o Brasil possui um potencial grande de ocorrência de mudanças em sua estrutura etária: se a taxa de fecundidade se mantiver inalterada ou diminuir mais lentamente do que o avanço no combate à mortalidade infantil, será grande o contingente de crianças e idosos, que são fortes demandantes de serviços de saúde. Além disso, espera-se que as condições de vida como abastecimento de água e esgoto, disponibilidade de vacinas e moradia sejam melhoradas, elevando a expectativa de vida do brasileiro.

Tendo em vista que o sistema público de saúde brasileiro é universal, é preciso planejamento de políticas de saúde frente às expectativas de queda da mortalidade e aumento da expectativa de vida no Brasil. Em alguns países, como os EUA, o sistema de saúde é parcialmente universal, existindo programas específicos para as populações desfavorecidas: o *Medicare* para idosos e portadores de deficiência, e o *Medicaid* para a população de baixa renda. O restante da população ou é coberta por seguros de saúde privados ou desembolsa diretamente

no pagamento de consultas e/ou internações. Desde a entrada em operação destes programas públicos de saúde, o orçamento americano destinado à saúde aumentou muito, o que resultou na implementação de mecanismos de gerenciamento dos gastos em saúde, o *Managed Care*<sup>2</sup>.

Nesse sentido, justifica-se uma análise aprofundada da estrutura do sistema de saúde brasileiro e das características da população. O presente trabalho tem por objetivo identificar e mensurar o efeito de características socioeconômicas e de saúde da população brasileira que atuam como determinantes na utilização de serviços de saúde públicos e privados, assim como comparar diferentes metodologias utilizadas para estimar a demanda por serviços de saúde.

Por meio da metodologia dos modelos de contagem aplicados aos dados da Pesquisa Nacional por Amostra de Domicílio (PNAD) de 2003, estimou-se o número de consultas médicas realizadas em um ano, por indivíduo, de acordo com as características da população e do sistema de saúde brasileiros. A escolha do indivíduo, ao invés da família, se justifica pelo fato de que a decisão de ir ao médico está intimamente relacionada ao estado de saúde do indivíduo, diferentemente dos gastos em saúde, que estão mais ligados ao orçamento familiar disponível.

A especificação dos modelos, em termos de variáveis explicativas, bem como a interpretação dos resultados teve como base teorias conhecidas na literatura de Economia da Saúde, tais como a teoria do capital humano aplicada à saúde (GROSSMAN, 1972) e as teorias de informação assimétrica aplicadas ao mercado de saúde suplementar (ZWEIFEL, 1981).

Existem muito estudos na literatura internacional sobre a aplicação de modelos de contagem na estimação da demanda por consultas médicas<sup>3</sup>. A maioria focaliza a questão econométrica, procurando identificar as vantagens e desvantagens de cada modelo. Na literatura nacional, contudo, há poucos trabalhos acerca do tema. Dentre o mais recentes, se destacam o trabalho de Andrade e Maia (2005), sobre o problema da seleção adversa nos planos de saúde, e o de

---

<sup>2</sup> Segundo Andrade e Lisboa (2000), o *Managed Care* é um sistema estruturado a partir de contratos estabelecidos entre as organizações que gerenciam a utilização e financiamento dos planos de saúde e os provedores de serviços de saúde com o objetivo de solucionar o problema do risco moral existente neste mercado que acaba por determinar elevação dos gastos.

<sup>3</sup> Ver, por exemplo, trabalhos de Pohlmeier e Ulrich (1995), Cameron *et al* (1988), Cameron e Johansson (1997) e Deb e Trivedi (1997).

Noronha e Andrade (2005), que utilizou um modelo *Hurdle* Binomial Negativo para estimar a demanda por serviços de saúde pelos idosos.

A análise realizada neste trabalho apresenta avanços em relação aos trabalhos relacionados, na medida em que introduz a variável preço dos serviços de saúde como um determinante importante para a decisão e a frequência de idas ao médico pelos indivíduos. Enquanto em outros trabalhos, freqüentemente, se utiliza a variável indicadora de presença ou ausência de plano de saúde privado como determinante da demanda por serviços de saúde, a variável preço, utilizada neste estudo, procura incorporar, além dessa característica, um preço sombra da utilização dos serviços de saúde pública, através de um procedimento que será detalhado no capítulo quatro.

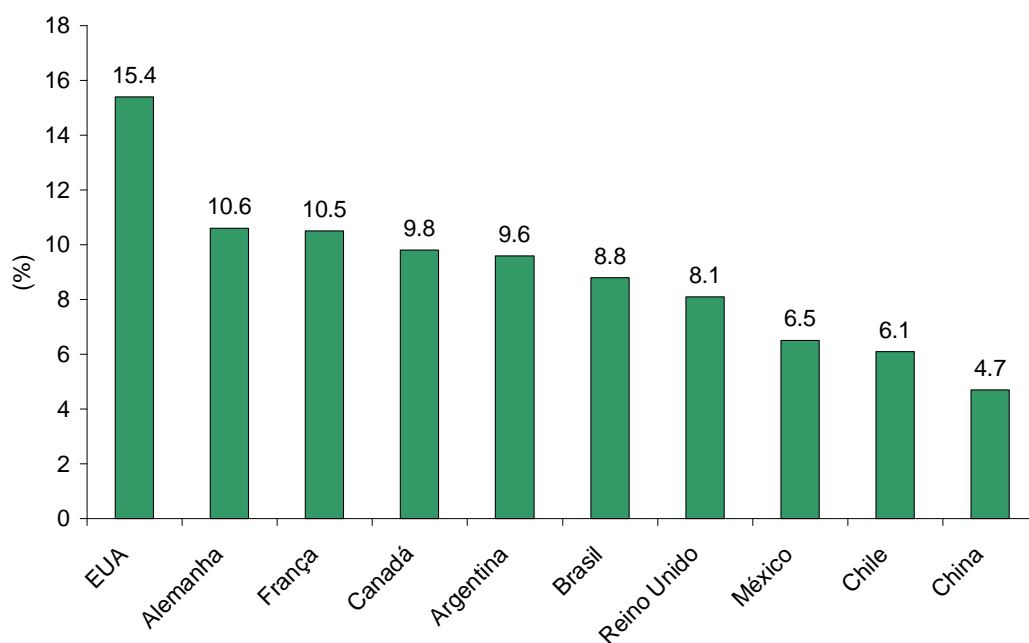
O trabalho será dividido em alguns pontos. Após essa introdução, o primeiro capítulo trata das características dos sistemas de saúde público e privado no Brasil e de argumentos teóricos para a intervenção do Estado no âmbito da saúde, introduzindo o debate entre universalização e focalização das políticas de saúde. O segundo capítulo descreve os dados utilizados na análise – a Pesquisa Nacional por Amostra de Domicílios (PNAD) de 2003. O terceiro capítulo apresenta a metodologia e traz uma revisão bibliográfica do uso de modelos de contagem na análise de determinantes da demanda por consultas médicas. Finalmente, o quarto capítulo apresenta os resultados, compara os diferentes modelos de contagem e traz conclusões da estimação do modelo econométrico. O último capítulo traz as conclusões gerais obtidas.



## 1 POLÍTICAS PÚBLICAS E O SISTEMA DE SAÚDE BRASILEIRO

Primeiramente, deve ser feita uma breve análise comparativa do Brasil com relação aos países desenvolvidos e aos países em desenvolvimento. Atualmente, o país que possui o maior percentual de gasto em saúde com relação ao produto interno bruto (PIB) é os EUA (15,4%). Neste quesito, o Brasil, situa-se atrás da Alemanha, França e Canadá e Argentina, com uma proporção de 8,8%, conforme o gráfico seguinte, em que é possível perceber que países como China e Reino Unido não comprometem uma alta proporção da renda com gastos em saúde (WHO, 2007):

**Figura 1 – gastos em saúde por país como proporção do PIB (%)**

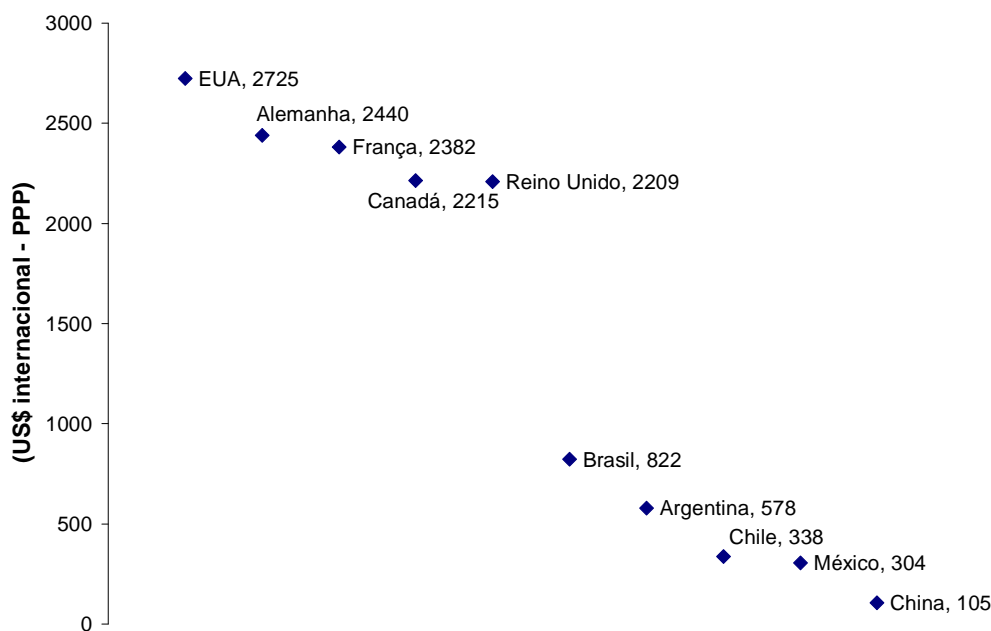


Fonte: World Health Organization (WHO), 2007 (dados de 2004)

Contudo, levando em consideração o PIB, a taxa de inflação, a taxa de câmbio internacional (segundo a metodologia do PPP- *Purchase Power Parity*) e a população de cada país, verifica-se que embora o Reino Unido não comprometa altas proporções do PIB, o gasto *per capita* em saúde é bastante elevado (2209 dólares internacionais), assim como na Alemanha, EUA, França e Canadá. O Brasil lidera este indicador dentro dos países selecionados da América Latina com um gasto *per capita* de 822 dólares internacionais, quase oito vezes

acima do gasto realizado pela China. O gráfico a seguir mostra a dispersão deste indicador para países selecionados.

**Figura 2 – gastos em saúde *per capita* por país em US\$ internacional**



Fonte: WHO, 2007 (dados de 2004)

Com relação ao tipo de financiamento dos gastos em saúde, o país que possui a maior proporção de gastos públicos é o Reino Unido (86,3%), uma vez que seu sistema de saúde é eminentemente centrado no financiamento público, sendo conhecido como *National Health System* (NHS).

Os EUA têm a maior proporção de gasto público em saúde com relação ao gasto público geral (18,9%), embora seu sistema de saúde seja predominantemente financiado pelo setor privado. A maior parte da população dos EUA é coberta por seguros-saúde privados: somente 23,8% dos gastos privados são gastos “*out-of-pocket*”, ou seja, gastos não cobertos pelo seguro como consultas a médicos não conveniados, além de *deductibles* e copagamentos<sup>4</sup>.

---

<sup>4</sup> *Deductibles* são uma espécie de quantia fixa cobrada pelas seguradoras de saúde para que seus clientes possam usufruir dos benefícios do seguro; similar a um “sinistro” no caso do seguro de automóvel.

Curiosamente, a China é o país com maior proporção de gastos em saúde privados, sendo ao mesmo tempo o que menos investe em saúde (em termos de gastos como percentual do PIB e de gastos *per capita*). A participação de gastos cobertos por seguros privados também é pequena na China: 86,5% dos gastos privados são do tipo “*out-of-pocket*”.

O Brasil possui maior participação do setor público nos gastos totais em saúde (54,1%), que representam 14,2% dos gastos públicos em geral. O setor privado brasileiro corresponde a uma parcela de 45,9% dos gastos totais em saúde, sendo 64,2% gastos não cobertos pelos planos de saúde.

O Chile, apesar de possuir maior participação do setor privado e maior cobertura de planos de saúde, é o país mais semelhante ao Brasil dentro dos países selecionados da América Latina. Assim, como no Brasil, o sistema de saúde do Chile é misto: enquanto o Fundo Nacional de Saúde (FONASA) gere os aportes dos assalariados que optaram pelo seguro público via contribuições mandatórias de 7% dos salários, as ISAPREs (companhias de seguro de saúde privado do Chile) se encarregam dos que optaram pela modalidade de seguro privado. A diferença está no fato de que, no Brasil, o Sistema Único de Saúde (SUS) é universal e toda a população contribui via impostos e contribuições.

As tabelas seguintes mostram as respectivas participações do setor público e do setor privado nos gastos em saúde para alguns países<sup>5</sup>:

---

<sup>5</sup> Seleção própria a partir dos dados do relatório “The World Health Statistics Report”, divulgado em 2007 pela WHO. Os dados referem-se ao último ano disponível no banco de dados da instituição.

**Tabela 3 – proporção de gastos públicos em saúde (%)**

<i>País</i>	<i>Gastos públicos em saúde (% do gasto total)</i>	<i>Gastos públicos em saúde (% gasto público geral)</i>
<b>Reino Unido</b>	86.3	15.9
<b>França</b>	78.4	15.4
<b>Alemanha</b>	76.9	17.3
<b>Canadá</b>	69.8	17.1
<b>Brasil</b>	54.1	14.2
<b>Chile</b>	47.0	13.1
<b>México</b>	46.4	12.9
<b>Argentina</b>	45.3	15.1
<b>EUA</b>	44.7	18.9
<b>China</b>	38.0	10.1

Fonte: WHO, 2007 (dados de 2004)

**Tabela 4 – proporção de gastos privados em saúde (%)**

<i>País</i>	<i>Gastos privados em saúde (% do gasto total)</i>	<i>Gastos "out-of-pocket" (% gasto privado geral)</i>
<b>China</b>	62.0	86.5
<b>EUA</b>	55.3	23.8
<b>Argentina</b>	54.7	48.7
<b>México</b>	53.6	94.4
<b>Chile</b>	53.0	45.9
<b>Brasil</b>	45.9	64.2
<b>Canadá</b>	30.2	49.4
<b>Alemanha</b>	23.1	57.5
<b>França</b>	21.6	34.9
<b>Reino Unido</b>	13.7	91.8

Fonte: WHO, 2007 (dados de 2004)

### 1.1 Sistema de saúde público e falhas de mercado

A principal justificativa econômica para a existência de sistemas públicos de saúde é o que a literatura do setor público denomina falhas de mercado. Falhas de mercado são situações em que os mercados possuem características distintas de um mercado competitivo e, por essa razão, não funcionam de maneira eficiente (no sentido de Pareto), resultando em equilíbrios

subótimos, que não garantem a maximização do bem-estar social. Assim, argumenta-se que o governo deve intervir de modo a tentar corrigir as falhas existentes e melhorar a situação dos agentes econômicos. No setor da saúde essas falhas são freqüentes, uma vez que o “bem” saúde tem características de bens públicos, além de gerar externalidades e envolver informação assimétrica (MUSGROVE, 1996)

Bens públicos são bens não rivais e não excludentes, ou seja, o consumo do bem por parte de um agente não exclui outro agente de consumi-lo, e a quantidade disponível para consumo não é reduzida pelo consumo do outro agente. A valoração dos bens públicos pode diferir de acordo com o indivíduo, porém, como o consumo é não excludente existe o “problema do carona”, em que um indivíduo pode não contribuir para a provisão do bem público e mesmo assim usufruí-lo, já que os demais indivíduos contribuíram. É por este motivo que a provisão de um bem público tem que ser muitas vezes realizada pelo governo. O controle de vetores de doenças e a proteção sanitária de alimentos e água são exemplos que se aproximam do conceito de bem público. A ação individual pode ser ineficaz ou muito custosa, de forma que o Estado tem que intervir.

Além de bens públicos, a saúde envolve externalidades. Estas ocorrem quando a ação de um agente afeta diretamente o conjunto de ações possíveis dos demais agentes. Por exemplo, a imunização de uma região contra determinada doença gera externalidades positivas para as demais regiões, uma vez que se previne a transmissão da doença. As externalidades também acarretam ineficiência alocativa, uma vez que os agentes não internalizam a totalidade dos custos e benefícios de suas ações e, portanto, não têm incentivos a produzir quantidades socialmente ótimas. Por essa razão, argumenta-se, por exemplo, que os programas de vacinação devem ser universais.

A terceira falha, a assimetria de informação, relaciona-se principalmente com o sistema privado de saúde, particularmente ao seguro-saúde. A informação assimétrica afeta os contratos, na medida em que o segurador não sabe ao certo o “tipo” de consumidor que busca seu plano. O segurador pode ter informações epidemiológicas da população, sabendo em média a probabilidade de incidência de morbidades. No entanto, a probabilidade individual é desconhecida, de forma que o contrato deve ser único para todos os “tipos” (no caso do seguro-saúde o contrato varia segundo faixas etárias, porém é o mesmo dentro de cada faixa)

ao invés de um contrato para cada indivíduo (o que seria possível sem informação assimétrica).

As falhas de informação dão origem ao problema pré-contratual conhecido como seleção adversa. Este problema caracteriza-se pelo fato de que as pessoas com maior probabilidade de desenvolverem doenças ou que já possuem debilidades são as que mais demandam o seguro. Mas o fundo de um plano de saúde é a princípio solidário: os saudáveis (que não usam frequentemente os serviços oferecidos) financiam os doentes. Logo, a seleção adversa pode gerar dificuldades financeiras na gestão dos fundos, que devem ser compostos por um *mix* de segurados que na média seja sustentável. É por este motivo que as seguradoras fazem análises de risco e mais ainda, preferem os planos coletivos aos individuais. Nesse sentido, o SUS, visto como seguro social, não é afetado por seleção adversa, uma vez que o acesso é universal e não há discriminação de preço: todos pagam proporcionalmente sob a forma de impostos e contribuições.

Outro aspecto da falha de informação assimétrica é a tendência à sobre utilização dos planos de saúde, problema pós-contratual conhecido como risco moral. Na relação entre paciente e provedor (médico) este problema aparece porque, para o paciente, o custo marginal de mais uma consulta ou exame é zero se este possui o seguro de saúde pleno. Na relação entre provedor e segurador o risco moral acontece porque os médicos possuem mais informação sobre o estado de saúde do paciente, e por isso têm incentivos a uma maior requisição de exames e procedimentos. Este problema, também conhecido por demanda induzida pelos médicos (PAULY, 1968), é bastante relevante na estimação da frequência com que se utilizam serviços de saúde. Os provedores (médicos) podem sobre utilizar o volume de serviços tanto para elevar seus rendimentos, se o sistema de reembolso for por tarefa, quanto para garantir a certeza do diagnóstico (ANDRADE; LISBOA, 2001).

Além das falhas já apontadas, Jack (2000) chama atenção para a competição imperfeita derivada do poder de mercado possuído pelas firmas seguradoras. Tal poder se origina da informação imperfeita pelo lado da demanda (o que aumenta os custos de uma firma nova atrair clientes) e dos crescentes retornos sobre custos administrativos. No entanto, apesar de grande parte da literatura de organização industrial considerar o poder de mercado nocivo ao bem estar social, no mercado de saúde suplementar, a competição perfeita pode ter algumas vezes efeitos negativos sobre esse bem estar. Por exemplo, se as firmas não podem oferecer

diferentes planos para clientes com diferentes graus de risco, haveria um incentivo para baixar a qualidade dos planos de maneira que apenas fossem atraídas pessoas de baixo custo (baixa probabilidade de desenvolver doenças). Daí a necessidade de regulação do mercado de saúde suplementar.

## 1.2 Breve discussão sobre equidade

A equidade em saúde pode ser entendida em três dimensões: condições de saúde, acesso aos serviços de saúde e financiamento. A pergunta que se faz é: para quem e como se devem alocar os recursos escassos em saúde? A discussão sobre justiça alocativa será baseada em dois princípios dos paradigmas de justiça social: o princípio da igualdade e o princípio da equidade.

Pelo princípio da igualdade todos deveriam receber igual tratamento. A idéia que está implícita neste princípio é de que todos os indivíduos são iguais, têm o mesmo direito e, portanto, deveriam receber a mesma fração dos recursos. É uma idéia baseada na moralidade de direitos, diferente da idéia dos direitos de propriedade ou do trabalho, em que cada indivíduo receberia proporcionalmente à sua contribuição para a sociedade (MEDEIROS, 1999).

Já o princípio da equidade reconhece a não igualdade entre os indivíduos. Sendo assim, a orientação é de eliminação ou redução dessas desigualdades através de políticas públicas: “O tratamento desigual é justo quando é benéfico ao indivíduo mais carente” (RAWLS *apud* MEDEIROS, 1999). Este critério para justiça na desigualdade também é conhecido como axioma do *Maxmin*. Este axioma se opõe à idéia utilitarista em que o que importa é a maximização da soma do bem-estar social de todos os indivíduos, sem se preocupar com a hierarquia de condições entre os indivíduos.

Ambos os princípios se remetem às divergências concernentes às políticas sociais: universalização versus focalização. Nas políticas universais todos os indivíduos têm acesso a determinados serviços públicos, uma vez que isso é garantido por direitos de cidadania. No

caso do Brasil, a igualdade em saúde é, a princípio, garantida pela legislação. Segundo a Constituição Federal:

“Art. 196. A saúde é direito de todos e dever do Estado, garantido mediante políticas sociais e econômicas que visem à redução do risco de doença e de outros agravos e ao acesso universal e igualitário às ações e serviços para sua promoção, proteção e recuperação.”

Em contraste, as políticas de focalização visam concentrar os recursos disponíveis em populações carentes e/ou composta de minorias, baseando, portanto, no princípio da equidade. No âmbito da saúde, a Política Nacional de Saúde da Pessoa Idosa é um exemplo de focalização. Outros exemplos são: o *Medicare* (programa destinado a pessoas idosas) e o *Medicaid* (programa destinado a pessoas de baixa renda) norte-americanos.

O maior obstáculo à universalização é a escassez de recursos. Uma vez que o público a ser atendido é bastante extenso haverá uma sobrecarga no orçamento público. O envelhecimento da população pode gerar sobrecargas no sistema de saúde e na previdência social, já que se pretende o atendimento universal. Quando não é possível aumentar o orçamento, a alternativa é reduzir os custos dos serviços prestados, o que por sua vez, pode gerar qualidade insuficiente. Este é um problema que acomete o SUS, cujas enormes filas e má qualidade do atendimento são de conhecimento comum.

Já os programas sociais focalizados seriam, em princípio, de maior qualidade. No entanto, a literatura faz a ressalva de que são programas de maior custo de implementação, além do fato de que estes programas, geralmente, prejudicam as classes médias. Como este grupo de interesse tem bastante representatividade política, muitas vezes a focalização não é implementada (MEDEIROS, 1999).

O tema da equidade em saúde, entendido em sua versão horizontal - igualdade no acesso a indivíduos com mesma morbidade/ necessidade - tem sido muito discutido pela literatura nacional e internacional de economia da saúde, em diferentes abordagens. Mas quais necessidades seriam objetos de políticas públicas? Por exemplo, qual indivíduo deveria ser contemplado pelo governo: um indivíduo portador do vírus HIV em fase inicial ou um em fase terminal? Nesse sentido, o governo deve estabelecer quais as necessidades são



contempláveis ou elegíveis. No caso do SUS, a diretriz de atendimento integral garantiria, em princípio, que todas as necessidades são elegíveis.

### **1.3 Características gerais do sistema de saúde brasileiro**

No Brasil, o setor público e o setor privado coexistem no provimento e financiamento dos serviços de saúde. O sistema nacional de saúde brasileiro foi consolidado pela Constituição de 1988. Nesta época, foram definidas as diretrizes básicas do SUS – Sistema Único de Saúde – tais como a descentralização (a execução dos serviços é comandada pelos estados e municípios, cabendo à instância federal o planejamento, fiscalização e controle), o atendimento integral e a participação social. Os recursos do SUS fazem parte do Orçamento da Seguridade Social, cujas principais fontes de financiamento são, originalmente, as contribuições sociais (CONFINS, PIS, PASEP e CPMF).

Com a extinção da CPMF em fins de 2007, houve muito questionamento sobre qual seria a nova fonte para financiamento da saúde. Em junho de 2008, foi aprovada a Contribuição Social para a Saúde (CSS), que passa a incidir sobre toda movimentação financeira a partir de janeiro de 2009, com uma alíquota de 0,1% (não incide sobre aposentados, pensionistas e trabalhadores da ativa que ganham até o teto dos benefícios da Previdência). Como em outros países, o sistema público de saúde do Brasil caracteriza-se pela qualidade questionável e demora no atendimento dos seus usuários, efeitos de um sistema congestionado e ineficiente.

