

Conselho Administrativo de Defesa Econômica
Departamento de Estudos Econômicos

Documento de Trabalho

Nº 003/2020

Aplicação de modelos de disposição a pagar no estudo da competição na saúde suplementar

Tatiana de Macedo Nogueira Lima
(Especialista em Políticas Públicas e Gestão Governamental/Cade)

Brasília, junho de 2020



Ministério da Justiça e Segurança Pública
Conselho Administrativo de Defesa Econômica

**Aplicação de modelos de disposição a pagar no estudo da competição
na saúde suplementar**

Departamento de Estudos Econômicos – DEE

SEPN 515 Conjunto D, Lote 4, Ed. Carlos Taurisano

Cep: 70770-504 – Brasília-DF

www.cade.gov.br

Este é um trabalho do Departamento de Estudos Econômicos (DEE).

Tatiana de Macedo Nogueira Lima

(Especialista em Políticas Públicas e Gestão Governamental/Cade)

As opiniões emitidas nos Documentos de Trabalho são de exclusiva e inteira responsabilidade do(s) autor(es), não exprimindo, necessariamente, o ponto de vista do Conselho Administrativo de Defesa Econômica ou do Ministério da Justiça.

Ainda que este artigo represente trabalho preliminar, citação da fonte é requerida mesmo quando reproduzido parcialmente.

Sumário Executivo

O objetivo deste documento é analisar a aplicação de modelos de preços hedônicos e de disposição a pagar (*willingness to pay* – wtp) no estudo (i) do poder de mercado de prestadores de serviços de saúde e (ii) da influência da rede assistencial na dinâmica competitiva no setor de planos de saúde. A disposição a pagar (wtp) é o quanto um consumidor está disposto a pagar a mais por determinada característica de um produto. É possível defini-la, também, como a utilidade esperada, *ex ante*, associada à rede assistencial de um plano.

No caso da saúde suplementar, os modelos que estimam a wtp permitem verificar qual o valor para o consumidor da inclusão de determinado prestador em uma rede assistencial. Alguns modelos permitem, ainda, estimar o impacto nos preços dos planos de saúde da inclusão do prestador e o preço negociado entre operadora e prestador. O arcabouço desses modelos pode, assim, ser usado na avaliação de atos de concentração horizontal no setor de planos de saúde e de prestadores e atos verticais que envolvam agentes em ambos os setores.

Inicialmente, é feita uma breve descrição da dinâmica entre prestadores de serviços à saúde e operadoras de planos de saúde no Brasil. Mostra-se que a maior parte das despesas assistenciais das operadoras refere-se a eventos realizados em prestadores terceirizados de sua rede assistencial, que são, em grande parte, pagos por procedimento. Para os prestadores, isso significa que, quanto mais beneficiários puderem acessá-lo e quanto mais beneficiários de cada plano atrair, maiores serão suas receitas. A atratividade de um prestador para os consumidores, contudo, não pode ser aferida diretamente, já que os beneficiários não costumam fazer pagamentos aos prestadores. Um meio indireto de medir o valor atribuído a rede pelos consumidores é verificar o quanto a utilidade desses consumidores decresce se o hospital for desvinculado da rede assistencial de seu plano ou, paralelamente, o quanto estariam dispostos a pagar a mais por um plano se um prestador fosse incluído. Em outras palavras, mensurar, a wtp.

Nota-se, ainda, que, usualmente, não há direcionamento na prestação de serviços dentro das redes assistenciais dos planos – i.e. um beneficiário que precise de um serviço coberto pelo plano de saúde ao qual está vinculado pode escolher qualquer prestador da rede

assistencial do plano sem que haja intervenção da operadora na escolha ou que seja utilizado mecanismo que incentive a procura de prestadores específicos. A extensão da rede assistencial e sua qualidade, ademais, são variáveis importantes na escolha de um plano pelo consumidor. Esses fatores influenciam os pagamentos totais feitos pelas operadoras aos prestadores e os preços definidos para cada procedimento. Além disso, tornam a wtp de cada consumidor em relação aos diferentes prestadores uma variável relevante na demanda por planos de saúde.

Feitas essas considerações, o documento apresenta os principais modelos que permitem estimar a wtp. São apresentadas duas linhas de pesquisas principais. A primeira baseia-se no trabalho de Rosen (1974) sobre preços hedônicos e disposição a pagar (wtp). A literatura sobre preços hedônicos (preços implícitos de atributos de produtos diferenciados) é anterior ao trabalho citado, mas Rosen (1974) inovou ao derivar, a partir dos preços hedônicos, a disposição a pagar dos consumidores. Além desse trabalho, são revisados artigos que utilizam esse modelo em mercados de saúde.