Em contraste, o sistema privado de saúde brasileiro é composto por diversas formas institucionais como a medicina de grupo, as cooperativas médicas, os planos geridos por empresas públicas e/ou privadas (autogestão) e o seguro-saúde. A expansão deste sistema se deu de forma diferenciada entre as modalidades, com maior crescimento recente dos seguros-saúde, incentivados pela legislação. A segmentação do mercado entre usuários do seguro social (SUS) e usuários de planos de saúde foi estimulada pela lei que permite a dedução integral dos valores destinados ao pagamento de planos de assistência médica suplementar para efeito de cálculo do imposto de renda das empresas e pessoas físicas (PEREIRA, 2002). Atualmente, estima-se que cerca de 24,6% da população (PNAD, 2003) seja coberta por

algum tipo de assistência médica suplementar. Ao contrário do SUS, nem todas as pessoas têm acesso ao sistema privado. Este é, em sua maioria, utilizado pela população de nível sócio econômico mais elevado, sendo em geral mais eficiente e de melhor atendimento.

## 2 ANÁLISE DESCRITIVA DOS DADOS DA PNAD

### 2.1 Metodologia da PNAD

A Pesquisa Nacional por Amostras de Domicílios (PNAD) teve início no segundo semestre de 1967. Trata-se de um sistema de pesquisas que tem como foco o estudo da unidade domiciliar (domicílio particular ou unidades de habitação em domicílio coletivo) e da população residente nestes domicílios. Tem como objetivo mapear características socioeconômicas, algumas de caráter permanente (características gerais da população, educação, trabalho, rendimento e habitação) e outras com periodicidade variável (migração, nutrição, saúde, fecundidade e outros temas), dependendo da relevância à época.

A partir de 1970, a pesquisa passa a ser realizada anualmente, com data de realização no segundo semestre, durante uma semana (semana de referência). O ano de referência da pesquisa é contabilizado da data de referência (último dia da semana de referência) até 365 dias anteriores à referida data.

Em relação às informações sobre a saúde da população, a PNAD não possui nenhuma pergunta em seu questionário permanente, sendo essas informações contempladas somente em suplementos especiais. Nas décadas de 80 e 90, o tema saúde foi incluído em quatro suplementos da PNAD: 1981, 1986, 1988 e 1998. A última vez em que o suplemento saúde foi incluído foi no ano de 2003. Entretanto, os suplementos de 1981 a 1988 não abrangem aspectos individuais de saúde da população, restringindo-se ao tema de acesso aos serviços de saúde (ANDRADE, 2002). Sendo assim, o suplemento de saúde de 1998 é a primeira fonte de informações sobre a saúde do indivíduo mais abrangente realizada no Brasil.

Recentemente, o suplemento da PNAD (2003) está organizado em três partes principais. A primeira parte contém informações sobre as morbidades dos indivíduos, incluindo os três indicadores mais utilizados para avaliar o estado de saúde dos indivíduos: uma medida de percepção do estado de saúde (auto-avaliação), uma medida baseada na presença de dificuldade para realização de tarefas habituais e uma medida de presença de doença crônica.

A segunda parte diz respeito à cobertura privada de serviços de saúde, abrangendo perguntas sobre a caracterização de planos de saúde. Finalmente, a terceira parte diz respeito ao acesso aos serviços de saúde (divididos em serviços ambulatoriais e de internação), sendo possível identificar o tipo de serviço, o local e o motivo de atendimento, bem como as dificuldades encontradas no acesso<sup>6</sup>.

## **2.2 Perfil socioeconômico e de saúde da população brasileira segundo a PNAD**

Primeiramente, serão explicitadas as informações socioeconômicas dos indivíduos que serão utilizadas como variáveis explicativas no modelo econométrico de demanda por serviços de saúde. Em seguida, será traçado um perfil específico das condições de saúde dos indivíduos, com base nas questões presentes no suplemento de saúde da PNAD (2003), demonstrando a importância e as limitações de cada uma para a análise empírica que será feita no capítulo posterior.

### **2.2.1 Características socioeconômicas da população brasileira**

Segundo a PNAD (2003), residiam no Brasil aproximadamente 176 milhões de habitantes, sendo que 85% residiam em áreas urbanas. O grupo dos idosos representava aproximadamente 9,1% da população brasileira no ano de referência da pesquisa. Já as crianças com idade de zero a nove anos representavam 18,2%.

Do total da população brasileira, aproximadamente 51,3% eram do sexo feminino. Na faixa etária dos idosos, o percentual de mulheres era de 56%, indicando que as mulheres têm uma expectativa de vida superior aos homens.

---

<sup>6</sup> No suplemento de saúde da PNAD (1998) havia mais um bloco de perguntas referente aos gastos privados com saúde, que foi extinto. No entanto, este bloco não será necessário para a análise. O suplemento de 2003 inclui ainda um novo bloco de investigação relativo ao acesso a serviços preventivos de saúde para mulheres de 25 anos ou mais de idade.

Em relação ao perfil de rendimento, nota-se que, ao analisarmos a variável rendimento da atividade principal<sup>7</sup>, 21,7% ganhavam menos do que meio salário mínimo (cujo valor na época era de R\$ 240). Os idosos, por sua vez, tinham um rendimento menor, pois 41,8% se encontravam na menor faixa de rendimento da atividade principal. No caso desta faixa etária, parece mais realista pensar em rendimento de todas as fontes ou no rendimento domiciliar, uma vez que o percentual de idosos que não estavam trabalhando na semana de referência (74,4%) é superior ao mesmo percentual para a população total (47,4%), e o rendimento de todas as fontes inclui aposentadoria. As crianças não possuem renda do trabalho e o restante da população ganhava em média R\$ 575 provenientes do trabalho principal, com maior concentração na faixa de um a dois salários mínimos.

A tabela cinco, abaixo, compara as distribuições de renda para os idosos e para a população total considerando o rendimento total de todas as fontes e o rendimento da atividade principal. Note que ao considerarmos o rendimento de todas as fontes, é pequena a participação dos idosos que recebem menos do que meio salário mínimo e é bastante expressiva a participação da faixa de meio a um salário.

**Tabela 5 - perfil de rendimentos de idosos e população total (%)**

	Rendimento da atividade principal		Rendimento de todas as fontes	
	<i>Idosos</i>	<i>Total</i>	<i>Idosos</i>	<i>Total</i>
Menos que 1/2 salário	41.8	21.7	12.6	44.1
De 1/2 a 1 salário	17.1	19.4	39.5	15.8
De 1 a 2 salários	16.4	26.4	18.9	17.1
De 2 a 3 salários	8.0	12.4	9.0	8.2
De 3 a 5 salários	6.0	9.6	8.2	6.7
De 5 a 10 salários	4.2	5.6	6.0	4.3
De 10 a 20 salários	2.4	2.6	2.7	2.0
Mais de 20 salários	4.2	2.3	3.1	1.8

Fonte: micro dados PNAD (2003)

Sobre a condição na família, 30,4% das pessoas eram chefes ou pessoa de referência na família. Espera-se que a pessoa de referência seja aquela que mais contribua na renda familiar e que, portanto, seja o titular dos planos de saúde. Mais ainda, do total de homens na amostra, 43% são pessoas de referência na família; do total de mulheres, somente 17,8% têm essa

<sup>7</sup> Somente para pessoas a partir de 10 anos de idade.

condição. Já no total de homens idosos, 89,7% são chefes de família; para as mulheres idosas esse percentual cai para 46,2%.

Sobre o nível de escolaridade da população, a faixa mais representativa é a de 4 a 7 anos de estudo, conforme a tabela seis a seguir, que mostra a distribuição da escolaridade – medida em anos de estudo – da população em geral.

**Tabela 6 - anos de estudo da população (%)**

	<b>População</b>
Sem instrução e menos de 1 ano	24.1
1 a 3 anos	15.5
4 a 7 anos	26.3
8 a 10 anos	13.5
11 a 14 anos	16.2
15 anos ou mais	4.2
Não determinado e sem declaração	0.3

Fonte: micro dados PNAD(2003)

O nível de escolaridade é uma variável importante na determinação da utilização de serviços de saúde por dois motivos. Em primeiro lugar, pessoas mais escolarizadas são mais informadas sobre doenças e possíveis tratamentos, elas tendem a adotar hábitos de vida mais saudáveis e a procurar mais os serviços médicos, especialmente os cuidados preventivos (GROSSMAN, 1972). O segundo motivo refere-se à disponibilidade de recursos para investir na produção de saúde: quanto maior o nível de escolaridade, maior tende a ser o nível de renda e, conseqüentemente, melhores as condições de vida e de acesso aos serviços de saúde.

De fato, os dados da PNAD (2003) corroboram esta hipótese, uma vez que existe uma correlação positiva entre a média de anos de estudo e a classe de renda domiciliar. Faz sentido pensar em renda do domicílio não apenas no investimento em educação, mas também em saúde, tendo em vista que muitos indivíduos não possuem renda própria, mas atuam como dependentes de outros membros do domicílio. A tabela sete, abaixo, mostra a participação das classes de rendimento domiciliar na população e a média de escolaridade para cada classe.

**Tabela 7 – rendimento domiciliar e escolaridade da população**

	<b>Participação da classe</b>	<b>Média de anos de estudo</b>
Menos que 1/2 salário	4.2%	3.96
De 1/2 a 1 salário	8.7%	3.94
De 1 a 2 salários	20.2%	4.58
De 2 a 3 salários	16.7%	5.43
De 3 a 5 salários	20.0%	6.43
De 5 a 10 salários	16.5%	7.94
De 10 a 20 salários	7.8%	9.81
Mais de 20 salários	6.0%	10.16

Fonte: micro dados PNAD (2003)

Outra característica socioeconômica relevante para a análise da demanda de serviços de saúde é a cor ou raça. Sabe-se que a raça pode ser influente no nível de renda do indivíduo, o que por sua vez, afeta a disponibilidade de recursos para investimento em saúde. Segundo estudo de Trujillo *et al* (2005), os idosos de raça negra reportaram<sup>8</sup> ter menos saúde do que os idosos brancos, mesmo com controles demográficos, socioeconômicos e de suporte familiar. Considerando o total da amostra da PNAD, correspondem à cor branca 47,7%, e as cores parda e preta 45,4% e 6,3%, respectivamente. Índios e indivíduos de cor amarela somam 0,6%.

Por último, é preciso notar as distribuições geográficas da população. As regiões sudeste e nordeste abrangem grande parte da população em geral e também dos idosos e crianças, mas é possível notar que há uma maior concentração dos idosos na região sudeste. A tabela oito, a seguir, mostra a dispersão geográfica da população, dos idosos e crianças (até nove anos).

---

<sup>8</sup> A variável cujo indivíduo reporta sua saúde na PNAD será exposta mais à frente, correspondendo a uma auto-avaliação em uma escala composta de cinco níveis de saúde.

**Tabela 8 - distribuição dos indivíduos por região geográfica (%)**

<b>Região</b>	<b>Total (%)</b>	<b>Idosos (%)</b>	<b>Crianças (%)</b>
Norte	6.0	3.8	7.6
Nordeste	28.4	27.1	32.2
Sudeste	43.5	47.4	39.0
Sul	15.0	16.2	13.8
Centro-Oeste	7.1	5.5	7.5

Fonte: micro dados expandidos PNAD (2003)

### 2.2.2 Características de saúde dos indivíduos segundo a PNAD (2003)

A auto-avaliação da condição de saúde é uma medida adotada pela PNAD em que o indivíduo reporta seu estado de saúde classificando-o dentre cinco níveis disponíveis: muito bom, bom, regular, ruim e muito ruim. Embora exista uma série de críticas a respeito desta variável, devido ao seu caráter subjetivo e também ao fato de ser influenciada por outras variáveis como grau de instrução e renda, esta medida ainda é a única pela qual se pode inferir o estado geral de saúde do indivíduo na PNAD.

Entre os estudos que questionam esta variável está o de Windmeijer e Santos Silva (1997). Os autores afirmam que a variável em questão é bastante subjetiva e não é necessariamente informativa do estado de saúde de um indivíduo, além de sofrer do problema de endogeneidade. Além desses autores, Rice *et al* (2005) apontam para o fato de que a variável auto-avaliação de saúde é viesada, uma vez que diferentes grupos socioeconômicos interpretam a pergunta e a respondem de maneira distinta. Em particular, o estudo procura avaliar uma mudança em determinada pergunta do questionário do *British Household Panel Survey* (BHPS) e seu efeito nas respostas para as diferentes classes de indivíduos.

Na população total, o percentual dos que consideram sua saúde como boa ou muito boa é 78,1%; já na população idosa, este percentual cai para 42,9%. É possível perceber ainda que, dentre os idosos, à medida que a idade avança, a avaliação da saúde piora: 17,3% dos idosos com 70 anos ou mais consideram sua saúde ruim ou muito ruim. A tabela nove, abaixo, mostra a distribuição da auto-avaliação da condição de saúde para a população e para o grupo dos idosos.



**Tabela 9 - auto-avaliação da saúde: população e idosos (%)**

<i>Níveis de avaliação</i>	<b>População</b>	<b>Idosos (total e subgrupos)</b>			
	Total	Total	60 a 64	65 a 70	70 - mais
Muito bom	23.5	7.7	9.4	8.2	6.3
Bom	54.6	35.2	39.7	35.6	31.9
Regular	18.4	43.2	40.3	44.3	44.5
Ruim	2.9	11.3	8.7	9.7	14.0
Muito ruim	0.6	2.6	1.9	2.2	3.3

Fonte: micro dados PNAD (2003)

Não apenas se espera que a auto-avaliação da saúde seja pior com a idade, mas também que as medidas de limitação de atividades habituais e medidas de incidência de morbidades sejam piores. Com relação à pergunta sobre limitação de atividades por motivo de saúde, 93% da população e 86,6% dos idosos responderam não possuir nenhum tipo de limitação. Dentre os motivos desta limitação chama atenção ao fato de que os problemas de coração e de pressão acometem proporcionalmente mais os idosos (20% responderam ser este o motivo contra 8,5% da população total).

Pode-se notar que o percentual de doenças crônicas é muito superior no caso dos idosos do que na população em geral. Nesta última, as maiores incidências foram hipertensão (11,8%) e problemas de coluna (13,2%). Já na população idosa, essas morbidades também são bastante significantes (48,6% e 36,5%, respectivamente). Além destas, há outras doenças frequentes nos idosos: doenças cardíacas (17,3%), diabetes (12,9%) e artrite (28,4%). Isso confirma a tese de que, por um lado, os idosos têm maior necessidade de recorrer a serviços de saúde, embora tenham proporcionalmente menos recursos disponíveis para este investimento. As crianças, por sua vez, não registram alta incidência de doenças crônicas (as respostas positivas somam 1,6% no caso de crianças com idade de zero a nove anos). A tabela dez, a seguir, mostra a incidência das diferentes morbidades na população total e entre os idosos.

**Tabela 10 - morbidades (% de respostas positivas)**

	<b>População</b>	<b>Idosos</b>
Coluna	13.2	36.5
Artrite	6.1	28.4
Câncer	0.4	2.1
Diabetes	2.5	12.9
Asma	0.4	2.1
Hipertensão	11.8	48.6
Coração	3.6	17.3
Insuficiência Renal	1.7	4.2
Depressão	4.0	9.6
Tuberculose	0.2	0.4
Tendinite	2.2	4.9
Cirrose	0.1	0.3

Fonte: micro dados PNAD (2003)

A segunda parte do suplemento de saúde da PNAD refere-se à saúde suplementar, visando traçar o perfil dos usuários de planos de saúde. A primeira pergunta é se o respondente possui ou não um plano de saúde: 24,6% da população responderam que “sim” (aproximadamente 5% são plano de servidores públicos e 20% privados), enquanto que no grupo dos idosos este percentual foi de 29,4%. Com relação à avaliação dos planos, cerca de 80% da população total consideram o plano “muito bom” ou “bom”.

Em relação à titularidade dos que possuem planos de saúde, constata-se que a maioria (56%) é dependente do titular na população geral, mas não no grupo dos idosos em que 59,3% são titulares. Este fato justifica-se pelo alto percentual de idosos que são considerados pessoas de referência na família. Do total de pessoas de referência na família cobertas por plano de saúde, 77,6% se situava na condição de titular. Para os idosos, esta proporção cai para 71%. A tabela 11, abaixo, mostra a distribuição da titularidade dos planos de saúde de acordo com a condição do indivíduo na família.

**Tabela 11 - pessoas de referência na família e titularidade dos planos (%)**

	Total	Pessoas de referência na família	Idosos	Idosos referência na família
Titular	44.0	77.6	59.3	71.1
Dependente ou agregado	56.0	22.4	40.7	28.9

Fonte: micro dados PNAD (2003)

Em relação ao pagamento da mensalidade do plano de saúde, nota-se que na população em geral, a maioria dos planos é paga pelos titulares através dos seus trabalhos (autogestão). Já no caso dos idosos, a maior parte paga diretamente ao plano, embora seja alta a participação dos quem pagam através dos seus trabalhos antigos (em certos casos, os titulares, ainda que se aposentem, continuam usufruindo do benefício). A tabela 12 mostra a distribuição dos responsáveis pelo pagamento das mensalidades dos planos de saúde.

**Tabela 12 - pessoa responsável pelo pagamento da mensalidade do plano (%)**

	Total	Idosos
Somente o empregador do titular	9.5	3.6
O titular, através do trabalho atual	41.3	9.5
O titular, através do trabalho anterior	6.2	22.5
O titular, diretamente ao plano	32.4	50.6
Outra morador do domicílio	6.3	3.3
Pessoa não-moradora	2.8	7.4
Outro tipo	1.5	3.1

Fonte: micro dados PNAD (2003)

No aspecto regional da cobertura por plano de saúde, a região Sudeste é a que possui maior número de pessoas seguradas, seguida da região Sul, conforme a tabela 13, abaixo.

**Tabela 13 - População coberta por plano de saúde (%)**

	<i>Não Coberta</i>	<i>Coberta</i>	<i>Total</i>	<i>Taxa cobertura</i>
<b>NORTE</b>	<b>9,027,565</b>	<b>1,572,561</b>	<b>10,600,126</b>	<b>14.8%</b>
Rondônia	798,739	182,722	981,461	18.6%
Acre	342,835	75,536	418,371	18.1%
Amazonas	2,048,422	318,996	2,367,418	13.5%
Roraima	264,384	28,187	292,571	9.6%
Pará	3,980,569	795,681	4,776,250	16.7%
Amapá	424,794	81,913	506,707	16.2%
Tocantins	1,167,822	89,526	1,257,348	7.1%
<b>NORDESTE</b>	<b>43,916,980</b>	<b>6,027,829</b>	<b>49,944,809</b>	<b>12.1%</b>
Maranhão	5,548,934	407,316	5,956,250	6.8%
Piauí	2,634,330	319,302	2,953,632	10.8%
Ceará	6,947,481	931,063	7,878,544	11.8%
Rio Grande do Norte	2,613,577	315,847	2,929,424	10.8%
Paraíba	3,103,629	442,765	3,546,394	12.5%
Pernambuco	6,957,819	1,292,874	8,250,693	15.7%
Alagoas	2,705,202	247,792	2,952,994	8.4%
Sergipe	1,647,850	260,270	1,908,120	13.6%
Bahia	11,758,158	1,810,600	13,568,758	13.3%
<b>SUDESTE</b>	<b>51,358,796</b>	<b>25,137,664</b>	<b>76,496,460</b>	<b>32.9%</b>
Minas Gerais	14,017,395	4,780,591	18,797,986	25.4%
Espírito Santo	2,488,895	818,200	3,307,095	24.7%
Rio de Janeiro	10,529,933	4,528,720	15,058,653	30.1%
São Paulo	24,322,573	15,010,153	39,332,726	38.2%
<b>SUL</b>	<b>19,007,881</b>	<b>7,354,841</b>	<b>26,362,722</b>	<b>27.9%</b>
Paraná	7,616,204	2,418,322	10,034,526	24.1%
Santa Catarina	4,149,244	1,547,817	5,697,061	27.2%
Rio Grande do Sul	7,242,433	3,388,702	10,631,135	31.9%
<b>CENTRO-OESTE</b>	<b>9,460,267</b>	<b>3,109,650</b>	<b>12,569,917</b>	<b>24.7%</b>
Mato Grosso do Sul	1,548,332	655,442	2,203,774	29.7%
Mato Grosso	2,241,213	464,759	2,705,972	17.2%
Goiás	4,151,722	1,267,121	5,418,843	23.4%
Distrito Federal	1,519,000	722,328	2,241,328	32.2%
<b>BRASIL</b>	<b>132,771,489</b>	<b>43,202,545</b>	<b>175,974,034</b>	<b>24.6%</b>

Fonte: micro dados expandidos PNAD (2003)

Comparando a idade com o valor da mensalidade paga pelo plano de saúde, verifica-se que a proporção dos indivíduos que pagam valores mais elevados cresce com a idade, como já era esperado. A tabela 14 mostra a distribuição dos valores das mensalidades para diferentes faixas etárias. No caso dos idosos, 6,3% responderam que pagam mais do que 500 reais como mensalidade, enquanto que para os demais grupos etários esse percentual é bem menor.

**Tabela 14 - faixas de mensalidade vs. faixas etárias (%)**

	0a 4	5 a 9	10 a 14	15 a 19	20 a 24	25 a 39	40 a 59	Idoso
Até 30 reais	11.3	11.7	12.2	29.4	34.1	22.9	15.0	13.3
Mais de 30 até 50 reais	34.0	33.3	23.9	24.8	22.3	16.6	12.3	8.3
Mais de 50 até 100 reais	42.6	44.1	48.3	26.9	23.4	25.1	22.2	15.5
Mais de 100 até 200 reais	10.1	9.3	11.7	9.0	8.4	19.1	23.1	24.6
Mais de 200 até 300 reais	0.3	0.2	0.9	1.2	1.5	5.4	10.8	15.3
Mais de 300 até 500 reais	0.0	0.0	0.4	0.0	0.4	2.5	7.3	11.0
Mais de 500 reais	0.0	0.0	0.0	0.0	0.0	0.6	2.8	6.3
Não sabe	1.7	0.9	2.6	8.8	9.8	7.8	6.5	5.5
Sem declaração	0.0	0.5	0.0	0.0	0.0	0.1	0.1	0.0

Fonte: micro dados PNAD (2003)

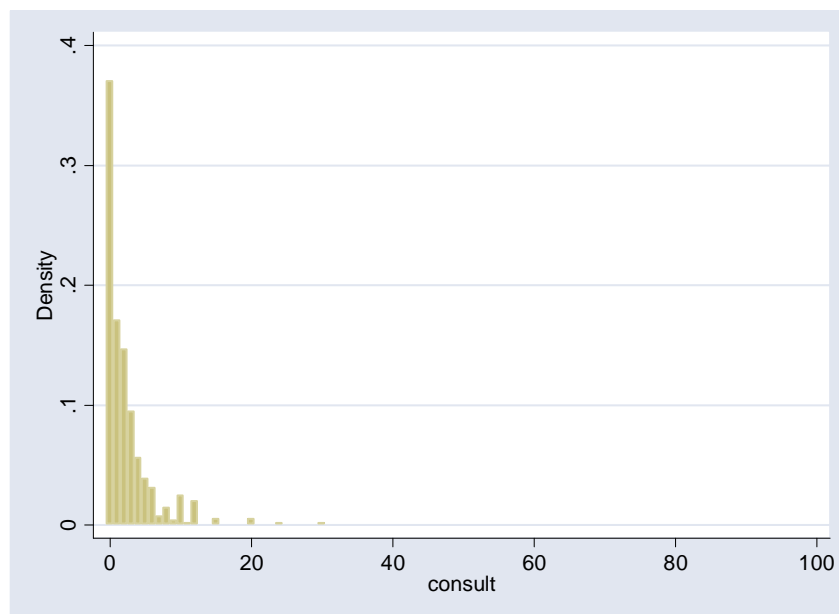
A terceira parte do suplemento da PNAD trata do acesso aos serviços de saúde. A primeira questão refere-se à fidelidade na relação médico-paciente: 78,3% responderam que costumam freqüentar o mesmo médico ou serviço de saúde. Sobre o local de atendimento que se costuma freqüentar, tem-se que 53,8% da população procuram os postos de saúde quando necessitam de atendimento médico, como pode ser visto na tabela 15, a seguir.

**Tabela 15 - local de atendimento que costuma freqüentar (%)**

Posto ou centro de saúde	53.8
Consultório particular	17.2
Ambulatório de hospital	17.0
Pronto-socorro ou emergência	5.0
Ambulatório ou consultório de clínica	4.5
Farmácia	1.4
Ambulatório ou consultório de empresa ou sindicato	0.9
Agente comunitário de saúde	0.2
Outro tipo de serviço (curandeiro, centro espírita, etc.)	0.1

Fonte: micro dados PNAD (2003)

Com relação à realização de consultas médicas no último ano, somente 63% dos indivíduos da amostra afirmaram terem feito alguma consulta. Destes indivíduos, a maior concentração (27%) é de apenas uma consulta realizada. A figura três abaixo, mostra o histograma do número de consultas médicas no ano de referência da PNAD (2003).