A segunda linha de artigos apresentados trata especificamente da dinâmica nos mercados de saúde. Consideram, assim, peculiaridades relativas não apenas à diferenciação dos produtos, mas à dinâmica da contratação de planos e da relação entre prestadores de serviços e seguradoras/operadoras de planos. São apresentados três artigos: (GAYNOR e TOWN, 2012), (HO e LEE, 2017) e (HO e LEE, 2019). Como o leitor verá, os três artigos estimam modelos empíricos, nos quais a wtp *ex-ante*, ou seja a utilidade do consumidor derivada da inclusão de um hospital na rede hospitalar antes de ter tido qualquer choque (doença ou outra condição que o leve a precisar de atendimento médico-hospitalar) na sua saúde, é utilizada como variável exógena na equação da demanda por planos. Em (HO e LEE, 2017) e (HO e LEE, 2019), os arcabouços utilizados para a análise das negociações entre prestadores e operadoras e entre contratantes de planos e operadoras são formalizados. Como se verá, esses artigos permitem, além de estimar a wtp, fazer diferentes simulações.

Para verificar a viabilidade de estimação de modelos da segunda linha analisada, é feita uma simulação do modelo de GAYNOR e TOWN (2012). A simulação compreende a estimação da probabilidade de visita de um consumidor a um determinado prestador e o cálculo da wtp por meio de um modelo logit. A estimação da probabilidade é feita com base na equação da utilidade *ex-ante* do consumidor associada à visita a um determinado

prestador, tendo sido determinado diagnóstico. A partir da estimação dessas probabilidades, é possível calcular a wtp referente à inclusão de um prestador na rede.

O procedimento para cálculo da wtp *ex-ante* é, a princípio, simples, contudo, necessita de um conjunto extenso de dados concernente a procedimentos realizados, valores pagos por operadoras a prestadores, preços de planos de saúde e beneficiários a eles vinculados, redes assistenciais de planos de saúde, prevalência de doenças em beneficiários e impacto financeiro, na saúde e no bem-estar de cada condição de saúde no indivíduo por ela afetado (peso do diagnóstico). Os dados coletados pela Agência Nacional de Saúde Suplementar referentes aos procedimentos realizados em ambiente hospitalar (e, se for o caso, ambiente ambulatorial) permitiriam a estimação da wtp por meio do modelo simulado se não fossem anonimizados os prestadores de serviços e as operadoras; e se fosse possível identificar o plano dos beneficiários. A principal dificuldade para estimação seria definir as probabilidades de ocorrência das doenças e os pesos dos diagnósticos. Ainda que, nos dados referentes aos procedimentos realizados, constem o diagnóstico do beneficiário (preenchimento facultativo) e os valores pagos por procedimento, seria preciso estabelecer um método para estimar essas duas variáveis. O uso da wtp para análise de atos de concentração depende, em geral, do estabelecimento da relação entre esta e os lucros ou preços dos prestadores de serviços, o que, diferentemente da estimação da wtp, pode ser mais complexo.

Em conclusão, os modelos e procedimentos revisados neste trabalho adequam-se à análise do setor de saúde suplementar brasileiro e poderiam ser usados, quando necessário, para simular o impacto de atos de concentração no setor de prestação de serviços de saúde e de planos de saúde, sendo mais uma ferramenta de análise disponível aos conselheiros do Cade e à Superintendência-Geral. Os dados necessários para análises desse tipo existem, mesmo que não estejam disponíveis publicamente.

Sumário

Introdução	8
1. Breve comentário sobre a dinâmica entre os elos da cadeia da saúde suplementar no Brasil ...	10
2. Revisão da literatura	16
3. Simulação do modelo apresentado em GAYNOR e TOWN (2012)	40
3.1. <i>Dados</i>	40
3.2. <i>Estimação</i>	42
3.3. <i>Resultados</i>	43
3.4. <i>Cálculo da disposição a pagar</i>	45
3.5. <i>Utilização da wtp na análise de atos de concentração</i>	46
4. Conclusão	49
Bibliografia	51
Anexo: Código da Simulação	54

Introdução

O objetivo deste documento é analisar a aplicação de modelos de preços hedônicos e de disposição a pagar (*willingness to pay* – wtp) no estudo (i) do poder de mercado de prestadores de serviços de saúde e (ii) da influência da rede assistencial na dinâmica competitiva no setor de planos de saúde. Como se sabe, as operadoras de planos de saúde contratam prestadores de serviços hospitalares, de diagnóstico, médicos e odontológicos, conforme a segmentação assistencial de seus planos, para o atendimento de seus beneficiários.