**Figura 3 - histograma do número de consultas**

Observando a distribuição do número de consultas por grupos etários, verifica-se que os grupos de idade que mais vão ao médico são os idosos e as crianças de 0 a 4 anos. Pode-se observar uma correlação negativa entre idade e número médio de visitas ao médico até os 19 anos e positiva a partir desta faixa etária em diante, conforme a tabela 16.

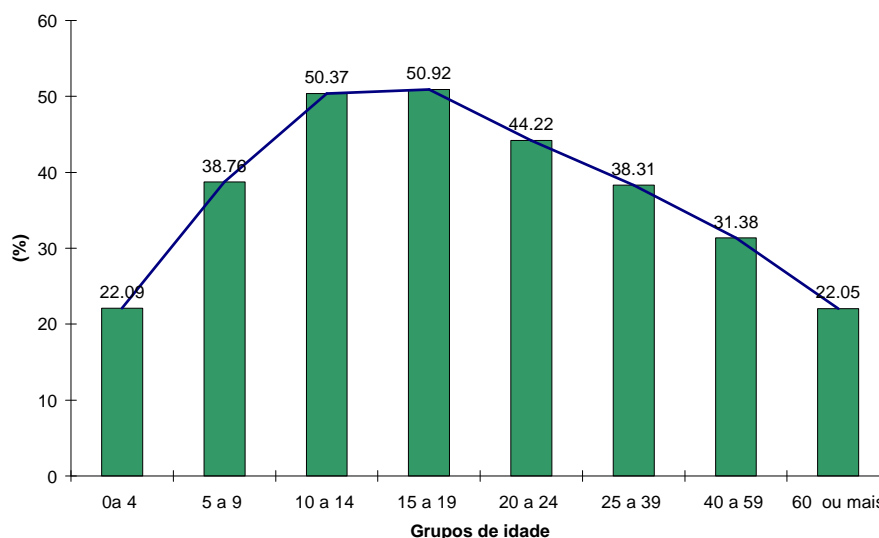
**Tabela 16 - média e desvio padrão do número de consultas por faixa etária**

	<b>Média</b>	<b>Desvio-Padrão</b>
0a 4	3.36	4.20
5 a 9	1.95	3.10
10 a 14	1.33	2.51
15 a 19	1.51	2.84
20 a 24	1.97	3.50
25 a 39	2.29	3.96
40 a 59	3.07	5.09
Idoso	4.32	5.99

Fonte: micro dados PNAD (2003)

A figura quatro, a seguir, faz uma análise complementar à tabela 16. Nela mostra-se a proporção de pessoas que não realizaram consultas médicas por grupos etários. Essa proporção é crescente até a faixa etária de 15 a 19 anos e decrescente a para as demais.

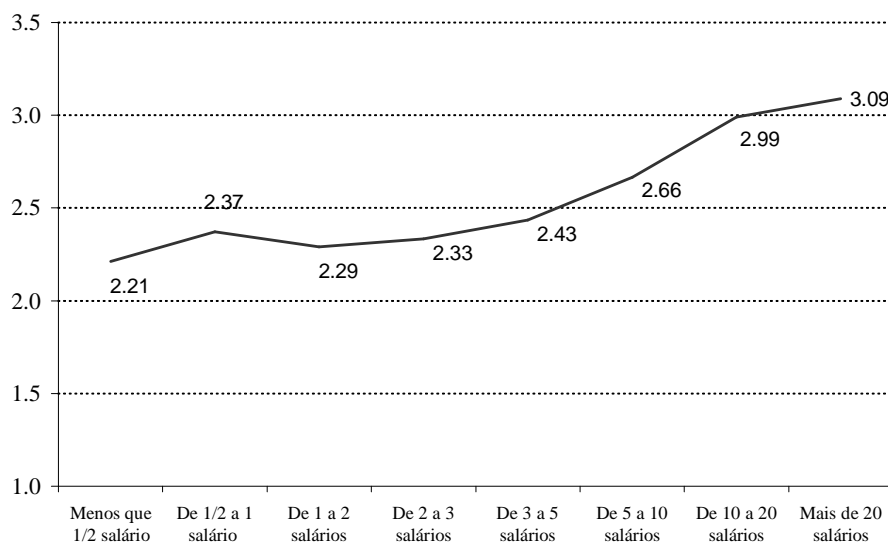
**Figura 4 - Proporção de pessoas que não realizaram consulta médica no último ano por faixa etária**



Fonte: PNAD (2003)

Com respeito ao rendimento domiciliar, conforme já ressaltado, existe uma correlação positiva entre o número de consultas e as faixas de rendimento, como mostra a figura cinco.

**Figura 5 – média de consultas realizadas no ano e faixas de rendimento domiciliar**



Fonte: PNAD (2003)

Dos indivíduos que não fizeram consultas ao médico, somente 39,8% são mulheres, o que pode indicar uma tendência dos indivíduos do sexo feminino a utilizar mais serviços de saúde

do que os do sexo masculino. Além do gênero, o fato de ser coberto por plano de saúde aumenta a probabilidade de utilizar serviços de saúde. A tabela 17 mostra a média de consultas no ano de acordo com a cobertura de plano de saúde e o sexo.

**Tabela 17 - média de consultas no ano segundo gênero e cobertura por plano**

	Sem plano	Com plano
Mulheres	2.68	4.20
Homens	1.63	2.83
Total	2.16	3.57

Fonte: micro dados PNAD (2003)

No questionário da PNAD saúde também existe um conjunto de perguntas sobre procura por atendimentos de saúde realizados nas duas semanas anteriores à pesquisa. Sobre esta questão, apenas 14,5% dos indivíduos afirmaram ter feito algum tipo de atendimento médico. Nesse grupo, 51% procuraram atendimento por motivo de doença e 26,3% com o intuito de cuidados preventivos. Também com relação às pessoas que procuraram cuidados médicos, 69,2% o fez apenas uma vez nas duas semanas anteriores à pesquisa. Além disso, semelhante à pergunta sobre consultas realizadas no último ano, aproximadamente 40,4% dos indivíduos responderam que procuraram o primeiro atendimento num posto de saúde, 19,2% em consultório particular e 18,3% em hospital. A tabela 18, abaixo, mostra os motivos da procura por atendimento para a população geral, idosos e crianças.

**Tabela 18 - motivo do principal atendimento realizado nas últimas duas semanas (%)**

	População	Idosos	Crianças
Acidente ou lesão	5.3	3.8	3.3
Problema odontológico	8.7	2.9	5.9
Reabilitação	2.3	2.9	1.3
Pré-natal	2.5	0.0	-
Puericultura	1.0	0.0	4.8
Vacinação	1.9	0.4	7.8
Outros atendimentos preventivos	26.2	29.4	18.7
Parto	0.4	0.0	0.0
Doença	51.1	60.4	58.1
Somente atestado de saúde	0.6	0.3	0.2
Sem declaração	0.1	0.0	0.0

Fonte: micro dados PNAD (2003)



Entretanto, nem todos os que procuraram atendimento de saúde foram atendidos na primeira vez que o fizeram: aproximadamente 4% dos indivíduos não foram atendidos. O motivo principal para o não atendimento (48,5%) foi o fato de não ter conseguido vaga/senha. O segundo motivo (26,2%) foi a falta de médicos.

O tipo de atendimento de saúde principal, nas duas últimas semanas, mais frequentemente apontado foi a consulta médica (71,5%), sendo realizado predominantemente na rede pública (61,4%). Do total de atendimentos principais, apenas 25,3% foram cobertos por plano de saúde e 58% das pessoas responderam que o atendimento principal foi realizado pelo Sistema Único de Saúde (SUS). A proporção de uso da rede pública e de financiamento via SUS é semelhante para os idosos: 60% e 59,2%, respectivamente. Para as crianças até dez anos, entretanto, estes percentuais são maiores: 71% e 69%. A tabela 19, a seguir, mostra os percentuais de tipo de atendimento para a população, idosos e crianças (de zero a nove anos).

**Tabela 19 – tipo do principal atendimento (%)**

	Total	Idosos	Crianças
Consulta médica	71.5	75.5	75.7
Consulta odontológica	8.7	2.8	5.7
Consulta de agente comunitário ou de parteira	0.1	0.2	0.2
Consulta de outros profissionais de saúde (fonoaudiólogos, psicólogos)	1.7	1.4	1.1
Consulta na farmácia	1.2	1.2	1.0
Quimioterapia, radioterapia, hemodiálise, hemoterapia	0.4	0.8	0.1
Vacinação, injeção, curativo, medição de pressão ou outro atendimento	4.2	3.7	9.6
Cirurgia em ambulatório	0.9	1.2	0.4
Gesso ou imobilização	1.0	0.7	0.6
Internação hospitalar	1.8	2.9	1.5
Exames complementares	6.1	7.1	2.6
Somente marcação de consulta	0.8	0.8	0.4
Outro atendimento	1.8	1.6	1.3
Sem declaração	0.0	0.0	0.0

Fonte: micro dados PNAD (2003)

Com respeito à avaliação deste atendimento principal, a proporção de “bom” e “muito bom” foi bastante alta – com um total de 85,2%. No entanto, se condicionarmos a pergunta somente aos atendimentos realizados pelo SUS, este percentual cai para 79,6%, com aumento da participação de respostas “regular”, “ruim” e “muito ruim” especialmente da primeira.

Finalmente, o último bloco de perguntas da PNAD Saúde (2003) refere-se à ocorrência de internação nos últimos 12 meses. Somente 7% da amostra total responderam que estiveram internados no último ano, sendo que 79,3% foram internados apenas uma vez, predominantemente com duração de um a três dias (representa 60,7% da pergunta sobre número de dias que esteve internado). Sobre o motivo do principal atendimento de internação, 54,2% responderam que estavam em tratamento clínico e 24,7% devido a cirurgias.

Assim como as consultas médicas, a maior parte (65,6%) da população utiliza a rede pública no caso de internações e apenas 23,2% destas foram cobertas por planos de saúde. Além disso, 68,5% dos que foram internados no último ano responderam que o atendimento foi realizado pelo SUS. Por último, semelhantemente às consultas, 86,7% avaliaram qualidade da internação sofrida como “muito bom” ou “bom”, sendo que este percentual cai levemente para 83,6% condicional à internação feita em rede pública.

### 3 MODELOS DE CONTAGEM<sup>9</sup>

Este capítulo tem por objetivo apresentar a metodologia dos modelos de contagem e bibliografia recente sobre o tema, os quais serão utilizados posteriormente na análise dos determinantes da demanda por consultas médicas. As seções de um a cinco mostram as distribuições de probabilidade e hipóteses utilizadas em cada modelo, assim como as vantagens e desvantagens de se utilizar cada um destes.

Os modelos de contagem possuem diversas aplicações no âmbito da economia, especialmente na Economia da Saúde. A estimação da demanda de serviços de saúde é um exemplo da utilização destes modelos, que tem como finalidade entender os determinantes sócio-econômicos que levam indivíduos a consultarem um profissional de saúde. Outros exemplos também no campo da saúde são: o consumo de cigarros durante um dia, o número de medicamentos receitados, e o número de dias de permanência no hospital.

Sob o ponto de vista econométrico, os modelos de contagem têm a peculiaridade de que a variável dependente é um número inteiro não negativo, sendo um evento a realização desta variável estocástica. Um evento pode ser interpretado como sendo a ocorrência de um ponto dentro de um processo em que há uma taxa de ocorrência específica daquele evento. Sendo assim, o modelo padrão de contagem é aquele que utiliza a distribuição Poisson, também conhecida como distribuição dos eventos raros<sup>10</sup>. O dual do evento é o intervalo em que estes ocorrem, o que leva ao estudo dos modelos de duração. No entanto, o foco dos modelos de

---

<sup>9</sup> A discussão dos modelos neste capítulo baseia-se em Cameron e Trivedi (1998).

<sup>10</sup> A lei dos eventos raros estabelece que o número total de eventos ocorridos em um dado intervalo de tempo segue aproximadamente a distribuição Poisson se qualquer evento pode acontecer em um número de tentativas grande e com uma probabilidade de ocorrência pequena. Formalmente, seja  $Y_{n,\pi}$  o número total de sucessos em um número  $n$  de tentativas de Bernoulli com probabilidade de ocorrência  $\pi$ . No caso limite em que  $n \rightarrow \infty$ ,  $\pi \rightarrow 0$  e  $n\pi = \mu > 0$ , tem-se:

$$\Pr[Y_{n,\pi} = k] = \binom{n}{k} \pi^k (1-\pi)^{n-k}, k = 0, 1, \dots, n$$

$$\lim_{n \rightarrow \infty} \left[ \binom{n}{k} \left( \frac{\mu}{n} \right)^k \left( 1 - \frac{\mu}{n} \right)^{n-k} \right] = \frac{\mu^k e^{-\mu}}{k!}$$

contagem, tema deste capítulo, é explicar o número de eventos que ocorrem em determinado período baseado em determinadas variáveis explicativas.

Dada a restrição sobre a variável dependente (o fato de ser discreta e não contínua), é preciso tratá-la de maneira adequada. Os modelos de regressão tradicionais (Método dos Mínimos Quadrados Linear) não são adequados para esse tipo de dado. Embora possa ser usado, o método de regressão linear é deficiente para prever dados de contagem, a menos que a média destes dados seja alta, caso em que a aproximação à distribuição normal seria satisfatória (CAMERON; TRIVEDI, 1998).

### 3.1 A distribuição Poisson

O modelo de contagem básico baseia-se na distribuição Poisson. Ainda que muitas vezes esta não se ajuste perfeitamente ao conjunto de dados – especialmente quando eles exibem sobredispersão (razão entre variância condicional e média condicional maior que um) – é importante salientar algumas características básicas desta distribuição.

Seja a variável discreta  $Y$  e a taxa de ocorrência (também conhecida como parâmetro de intensidade)  $\mu$ ,  $\mu > 0$ , e  $t$  o tempo de exposição definido como o período em que os eventos serão contados, então  $Y$  tem função de densidade:

$$\Pr[Y = y] = \frac{e^{-\mu t} (\mu t)^y}{y!}, y = 0, 1, 2, \dots \quad (3.1)$$

Onde  $E[Y] = V[Y] = \mu t$ , ou seja, a média é igual à variância. Esta propriedade é conhecida como equidispersão, mas é frequentemente violada na prática<sup>11</sup>.

---

<sup>11</sup> Outra propriedade chave da distribuição Poisson é a aditividade: seja  $Y_i \sim P[\mu_i]$ ,  $i=1, 2, \dots$  variáveis aleatórias independentes; se  $\sum_{i=1}^n \mu_i < \infty$ , então  $S_i = \sum_{i=1}^n Y_i \sim P[\sum_{i=1}^n \mu_i]$ . Nesse sentido, as distribuições Binomial e Multinomial podem ser derivadas da Poisson da seguinte forma:

### 3.1.1 Regressão Poisson – modelo de contagem padrão

O modelo de contagem padrão utiliza a distribuição Poisson, gerando um modelo de regressão não linear em que o parâmetro de intensidade ( $\mu$ ) depende de covariadas (regressores). Se esta dependência é parametricamente exata e envolve somente variáveis exógenas, o modelo é chamado de regressão Poisson padrão. Já se a função relacionada ao parâmetro  $\mu$  envolve também variáveis estocásticas não observáveis, obtém-se uma regressão Poisson mista.

Seja  $y_i$  o número de ocorrências de um evento de interesse e  $\mathbf{x}_i$  um vetor linearmente independente k-dimensional. O modelo de regressão de  $y_i$  em  $\mathbf{x}_i' = [x_{1i}, x_{2i}, \dots, x_{ki}]$  é obtido da determinação do parâmetro  $\mu(\mathbf{x}_i, \beta) = E[y_i/\mathbf{x}_i]$ . Ou seja:

$$f(y_i / x_i) = \frac{e^{-\mu_i} \mu_i^{y_i}}{y_i!}, y_i = 0, 1, 2, \dots \quad (3.2)$$

Na versão log-linear do modelo,  $\mu_i = \exp(\mathbf{x}_i' \boldsymbol{\beta})$  tal que a média condicional possui a forma multiplicativa  $E[y_i/\mathbf{x}_i] = \exp(x_{1i} \beta_1) \exp(x_{2i} \beta_2) \dots \exp(x_{ki} \beta_k)$ , além de assegurar que o parâmetro seja positivo.

### 3.2 Distribuição e modelo de regressão Binomial Negativo

A distribuição Poisson raramente ajusta-se aos dados de contagem, uma vez que é muito freqüente que estes exibam sobredispersão<sup>12</sup>. Além disso, a distribuição Poisson assume que os eventos ocorrem de maneira independente ao longo do tempo, isto é, que a probabilidade do paciente visitar o médico pela j-ésima vez é independente da (j+1)-ésima e (j-1)-ésima

$$\Pr\{Y_1 = y_1, Y_2 = y_2, \dots, Y_n = y_n / S_y = s\} = \left[ \prod_{j=1}^n \frac{e^{-\mu_j} \mu_j^{y_j}}{y_j!} \right] / \left[ \frac{(\sum \mu_i)^s e^{-\sum \mu_i}}{s!} \right]$$

, chegando-se à distribuição Multinomial (a Binomial é um caso específico em que  $n=2$ ).

<sup>12</sup> Diz-se que os dados são sobredispersos quando a variância condicional excede a média condicional. Uma indicação da magnitude da sobredispersão é simplesmente obtida comparando-se a média amostral com a variância da variável dependente.

visitas. Esta é uma hipótese que faz pouco sentido, pois uma vez tendo consultado determinado médico é bastante provável que ele sugira um retorno para verificar a eficácia do tratamento proposto. Sendo assim, a distribuição Binomial Negativa é mais adequada para tratar dos dados cuja variância é superior à média condicional, através da adição de um novo parâmetro ( $\alpha$ ) que reflete a heterogeneidade não observada. A função densidade da distribuição binomial negativa é dada pela seguinte equação:

$$f(y/\mu, \varphi) = \frac{\Gamma(y + \varphi)}{\Gamma(y + 1)\Gamma(\varphi)} \left(\frac{\varphi}{\varphi + \mu}\right)^{\alpha-1} \left(\frac{\mu}{\varphi + \mu}\right)^y, \varphi > 0, y = 0, 1, 2, \dots \quad (3.3)$$

Onde  $\Gamma(\cdot)$  é a distribuição Gamma. Se o parâmetro  $\alpha$  assume o valor zero, temos novamente a distribuição Poisson<sup>13</sup>.

Para efeito dos modelos de contagem, o que antes era  $E[y_i/x_i] = \exp(x_{i1}\beta_1) \exp(x_{i2}\beta_2) \dots \exp(x_{ik}\beta_k)$  ou  $\mu_i = \exp(\beta_0 + x_{i1}\beta_1 + x_{i2}\beta_2 + x_{ik}\beta_k)$  na distribuição Poisson, torna-se  $\mu_i' = \exp(\beta_0 + x_{i1}\beta_1 + x_{i2}\beta_2 + x_{ik}\beta_k + \varepsilon_i)$  na distribuição Binomial Negativa, onde acrescenta-se um termo de erro aleatório para captar a heterogeneidade latente<sup>14</sup>.

Para testar a presença de sobredispersão utiliza-se o seguinte teste. No modelo de regressão Poisson considerou-se que a média  $\mu_i = \exp(x_i\beta) =$  variância  $= \mu_i$ . Relaxando-se esta hipótese, assume-se genericamente que a variância da Binomial Negativa é:

$$\omega_i = V[y_i / x_i] = w(\mu_i, \alpha)$$

$$\omega_i = \mu_i + \alpha\mu_i^p \quad (3.4)$$

Ou seja, a variância condicional é função da média e do parâmetro de dispersão  $\alpha$ , além de uma constante  $p$  já especificada. Note que se  $\alpha = 0$ , tem-se a distribuição Poisson e se  $p = 1$  ou

<sup>13</sup> Existem dois tipos de distribuição Binomial Negativa, dependendo do valor do parâmetro  $\varphi$ . Se  $\varphi = 1/\alpha$ , tem-se a distribuição Binomial Negativa II (NB-II) e se  $\varphi = \mu/\alpha$ , tem-se a distribuição Binomial Negativa I (NB-I).

<sup>14</sup> Sendo  $\exp(\varepsilon_i) = \delta$ , então  $\mu_i' = \exp(\beta_0 + x_{i1}\beta_1 + x_{i2}\beta_2 + x_{ik}\beta_k + \varepsilon_i) = \exp(\beta_0 + x_{i1}\beta_1 + x_{i2}\beta_2 + x_{ik}\beta_k) \exp(\varepsilon_i) = \exp(\beta_0 + x_{i1}\beta_1 + x_{i2}\beta_2 + x_{ik}\beta_k) \delta$ . Para a identificação do modelo assume-se que  $E(\delta) = 1$ , o que significa que  $E(\varepsilon) = 0$ . Com esta hipótese, é fácil mostrar que  $E(\mu') = \mu E(\delta) = \mu$ . Ou seja, a média da distribuição Binomial Negativa é igual à da Poisson. Portanto, pode-se derivar a densidade da distribuição Binomial assumindo que os dados seguem uma distribuição Poisson, em que o parâmetro é  $\mu'$ . Dado que não se conhece o parâmetro  $\delta$ , assume-se que este tem distribuição Gamma, de maneira que após alguns desenvolvimentos chega-se à densidade exposta acima.

$p = 2$  tem-se as distribuições Binomial Negativa 1 (NB1) e Binomial Negativa 2 (NB2), respectivamente.

Sendo assim, um teste intuitivo é o teste LR da hipótese  $H_0: \alpha = 0$  contra  $H_1: \alpha > 0$ . Note que a restrição de que  $\alpha > 0$  implica achar a distribuição da estatística do teste, sendo necessário um experimento de Monte Carlo para determinar esta distribuição.

### 3.3 Modelos *Hurdle*

Os dados de utilização de serviços de saúde freqüentemente exibem uma distribuição fortemente concentrada no valor zero. Isso pode ocorrer pelo fato de que os zeros são originados de duas fontes. Caso seja feita uma pergunta do tipo “quantas vezes você foi ao médico nas duas últimas semanas?”, pode-se obter a resposta “nenhuma” ou porque a pessoa não vai ao médico nunca, ou ela até freqüenta, mas não fez nenhuma consulta nas últimas duas semanas.

Quando isso ocorre a amostra é considerada uma mistura (CAMERON; TRIVEDI, 1998). O modelo *Hurdle* trata dessa questão ao introduzir dois estágios na estimação, separando os processos geradores de zeros e de números positivos. Inicialmente, este modelo foi desenvolvido por Cragg (1971) no contexto do modelo Tobit, que trata a censura nos dados.

Cragg (1971) começa sua exposição com um exemplo de um evento que pode ou não ocorrer: a compra de um bem durável. Se um determinado evento ocorre, então se associa uma variável aleatória contínua e positiva, caso contrário, a variável toma o valor zero. Parte-se de um modelo Probit, em que a probabilidade média  $p(E_t)$  de que um evento ocorra na observação  $t$  é dada por:

$$p(E_t) = \int_{-\infty}^{x_t \beta} (2\pi)^{-1/2} \exp(-z^2 / 2) dz \quad (3.5)$$

Onde  $X_t$  é um vetor  $K \times 1$  de variáveis independentes na observação  $t$  e  $\beta$  os coeficientes associados.

Em seguida, Cragg (1971) menciona o modelo Tobit, que é uma extensão do modelo Probit em que se assume a existência de uma variável latente  $q_t$  gerada por:

$$q_t = X_t' \gamma + \varepsilon_t \quad (3.6)$$

Sendo  $\gamma$  um vetor de coeficientes e  $\varepsilon_t$  um erro aleatório normalmente distribuído com média zero e variância  $\sigma^2$ . Se  $q_t$  é negativa, então a variável observada  $y_t$  é zero. Se  $q_t$  é positiva, então  $y_t = q_t$ .

Entretanto, Cragg destaca o fato de que em algumas situações a decisão de adquirir um bem durável e a quantidade adquirida podem não ser intimamente relacionadas. Embora a aquisição de um bem só ocorra se existe o desejo de adquirir tal bem, há alguns fatores tais como os custos de transação que podem inibir que a compra se conclua de fato. Nesse sentido, podem-se obter zero de aquisição por ambos os motivos: não existe desejo e/ou há fatores impeditivos. Esta característica, segundo esse autor, não é incorporada pelo modelo *Tobit*.

Para contornar este problema, Cragg (1971) sugere um modelo alternativo estimado em duas partes: a primeira, relacionada à decisão de aquisição é um modelo de escolha binária (*Probit* ou *Logit*); a segunda, relacionada à quantidade adquirida é um modelo de regressão padrão. Note que para garantir a não negatividade de  $y_t$ , Cragg sugere a truncagem em zero da distribuição utilizada na segunda parte da estimação.

No âmbito da literatura sobre demanda por serviços de saúde, a distribuição que tem sido mais aplicada no segundo estágio do modelo *Hurdle* é a Binomial Negativa truncada em zero, dada a natureza dos dados em saúde, cuja variância é alta em relação à média.

O modelo *Hurdle* é, portanto, uma mistura (*finite mixture*) composta pela combinação de zeros gerados por uma densidade com zeros (a decisão de visitar um médico é negativa) e números positivos gerados por outra densidade truncada (frequência de visitas ao médico). Formalmente:



$$E[y/x] = \Pr[y > 0/x]E_{y>0}[y/y > 0, x] \quad (3.7)$$

Sendo a segunda esperança obtida a partir da densidade truncada<sup>15</sup>.

Mais especificamente no caso do modelo *Hurdle* Binomial Negativo, e utilizando a distribuição Binomial Negativa 2 (NB2) - que possui função de variância quadrática -, seja  $\mu_{i1} = \exp(x_i' \beta_1)$  o parâmetro da distribuição para o caso de zeros e  $\mu_{i2} = \exp(x_i' \beta_2)$  para o caso de um conjunto de números positivos:  $J=\{1,2,\dots\}$ . Definindo uma função indicadora  $I[y_i \in J]=1$  se  $y_i \in J$  e  $[y_i \in J]=0$  se  $y_i = 0$ . Pode-se obter da distribuição NB2 que:

$$\Pr[y_i = 0/x_i] = (1 + \alpha_1 \mu_{i1})^{-1/\alpha_1} \quad (3.8)$$

$$\Pr[y_i / x_i, y_i > 0] = \frac{\Gamma(y_i + \alpha_2^{-1})}{\Gamma(\alpha_2^{-1})\Gamma(y_i + 1)} \left( \frac{1}{(1 + \alpha_2 \mu_{i2})^{1/\alpha_2} - 1} \right)^{-\alpha_2^{-1}} \left( \frac{\mu_{i2}}{\mu_{i2} + \alpha_2^{-1}} \right)^{y_i} \quad (3.9)$$

Onde  $x$  refere-se à probabilidade de se obter zeros e  $y$  à probabilidade condicional da distribuição NB2 truncada<sup>16</sup>.

Na literatura sobre demanda de cuidados médicos, o trabalho de Pohlmeier e Ulrich (1995) é referência, tendo sido utilizado um modelo *Hurdle* Binomial Negativo para analisar os dois processos inerentes à utilização de serviços de saúde: a decisão de consultar um médico e a frequência das consultas. A base de dados utilizada foi o Painel Socioeconômico Alemão de 1985. O intuito era estimar a demanda por consultas a médicos generalistas e especialistas, no contexto do sistema de saúde alemão, caracterizado pelo uso de *vouchers*<sup>17</sup>.

---

<sup>15</sup> Pode-se demonstrar que a variância condicional é: a seguinte  $V[y/x] = \Pr[y > 0/x]V_{y>0}[y/y > 0, x] + \Pr[y = 0/x]E_{y>0}[y/y > 0/x]$

<sup>16</sup> Note que o modelo *Hurdle* Poisson e *Hurdle* Geométrico, ambos examinados por Mullahy (1986) podem ser obtidos de (4.8) através de (4.9) fazendo  $\alpha_1 = \alpha_2 = 0$  e  $\alpha_1 = \alpha_2 = 1$ , respectivamente. Quando  $\alpha_1 = 1$ ,  $\Pr[y_i = 0/x_i] = (1 + \mu_{i1})^{-1}$ , tal que se  $\mu_{i1} = \exp(x_{i1}' \beta_1)$ , o primeiro estágio do modelo *Hurdle* é um modelo Logit de escolha binária.

<sup>17</sup> Os indivíduos compram *vouchers* ou cupons emitidos pelo governo que serão enviados às companhias seguradoras. Estas, por sua vez, pagam aos médicos.

Para a estimação da primeira parte do processo de utilização de serviços de saúde, Pohlmeier e Ulrich (1995) seguem o modelo proposto por Grossman (1972). Neste modelo, a saúde é vista como um estoque de capital durável que gera como produto “tempo de vida saudável”. Assume-se que os indivíduos possuem um estoque inicial de saúde que vai se depreciando com o tempo e que pode ser aumentado com o investimento em saúde (bons hábitos alimentares, exercícios físicos, visitas ao médico). Nesse sentido, o preço sombra da saúde depende não apenas do preço observado dos cuidados médicos, mas também da idade, nível de educação (espera-se que o preço sombra se reduza com mais anos de estudo, uma vez que pessoas mais educadas produzem mais eficientemente saúde), salário e outros fatores.

Segundo o trabalho de Grossman (1972), os indivíduos têm preferências bem-comportadas para o bem de consumo “boa saúde”, uma vez que ficar doente é considerado uma desutilidade (a curva de demanda por saúde é negativamente correlacionada com seu preço sombra). Por outro lado, a boa saúde também é um bem de investimento, pois ela determina o tempo total disponível para todos os tipos de atividade. Ou seja, dias com saúde debilitada reduzem os dias de trabalho, e esta redução monetária pode ser contabilizada no índice de retorno do investimento em saúde.

Sendo assim, considera-se que o consumidor típico maximiza intertemporalmente sua utilidade, sendo o tempo de vida ( $n$ ) de cada consumidor uma variável endógena: a morte ocorre quando se atinge o nível mínimo de saúde, isto é quando  $H_i = H_{\min}$ . Seja a utilidade intertemporal de um indivíduo:

$$U = U(\phi_0 H_0, \dots, \phi_n H_n, Z_0, \dots, Z_n) \quad (3.10)$$

Onde  $h_i = \phi_i H_i$  é o total consumido de serviços de saúde e  $Z_i$  é o total consumido de outros bens no  $i$ -ésimo período.

No modelo de Grossman (1972), ao longo do ciclo de vida do indivíduo, a taxa de depreciação da saúde eleva-se, mas a demanda por cuidados de saúde pode aumentar ou não, dependendo do custo do investimento. O indivíduo “escolheria” a morte (nível mínimo de saúde) quando o custo de investir em mais saúde fosse demasiado elevado e/ou impraticável.

Sob esse referencial teórico, Pohlmeier e Ulrich (1995) definiram a seguinte especificação para a primeira parte da estimação:

$$\ln M(t) = \beta_0 + \ln H(t) + \beta_1 \ln w(t) - \beta_2 \ln P_m(t) + \beta_3 t + \beta_4 X_1 + \beta_5 E + u(t) \quad (3.11)$$

Onde  $M(t)$  é a demanda por serviços médicos,  $H(t)$  o status de saúde,  $w(t)$  o salário do indivíduo,  $P_m(t)$  o preço dos serviços médicos,  $t$  uma tendência,  $X$  são fatores ambientais,  $E$  o nível de educação e  $u(t)$  o erro aleatório. Espera-se que o salário seja positivamente correlacionado com a demanda, uma vez que quanto maior o salário que a pessoa ganha, maior é a disposição para investir em saúde (custo de oportunidade é menor).

O modelo de Grossman (1972) focaliza principalmente o indivíduo, ou paciente em si, como determinante do investimento em saúde. No entanto, existe outra abordagem na literatura que também vê o médico, ou provedor, como um agente importante na demanda por cuidados médicos: ele determina, pelo menos em parte (pode haver alguma restrição por parte do paciente), a frequência de consultas. Esta é a perspectiva da teoria do agente-principal, em que se considera que a relação médico-paciente envolve assimetria de informação, gerando um risco moral de ambas as partes. Os pacientes desejam manter ou restaurar um determinado estoque de saúde, e por isso decidem procurar um serviço de saúde. No entanto, os médicos também influenciam a demanda do paciente, uma vez que são quem avaliam e detêm informação privilegiada sobre o status de saúde deste, determinando assim a frequência com que o paciente consome seus serviços.

A teoria do agente principal na demanda de cuidados médicos foi apontada por Zweifel (1981). Para esse autor, além de seguirem critérios estritamente clínicos, os médicos respondem a incentivos econômicos. Nesse sentido, a frequência de consultas não é mais um resultado gerado unicamente pela maximização da utilidade dos pacientes, mas também é resultante da maximização da utilidade dos provedores.

Quando esta relação envolve ainda um terceiro agente - o segurador - pode ser incentivada uma demanda por consultas induzida pelos médicos, que tendem a pedir mais exames e tratamentos desnecessários, especialmente se o reembolso for por consulta/ procedimento. Desta forma, os médicos tenderiam a compensar a perda de renda que eles obteriam com a

alta do paciente e a redução dos incentivos à primeira consulta em decorrência da introdução de mecanismos de copagamento pelos planos de saúde. Este mecanismo, por sua vez, é justamente uma das formas com que os seguros-saúde procuram lidar com o *moral hazard* inerente ao processo de utilização de serviços médicos.

Um último aspecto do risco moral existente nas relações entre paciente, provedor e segurador é o fato de que o paciente segurado não hesita em utilizar os serviços oferecidos pelo plano, uma vez que o custo marginal de mais uma consulta é muito reduzido caso o paciente seja coberto por um plano de saúde integral. Sendo assim, é natural que pessoas com maior probabilidade de ficarem doentes busquem o máximo de cobertura pelos planos e saúde, problema conhecido como seleção adversa.

O modelo *Hurdle* também foi aplicado no estudo do mercado de saúde suplementar. Cameron *et al* (1988) procuraram tratar da relação entre seguro saúde e demanda utilizando duas estratégias de estimação: modelos de contagem (Binomial Negativo) e variáveis instrumentais<sup>18</sup>. Embora a estimação do modelo *Hurdle* Binomial Negativo incorpore o fato de que as variáveis dependentes são números inteiros não negativos, ele não comporta a suspeita de endogeneidade do seguro saúde - a compra do seguro depende da expectativa de uso futuro de serviços de saúde (auto-seleção). Nesse sentido, foi utilizado o método de variáveis instrumentais para corrigir este problema, empregando como instrumento as previsões de um modelo de escolha binária (Logit) estimado previamente para as *dummies* de seguro saúde. Estas previsões atuaram como instrumentos das próprias *dummies* de seguro no modelo de estimação da demanda.

O problema de seleção adversa também foi discutido em Cutler e Zeckhauser (1998). Para esses autores, a seleção adversa pode levar a três classes de ineficiências: (i) os preços podem não serem iguais aos custos marginais, o que pode fazer com que os indivíduos escolham

---

<sup>18</sup> A base de dados utilizada foi a *Australian Health Survey* (AHS) realizada pela *Australian Bureau of Statistics* (ABS) entre 1977 e 1978. Foram usadas como variáveis dependentes: o número de consultas realizadas nas últimas duas semanas a contar da semana de referência da pesquisa, o número de internações em hospital no último ano, o número de consultas feitas a profissionais não médicos (fisioterapia, enfermagem, etc) nas últimas quatro semanas e o número de medicamentos utilizados com e sem prescrição médica nos últimos dois dias. Como variáveis explicativas foram usadas *dummies* de seguros saúde, *dummy* de auto avaliação da saúde, número de condições crônicas, *dummies* de limitação de atividade, duração da doença, grau de preocupação devido a doença, número de dias de atividade limitada pela doença e tipo de médico consultado, além de variáveis sócio econômicas como sexo, idade, ocupação, região, origem, classe de renda, educação.

planos inadequados; (ii) a dispersão do risco pelos usuários dos planos pode ser prejudicada; (iii) os planos de saúde podem manipular suas ofertas de modo a atrair somente pessoas saudáveis.

Por exemplo, se um indivíduo de baixo risco deve pagar \$40 pelo Plano Básico e \$60 pelo Plano Superior, ele deverá escolher o Plano Básico, uma vez que a diferença de benefício se ele escolher o Plano Superior é menor do que a diferença no custo. Já o indivíduo de alto risco deverá escolher o Plano Superior pela mesma razão com sentido inverso.

**Tabela 20 - exemplo custo/benefício de um plano de saúde**

Custo/Benefício (\$)	Plano Básico	Plano Superior	Diferença de Custo	Diferença de Benefício (Superior-Básico)
Baixo risco	40	60	20	15
Alto risco	70	100	30	40

Fonte: CUTLER E ZECKHAUSER (1998)

Se um empregador oferece esses dois planos para seus funcionários, por questões legais ou porque o empregador não consegue distinguir o tipo de funcionário, partindo do equilíbrio eficiente (indivíduo de baixo risco escolhe o Plano Básico e indivíduo de alto risco escolhe o Plano Superior), ele deverá cobrar \$40 pelo Plano Básico e \$100 pelo Plano Superior. Mas, desta maneira, todos os indivíduos de alto risco irão migrar para o Plano Básico, uma vez que a diferença em termos de benefício (\$40) não compensa a diferença de custo (\$60). A ineficiência relacionada aos preços surge em razão deste fato.

Arrow (1963), em seu *paper* seminal, já apontava o problema da assimetria de informação entre pacientes e médicos como relevante. Mais ainda, ressaltou a incerteza inerente à demanda por cuidados médicos. Segundo esse autor, os cuidados de prevenção são muito mais periódicos e constantes do que os cuidados médicos, nos quais o paciente só procura o atendimento em função de uma doença que se desenvolveu recentemente.

Além disso, Arrow (1963) também discute a questão do *moral hazard* no caso do seguro saúde. Segundo ele, a expansão dos seguros gerou uma expansão da demanda por serviços de saúde, motivada pelo comportamento dos médicos, que passaram a prescrever drogas mais caras, exames complementares diversos e maior frequência de consultas. Adicionalmente,

para ele, os seguros geram ineficiência no sentido de que tiram os incentivos a procurar preços melhores para cirurgias, exames, consultas, etc.

Uma outra aplicação do modelo *Hurdle* para estimar a demanda por serviços de saúde é o trabalho de Gurmu (1997), em que se utilizou o modelo *Hurdle* para estudar a utilização do *Medicaid recipients* através de uma medida do número de visitas a médicos e centros de saúde durante um período de quatro meses<sup>19</sup>. Assim como os demais autores, a justificativa para a escolha do modelo foi de que a estimação em dois estágios seria condizente com o processo dual da decisão de contato e frequência de consultas médicas.

### 3.4 Modelos *Finite Mixture*

O modelo *Hurdle*<sup>20</sup>, também conhecido como “*two-part model*” tem como motivação a teoria do agente principal aplicada à demanda de cuidados médicos, a qual sugere que o paciente inicialmente escolhe se vai ou não procurar tratamento, sendo esta uma escolha individual. Após decisão positiva de consultar um médico, este último influenciaria a frequência das consultas.

Os modelos conhecidos como *Finite Mixture*, entretanto, não necessariamente trabalham com a dicotomia usuários e não usuários, sendo mais flexíveis, uma vez que admitem várias subpopulações. Estes modelos caracterizam-se por terem funções de densidade compostas, em que a variável aleatória é gerada através de uma mistura de C populações distintas, cujas proporções são  $\pi_1, \pi_2, \dots, \pi_C$ , onde  $\sum_{j=1}^C \pi_j = 1$ ,  $\pi_j > 0$ . A densidade mista é definida por:

---

<sup>19</sup> Entretanto, a maior inovação de Gurmu (1997) foi estimar um modelo *Hurdle* semi paramétrico usando expansão de séries (Laguerre) para analisar a demanda de cuidados de saúde por usuários do programa *Medicaid*. A base de dados para o estudo foi a *Medicaid Consumer Survey*, realizada em 1986. Foram utilizadas como variáveis explicativas: número de crianças no domicílio, idade, renda, limitações funcionais, doenças crônicas, sexo, raça, educação, estado civil, etc. As conclusões foram de que não há evidência de que a migração de usuários do *Medicaid* para o *managed care* ( programa designado para contenção de custos, em que o usuário escolhe um provedor de cuidados primários, que autoriza hospitalizações e outros procedimentos) reduza a utilização de serviços de saúde.

<sup>20</sup> Embora possa ser considerado uma mistura, o modelo *Hurdle* difere estruturalmente dos modelos *finite mixture*, que se caracterizam por constituírem uma adição de densidades de C distintas populações.

$$f(y_i / \Theta) = \sum_{j=1}^{C-1} \pi_j f_j(y_i / \theta_j) + \pi_C f_C(y_i / \theta_C) \quad (3.12)$$

Para identificação do modelo, em que  $(\pi_C = (1 - \sum_{j=1}^{C-1} \pi_j))$ , assume-se que  $\pi_1 \geq \pi_2 \geq \dots \geq \pi_C$ .

É importante notar que os modelos *Finite Mixture*, são modelos semi paramétricos: não há necessidade de se fazer hipóteses sobre as distribuições de cada subpopulação. Contudo, algumas vezes, é realizada uma parametrização usando, por exemplo, a função logística tal que  $\pi_j = \frac{\exp(\lambda_j)}{(1 + \exp(\lambda_j))}$  e, por sua vez,  $\lambda_j$  é parametrizado em termos de variáveis observadas.

### 3.4.1 Modelos *Zero-Inflated*

Assim como o modelo *Hurdle*, o modelos *Zero-Inflated* também tratam do excesso de zeros, porém considera outra alternativa ao invés da truncagem. Considere:

$$\Pr[y_i = 0] = \varphi_i + (1 - \varphi_i)e^{-\mu_i} \quad (3.13)$$

$$\Pr[y_i = r] = (1 - \varphi_i) \frac{e^{-\mu_i} \mu_i^r}{r!}, \quad r = 1, 2, \dots \quad (3.14)$$

Onde  $\varphi_i$  é a proporção de zeros e a distribuição pode ser considerada um caso especial de *Finite Mixture* com uma distribuição degenerada cuja massa está concentrada no zero.

Pode-se notar que o excesso de zeros implica sobredispersão, uma vez que a variância é maior que a média:

$$V[y_i] = (1 - \varphi_i)(\mu_i + \varphi_i \mu_i^2) > \mu_i(1 - \varphi_i) = E[y_i] \quad (3.15)$$

A proporção  $\varphi_i$  pode ser parametrizada por uma função logística do vetor de covariadas  $z_i$ , assegurando a não negatividade de. Assim:

$y_i = 0$ , com probabilidade  $\varphi_i$

$y_i \sim P[\mu_i]$ , com probabilidade  $(1 - \varphi_i)$

$$\varphi_i = \frac{\exp(z_i' \gamma)}{1 + \exp(z_i' \gamma)} \quad (3.16)$$

Os modelos *Zero-inflated* são variantes das regressões Poisson e Binomial Negativa tradicionais em que se tratam os zeros da amostra: os zeros podem ser originados parcialmente das distribuições Poisson (*Zero-inflated Poisson*) ou Binomial Negativa (*Zero-inflated Binomial Negativo*) com probabilidade  $(1 - \varphi_i)$  e parcialmente da distribuição de não usuários com probabilidade  $\varphi_i$ .

### 3.4.2 Modelos *Finite Mixture* genéricos

Os modelos *Zero-inflated* são casos particulares da abordagem *Finite Mixture*. Estes modelos são baseados em distribuições de probabilidade compostas, conferindo flexibilidade à forma funcional, uma vez que cada componente pode ser interpretado como uma aproximação local de alguma parte da verdadeira distribuição da variável dependente<sup>21</sup>.

---

<sup>21</sup> Uma outra característica é que estes modelos não pressupõem hipóteses de distribuição da variável ou mistura. As distribuições dos componentes, contudo, são comumente especificadas – embora possam diferir, geralmente os componentes são igualmente parametrizados, utilizando as mesmas funções de densidade. Um caso específico é o modelo em que o  $j$ -ésimo componente possui intercepto  $\theta_j$  e a inclinação é a mesma para todos os componentes. Ou seja, as subpopulações diferem aleatoriamente somente com respeito ao parâmetro de localização (inclinação). Esta forma de representação da heterogeneidade é comumente chamada de representação semiparamétrica, já que não há parametrização dos componentes. Um outro modelo, ainda mais genérico, é aquele em que não só os parâmetros de inclinação são aleatórios, mas também os de intercepto. (CAMERON; TRIVEDI, 1998).



O modelo *Finite Mixture* é intimamente relacionado com a análise *Latent Class* (CAMERON; TRIVEDI, 1998). Na prática, o número de componentes deve ser estimado<sup>22</sup>, porém assumindo que o número de componentes é igual a C, sejam  $d_i = (d_{i1}, \dots, d_{iC})$  *dummies* tal que  $d_{ij} = 1; \sum_j d_{ij} = 1$ , indicando que  $y_i$  foi gerado pelo j-ésimo grupo ou classe latente para  $i=1,2,\dots,n$  observações. O modelo *Finite Mixture* especifica que  $(y_i / d_i, \theta, \pi)$  são independentemente distribuídos com densidades iguais a:

$$\sum_{j=1}^C d_{ij} f(y_i / \theta_j) = \prod_{j=1}^C f(y_i / \theta_j)^{d_{ij}} \quad (3.17)$$

E  $d_{ij} / \theta, \pi$  são independentes e identicamente distribuídos com distribuição multinomial:

$$\prod_{j=1}^C \pi_j^{d_{ij}}, \quad 0 < \pi_j < 1, \quad \sum_{j=1}^C \pi_j = 1 \quad (3.18)$$

A probabilidade de que  $y_i$  pertença à subpopulação j é igual a:

$$z_{ij} = \frac{\pi_j f_j(y_i / x_i', \theta_j)}{\sum_{j=1}^C \pi_j f_j(y_i / x_i', \theta_j)} \quad (3.19)$$

Um trabalho a ser destacado na linha dos modelos *finite mixture* é o de Creel e Farell (2005). Os autores estimaram e compararam a performance de oito modelos de contagem diferentes todos através do método de máxima verossimilhança. Os dados utilizados foram obtidos do *Medical Expenditure Panel Survey* de 1996-2000, nos EUA. As variáveis dependentes utilizadas foram o número de visitas (ao médico, ao dentista, a salas de emergência) e o número de drogas prescritas. Já como variáveis explicativas foram testadas as coberturas por

---

22 Para inferência sobre o número de componentes C, é possível utilizar dois procedimentos: teste LR entre as alternativas  $C=C^*$  versus  $C=C^*+1$  ou estimar o modelo para  $C \leq C^*$ , e daí usar critérios de informação para decidir entre as especificações. O problema do primeiro teste é que a estatística não tem distribuição definida, portanto, os valores críticos devem ser derivados via simulações de Monte Carlo. No modelo apresentado no próximo capítulo seguiu-se a abordagem de Deb e Trivedi (1997), adotando uma especificação com dois componentes.

seguro privado e por seguro público e variáveis sócio econômicas como sexo, idade, educação e renda familiar.

Primeiramente, foi utilizada uma estratégia paramétrica: modelo de regressão Binomial Negativo (NB) e modelo *Hurdle* Binomial Negativo (HNB). Após estimar modelos paramétricos, Creel e Farell (2005) passam à abordagem semi paramétrica de Gurmu (1997), em que também se assume a existência de heterogeneidade latente, permitindo que esta siga uma densidade não paramétrica, semelhante ao trabalho de Gallant e Nychka (1987). A diferença é que ao invés de polinômios de Hermite, utilizados por estes últimos, empregou expansão de polinômios de Laguerre<sup>23</sup>

Para a modelagem da densidade  $g_v(v, \phi)$  Gurmu utilizou uma expansão de séries polinomiais de Laguerre em torno de uma densidade Gamma, obtendo o modelo Poisson semi-paramétrico (PSP). De forma semelhante, foram feitas extensões para o modelo *Hurdle* Binomial Negativo, tal que se obteve o modelo *Hurdle* Poisson Semi Paramétrico (HPSP).

Em seguida, Creel e Farell (2005) introduziram modelos semi não paramétricos, já trabalhados por Cameron e Johansson (1997), através de um rearranjo da função de densidade Poisson utilizando um polinômio quadrático. Nesse sentido, foram estimados os modelos Poisson Semi Não Paramétrico (PSNP) e Binomial Negativo Semi Não Paramétrico (NBSNP). De forma semelhante, foram feitas extensões para o modelo *Hurdle* Binomial Negativo, tal que se obteve o modelo *Hurdle* Poisson Semi Paramétrico (HPSP).

---

<sup>23</sup> No modelo Poisson semi paramétrico (PSP), a hipótese é de que a média é estocástica tal que  $E[Y/V=v] = \lambda v$ . Através da integração da função densidade Poisson com média  $\lambda v$ , tem-se:

$$f_y(y, \lambda, \phi) = \int \frac{e^{-\lambda v} (\lambda v)^y}{y!} g_v(v, \phi) dv$$

$$= \frac{\lambda^y}{y!} M_v^y(-\lambda)$$

Onde  $M_v^y(-\lambda)$  é a derivada de y-ésima ordem da função geradora de momento de V avaliada em  $-\lambda$ .  $M_v^0(-\lambda) = M_v(-\lambda)$  é a função geradora em si.

Para a modelagem da densidade  $g_v(v, \phi)$  Gurmu utilizou uma expansão de séries polinomiais de Laguerre em torno de uma densidade Gamma, obtendo o modelo Poisson semi-paramétrico (PSP).