A rede assistencial formada influencia a percepção dos potenciais contratantes em relação à qualidade dos planos e sua disposição a pagar por eles. O valor atribuído por contratante ou beneficiário à presença de determinado prestador na rede assistencial de um plano, contudo, não pode ser inferido diretamente, já que os valores pagos se referem ao plano e tudo que ele compreende.

Por essa razão, os modelos usuais para inferência de poder de mercado podem não captar adequadamente a estrutura concorrencial e dinâmica nos setores de prestação de serviços de saúde vinculados a planos de saúde ou a influência de determinado prestador na dinâmica do setor de planos de saúde. Os modelos que permitem a estimação da disposição a pagar dos consumidores de planos por um determinado prestador de saúde preenchem essa lacuna.

A disposição a pagar (wtp), como pode ser deduzido, é o quanto um consumidor está disposto a pagar a mais por determinada característica de um produto. É possível defini-la, também, como a utilidade esperada, *ex ante*, associada à rede assistencial de um plano. No caso da saúde suplementar, os modelos que estimam a wtp permitem verificar qual o valor para o consumidor da inclusão de determinado prestador em uma rede assistencial. Alguns modelos permitem, ainda, estimar o impacto nos preços dos planos de saúde da inclusão do prestador ou/e o preço negociado entre operadora e prestador a partir da wtp.

Assim, modelos nos quais essa variável é estimada podem ser usados para avaliar o poder de mercado dos prestadores de serviços de saúde, conseqüentemente, os efeitos de concentrações horizontais nesse setor. Também podem ser usados na avaliação de atos de concentração no setor de planos de saúde que envolvam operadoras verticalizadas ou que

atuam em mercados nos quais há prestadores verticalizados. À medida que aumenta a participação de grandes grupos hospitalares na prestação de serviços e que a verticalização passa a ser adotada por algumas das maiores operadoras como estratégia de negócios, a melhor compreensão das relações entre rede assistencial e planos de saúde torna-se imprescindível para a inferência dos efeitos de atos de concentração no setor.

As discussões relativas ao setor de saúde suplementar no Conselho Administrativo de Defesa Econômica (Cade) revelam o entendimento do Conselho de que a concorrência no setor de prestação de serviços de saúde é relevante para a configuração do mercado a jusante, sendo a verticalização uma das principais preocupações na análise de atos que envolveram operadoras de planos (CADE, 2018). Ainda assim, mesmo em casos complexos, não é comum a utilização de modelos formais para o entendimento das relações entre prestadores e operadoras. Esse documento de trabalho pode ser o primeiro passo para a utilização de modelos e métodos específicos para o setor, que podem gerar conclusões mais robustas, em atos de concentração no setor de prestação de serviços de saúde ou de planos de saúde.

O documento está dividido em três seções além desta introdução. Na primeira, é descrita a dinâmica entre os elos da cadeia da saúde suplementar no Brasil. Essa descrição facilita a compreensão de como os modelos apresentados na seção seguinte podem ser usados no estudo do setor de saúde suplementar e de prestação de serviços de saúde. Na segunda seção, é feita uma revisão da literatura que pode ser dividida em duas linhas de pesquisas principais. A primeira baseia-se no trabalho de Rosen (1974) sobre preços hedônicos e disposição a pagar (wtp). A literatura sobre preços hedônicos (preços implícitos de atributos de produtos diferenciados) é anterior ao trabalho citado, mas Rosen (1974) inovou ao derivar, a partir dos preços hedônicos, a disposição a pagar dos consumidores. Além desse trabalho, são revisados artigos que utilizam esse modelo em mercados de saúde.

A segunda linha de artigos apresentados trata especificamente da dinâmica nos mercados de saúde. Consideram, assim, peculiaridades relativas não apenas à diferenciação dos produtos, mas à dinâmica da contratação de planos e da relação entre prestadores de serviços e seguradoras/operadoras de planos. O foco de atenção, neste caso, é a wtp *ex-ante*, ou seja a utilidade do consumidor associada à inclusão de um hospital na rede hospitalar antes de o beneficiário ter usado os serviços daquele prestador ou mesmo de saber que

precisará utilizá-lo (em outras palavras, antes de ter tido qualquer choque em sua saúde que o leve a precisar de atendimento médico-hospitalar). O primeiro artigo revisado é (GAYNOR e TOWN, 2012). Apesar de não desenvolverem um modelo teórico formal, os autores descrevem a dinâmica entre prestadores e operadoras subjacente às estimações realizadas e constroem um procedimento que permite estimar a wtp. Já HO e LEE (2017 e 2019) formalizam a dinâmica de interações na saúde suplementar, tanto entre prestadores e operadoras quanto entre contratantes e prestadores. Apesar de os artigos terem como foco o mercado de planos de saúde empresarial norte-americano, o modelo pode ser adaptado e utilizado no Brasil.