Por último, Creel e Farell (2005) usam a abordagem dos modelos *Finite Mixture*, bastante conhecidos na literatura de econometria da saúde. Introduzida por Deb e Trivedi (1997), conforme explicitado anteriormente, esta abordagem tem o apelo intuitivo de permitir subgrupos da população com diferentes estados de saúde, captando a heterogeneidade latente de maneira mais realista.

No estudo de Creel e Farell (2005) sobre a abordagem *Finite Mixture* foi utilizado o modelo Binomial Negativo Misto (MNB), o qual possui a seguinte função de densidade:

$$f_y = (y, \phi_1, \dots, \phi_p, \pi_1, \dots, \pi_{p-1}) = \sum_{i=1}^{p-1} \pi_i f_y^{(i)}(y, \phi_i) + \pi_p f_y^{(p)}(y, \phi_p)$$

$$\text{Onde } \pi_i > 0, i=1,2,\dots,p; \pi_p = 1 - \sum_{i=1}^{p-1} \pi_i \text{ e } \sum_{i=1}^p \pi_i = 1. \quad (3.20)$$

As funções  $f_y^{(i)}(y, \phi_i), \phi_i = \{\lambda_i, \varphi_i\}$  são  $p$  densidades separadas, podendo ser dos tipos Binomial Negativa 1 (NB1) ou Binomial Negativa 2 (NB2).

A conclusão do trabalho Creel e Farell (2005) foi de que os resultados estatísticos dos oito modelos estimados foram muito semelhantes.

Uma outra comparação entre modelos de contagem também foi elaborada por Gerdtham e Trivedi (2000). Os autores utilizaram dados da pesquisa *Level of Living* de 1991 (Suécia) para estimar o número de visitas ao médico em um ano e o número de semanas de permanência no hospital por ano. As conclusões obtidas, no entanto, foram diferentes das de Creel e Farell (2005). Ao compararem o modelo em duas partes (*Hurdle*) com o *Finite Mixture*, Gerdtham e Trivedi (2000) verificaram que este último teve melhor ajuste aos dados do que o primeiro. Para esta comparação foram utilizados critérios de informação de *Akaike* (AIC) e *Bayesian* (BIC). Segundo os autores, o modelo misto é mais flexível que o modelo em suas partes, pois não se restringe à dicotomia usuário freqüente ou não freqüente. Ao contrário, acomoda diferentes subpopulações, com níveis de saúde e uso dos serviços médicos distintos.



## 4 MODELO ECONOMETRICO<sup>24</sup>

O objetivo deste capítulo é aplicar o conhecimento exposto anteriormente sobre os vários modelos de contagem utilizados na estimação da demanda por consultas médicas. Sendo assim, serão expostos e comparados os resultados dos seguintes modelos de contagem: Poisson (PRM), Binomial Negativo (NBRM), *Hurdle* Poisson (HP), *Hurdle* Binomial Negativo (HNB), Poisson *Zero-inflated* (ZIP), Binomial Negativo *Zero-inflated* (ZINB) e *Finite Mixture* (FMM). Conforme mencionado anteriormente, a base de dados utilizada foi a PNAD (2003), incluindo as variáveis específicas do suplemento sobre saúde. Em cada seção serão confrontados dois modelos, buscando aquele com melhor ajuste. A última seção traz uma comparação geral de todos os modelos, indicando o mais adequado segundo critérios econométricos específicos. Escolhido o modelo, as conclusões finais serão explicitadas no capítulo seguinte.

### 4.1 Estimação do preço dos serviços de saúde

O processo de estimação teve início com a determinação dos direcionadores de preço dos serviços de saúde. Para o sistema de saúde privado, o preço pago pelo indivíduo é explícito: ou se desembolsa diretamente, através do pagamento da consulta em si, ou indiretamente, através de uma mensalidade paga ao plano de saúde. No questionário da PNAD, existe a pergunta se o atendimento principal realizado nas duas últimas semanas anteriores à pesquisa foi pago e outra pergunta se este atendimento foi financiado pelo SUS, mas sem discriminar o valor efetivamente pago pela consulta. Nesse sentido, não é possível saber objetivamente o valor desembolsado pelas pessoas que geralmente são usuárias do sistema privado, mas não possuem plano de saúde.

---

<sup>24</sup>Os modelos foram estimados usando o programa STATA versão 9.2. Os comandos e programas foram retirados de Long & Freese (2001) e Jones (2007).

Numa perspectiva simplista, poderia ser considerado como usuário do sistema privado todo indivíduo que não usou o SUS para seus atendimentos de saúde. No entanto, é preciso distinguir entre aquele indivíduo que tem plano de saúde do indivíduo que não tem plano. Além disso, é possível que existam casos mistos: pessoas que ora usam o SUS, ora pagam para consultar um médico de sua preferência. No mesmo sentido, existem pessoas que geralmente usam os serviços do plano, mas que podem ter utilizado o SUS ocasionalmente, por exemplo, em caso de acidente com veículos ou para um serviço não coberto pelo plano. Entretanto, para captar esses efeitos não é possível utilizar as seguintes variáveis: a *dummy* indicadora de pagamento pelo atendimento principal realizado nas duas semanas anteriores e a *dummy* indicadora de uso do SUS no atendimento principal nas duas últimas semanas. O motivo é que o modelo proposto tem como variável dependente o número de consultas realizado no ano e não necessariamente que se consultou no ano se consultou nas duas últimas semanas, o que reduziria muito o tamanho da amostra.<sup>25</sup>

Apesar de não haver discriminação do preço pago pelo atendimento em si, é possível derivar a variável de mensalidade paga pelo plano de saúde<sup>26</sup>. De forma a utilizar valores ao invés de números indicadores de cada faixa, substituíram-se os números correspondentes pelos valores médios para cada uma das sete faixas de pagamento designadas pela pesquisa: até 30 reais, de 30 a 50, de 50 a 100, de 100 a 200, de 200 a 300, de 300 a 500, mais de 500. Por exemplo, a primeira faixa corresponde ao valor de 15 reais. No entanto, um problema de construir a variável desta forma é que não se sabe qual o limite superior: na faixa dos que pagam mais do que 500 reais. Podem existir, por exemplo, pessoas que pagam mensalidades superiores a 1000 reais pelo plano de saúde. Para contornar este problema, assume-se que o valor máximo pago como mensalidade é seis vezes o valor do plano mais básico, como definido pela lei 9656, de 3 de junho de 1998. Portanto, a última faixa tem como valor médio 630 reais.

Deve-se notar ainda que só existe um valor para aqueles que são clientes de algum plano, o que remete a um problema de seleção amostral. Além disso, dentre aqueles que reportaram serem clientes (de acordo com a pergunta se o indivíduo tem ou não plano de saúde), nem todos responderam à questão sobre a mensalidade. Dos 90.848 indivíduos que responderam à

---

<sup>25</sup>Foram testados modelos que utilizaram estas variáveis, porém houve perda significativa no poder explicativo em função da redução do tamanho da amostra.

<sup>26</sup>A mensalidade na PNAD é definida como uma variável *dummy*, que associa um número à faixa de mensalidade correspondente.

pergunta sobre o plano, 39.955 são titulares do plano, e somente 34.446 declaram qual era a mensalidade paga à época.

Portanto, é preciso saber quem são os potenciais “consumidores” de planos de saúde. Para tanto, utilizou-se o procedimento de Heckman para seleção<sup>27</sup> (*Heckit Model*). Este modelo, por sua vez, é composto de duas equações: a primeira, em que se determina se o indivíduo faz parte da seleção, isto é, se ele possui plano de saúde; e a segunda, uma equação linear condicionada à amostra dos respondentes. Na prática, a primeira equação é um modelo de escolha binária do tipo *Probit*, de onde se extrai a variável *Inverse Mills Ratio* (IMR), que é o inverso da razão entre a função de distribuição acumulada e a função de densidade da normal padrão. Em seguida, esta variável torna-se um regressor da segunda equação, a qual estima os efeitos de determinadas variáveis sobre a variável dependente. Este procedimento, realizado em dois passos, pode ser feito automaticamente em pacotes estatísticos adequados, usando o método de máxima verossimilhança. Genericamente:

Equação de interesse:  $y_i = x_i' \beta + \varepsilon_i$

Equação de seleção:  $z_i^* = w_i' \gamma + u_i$

Regra de seleção:  $y_i$  é observado quando  $z_i^* > 0$

Procedimento *Heckit*:

(i) estimar um modelo *Probit* com a equação x e obter a IMR  $\hat{\lambda}_i = \phi(w_i' \hat{\gamma}) / \Phi(w_i' \hat{\gamma})$ ;

(ii) estimar a regressão de y sobre x e  $\hat{\lambda}_i$ .

Para equação de seleção foram consideradas as variáveis: gênero, idade, renda domiciliar, anos de estudo, *dummy* se o indivíduo é chefe do domicílio, *dummies* de região, proporção de idosos no domicílio, proporção de crianças no domicílio e presença de doenças crônicas. Conforme esperado, o fato de um indivíduo ser mais velho, ter uma renda maior no domicílio<sup>28</sup>, ter parentes idosos no domicílio e ter alguma doença crônica contribui positivamente para adquirir um plano de saúde, em que o parente idoso poderá ser

<sup>27</sup> Ver GREENE (2003)

<sup>28</sup> A renda domiciliar foi escolhida em função do fato de que muitos indivíduos têm planos de saúde, mas quem efetivamente paga a mensalidade é o titular e/ou o chefe do domicílio.

dependente. As *dummies* de região indicaram que a probabilidade de ter plano de saúde é maior no sudeste (referência), devido aos coeficientes negativos, com menos intensidade para a região sul, a qual possui uma condição socioeconômica semelhante ao sudeste.

Para explicar o valor da mensalidade paga, na equação de interesse foram utilizadas variáveis: *dummies* que caracterizam os planos de saúde (existência de dependentes no plano, de cooperativa médica, de lista de credenciados, de reembolso, de cobertura em outra UF, de consultas, de exames, de internações, de medicamentos, de odontologia e de copagamento), além do gênero, idade, renda domiciliar, anos de estudo e *dummies* regionais. A variável dependente mensalidade foi utilizada em sua forma logarítmica para evitar a obtenção de valores estimados negativos. Os resultados da estimação são os seguintes:



**Tabela 21 - resultados do procedimento de Heckman para estimação do preço**

Número de obs =		366299				
Número de obs censuradas =		341727				
Número de obs não censuradas =		24572				
<i>lmenalplan</i>	Coef.	Desvio Padrão	z	P> z	Intervalo de confiança 95%	
homem	-0.055	0.013	-4.360	0	-0.079	-0.030
idade	0.015	0.001	25.090	0	0.014	0.016
renddom	0.000	0.000	29.660	0	0.000	0.000
anosestudo	0.012	0.004	3.280	0	0.005	0.019
norte	-0.154	0.023	-6.700	0	-0.199	-0.109
nordeste	0.068	0.015	4.590	0	0.039	0.097
Sul	-0.027	0.015	-1.830	0.067	-0.055	0.002
centroeste	-0.012	0.021	-0.550	0.581	-0.053	0.030
ddepend	0.220	0.011	19.280	0	0.198	0.243
dcooper	0.086	0.013	6.460	0	0.060	0.113
dlistacred	0.295	0.027	11.010	0	0.242	0.347
dreembolso	0.135	0.011	11.850	0	0.113	0.157
dcobertuf	0.319	0.012	26.840	0	0.295	0.342
dcobertcons	-0.124	0.061	-2.020	0.043	-0.244	-0.004
dcobertexam	0.130	0.042	3.120	0.002	0.048	0.211
dcobertint	0.623	0.025	25.360	0	0.575	0.671
dcobertmed	-0.101	0.020	-5.040	0	-0.140	-0.061
dcobertodonto	-0.281	0.012	-23.420	0	-0.305	-0.258
dcopagmto	-0.244	0.013	-18.920	0	-0.269	-0.219
_cons	2.528	0.131	19.270	0	2.271	2.785
<i>select</i>						
homem	0.047	0.009	5.430	0	0.030	0.064
idade	0.010	0.000	30.970	0	0.010	0.011
renddom	8.150E-05	0.000	54.930	0	0.000	0.000
anosestudo	0.121	0.001	126.380	0	0.119	0.123
chefedom	0.471	0.010	46.960	0	0.451	0.491
norte	-0.285	0.015	-19.600	0	-0.313	-0.256
nordeste	-0.189	0.010	-19.200	0	-0.208	-0.169
centroeste	-0.286	0.014	-20.730	0	-0.313	-0.259
Sul	-0.035	0.011	-3.350	0	-0.056	-0.015
pidososdom	0.067	0.019	3.450	0	0.029	0.105
pcriancadom	-0.089	0.021	-4.310	0	-0.129	-0.048
dcronica	0.134	0.009	15.080	0	0.117	0.152
_cons	-3.116	0.017	-182.790	0	-3.149	-3.082

Fonte: elaboração própria.

É possível notar que idade, anos de estudo, renda domiciliar, o fato de possuir dependente, cobertura nacional e reembolso pelo plano contribuem positivamente para a decisão de adquirir um plano de mensalidade mais elevada.

De acordo com esses resultados, a mensalidade paga pelo plano de saúde privado estimada é, em média, 119 reais<sup>29</sup>, com um valor máximo de 5500 reais. A princípio este valor máximo pode parecer elevado, mas analisando caso a caso as estimativas superiores a 2000 reais, verifica-se que se trata de apenas seis indivíduos com renda domiciliar muito acima da média (variando de 38775 a 56600 reais), e que são, em sua maioria, homens que são chefes de domicílio (cinco casos), com idade variando de 39 a 83 anos. Outro fato interessante é que em dois terços destes casos o indivíduo possui 16 anos de estudo.

Ao compararmos os dados de mensalidade estimada para os planos de saúde privados com os dados do banco de despesas individuais da Pesquisa de Orçamento Familiares de 2002/2003, nota-se uma pequena diferença, provavelmente em função de diferenças metodológicas no processo de amostragem. De acordo com a estimação da mensalidade a partir dos dados disponíveis na PNAD, os indivíduos pagam, em média, 119 reais por um plano de saúde. Entretanto, pela POF, a mensalidade média é da ordem de 98 reais para um período de referência semelhante.

Entretanto, algumas ressalvas devem ser feitas. Em primeiro lugar, na POF, o valor mínimo da despesa deflacionada e anualizada com planos de saúde é de doze reais por indivíduo, o que equivaleria a gastar apenas um real por mês. Ainda que existam instituições de caridade e/ou sem fins lucrativos que atuam na assistência à saúde da população, este valor é muito baixo se comparado ao valor mínimo estimado usando a PNAD (doze reais mensais), ainda mais considerando que a renda média familiar *per capita* é superior na POF. Nesse sentido, é possível perceber as diferenças de metodologia das duas pesquisas, que se reflete em valores distintos para a mensalidade. Caso tirássemos da amostra da POF as observações com valores inferiores ao mínimo estimado pela PNAD (doze reais), e da amostra da PNAD os valores superiores a 2030 reais (máximo observado na POF), chegaríamos às médias 113 e 118 para a POF e para a PNAD, respectivamente.

---

<sup>29</sup> Valor aproximado, calculado com base no exponencial do valor previsto.

Um outro ponto de ressalva nesta comparação é o fato de que o número de indivíduos que registraram terem gasto com plano de saúde na POF é aproximadamente um terço da amostra de respondentes na PNAD. Mesmo considerando os pesos amostrais, a diferença nos valores persiste, principalmente nas médias por região, conforme a tabela 22 seguinte. Ainda que haja diferenças nos valores, a hierarquia entre regiões é semelhante, com exceção da região Nordeste, que na POF aparece como sendo a maior média nacional. Ainda assim, nota-se que o Nordeste é superior ao Sul em ambas as fontes de dados.

**Tabela 22 - médias regionais - mensalidade paga a um plano de saúde privado**

<b>Região</b>	<b>PNAD (estimação)</b>	<b>POF</b>
<b>Norte</b>	92.8	69.0
<b>Nordeste</b>	124.8	113.1
<b>Sudeste</b>	129.2	100.9
<b>Sul</b>	108.7	82.0
<b>Centro-oeste</b>	110.8	94.5

Por último, deve-se notar que o questionário da PNAD admite que a assistência odontológica possa estar inclusa no plano de assistência médico-hospitalar, na medida em que existe a pergunta sobre o plano cobrir despesas com tratamento odontológico ou não (existe ainda outra pergunta sobre a existência de um plano odontológico em separado). Já na POF, as despesas com dentistas são registradas em outro item. Sendo assim, o valor mais elevado da PNAD pode estar viesado por incluir esse tipo de cobertura.

Feitas as devidas considerações sobre a mensalidade estimada, prosseguiu-se com o processo de construção da variável preço dos serviços de saúde com a extrapolação dos coeficientes gerados para o resto da população. Dado que somente para titulares de planos havia informações sobre o valor pago na mensalidade (embora nem todos os titulares, tendo em vista que em alguns casos outra pessoa do domicílio é responsável pelo pagamento), foram gerados valores previstos para as demais pessoas que afirmaram serem cobertas por algum plano, mas que atuam como dependentes. Para isto, foram utilizados os coeficientes estimados no segundo estágio mais a constante, exclusive os referentes às características do plano, uma vez que não há estas informações para quem não paga a mensalidade. Com isso, o objetivo é gerar um preço para todos os indivíduos que utilizam planos de saúde, ainda que este preço seja não um valor pago de fato, mas um preço sombra, semelhante ao procedimento realizado em Alves e Timmins (2003).

De maneira semelhante à geração de um preço para os dependentes do setor privado, os coeficientes obtidos no segundo estágio também foram usados para estimar um preço para aqueles que utilizam o sistema de saúde público (que não possuem plano de saúde privado). Uma vez que este preço também não é explícito, mas sim um preço sombra, adotou-se a seguinte hipótese: o preço público equivale ao preço do plano de saúde mais simples, que oferece poucas coberturas, como ocorre de fato no sistema público. Sendo assim, foi estimado um preço para cada indivíduo, de acordo com os parâmetros referentes a características socioeconômicas e com os coeficientes das seguintes *dummies*: cobertura em outra UF, cobertura de exames e consultas, cobertura de internações e a constante. Com isso, o preço público obteve uma média de 65 reais.

O último passo na construção de um preço para os serviços de saúde foi a agregação dos preços privados com os preços públicos estimados para cada indivíduo, gerando a variável preço, a qual foi posteriormente utilizada no modelo de demanda por consultas.

#### **4.1.1 Gastos com consultas médicas segundo a POF 2002/03**

Conforme mencionado anteriormente, no banco de dados da PNAD (2003) não é possível obter a despesa gasta com consultas médicas diretamente. Contudo, na POF 2002/2003 é possível obter não apenas a despesa com planos de saúde, mas também os gastos individuais com consultas médicas<sup>30</sup>. Embora não seja possível utilizar estes dados no modelo de estimação do número de consultas médicas realizadas em um ano, haja vista as diferenças metodológicas entre as duas pesquisas, é relevante analisar algumas estatísticas descritivas.

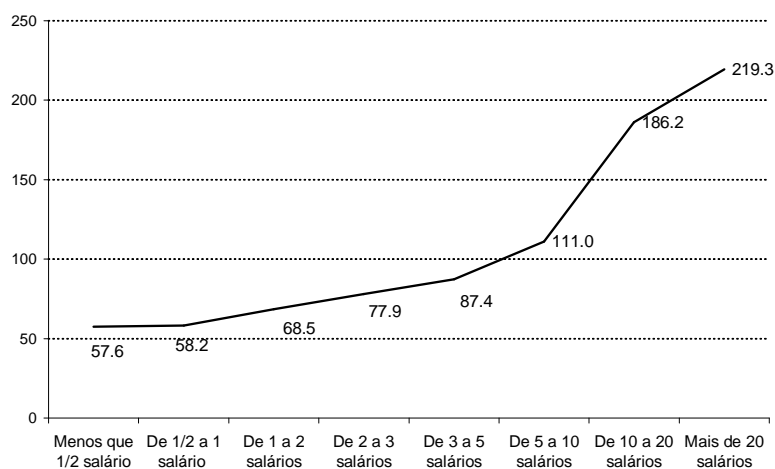
Conforme esperado, verifica-se que a renda possui correlação positiva em relação aos gastos com consultas médicas, principalmente para as classes de renda mais elevadas, como pode ser visto no gráfico seguinte, em que a curva torna-se mais inclinada entre as faixas de cinco a

---

<sup>30</sup> Nas modalidades odontologia, ginecologia, pediatria, cardiologia, oftalmologia, ortopedia, neurologia, psiquiatria, dermatologia, alergia, gastroenterologia, geral, geriatria, homeopatia, nefrologia, nutricionista, obstetrícia, oncologia, otorrino, pneumologia, reumatologia, endocrinologia, proctologia, urologia.

vingte salários. Um resultado semelhante foi encontrado no trabalho de Andrade *et al* (2006), em que foram estimados os determinantes dos gastos em saúde das famílias utilizando a PNAD 1998, através de um modelo Tobit. Neste estudo, verificou-se que os gastos com planos de saúde possuem caráter regressivo: a população mais pobre tende a alocar uma parcela maior de sua renda para os gastos com cuidados curativos, tais como consultas médicas e medicamentos, enquanto que entre os mais ricos, o principal componente dos gastos privados é a mensalidade dos planos de saúde os quais têm caráter preventivo. Nesse mesmo estudo, também foi demonstrado que a elasticidade renda é a mais elevada dentre todos os gastos em saúde (medicamentos, hospitalares, etc.), reforçando a tese de que a população de baixa renda tem dificuldades para adquirir um plano ou seguro-saúde. Ainda segundo Andrade *et al* (2006, p. 497): “O menor gasto observado entre as famílias de baixa renda parece refletir mais a falta de recursos financeiros para despende com a saúde do que um indício de que o setor público estaria atendendo a demanda desses indivíduos”, refletindo assim a desigualdade social na utilização dos serviços de saúde no Brasil.

**Figura 6 - relação entre renda *per capita* e valor gasto com consultas médicas**



Fonte: micro dados POF (2002/03)

Um outro dado interessante apontado pelo mesmo estudo é de que a escolaridade do chefe de família e a renda possuem correlação positiva, de maneira que famílias mais bem instruídas gastam mais com saúde. Nesse sentido, se considerarmos o aspecto regional, o Sudeste seria a região com maior nível de gastos, na medida em que a proporção das classes A e B é maior. Para os gastos com plano de saúde, os dados da POF não corroboram este resultado. Entretanto, no caso de gastos com consultas médicas fica comprovado que o Sudeste possui a

maior média, e o Nordeste a menor (pela POF, esta última região é a maior média de mensalidade do plano de saúde vide Tabela 22), conforme a tabela abaixo.

**Tabela 23 - média regional do valor pago em uma consulta segundo a POF 2002/03**

<b>Região</b>	<b>Média consulta (R\$)</b>
<b>Norte</b>	90.4
<b>Nordeste</b>	76.9
<b>Sudeste</b>	113.0
<b>Sul</b>	94.6
<b>Centro-oeste</b>	99.1

Fonte: micro dados POF (2002/03)

Finalmente, com relação à idade também é possível perceber na POF que os idosos gastam mais com consultas médicas (média de R\$ 108,1), seguido das crianças de 10 a 14 anos (média 107,9), da mesma forma que os planos de saúde são mais caros para estas faixas etárias, conforme verificado pelos coeficientes da estimação supracitada.

#### **4.2 Estimação do uso de serviços de saúde: consultas**

Na PNAD Saúde (2003) existem algumas variáveis que são números inteiros não negativos, as quais poderiam ser utilizadas como variável dependente de um modelo de contagem: o número de consultas realizadas no ano, o número de atendimentos realizados nas duas semanas anteriores à semana de referência, o número de vezes em que o indivíduo esteve internado no ano e o tempo de internação (número de meses e dias). Neste trabalho utilizou-se como variável dependente o número de consultas no ano, uma vez que esta é a mais abrangente, tanto em termos de horizonte temporal como em termos de frequência (indivíduos fazem mais consultas do que internações). O conjunto de variáveis explicativas é descrito na tabela a seguir.

**Tabela 24 - descrição das variáveis explicativas utilizadas**

<b>Variáveis</b>	<b>Descrição</b>	<b>Média</b>	<b>Desvio Padrão</b>	<b>Min</b>	<b>Max</b>
lpreco_est	logaritmo do preço dos serviços de saúde	4.00	0.55	2.33167	8.61416
lrenddom	renda domiciliar	6.56	1.02	0	10.9438
metro	se reside em região metropolitana=1	0.37	0.48	0	1
urbano	se reside em área urbana=1	0.85	0.36	0	1
homem	se é do sexo masculino=1	0.49	0.50	0	1
casado	se é casado=1	0.39	0.49	0	1
idade01	se tiver de 0 a 4 anos=1	0.09	0.28	0	1
idade02	se tiver de 5 a 9 anos=1	0.10	0.30	0	1
idade04	se tiver de 15 a 19 anos=1	0.10	0.30	0	1
idade05	se tiver de 20 a 24 anos=1	0.10	0.29	0	1
idade06	se tiver de 25 a 39 anos=1	0.22	0.42	0	1
idade07	se tiver de 40 a 59 anos=1	0.20	0.40	0	1
idoso	se for idoso=1	0.09	0.29	0	1
preta	se for de cor preta=1	0.06	0.24	0	1
amarela	se for de cor amarela=1	0.00	0.06	0	1
parda	se for de cor parda=1	0.46	0.50	0	1
indio	se for indio=1	0.00	0.05	0	1
norte	se reside na região norte=1	0.11	0.32	0	1
nordeste	se reside na região nordeste=1	0.33	0.47	0	1
centroeste	se reside na região centroeste=1	0.11	0.31	0	1
sul	se reside na região sul=1	0.15	0.36	0	1
anosestudo	número de anos de estudo	6.16	4.45	1	16
autob	se considera a saúde boa=1	0.55	0.50	0	1
autoreg	se considera a saúde regular=1	0.18	0.39	0	1
autoruim	se considera a saúde ruim=1	0.03	0.17	0	1
automruim	se considera a saúde muito ruim=1	0.01	0.08	0	1
dinativ	se possui alguma limitação de atividade=1	0.07	0.26	0	1
drcolumna	se possui problema de coluna=1	0.13	0.34	0	1
drcartrite	se possui artrite=1	0.06	0.24	0	1
drcancer	se possui cancer=1	0.00	0.06	0	1
dcrdiabetes	se possui diabetes=1	0.02	0.15	0	1
dcrhipert	se possui hipertensão=1	0.12	0.32	0	1
drcora	se possui problemas de coração=1	0.04	0.18	0	1
dcrrenal	se possui doença renal=1	0.02	0.13	0	1
dcrdepres	se possui depressão=1	0.04	0.19	0	1
dcr tuberc	se possui tuberculose=1	0.00	0.04	0	1
dcr tend	se possui tendinite=1	0.02	0.14	0	1
dcirr	se possui cirrose hepática=1	0.00	0.03	0	1
dcama	se esteve acamado recentemente=1	0.04	0.20	0	1

#### 4.2.1 Modelos Poisson e Binomial Negativo

Os modelos Poisson e Binomial Negativo mostraram-se semelhantes em termos dos coeficientes estimados. Em geral, não houve diferenças de sinal, com exceção da variável idade04, que corresponde à faixa etária dos 15 aos 19 anos e que, conforme detalhado no capítulo três, possui a maior proporção de zeros na variável dependente número de consultas realizadas no ano em comparação às demais faixas. No modelo Binomial Negativo, ao contrário do modelo Poisson, a variável idade04 é significativa, embora possua um coeficiente negativo e baixo em valor absoluto além de um efeito marginal também reduzido. As outras diferenças entre os dois modelos ocorreram principalmente no sentido de reduzir os efeitos marginais das variáveis de idade e de aumentar estes efeitos das variáveis indicadoras de doenças crônicas no modelo Binomial Negativo. A variável de raça índio, também se tornou mais intensa neste último modelo.

Em ambos os modelos o preço dos serviços de saúde, em sua forma logarítmica<sup>31</sup>, tem um efeito negativo sobre a quantidade de consultas, conforme esperado. A renda domiciliar<sup>32</sup>, ao contrário, contribui positivamente, já que muitas vezes um membro do domicílio auxilia outro, tanto para o pagamento da mensalidade do plano, quanto no caso de ter que pagar pela consulta em si. Além disso, a renda pode estar associada à escolaridade, que teria um efeito positivo no sentido de um maior investimento em saúde, conforme Grossman (1972).

As *dummies* indicativas de região e área censitária mostraram-se significativas e com efeito positivo: pessoas das áreas urbanas e metropolitanas vão mais ao médico. Um dos motivos, além da renda e escolaridade mais alta, é que nestas áreas a oferta de médicos, hospitais e clínicas especializadas é maior, fato que infelizmente não pode ser captado diretamente por ausência de informações sobre oferta na PNAD Saúde (2003).

O gênero é uma variável importante na determinação das visitas ao médico: os homens tendem a ir menos ao médico do que as mulheres, talvez por um motivo fisiológico ou ainda porque não se preocuparem com cuidados preventivos da mesma forma que as mulheres.

---

<sup>31</sup>O preço foi colocado em logaritmo natural, uma vez que o procedimento de Heckman gerou valores previstos também sob esta forma.

<sup>32</sup>A renda foi colocada em logaritmo natural para amenizar a alta variância desta variável em nível.



Contudo, o fato de serem casados influencia positivamente na quantidade de consultas realizadas pelos homens.

Com relação à idade, tendo como base de comparação a faixa dos 10 aos 14 anos (idade03), verifica-se que crianças e idosos tem maior necessidade de freqüentar médicos, resultado naturalmente esperado. Da idade05 (20 a 24 anos) em diante, os efeitos são positivos e crescentes.

As variáveis de raça, com relação à raça branca, obtiveram coeficientes negativos, com exceção da variável índio. No modelo Binomial Negativo esta se mostrou bastante intensa, em contraposição às expectativas. O motivo para isto provavelmente tem relação com a amostra. As tabulações cruzadas da variável índio com o número de consultas no ano mostraram que de 814 observações de indivíduos indígenas, 88 fizeram 12 consultas ao ano (uma vez ao mês), o que equivale a 11%. Enquanto isso, para os 17.8492 brancos, somente 4.010 (2%) foram ao médico uma vez ao mês em média. Além disso, a média de consultas realizadas no ano é 3,44 para índios e 2,7 para brancos.

As *dummies* regionais indicaram que os indivíduos da região Sudeste freqüentam mais o médico, seguidos das regiões Nordeste, Sul, Centro-oeste e Norte. A renda e a escolaridade mais elevada podem ser os motivos pelos quais as pessoas do Sudeste investem mais em saúde.

A escolaridade, apesar de significativa, não exerce tanta influência direta na utilização de serviços de saúde. Pode-se levantar a questão de que o maior nível educação atua de forma positiva para que o indivíduo resolva se cuidar e manter rotinas preventivas periódicas, devido a um maior esclarecimento sobre a importância destes cuidados, influenciando ainda a decisão de adquirir um seguro-saúde. No entanto, a parcela dos que investe em cuidados preventivos ainda é baixa: enquanto 51% procuraram atendimento de saúde por motivo de doença, somente 26% o fizeram por cuidados preventivos. Esta diferença diminui para aqueles que possuem plano de saúde: 33% procuraram atendimento devido a cuidados preventivos, e 43% devido a doenças. Nesse sentido, conclui-se que ter plano de saúde ou estimula o maior número de exames com fins de checagem do estado de saúde e/ou as pessoas que são mais “cuidadasas” e mais ricas - o que pode ser resultado de mais educação - se auto selecionam para adquirir um plano.

Com relação às variáveis de auto-avaliação da saúde, os resultados mostram-se dentro do esperado – indivíduos que se consideram menos saudáveis vão mais ao médico. Os efeitos, em comparação aos que consideram sua saúde como muito boa são positivos e crescentes à medida que se torna mais pessimista com relação à saúde.

As *dummies* indicadoras de doenças crônicas e a *dummy* que indica que o indivíduo possui alguma limitação em suas atividades, se mostraram bastante significantes e positivas com destaque para a *dummy* de câncer. Como mencionado anteriormente, estas possuem um efeito maior no modelo Binomial Negativo. Por último, a *dummy* indicativa de que o indivíduo esteve acamado nas duas últimas semanas se mostrou significativa, cujo efeito é de elevar o número de consultas.

Os sinais e magnitudes dos coeficientes foram explicitados acima, contudo, em se tratando de modelos não lineares, a simples constatação do valor dos coeficientes não é o bastante para interpretá-los de maneira correta. Para isso, há outros critérios mais relevantes, que medem a contribuição de cada variável explicativa para o número de consultas. Segundo Long & Freese (2001), o efeito marginal é uma das formas de medir esta contribuição, em que se calcula o efeito de cada variável em separado, mantidas as demais em seus valores médios.

Uma outra maneira, citada pelos autores, é o cálculo da mudança discreta no número esperado da variável dependente dado que a variável explicativa mudou de  $x_E$  para  $x_S$ <sup>33</sup>. Esta mudança, por sua vez, pode ser computada de algumas formas, como a variação desde o valor mínimo até o valor máximo da variável ou ainda, uma mudança centrada, isto é, verificar o efeito desde a média da variável mais e/ou menos o seu desvio-padrão.

Por último, os autores também propõem uma análise das probabilidades estimadas, ou seja, verificar a proporção observada de um determinado evento (por exemplo, número de consultas igual a zero) e comparar com a probabilidade estimada pelo modelo para este evento. Se ambas estiverem próximas significa que o modelo gerou um bom ajuste.

---

<sup>33</sup>  $\frac{\Delta E(y/x)}{\Delta x_k} = E(y/x, x_k = x_E) - E(y/x, x_k = x_S)$

As tabelas a seguir mostram o coeficiente estimado, a estatística z, a probabilidade associada (p-valor), o efeito discreto variando do mínimo até o máximo e o efeito marginal para cada variável nos modelos Poisson e Binomial Negativo.

**Tabela 25 - resultados do modelo de regressão Poisson**

<i>nconsult</i>	<i>Coef.</i>	<i>z</i>	<i>P&gt; z </i>	<i>min-&gt;max</i>	<i>Efeito marginal</i>
lpreco_est	-0.193	-81.75	0	-1.97	-0.39
lrenddom	0.065	52.67	0	0.36	0.36
metro	0.175	77.19	0	1.37	0.13
urbano	0.204	55.77	0	0.39	0.42
homem	-0.369	-163.88	0	-0.75	-0.75
casado	0.163	59.19	0	0.34	0.33
Idade01	0.950	166.71	0	2.96	1.93
Idade02	0.417	68.64	0	1.01	0.85
Idade04	0.009	1.37	0.171	0.02	0.02
Idade05	0.154	24.46	0	0.33	0.31
Idade06	0.185	31.7	0	0.40	0.38
Idade07	0.223	36.77	0	0.49	0.45
Idoso	0.352	53.37	0	0.83	0.72
Preta	-0.032	-7.05	0	-0.06	-0.06
amarela	-0.138	-7.65	0	-0.26	-0.28
Parda	-0.061	-24.55	0	-0.12	-0.12
Índio	0.237	12.18	0	0.54	0.48
Norte	-0.273	-66.91	0	-0.50	-0.55
nordeste	-0.094	-33.02	0	-0.19	-0.19
centroeste	-0.196	-49.89	0	-0.37	-0.40
Sul	-0.097	-29.9	0	-0.19	-0.20
anosestudo	0.021	60.73	0	0.66	0.04
Autob	0.172	54.92	0	0.35	0.35
autoreg	0.654	174.7	0	1.66	1.33
autoruim	0.963	174.38	0	3.20	1.96
automruim	1.012	115.92	0	3.54	2.06
dinativ	0.424	102.64	0	1.04	0.86
drcroluna	0.170	56.58	0	0.37	0.35
dcrartrite	0.055	14.46	0	0.11	0.11
dercancer	0.499	54.53	0	1.31	1.02
dcrdiabetes	0.258	55	0	0.60	0.53
dcrhipert	0.303	92.94	0	0.69	0.62
drcora	0.187	44.4	0	0.41	0.38
dcrrenal	0.102	16.98	0	0.22	0.21
dcrdepres	0.245	61.76	0	0.56	0.50
dcr tuberc	0.262	14.53	0	0.61	0.53
dcr tend	0.144	26.95	0	0.31	0.29
Dcirr	0.180	8.21	0	0.40	0.37
dcama	0.086	17.45	0	0.18	0.17
_cons	0.304	21.74	0		
N obs =	361515				
Log verossim. =	-931416				

Tabela 26 - resultados do modelo de regressão Binomial Negativo

<i>nconsult</i>	<i>Coef.</i>	<i>z</i>	<i>P&gt; z </i>	<i>min-&gt;max</i>	<i>Efeito marginal</i>
lpreco_est	-0.220	-44.71	0	-2.15	-0.44
lrenddom	0.064	26.34	0	0.37	0.36
metro	0.182	38.72	0	1.33	0.13
urbano	0.229	33.51	0	0.42	0.46
homem	-0.442	-99.89	0	-0.88	-0.88
casado	0.206	36.49	0	0.42	0.41
idade01	1.005	95.66	0	3.14	1.99
idade02	0.440	41.97	0	1.05	0.87
idade04	-0.044	-4.11	0	-0.09	-0.09
idade05	0.082	7.38	0	0.17	0.16
idade06	0.108	10.54	0	0.22	0.22
idade07	0.142	12.88	0	0.29	0.28
idoso	0.343	26.85	0	0.79	0.68
preta	-0.026	-2.78	0.005	-0.05	-0.05
amarela	-0.119	-3.25	0.001	-0.22	-0.24
parda	-0.061	-12.08	0	-0.12	-0.12
indio	0.643	14.08	0	1.79	1.28
norte	-0.258	-32.25	0	-0.46	-0.51
nordeste	-0.081	-13.91	0	-0.16	-0.16
centroeste	-0.176	-22.42	0	-0.33	-0.35
sul	-0.108	-15.76	0	-0.21	-0.21
anosestudo	0.025	35.73	0	0.80	0.05
autob	0.180	31.8	0	0.35	0.36
autoreg	0.672	89.93	0	1.68	1.33
autoruim	1.017	74.66	0	3.40	2.02
automruim	1.033	38.95	0	3.57	2.05
dinativ	0.495	47.8	0	1.23	0.98
drcolumna	0.217	31.92	0	0.47	0.43
dcrartrite	0.091	9.71	0	0.19	0.18
drcancer	0.593	19.17	0	1.60	1.18
dcrdiabetes	0.345	25.97	0	0.81	0.68
dcrhipert	0.354	47.18	0	0.81	0.70
drcora	0.247	21.62	0	0.55	0.49
dcrrenal	0.133	8.23	0	0.28	0.26
dcrdepres	0.325	30.37	0	0.75	0.65
dcrtuberc	0.396	7.76	0	0.96	0.79
dcrtrend	0.220	15.46	0	0.49	0.44
dcirr	0.186	3.08	0.002	0.41	0.37
dcama	0.089	6.8	0	0.18	0.18
_cons	0.354	12.96	0		
/lnalpha	0.142				
alpha	1.152				
N obs =	361515				
Log verossim. =	-696629				

As informações sobre magnitude e sinal dos parâmetros, assim como os efeitos marginal e discreto contribuem, mas não permitem uma comparação mais apurada entre os modelos

Poisson e Binomial Negativo, uma vez que os resultados são semelhantes. Conforme visto no capítulo anterior, a grande concentração de zeros dificulta o uso de modelos que não levam em consideração componentes latentes, como é o caso do modelo de regressão Poisson. A análise das probabilidades estimadas, conforme Long e Freese (2001), os critérios de informação<sup>34</sup> e o teste de sobredispersão<sup>35</sup> constituem instrumentos úteis para a escolha do melhor modelo, tendo em vista as características dos dados – muitos zeros e alta variância no número de consultas.

As probabilidades estimadas<sup>36</sup> são favoráveis à escolha do modelo Binomial Negativo. Conforme a tabela e os gráficos abaixo, o modelo Poisson não consegue prever bem o evento “zero consulta”.

**Tabela 27 - probabilidades estimadas pelos modelos Poisson e Binomial Negativo**

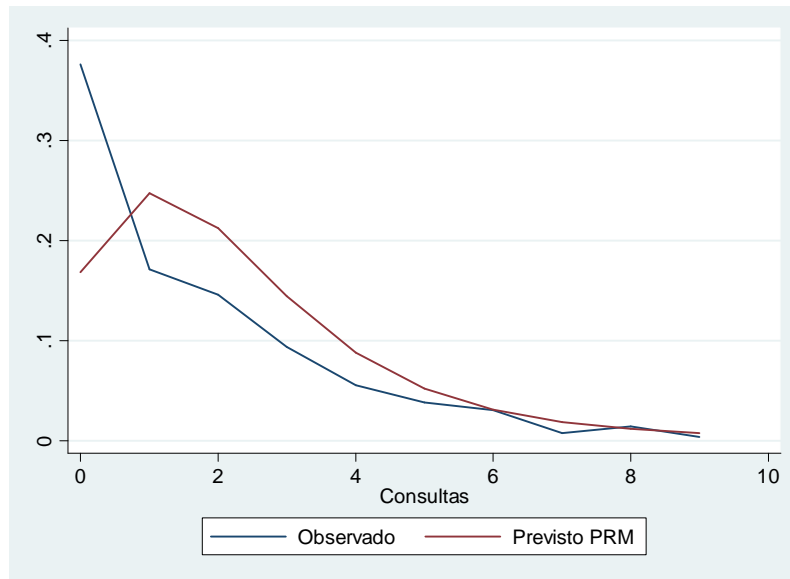
<i>Evento</i>	<i>Observado</i>	<i>Previsto Binomial Negativo</i>	<i>Previsto Poisson</i>
0	0.376	0.369	0.168
1	0.171	0.205	0.247
2	0.146	0.126	0.212
3	0.094	0.082	0.144
4	0.055	0.055	0.088
5	0.038	0.038	0.052
6	0.030	0.027	0.031
7	0.007	0.020	0.019
8	0.014	0.015	0.012
9	0.004	0.011	0.008

<sup>34</sup> Akaike information criterion e Bayesian information criterion. Os critérios de informação são úteis na comparação entre modelos não aninhados. Quanto mais baixos mais preferíveis serão os modelos.

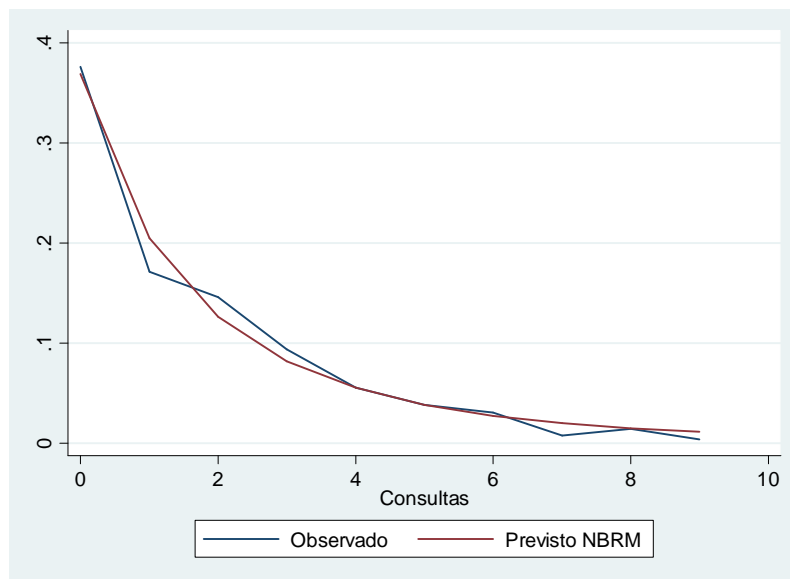
<sup>35</sup> Teste de razão de verossimilhança, cuja estatística é igual a  $LR_{estat} = 2 * (LR_{NBRM} - LR_{PRM})$

<sup>36</sup> Por motivos práticos, só foram geradas probabilidades até N = 9, equivalente a nove consultas realizadas no ano.

**Figura 7 - ajuste gráfico das probabilidades estimadas pelo modelo Poisson**



**Figura 8 - ajuste gráfico das probabilidades estimadas pelo modelo Binomial Negativo**



O teste LR confirmou a existência de sobredispersão nos dados, rejeitando a hipótese de que  $\alpha^{37}$  é igual a zero e favorecendo, novamente, o modelo Binomial Negativo.

---

<sup>37</sup> Vide equação 4.4.

**Tabela 28 - teste de sobredispersão Poisson e Binomial Negativo**

ln(alpha)	0.142
Alpha	1.152
chibar2(01) =	469574
teste LR Prob>0 =	0
Log verossim. PRM =	-931416
Log verossim. NBRM =	-696629

Também os critérios de informação deram preferência ao modelo de regressão Binomial Negativo, conforme visto na tabela seguinte.

**Tabela 29 - critérios de informação Poisson e Binomial Negativo**

	AIC	BIC
PRM	1862911	1863343
NBRM	1393339	1393782

#### 4.2.2 Modelos *Hurdle*

As estimações dos modelos *Hurdle* Poisson (HP) e *Hurdle* Binomial Negativo (HNB) geraram coeficientes semelhantes em termos do sentido (negativo ou positivo) do efeito de cada variável. Na estimação do modelo com a distribuição Binomial Negativa, as variáveis de raça preta e amarela foram significantes, e o coeficiente da variável índio foi alto, semelhantemente ao modelo Binomial Negativo. A tabela seguinte mostra os coeficientes, estatísticas *Z* e *p*-valor de cada variável para cada modelo. Os resultados do modelo *Logit*, primeiro estágio do *Hurdle* são os mesmos para ambos.





Assim como para os modelos Poisson e Binomial Negativo, o teste de sobredispersão e os critérios de informação<sup>38</sup> foram favoráveis à escolha do modelo *Hurdle* Binomial Negativo, conforme as tabela seguintes:

**Tabela 31 - teste de sobredispersão *Hurdle* Poisson e *Hurdle***

ln(alpha)	0.1785925
Chibar2(01) =	279808.26
teste LR Prob>0 =	0
Log verossim. HP =	-829056
Log verossim.HNB =	-689152

**Tabela 32 - critérios de informação modelos *Hurdle***

	AIC	BIC
HP	1658272	1659135
HNB	1378465	1379340

### 4.2.3 Modelos *Finite Mixture*

#### 4.2.3.1 Modelos *Zero-inflated*

A estimação dos modelos *Zero-inflated*, cuja idéia principal é estimar a probabilidade de se obter observações iguais à zero, gerou conclusões semelhantes às considerações feitas anteriormente sobre a escolha entre os modelos Poisson e Binomial Negativo. Não houve qualquer mudança no sinal dos parâmetros e na significância dos mesmos entre os modelos *Zero-Inflated* Poisson e *Zero-Inflated* Binomial Negativo. Contudo, no modelo *Zero-Inflated* Binomial Negativo, os valores absolutos dos coeficientes, em geral, aumentaram com relação ao modelo *Zero-Inflated* Poisson, provavelmente em função das diferenças ocorridas na estimação do “Alyaws-zero group” e da própria distribuição de probabilidade utilizada. No

---

<sup>38</sup> As probabilidades estimadas não foram calculadas para o *Hurdle* devido à inexistência de programas para este modelo no pacote SPost para STATA (desenvolvido pela Universidade de Indiana, EUA).

modelo *Zero-Inflated* Poisson somente se mostraram insignificantes a 10% a *dummy* indicadora de cirrose, a *dummy* indicadora de região metropolitana e a *dummy* da raça indígena. Já no modelo *Zero-Inflated* Binomial Negativo, que utiliza a distribuição binomial negativa, e que como tal tem maior aderência no caso de haver sobredispersão nos dados, as variáveis não significantes a 10% na estimação do grupo que não realiza consultas foram: idade01 (faixa de 0 a 4 anos), idoso, *dummies* das raças preta e amarela, *dummies* indicadoras de doenças do coração, câncer e cirrose e a *dummy* indicadora de que o indivíduo esteve acamado recentemente. Com efeito, a segunda opção (modelo *Zero-Inflated* Binomial Negativo) possui fundamentação consistente, no sentido que idosos e crianças recém nascidas são notadamente grupos em que o número de consultas é alto, sendo pouco provável que estes estejam no grupo dos que nunca realiza consultas médicas. A tabela abaixo mostra as diferenças nas estimativas do grupo que não se consulta<sup>39</sup>:

---

<sup>39</sup>A introdução da renda domiciliar no 1º. estágio de estimação dos modelos *Zero-inflated* inviabilizou a convergência dos mesmos. Portanto, a variável não entrou no cálculo da probabilidade de nunca ir ao médico. Tendo em vista que o SUS é uma alternativa sempre disponível, e que este sistema não discrimina pessoas por renda, a retirada da variável não parece trazer prejuízo às estimativas.

Tabela 33 - resultados da estimação do “Always-zero group”

Modelo	ZIP			ZINB		
	Coef.	z	P> z	Coef.	z	P> z
lpreco_est	0.469	44.8	0	0.621	23.15	0
metro	-0.014	-1.59	0.113	0.171	7.67	0
urbano	-0.309	-25.03	0	-0.367	-12.61	0
homem	0.655	78.78	0	1.406	46.52	0
casado	-0.302	-27.3	0	-0.418	-16.63	0
idade01	-1.374	-66.39	0	-17.602	-0.03	0.978
idade02	-0.618	-32.69	0	-1.186	-17.48	0
idade04	0.363	19.76	0	1.054	21.94	0
idade05	0.293	15.09	0	1.062	20.88	0
idade06	0.107	5.77	0	0.740	14.48	0
idade07	-0.042	-2.08	0.038	0.510	9.12	0
idoso	-0.427	-16.91	0	-0.077	-1.05	0.294
preta	0.074	4.2	0	0.051	1.22	0.224
amarela	0.187	2.66	0.008	0.106	0.55	0.581
parda	0.075	7.85	0	0.072	3.02	0.003
índio	-0.030	-0.35	0.73	0.298	1.71	0.087
norte	0.147	9.51	0	-0.111	-2.69	0.007
nordeste	0.059	5.39	0	-0.043	-1.65	0.099
centroeste	0.051	3.36	0.001	-0.213	-5.12	0
sul	0.139	10.61	0	0.095	2.94	0.003
anosestudo	-0.072	-55.14	0	-0.123	-40.77	0
autob	-0.192	-19.06	0	-0.231	-9.72	0
autoreg	-0.678	-46.07	0	-0.976	-23.72	0
autoruim	-0.897	-26.52	0	-1.422	-11.18	0
automruim	-0.756	-10.1	0	-0.757	-3.23	0.001
dinativ	-1.122	-37.84	0	-35.758	0	1
drcoluna	-0.490	-32.05	0	-1.195	-18.96	0
dcrartrite	-0.298	-12.73	0	-1.068	-8.07	0
drcancer	-0.948	-8.9	0	-18.546	0	0.999
dcrdiabetes	-0.774	-19.26	0	-13.357	0	0.997
dcrhipert	-0.653	-37.14	0	-1.959	-15.05	0
drcora	-0.567	-17.15	0	-4.414	-1.35	0.177
dcrrenal	-0.208	-5.2	0	-0.784	-4.48	0
dcrdepres	-0.631	-20.62	0	-2.583	-5.82	0
dcr tuberc	-0.260	-2.18	0.03	-0.899	-1.9	0.058
dcr tend	-0.542	-13.91	0	-1.955	-5.72	0
dcirr	-0.028	-0.2	0.838	-0.598	-1.24	0.215
dcama	-0.375	-9.45	0	-21.871	0	0.999
_cons	-1.544	-35.66	0	-3.636	-33.15	0

Com relação aos resultados da estimação do número de consultas em si, verifica-se que no modelo *Zero-Inflated Binomial Negativo*, a renda domiciliar e o preço possuem menor importância (em termos de efeito marginal) como fator positivo e negativo, respectivamente. Outro fato interessante, como ocorrido entre os modelos Binomial Negativo e Poisson, é que a variável índio torna-se mais intensa na estimação *Zero-Inflated Binomial Negativo*. Além

disso, as variáveis de auto-avaliação da saúde perdem importância no modelo que utiliza a distribuição binomial negativa, assim como as *dummies* de doenças crônicas. A educação também teve seu efeito marginal reduzido. Já as *dummies* de idade a partir de 20 anos (idade05 em diante) ganharam intensidade no modelo *Zero-Inflated* Binomial Negativo, em relação ao modelo *Zero-Inflated* Poisson.

Por último, é preciso salientar que, apesar das alterações ocorridas entre estimar um modelo de regressão Poisson ou Binomial Negativo e estimar modelos *Zero-inflated*, o sentido do efeito de cada variável continuou constante: os sinais dos coeficientes e os efeitos marginais se mantiveram, com exceção da variável idade04. Portanto, as mesmas motivações teóricas sobre o sentido do efeito de cada variável, citadas na seção anterior, valem no caso dos modelos *Zero-inflated*.

As tabelas seguintes mostram os valores dos coeficientes, os efeitos discretos (mudança do mínimo para o máximo de cada variável) e efeitos marginais de cada variável nos modelos *Zero-Inflated* Poisson e *Zero-Inflated* Binomial Negativo.

**Tabela 34 - resultados do modelo *Zero-inflated Poisson***

<i>nconsult</i>	<i>Coef.</i>	<i>z</i>	<i>P&gt; z </i>	<i>min-&gt;max</i>	<i>Efeito Marginal</i>
lpreco_est	-0.109	-45.16	0	-2.74	-0.55
lrenddom	0.036	28.75	0	0.84	0.08
metro	0.173	72.66	0	0.40	0.40
urbano	0.119	30.14	0	0.44	0.44
homem	-0.189	-78.58	0	-0.85	-0.85
casado	0.091	31.79	0	0.40	0.40
idade01	0.577	91.32	0	2.73	2.73
idade02	0.224	32.82	0	0.98	0.98
idade04	0.105	14.68	0	-0.04	-0.04
idade05	0.207	29.32	0	0.25	0.25
idade06	0.175	26.77	0	0.32	0.32
idade07	0.178	26.38	0	0.44	0.44
idoso	0.234	32.36	0	0.88	0.88
preta	-0.020	-4.11	0	-0.09	-0.09
amarela	-0.078	-4.09	0	-0.28	-0.28
parda	-0.045	-16.92	0	-0.15	-0.15
indio	0.261	13.06	0	0.67	0.67
norte	-0.236	-54.09	0	-0.55	-0.55
nordeste	-0.077	-25.87	0	-0.20	-0.20
centroeste	-0.177	-42.16	0	-0.39	-0.39
sul	-0.068	-19.98	0	-0.23	-0.23
anosestudo	0.006	17.83	0	0.89	0.06
autob	0.101	29.53	0	0.35	0.35
autoreg	0.439	109.74	0	1.65	1.65
autoruim	0.727	127.71	0	3.25	3.25
automruim	0.804	90.86	0	3.61	3.61
dinativ	0.275	66.57	0	1.42	1.42
drcolumna	0.088	28.7	0	0.52	0.52
dcrartrite	0.038	9.88	0	0.28	0.28
drcancer	0.451	49.17	0	2.03	2.03
dcrdiabetes	0.210	44.56	0	1.04	1.04
dcrhipert	0.191	57.86	0	0.90	0.90
drcora	0.160	37.98	0	0.77	0.77
dcrrenal	0.086	14.06	0	0.34	0.34
dcrdepres	0.208	52.32	0	0.95	0.95
dcr tuberc	0.233	12.76	0	0.78	0.78
dcr tend	0.114	21.1	0	0.62	0.62
dcirr	0.154	6.94	0	0.38	0.38
dcama	0.069	13.95	0	0.40	0.40
_cons	0.850	58.68	0		

**Tabela 35 - resultados do modelo *Zero-inflated Binomial Negativo***

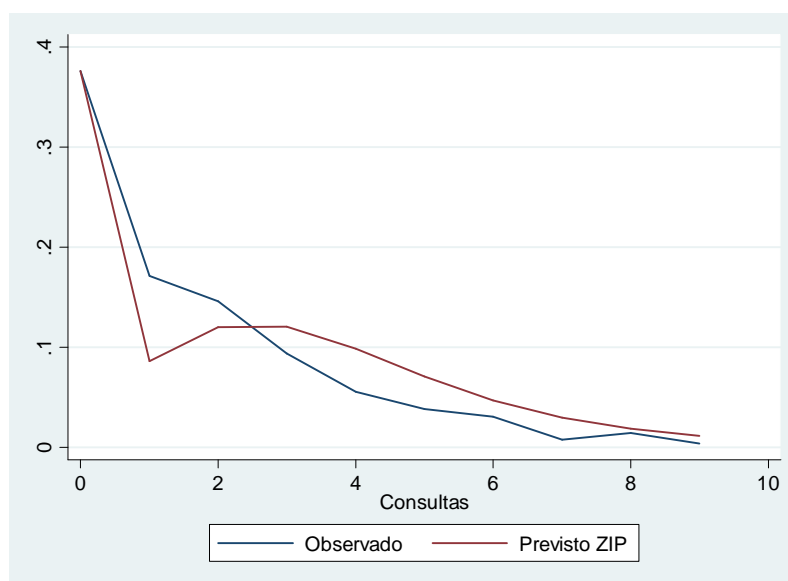
<i>nconsult</i>	<i>Coef.</i>	<i>z</i>	<i>P&gt; z </i>	<i>min-&gt;max</i>	<i>Efeito marginal</i>
lpreco_est	-0.173	-35.8	0	-2.09	-0.41
lrenddom	0.061	25.6	0	1.49	0.14
metro	0.200	41.34	0	0.48	0.48
urbano	0.177	24	0	0.39	0.39
homem	-0.259	-52.85	0	-0.61	-0.61
casado	0.140	23.94	0	0.34	0.33
idade01	0.762	66.28	0	2.51	2.51
idade02	0.293	24.67	0	0.78	0.78
idade04	0.153	12.2	0	0.38	0.38
idade05	0.279	21.59	0	0.73	0.73
idade06	0.252	21.04	0	0.63	0.63
idade07	0.250	19.99	0	0.63	0.63
idoso	0.375	27.04	0	1.03	1.03
preta	-0.022	-2.26	0.024	-0.05	-0.05
amarela	-0.114	-3.01	0.003	-0.25	-0.25
parda	-0.053	-10.15	0	-0.12	-0.12
indio	0.612	12.87	0	1.98	1.98
norte	-0.274	-33.27	0	-0.58	-0.58
nordeste	-0.082	-13.55	0	-0.19	-0.19
centroeste	-0.203	-25.01	0	-0.44	-0.44
sul	-0.102	-14.46	0	-0.23	-0.23
anosestudo	0.010	13.28	0	0.35	0.02
autob	0.138	22.42	0	0.32	0.32
autoreg	0.539	69.38	0	1.52	1.52
autoruim	0.859	65.66	0	3.12	3.12
automruim	0.911	37.14	0	3.47	3.47
dinativ	0.404	42.85	0	1.15	1.15
drcolumna	0.135	20.3	0	0.33	0.33
dcrartrite	0.075	8.48	0	0.18	0.18
drcancer	0.547	19.92	0	1.71	1.71
dcrdiabetes	0.295	24.74	0	0.80	0.80
dcrhipert	0.257	35.07	0	0.67	0.67
drcora	0.216	20.55	0	0.56	0.56
dcrrenal	0.101	6.67	0	0.25	0.25
dcrdepres	0.284	28.84	0	0.76	0.76
dcr tuberc	0.307	6.37	0	0.84	0.84
dcr tend	0.178	13.56	0	0.46	0.46
dcirr	0.128	2.26	0.024	0.32	0.32
dcama	0.087	7.44	0	0.22	0.22
_cons	0.456	16.76	0		
/lnalpha	-0.115	-22.31	0		
alpha	0.892				

Como realizado para os modelos Poisson e Binomial Negativo, foram comparadas as probabilidades estimadas por cada modelo para cada evento (até nove consultas), os critérios de informação e o teste de sobredispersão. As probabilidades estimadas novamente são favoráveis ao modelo que utiliza a distribuição binomial negativa (ZINB), como pode ser visto na tabela e nos gráficos abaixo. Na comparação entre os modelos Binomial Negativo e Poisson, o fator crítico era a previsão dos zeros. Já nos modelos *Zero-inflated*, os quais tratam os zeros adequadamente, o que passa a ser problemático é a previsão do evento uma consulta realizada no ano.

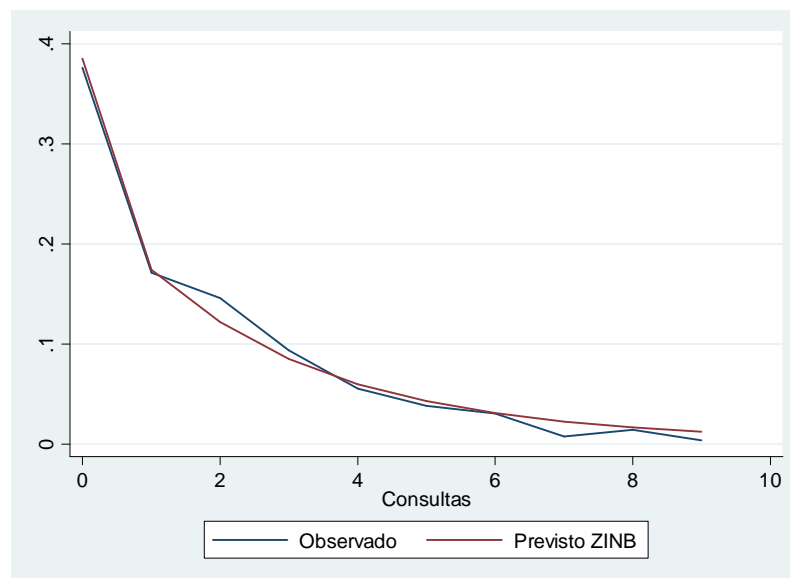
**Tabela 36 - probabilidades estimadas pelos modelos *Zero-inflated***

<i>Evento</i>	<i>Observado</i>	<i>Previsto ZINB</i>	<i>Previsto ZIP</i>
0	0.376	0.385	0.376
1	0.171	0.174	0.086
2	0.146	0.122	0.120
3	0.094	0.085	0.120
4	0.055	0.060	0.098
5	0.038	0.043	0.070
6	0.030	0.031	0.047
7	0.007	0.022	0.029
8	0.014	0.017	0.018
9	0.004	0.012	0.011

**Figura 9 - ajuste gráfico das probabilidades estimadas pelo modelo *Zero-Inflated* Poisson**



**Figura 10 - ajuste gráfico das probabilidades estimadas pelo modelo *Zero-inflated Binomial Negativo***



Os critérios de informação e o teste de sobredispersão também favoreceram o modelo *Zero-Inflated Binomial Negativo*, conforme as tabelas que se seguem:

**Tabela 37 - teste de sobredispersão modelos *Zero-inflated***

$\ln(\alpha)$	-0.115
$\alpha$	0.892
chibar2(01) =	277389
teste LR Prob>0 =	0
Log verossim. ZIP =	-829335
Log verossim.ZINB =	-690641

**Tabela 38 - critérios de informação modelos *Zero-inflated***

	AIC	BIC
ZIP	1658828	1659682
ZINB	1381441	1382305

Portanto, o modelo *Zero-Inflated Binomial Negativo* mostrou-se superior ao modelo *Zero-Inflated Poisson*. Na seção seguinte dar-se-á continuidade a estimação dos modelos do tipo *finite mixture*, com a estimação do modelo *Finite Mixture Binomial Negativo* (FMM-NB2)



com dois componentes latentes, utilizando uma mistura da distribuição binomial negativa 2 (NB2). A diferença entre o modelo *Finite Mixture* e os modelos *Zero-inflated* está no fato de que o primeiro admite misturas de zeros e valores positivos, enquanto os últimos não: o modelo FMM na sua formulação com duas subpopulações permite a separação entre usuários “altos” e “baixos”, enquanto que os modelos *Zero-inflated* fazem a distinção entre usuários e não usuários.

#### 4.2.3.2 Modelo *Finite Mixture* Binomial Negativo

A estimação do modelo *Finite mixture* com dois componentes<sup>40</sup> utilizando uma mistura da distribuição Binomial Negativa 2 (NB2)<sup>41</sup> gerou resultados diversos para cada uma das subpopulações. A primeira obteve um valor previsto médio de 1,81 consultas ao ano, enquanto a segunda uma média de 5,05 consultas. Nesse sentido, pode-se definir o primeiro componente como sendo referente ao grupo das pessoas saudáveis ou “*low-users*” e o segundo como o das pessoas mais suscetíveis a adoecerem ou “*heavy-users*”.

Conforme explicitado no capítulo anterior, a metodologia dos modelos *finite mixture* possui uma formulação mais flexível por dois motivos (DEB; TRIVEDI, 1997): a escolha do número de componentes determina o número de “tipos” e a escolha da distribuição permite a acomodação de heterogeneidade em cada componente, conferindo possibilidade de se retratar um contínuo de estados/situações. Além disso, a dicotomia entre “usuários” e “não usuários” pode não ser válida no caso da demanda por consultas: mesmo um indivíduo tipicamente saudável pode freqüentar o médico, especialmente para cuidados preventivos. O fato de que existe incerteza no âmbito da saúde – não se pode prever perfeitamente um acidente de modo que um indivíduo potencialmente “*low-user*” tenha que freqüentar o médico, por exemplo – também reforça a idéia de que a dicotomia supracitada é uma hipótese forte.

---

<sup>40</sup>É importante notar novamente que o uso de dois componentes foi adotado de acordo com o trabalho de Deb & Trivedi (1997), em que os modelos com três componentes foram preteridos em favor dos modelos com dois componentes, de acordo com os critérios de informação AIC e BIC.

<sup>41</sup>O uso da distribuição NB2 se deve ao fato de que a regressão NBRM tem como default esta mesma distribuição de probabilidade.

Os resultados mostraram que o grupo de usuários de baixa frequência não é influenciado pela *dummy* indicadora de cirrose, já no grupo de usuários “pesados” as *dummies* das raças preta e amarela, a *dummy* de atrite e a *dummy* indicadora de que o indivíduo esteve acamado recentemente não são significantes, como pode ser visto nas tabelas seguintes. Também é possível perceber que índios, pessoas com idade avançada, recém-nascidos e pessimistas com relação à saúde pertencem predominantemente ao grupo de usuários de alta frequência. A escolaridade dos indivíduos, ainda que tenha se mostrado pouco representativa em ambos os grupos, é proporcionalmente mais influente no grupo de “*low users*”. O motivo pode ser o fato de que para ser um usuário mais frequente basta que o indivíduo seja acometido por várias doenças ou que seja idoso ou que julgue sua saúde de forma negativa. Contudo, para um indivíduo que não necessariamente tem que ir ao médico, a educação pode ser decisiva para incentivar cuidados preventivos, colocando-o na subpopulação de indivíduos frequentadores não tão assíduos.

Tabela 39 - resultados do modelo *Finite Mixture*: “low-users”

<i>nconsult</i>	<i>Coef.</i>	<i>z</i>	<i>P&gt; z </i>	<i>Efeito marginal</i>
lpreco_est	-0.299	-41.68	0	-0.39
lrenddom	0.103	28.19	0	0.13
metro	0.134	20.37	0	0.18
urbano	0.273	26.6	0	0.32
homem	-0.535	-82.25	0	-0.69
casado	0.243	29.15	0	0.32
idade01	1.303	81.32	0	3.07
idade02	0.603	38.73	0	1.00
idade04	-0.179	-10.92	0	-0.22
idade05	-0.030	-1.8	0.071	-0.04
idade06	0.093	6.12	0	0.12
idade07	0.165	10.22	0	0.22
idoso	0.403	22.04	0	0.62
preta	-0.054	-4.13	0	-0.07
amarela	-0.193	-3.71	0	-0.23
parda	-0.059	-8.41	0	-0.08
indio	-0.159	-2.45	0.014	-0.19
norte	-0.260	-23.05	0	-0.30
nordeste	-0.088	-10.75	0	-0.11
centroeste	-0.212	-19.41	0	-0.25
sul	-0.171	-18.07	0	-0.21
anosestudo	0.039	38.05	0	0.05
autob	0.231	28.17	0	0.30
autoreg	0.773	72.03	0	1.30
autoruim	1.048	56.47	0	2.32
automruim	1.055	31.29	0	2.40
dinativ	0.589	42.34	0	0.99
drcoluna	0.295	31.93	0	0.43
dcrartrite	0.144	11.64	0	0.20
drcancer	0.568	14.79	0	0.98
dcrdiabetes	0.414	24.01	0	0.65
dcrhipert	0.418	41.16	0	0.64
drcora	0.275	18.58	0	0.40
dcrrenal	0.101	4.86	0	0.14
dcrdepres	0.279	20.02	0	0.41
dertuberc	0.317	4.37	0	0.48
dertend	0.231	12.81	0	0.33
dcirr	0.117	1.45	0.148	0.16
dcama	0.138	8	0	0.19
_cons	-0.176	-4.23	0	
/lnalpha1	-0.384	-34.17	0	
alpha1	0.681		0	
pi1	0.783			
N obs =	361515			
Log verossim.	-690478.2			

Tabela 40 - resultados do modelo *Finite Mixture*: “heavy-users”

<i>nconsult</i>	<i>Coef.</i>	<i>z</i>	<i>P&gt; z </i>	<i>Efeito marginal</i>
lpreco_est	-0.118	-11.36	0	-0.52
lrenddom	0.028	5.74	0	0.12
metro	0.235	23.95	0	1.07
urbano	0.197	14.25	0	0.81
homem	-0.277	-27.86	0	-1.22
casado	0.139	12.17	0	0.62
idade01	0.608	25.8	0	3.48
idade02	0.249	11.56	0	1.21
idade04	0.116	5.44	0	0.53
idade05	0.235	10.41	0	1.14
idade06	0.179	8.56	0	0.83
idade07	0.165	7.33	0	0.76
idoso	0.284	10.87	0	1.41
preta	0.003	0.16	0.871	0.01
amarela	-0.065	-0.85	0.393	-0.28
parda	-0.061	-5.83	0	-0.27
indio	0.975	11.49	0	7.24
norte	-0.274	-16.22	0	-1.09
nordeste	-0.078	-6.46	0	-0.34
centroeste	-0.156	-9.41	0	-0.64
sul	-0.039	-2.71	0.007	-0.17
anosestudo	0.008	5.25	0	0.03
autob	0.118	10.05	0	0.52
autoreg	0.528	33.65	0	2.77
autoruim	0.918	33.4	0	6.43
automruim	0.970	17.66	0	7.16
dinativ	0.352	16.28	0	1.81
drcoluna	0.112	7.81	0	0.51
dcrartrite	0.019	0.93	0.354	0.08
drcancer	0.567	8.79	0	3.35
dcrdiabetes	0.212	7.29	0	1.03
dcrhipert	0.256	16.27	0	1.24
drcora	0.183	7.59	0	0.88
dcrrenal	0.163	4.85	0	0.77
dcrdepres	0.324	14.73	0	1.66
dertuberc	0.368	3.66	0	1.96
dertend	0.169	5.51	0	0.81
dcirr	0.228	1.9	0.058	1.13
dcama	0.012	0.44	0.656	0.05
_cons	1.140	19.93	0	
/lnalpha2	-0.161	-6.2	0	
alpha2	0.852			
pi2	0.217			
N obs =	361515			
Log verossim.	-690478.2			

### 4.3 Comparação entre modelos

Os testes de sobredispersão e critérios de informação mostraram, sem exceção, que os modelos parametrizados pela distribuição Binomial Negativa são superiores aos que utilizam a distribuição Poisson, uma vez que os dados exibem uma grande concentração de zeros e variância elevada. Portanto, nesta seção, serão comparados somente os modelos Binomial Negativo, *Hurdle* Binomial Negativo, *Zero-Inflated* Binomial Negativo e *Finite Mixture* Binomial Negativo.

Os critérios utilizados para avaliação entre modelos foram: os testes de razão de verossimilhança (teste LR), os critérios de informação AIC e BIC e os valores previstos por cada modelo. Os testes LR do modelo Binomial Negativo contra os modelos *Hurdle* Binomial Negativo, *Zero-Inflated* Binomial Negativo e *Finite Mixture* Binomial Negativo não foram favoráveis ao primeiro.

Os critérios de informação também apontaram para a mesma conclusão. Já entre os modelos *Hurdle* Binomial Negativo, *Zero-Inflated* Binomial Negativo e *Finite Mixture* Binomial Negativo, o modelo selecionado foi o *Hurdle* Binomial Negativo, conforme a tabelas seguintes, em que se verifica que os valores do AIC e BIC são ligeiramente menores para este último modelo e que o valor médio previsto é mais próximo do observado para este último modelo.

**Tabela 41 - resumo critérios de informação**

	AIC	BIC
NBRM	1393339	1393782
ZINB	1381441	1382305
HNB	1378465	1379340
FMM-NB2	1381122	1382019

**Tabela 42 - valores previstos para o número de consultas por modelo**

<b>Modelo</b>	<b>Média</b>	<b>Desvio-padrão</b>	<b>Mín</b>	<b>Máx</b>
NBRM	2.541	2.600	0.415	125.941
HNB	2.449	1.789	0.637	82.248
ZINB	2.478	2.128	0.184	79.346
FMM-NB2	2.516	2.489	0.464	88.962
FMM-NB2 (comp1)	1.812	2.249	0.166	89.149
FMM-NB2 (comp2)	5.050	3.595	1.301	208.935
Observado	2.450	4.181	0.000	98.000

#### 4.4 Resultados

De acordo com o modelo selecionado (*Hurdle* Binomial Negativo), o brasileiro faz, em média, 2,45 consultas ao ano. Para os idosos, este valor é quase o dobro, aproximadamente, 4,21 consultas ao ano e para as crianças (até 14 anos), uma média de 2,13.

Foi observado que, fundamentalmente, contribuem de forma positiva para a decisão de consultar um médico os seguintes fatores: renda domiciliar, residência em região urbana e metropolitana, estado civil casado, faixa etária avançada (idoso) ou inicial (criança), raça indígena (podendo este fato ser decorrente da oferta – mais médicos e/ou Programa Saúde da Família em áreas indígenas), nível educacional, auto-avaliação pessimista com relação ao estado de saúde, presença de limitações de atividade e ser portador de doenças crônicas. De forma negativa, contribuem o preço da consulta médica (incluindo o custo de oportunidade de frequentar o sistema público de saúde), sexo masculino, faixa etária intermediária (adolescentes ou adultos até 39 anos), cor preta, parda ou amarela e residência fora da região sudeste.

Com relação à frequência das idas ao médico (segundo estágio do *Hurdle*) também se constatou que preço, sexo e região são fatores limitantes do número de consultas realizadas. Além disso, residência em área urbana e metropolitana, ser criança e/ou idoso, ter renda e escolaridade elevadas, ser casado, ser pessimista com a saúde e ter doenças crônicas (principalmente câncer, tuberculose e diabetes) contribuem para aumentar esta frequência.

Alguns estudos como os de Andrade *et. al* (2006) e de Médici (2002) atestam para uma regressividade nos gastos com saúde: famílias de baixa renda comprometem alta proporção de sua renda no pagamento de medicamentos, consultas médicas e, principalmente, planos de saúde (cujas elasticidades de renda calculadas pelo primeiro estudo foram as mais elevadas dentre todos os outros gastos em saúde). O presente trabalho incorporou este caráter regressivo através do procedimento de estimação da mensalidade privada e do preço-sombra de utilizar o setor público de assistência à saúde, cujo efeito marginal na estimação foi negativo, conforme esperado. Os efeitos marginais de ambas as etapas (decisão de contato e frequência) de estimação do *Hurdle* se encontram na tabela 43, no Apêndice.

Outro resultado interessante é que, ao fazer o exercício de restringir a amostra para somente os indivíduos que têm plano de saúde, e utilizando a mesma especificação e o mesmo modelo, o preço, apesar de ainda ser um fator limitante na decisão de contato, torna-se um fator positivo na frequência de consultas. A renda, por sua vez, torna-se um determinante negativo para o contato e continua sendo positivo no segundo estágio. Já a escolaridade não é significativa.

Por outro lado, ao se restringir a amostra somente para indivíduos que não têm plano de saúde, e que, por hipótese, usam o sistema público de saúde, o preço passa a ser um determinante positivo no primeiro estágio e negativo no segundo. A renda, por sua vez, é negativa em ambos e a escolaridade também.

Algumas conclusões podem ser elaboradas a partir desses resultados. Em primeiro lugar, quem usa o sistema privado geralmente o faz porque possui renda e escolaridade suficiente para permitir que o indivíduo adquira um seguro-saúde. Uma vez tomada esta decisão de adquirir um plano, quanto mais coberturas este oferece mais caro é o plano. Assim, se o indivíduo escolhe um plano superior é porque ele precisa por motivo de saúde ou porque tem um grande número de dependentes (seleção adversa) ou ainda, dado que o plano oferece ampla cobertura, há uma tendência a querer usufruir de todas as funcionalidades que este oferece (risco moral).

Por outro lado, para quem utiliza o sistema público de saúde, o preço significa o custo de oportunidade de consultar o médico. A princípio, o paciente tem necessidade de recorrer a um médico. No entanto, uma vez realizada a consulta, o preço torna-se um fator limitante para a

freqüência, pois o indivíduo terá que enfrentar novamente filas, falta de médicos, etc. A renda neste caso não é importante, uma vez que o SUS é universal, mas geralmente quem mais utiliza o sistema público é a população de baixa renda.

Nesse sentido, pode-se concluir que existe desigualdade no acesso aos serviços de saúde, em particular nas consultas médicas. Enquanto a população de renda e escolaridade mais elevada tende a freqüentar serviços privados de saúde, sob cobertura de plano de saúde suplementar, a população de baixa renda necessita do SUS, submetendo-se aos notórios problemas na qualidade, infra-estrutura e organização dos serviços públicos.

Por último, cabe fazer algumas ressalvas. Embora, o modelo proposto baseia-se, teoricamente, em características da demanda, os resultados ora apresentados podem refletir também características da oferta. Ainda que não existam variáveis de oferta na PNAD (por exemplo, número de leitos, número de hospitais, número de médicos, tipos de hospitais, etc.), algumas variáveis podem internalizar algumas características de oferta, por exemplo, o fato da *dummy* indicadora de raça indígena ter um efeito marginal alto: pode haver algum programa de saúde especialmente voltado para esta população.



## 5 CONSIDERAÇÕES FINAIS

Neste trabalho ficou demonstrado que existem grandes diferenças nos padrões de acesso à saúde de indivíduos pertencentes a grupos socioeconômicos distintos. Contudo, cabe notar que o maior problema não está no fato dos mais pobres e/ou os idosos utilizarem mais o setor público de saúde, mas no fato de que este setor ainda não consegue prover serviços de saúde adequados para tais grupos, principalmente no que tange à realidade regional.

Segundo Médici (2003), mesmo com o processo de descentralização da gestão dos recursos da saúde da esfera federal para os níveis estaduais e municipais, alguns municípios ainda não cumprem os requisitos mínimos para o recebimento da transferência. Com o intuito de aumentar a eficiência dos gastos, o Ministério da Saúde substituiu alguns mecanismos de reembolso por procedimento (*fee-for-service*) e pré-pagamentos no âmbito do SUS, criando pacotes para atenção primária em saúde (transferências de recursos baseadas em indicadores *per capita* de custo efetividade a serem atingidos). Contudo, segundo o autor, esta alocação de recursos ainda carece de critérios epidemiológicos, geográficos e socioeconômicos. Mais ainda, critérios políticos muitas vezes são decisivos nesta alocação de recursos, criando anomalias como o fato de existirem estados/municípios com PIB *per capita* bem menor, mas com gasto público *per capita* em saúde maiores.

Além disso, Médici (2003) também ressalta que a universalidade do SUS é muitas vezes perversa do ponto de vista socioeconômico, na medida em que muitas pessoas de renda mais alta utilizam o SUS para procedimentos mais complexos e custosos, não cobertos pelos planos de saúde. Nesse sentido, enquanto os prêmios dos seguros privados são reduzidos por não cobrirem determinados procedimentos, os custos do SUS aumentam, impossibilitando a expansão e melhoria dos serviços para quem realmente necessita.

Ainda que o governo tenha criado mecanismos para inibir este comportamento oportunista por parte de segurados do sistema privado, isto não foi o bastante, na opinião do autor. A criação da Agência Nacional de Saúde Suplementar (ANS), que criou regras garantidoras do direito do consumidor - tentando inibir a discriminação de clientes por parte dos planos de saúde por idade e a não cobertura de certos procedimentos - foi um desses mecanismos.

Outro mecanismo foi a criação de uma lista de reembolso de procedimentos realizados pelo SUS a preços de mercado, a Tabela Única Nacional de Equivalência de Procedimentos (TUNEP). O problema, no entanto, é que em muitos procedimentos realizados pelo SUS, os provedores ainda são reembolsados pela Autorização de Internação Hospitalar (AIH), lista com preços inferiores aos de mercado, sendo a mesma praticada quando o SUS reembolsa instituições privadas de saúde.

Ainda de acordo com Médici (2002), verifica-se que a distribuição regional dos recursos públicos necessita de melhorias. Embora a orientação de descentralização seja uma diretriz do SUS, os governos locais nem sempre têm autonomia. Um caso que exemplifica esta dificuldade é o extinto Programa de Atenção à Saúde (PAS), um programa orientado pelo lado da demanda, criado pelo município de São Paulo em 1994. O programa constituía-se de redes dirigidas por hospitais públicos e geridas por cooperativas de médicos que competiam entre si para atrair pacientes do sistema público num mecanismo de livre escolha. As cooperativas eram pagas em bases *per capita* (de acordo com um número determinado de pacientes que deveriam atender), ao invés do sistema de reembolso por procedimento. O sistema funcionou bem, com grande aceitação popular, até ser considerado inconstitucional e uma “privatização” do setor público.

Atualmente, os médicos do sistema público continuam sendo reembolsados por procedimento. Esse tipo de pagamento aos provedores, além de não gerar incentivo à qualidade dos serviços; produz, ao contrário, incentivos perversos à sobreutilização dos serviços e aumento dos custos. Segundo Médici (2002), ainda há poucas experiências no Brasil para pagamentos baseados em *performance* ou incentivos orientados pela demanda. Situações em que recursos públicos foram usados para contratar seguros privados de saúde em regime de capitação, como é o caso do *Medicaid* e do *Medicare* norte-americanos também são raras. A experiência de São Paulo, inovadora em termos de formato de gestão, acabou não prosseguindo.

Portanto, o que se tem observado até então é que o sistema brasileiro de saúde, embora sob a denominação de “sistema único”, e mesmo com as tentativas de integração entre seus componentes público e privado ainda funciona sob mecanismos da administração tradicional, que organiza os serviços pelo lado da oferta. Ou seja, os hospitais e postos de saúde estão disponíveis e as pessoas que os necessitam devem procurá-los.

Segundo Médici (2003), o grande problema dos serviços organizados pelo lado da oferta é acarretar um baixo grau de competitividade, uma vez que os pacientes não têm outra opção a não ser utilizar serviços que nem sempre possuem qualidade satisfatória, na medida em que também não há incentivos para buscar esses parâmetros de aceitação por parte do público. Outros problemas destacados por esse autor são: o multi-emprego dos profissionais de saúde, gerando um absenteísmo generalizado e o baixo grau de prevenção, uma vez que também não há estímulos para isso (pois estreitaria o mercado, tendo em vista que o risco de utilização dos serviços de saúde cairia). A incorporação acrítica da tecnologia por parte dos médicos é outra questão levantada pelo autor: poucos incentivos à prevenção gerariam deformidades nos diagnósticos, levando a sobreutilização de aparelhos de alto custo para sistema. Portanto, a persistência deste modelo regido pelo lado da oferta implicaria altos custos, baixa qualidade e conseqüente permanência da dificuldade no acesso aos serviços de saúde pela população.

Nesta mesma direção aponta Mendes (2002), que assinala que hoje em dia os gestores estão se movendo das formas de pagamento por subsídio à oferta para formas de pagamento por subsídio à demanda. O autor discute que as formas de pagamento apresentam, todas, fortalezas e debilidades. Assim, o pagamento por capitação é forte no incentivo de medidas promocionais e preventivas e na contenção de custos, mas pode induzir os prestadores a não oferecer certos serviços necessários; por outro lado, o pagamento por unidade de serviços incentiva a sobreprestação de serviços, mas é o sistema que mais satisfaz a população. Mas, a seguir o autor salienta alguns pontos consensuais, dentre os quais: é necessário algum nível de capacitação gerencial para tornar qualquer método eficaz; o método de pagamento por unidade de serviços deve ser evitado porque suas debilidades são superiores às suas fortalezas; a tendência dos gestores, no plano internacional, é no sentido das formas de subsídio à demanda; a forma de pagamento deve induzir a algum compartilhamento de riscos financeiros com os prestadores de serviço; e, a forma de pagamento deve fazer parte de um contrato explícito entre os gestores e os prestadores de serviço.

Diante das dificuldades encontradas no atual sistema de saúde supracitadas, deve-se retomar à discussão entre universalização e focalização de políticas públicas: será mesmo que o SUS, universal e integral como pretende ser, garante a equidade? Será que esse sistema é o mais eficiente com relação à alocação dos recursos? De fato, a concepção do SUS como um sistema universal não é em si a maior inadequação. A questão, entretanto, reside no fato de que, efetivamente, o sistema como tal ainda não conseguiu superar alguns problemas tais

como garantir o acesso de qualidade para pessoas com maior necessidade (conceito de equidade apresentado no capítulo um deste trabalho) e racionalizar os gastos de maneira a atingir este objetivo.

Para atingir o objetivo de racionalização dos gastos, a política de saúde pública no Brasil deve seguir a tendência mundial da assistência médica gerenciada (conhecida como *managed care*) a fim de reduzir a ineficiência. Também poderia ser considerada a imposição de parâmetros de *performance* aos médicos, que deverão ter seus pagamentos orientados a estes parâmetros e/ou seguindo um regime de capitação (pagamento por grupo de pacientes que atende). O incentivo à prevenção também deve ser realçado, de forma a reduzir custos futuros de tratamentos e assim garantir o acesso a quem realmente tem necessidade.

Além disso, o Brasil necessita introduzir mecanismos de gerenciamento baseados não apenas na oferta, mas principalmente na demanda, respeitando os aspectos regionais, sociais, econômicos e epidemiológicos na ocasião de transferências intergovernamentais, coibindo a atuação de *lobbies* políticos. O presente trabalho, cuja análise é pautada na demanda por serviços de saúde constitui uma ferramenta importante na orientação das políticas de saúde, permitindo a identificação e a mensuração dos fatores sociais, regionais, demográficos, econômicos e de saúde direcionadores desta demanda, bem como estimar a quantidade de consultas médicas que a população demanda.

De acordo com os resultados do modelo econométrico apresentado no capítulo quatro, o brasileiro médio vai ao médico cerca de 2,45 vezes ao ano. As mulheres freqüentam mais os consultórios: enquanto que os homens fazem, em média, 1,97 consultas ao ano, as mulheres consultam-se 2,89 vezes no mesmo período. A média de consultas também pode variar de acordo com a região do país: para o norte, o nordeste, o sul, o sudeste e o centro-oeste as médias são 2,31, 2,01, 2,67, 2,75 e 2,15, respectivamente. O perfil de rendimento também é relevante na demanda por consultas: para a população cuja renda do domicílio não supera um salário mínimo, a média de visitas ao médico é de 2,14; já para aqueles cuja renda supera este valor é de 2,46. É este perfil, assim como a escolaridade, que também determinam se o indivíduo possui ou não condições de adquirir um plano ou seguro-saúde. Com efeito, pessoas que possuem assistência médica suplementar vão 2,88 vezes ao médico, enquanto que pessoas que dependem do sistema público de saúde vão somente 2,33.

É importante ressaltar que os idosos realizam, em média, 4,21 consultas ao ano, segundo o modelo estimado. Este fato chama atenção, na medida em que a perspectiva de envelhecimento progressivo da população brasileira tende a trazer problemas futuros para a sustentabilidade do sistema único, caso a ineficiência alocativa e os desperdícios não sejam solucionados. Infelizmente, no caso dos idosos, não apenas os mecanismos de gestão devem ser alterados, mas também mudanças no comportamento deverão fazer parte do processo de reformulação das políticas de saúde. Segundo Bahia (2008), nos EUA as desproporções entre a quantidade de idosos e gastos assistenciais (13% da população tem mais de 65 anos, correspondendo a 36% dos gastos com saúde) e a preponderância do financiamento do *Medicare* estão mais do que decantadas. Entretanto, no Brasil, o envelhecimento populacional não parece gerar a mesma dinâmica de gastos. Em parte, isto pode ser justificado em decorrência de um grande contingente de pessoas abaixo da linha de pobreza que também necessita de atenção. Por outro lado, há um problema comportamental, em que existem parâmetros de priorização relativa dos jovens em detrimento dos idosos, reveladas em práticas não regulamentadas, como a seleção dos pacientes mais jovens para ingresso em unidades de terapia intensiva da rede pública.

Seguindo estas informações apresentadas, além de outros possíveis cruzamentos entre variáveis do modelo econométrico considerado neste trabalho, é possível medir o potencial da demanda para cada região do Brasil, de acordo com o perfil socioeconômico e de saúde da população. Adicionando a estas medidas, dados epidemiológicos, poderia ser feito um planejamento dos gastos e transferências governamentais seguindo critérios objetivos, baseados na demanda por consultas médicas destas regiões. Procedimentos semelhantes poderiam ser utilizados para medir o número de internações (que são bastante custosas para o SUS), e até mesmo o número de dias em se permanece internado.

Por último, deve-se lembrar que o trabalho considerado seria ainda mais útil do ponto de vista prático se houvesse uma base de dados com corte municipal, tendo em vista que a descentralização da gestão do sistema público preconiza maior autonomia para os municípios. Esta é, portanto, outra proposição importante para gerar melhoria nas políticas públicas: investir em sistemas de informação e em estatísticas regulares para o setor saúde.

## APÊNDICE

Tabela 43 - efeitos marginais dos estágios do *Hurdle*

	<i>Logit</i>	<i>Zero-truncated Binomial Negativo</i>
lpreco_est	-0.441	-0.340
lrenddom	0.122	0.092
metro	0.085	0.550
urbano	0.307	0.350
homem	-0.681	-0.647
casado	0.323	0.331
idade01	1.492	2.398
idade02	0.647	0.751
idade04	-0.259	0.302
idade05	-0.157	0.625
idade06	-0.009	0.478
idade07	0.089	0.465
idoso	0.433	0.825
preta	-0.049	-0.035
amarela	-0.218	-0.186
parda	-0.066	-0.130
indio	0.162	2.433
norte	-0.231	-0.606
nordeste	-0.056	-0.190
centroeste	-0.132	-0.458
sul	-0.163	-0.204
anosestudo	0.059	0.019
autob	0.235	0.296
autoreg	0.823	1.515
autoruim	1.098	3.380
automruim	0.959	3.885
dinativ	1.149	1.000
drcoluna	0.487	0.310
dcrartrite	0.296	0.159
drcancer	1.007	1.898
dcrdiabetes	0.784	0.818
dcrhipert	0.671	0.676
drcora	0.583	0.578
dcrrenal	0.232	0.290
dcrdepres	0.665	0.808
dcrtuberc	0.315	1.074
dcrvend	0.542	0.481
dcirr	0.081	0.486
dcama	0.373	0.165

## REFERÊNCIAS

ALVES, Denisard Cneio de Oliveira; TIMMINS, Christopher. *Social exclusion and two-tiered healthcare systems of Brazil*. Em: BEHRMAN, Jere; R.; GAVIRIA, Alejandro; SZÉKELY, Miguel. (Org.). *Who's in and Who's Out*. Washington, D. C., p. 45-71, 2003.

ANDRADE, Mônica Viegas; LISBOA, Marcos de Barros. A economia da saúde no Brasil. *In* LISBOA, Marcos de Barros; MENEZES, Naércio Aquino. (Org.). *Microeconomia e Sociedade no Brasil*. Rio de Janeiro, 2001.

\_\_\_\_\_. Sistema privado de seguro-saúde: lições do caso americano. *Ensaio Econômico EPGE*. Rio de Janeiro, n. 382, 2000.

ANDRADE, Mônica Viegas; OLIVEIRA, Thiago Bastos; NORONHA, Kênia V. M. S. Determinantes dos gastos das famílias com saúde no Brasil. *Economia*, v.7, n.3, p. 485-508, 2006.

ANDRADE, Mônica Viegas; MAIA, Ana Carolina. A seleção adversa no sistema de saúde suplementar brasileiro. *In*: SDE/ IPEA/ ANPEC. *Estudos de Métodos Quantitativos Aplicados a Defesa da Concorrência e a Regulação Econômica (Relatório Final)*. Brasília, p.33-64, 2005.

ARROW, Kenneth J. *Uncertainty and the welfare economics of medical care*. *American Economic Review*, v. 53, n. 5 p. 941-973, 1963.

BAHIA, Ligia. A janela demográfica, crescimento econômico e as políticas de saúde e proteção social. *Ciência e Saúde Coletiva*, v. 13, n. 4., 2008.

BARROS, Ricardo Paes de; MENDONÇA, Rosane; SANTOS, Daniel. Incidência e natureza da pobreza entre idosos no Brasil. *In*: CAMARANO, Ana Amélia (Org.). *Muito além dos 60: os novos idosos brasileiros*. Rio de Janeiro, p. 221-249, 1999.

CAMARANO, Ana Amélia. Envelhecimento da população brasileira: uma contribuição demográfica. *Texto para discussão IPEA n. 858*, 2002.

CAMERON, A. Colin; TRIVEDI, Pravin .K. *Regression analysis of count data*. *Econometric Society Monographs*. n. 30. New York: Cambridge University Press, 1998.

CAMERON, A. Colin; JOHANSSON, Per. *Count data regression using series expansions: With applications*. Journal of Applied Econometrics, v. 12, n. 3, p. 203-23, 1997.

CAMERON, A. Colin; TRIVEDI, Pravin K.; MILNE, Frank; PIGGOTT, John. *A Microeconomic model of demand of healthcare and health insurance in Australia*. Review of Economic Studies, v. LV(1), n.181, p. 85-106, 1988.

CONSTITUIÇÃO FEDERAL, art. 196.

CRAGG, John. G. *Some statistical models for limited dependent variables with applications to the demand for durable goods*. Econometrica, v. 39, n. 39, p. 829-44, 1971.

CREEL, Montserrat; FARELL, Michael. *Modeling usage of medical care services: the medical expenditure panel survey data: 1996-2000*. UFAE and IAE Working Papers, n. 646.05. Barcelona, 2005.

CUTLER, David; ZECKHAUSER, Richard. *Adverse Selection in Health Insurance*. Forum for Health Economics & Policy, v.1, n.1. California: Berkeley Electronic Press, 1998.

DEB, Partha; TRIVEDI, Pravin K. *Demand for medical care by the elderly: a finite mixture approach*. Journal of Applied Econometrics, v. 12, n. 3, p. 313-36, 1997.

GALLANT, A. Ronald; NYCHKA, Douglas W. *Semi-nonparametric maximum likelihood estimation*. Econometrica, v. 55: 363-90, 1987.

GERDTHAM, Ulf-G; TRIVEDI, Pravin. K. *Equity in Swedish health care reconsidered: new results based on the finite mixture model*. Working Paper Series in Economics and Finance n. 365. Stockholm, 2000.

GREENE, William H. *Econometric Analysis*, 5th. ed. Prentice Hall, 2003.

GROSSMAN, Michael. *On the concept of health capital and the demand for health*. The Journal of Political Economy, v. 80, n. 2, p. 223-225, 1972.



GURMU, Shiferaw. *Semi-parametric estimation of count regression models*. Journal of Econometrics, v. 88, p. 123-50, 1999.

JACK, William. *Health insurance reform in four Latin American countries: theory and practice*. World Bank Policy Research Working Paper Series. Washington, D. C., 2000.

JONES, Andrew M. *Applied health econometrics for health economists*, 2nd ed. Radcliffe, 2007.

LONG, J.S.; FREESE, Jeremy. *Regression models for categorical dependent variables using STATA*. College Station: Stata Press Publication, 2001.

MEDEIROS, Marcelo. Princípios de justiça na alocação de recursos em saúde. Texto para discussão IPEA, n. 687. Rio de Janeiro, 1999.

MÉDICI, André Cezar. *Financing health policies in Brazil: achievements, challenges and proposals*. Sustainable Development Technical Department Papers Series. Banco Interamericano de Desenvolvimento, 2002.

\_\_\_\_\_. Tendências da gestão em saúde ao nível mundial: o caso da assistência médica gerenciada (AMG). Fórum de Saúde Suplementar da ANS. Rio de Janeiro, 2003.

MENDES, Eugênio Vilaça. Os sistemas de serviços de saúde: o que os gestores deveriam saber sobre estas organizações complexas. Escola de Saúde Pública do Ceará. Fortaleza, p. 47-50, 2002.

MULLAHY, John. *Specification and testing of some modified count data models*. Journal of Econometrics, v. 33, p. 341-65, 1986.

NORONHA, Kênia V. M. S.; ANDRADE, Mônica Viegas. Desigualdades Sociais em saúde e na utilização dos serviços de saúde entre os idosos da América Latina. Pan American Journal of Public Health. Washington, D.C., 2005.

PAULY, Mark V. *The economics of moral hazard: comment*. American Economic Review, v. 58, p. 531-37, 1968.

PEREIRA, M. D. R. C. Fraudes e má utilização no sistema de saúde suplementar e os reflexos para o consumidor. Rio de Janeiro: Fundação Escola Nacional de Seguros, 2002.

PESQUISA DE ORÇAMENTOS FAMILIARES (POF). Micro dados, IBGE, 2002/03.

PESQUISA NACIONAL POR AMOSTRA DE DOMICÍLIOS (PNAD). Micro dados, IBGE, 2003.

POHLMEIER, Winfried; ULRICH, Volker. *An econometric model of the two-part making process in the demand for health care*. Journal of Human Resources, v. 30, p. 339-361, 1995.

HERNÁNDEZ-QUEVEDO, Cristina; JONES, Andrew, M; RICE, Nigel. *Reporting bias and heterogeneity in self-assessed health: evidence from the British Household Panel Survey*. HEDG Working Paper 05/04, 2005.

TRUJILLO, Antonio J.; VERNON, John A.; WONG, Laura Rodriguez; ANGELES, Gustavo. *Race and Health Disparities Among Seniors in Urban Areas in Brazil*. NBER Working paper n. 11690, 2005.

WINDMEIJER, Frank. A. G.; SANTOS SILVA, João. *Endogeneity in count data models: an application to the demand for health care*. Journal of Applied Econometrics, v. 12, n. 3: 281-94, 1997.

WORLD HEALTH ORGANIZATION. *The world health statistics report, 2007*.

ZWEIFEL, Peter. *Supplier-induced demand in a model of physician behaviour*. In: Health, Economics, and Health Economics, p. 245-67, 1981.