Na terceira seção, é feita simulação com base no artigo de Gaynor e Town (2012), sumariado na seção anterior¹. Nesse artigo, os autores desenvolvem modelo específico para a saúde suplementar. Os dados públicos disponibilizados pela Agência Nacional de Saúde Suplementar (ANS) não são suficientes para a estimação do modelo. Por isso, foram criados conjuntos de dados simulados simplificados, baseados nos dados públicos disponíveis, para a estimação. Por fim, são apresentadas as conclusões.

1. Breve comentário sobre a dinâmica entre os elos da cadeia da saúde suplementar no Brasil

Há, no Brasil, aproximadamente 47 milhões de beneficiários² de planos médico-hospitalares e 26 milhões de beneficiários de planos odontológicos³, sendo 67,5% dos contratos de planos de assistência médica e 73% dos planos odontológicos⁴ coletivos empresariais. Os demais beneficiários estão em planos coletivos por adesão (13,2% dos beneficiários de planos de assistência médica e 9,8% dos beneficiários de planos

¹ O presente documento de trabalho inclui um arquivo, na linguagem R, referente a esta simulação, para replicabilidade dos resultados, que está no Anexo deste documento.

² Dados coletados na sala de situação da ANS (<http://www.ans.gov.br/perfil-do-setor/dados-e-indicadores-do-setor/sala-de-situacao>) em 04/03/2020, referentes a dezembro de 2019. O termo beneficiário refere-se a vínculos com planos de saúde, podendo haver mais de um vínculo para um mesmo indivíduo.

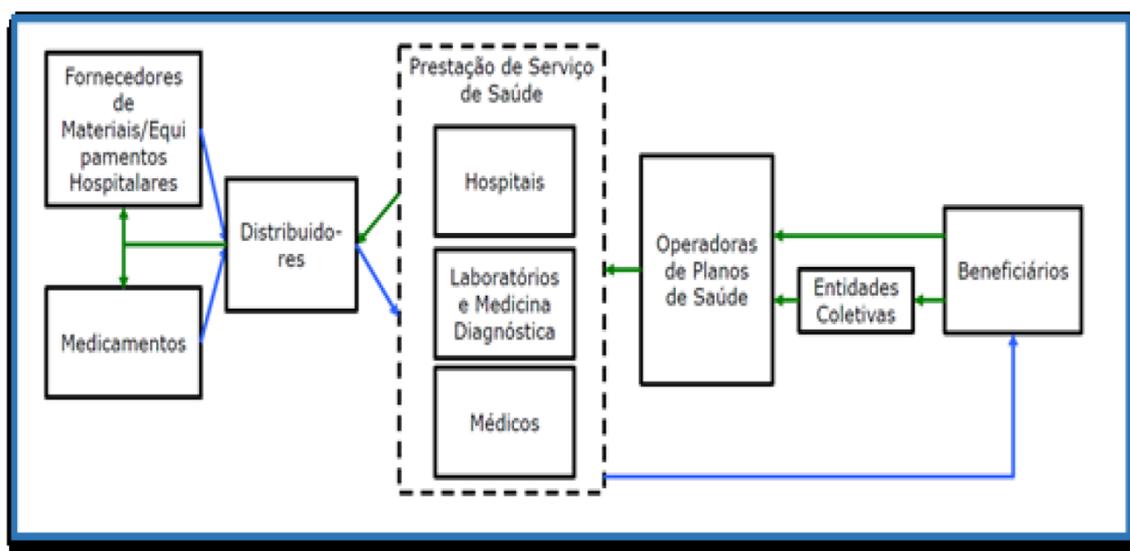
³ Dados coletados na sala de situação da ANS (<http://www.ans.gov.br/perfil-do-setor/dados-e-indicadores-do-setor/sala-de-situacao>) em 04/03/2020, referentes a dezembro de 2019. O termo beneficiário refere-se a vínculos com planos de saúde, podendo haver mais de um vínculo para um mesmo indivíduo.

⁴ Porcentagens calculadas a partir de dados coletados na sala de situação da ANS (<http://www.ans.gov.br/perfil-do-setor/dados-e-indicadores-do-setor/sala-de-situacao>) em 04/03/2020, referentes a dezembro de 2019.

odontológicos⁵) e em planos individuais/familiares (19,3% dos beneficiários de planos de assistência médica e 17,2% dos beneficiários de assistência odontológica⁶).

A provisão da assistência aos beneficiários de planos de saúde é feita majoritariamente por prestadores terceirizados (93,07%)⁷, contratados pelas operadoras (apenas 0,32% das despesas assistenciais pagas foram reembolsadas diretamente aos beneficiários⁸). A dinâmica da cadeia da saúde suplementar brasileira, no segmento médico-hospitalar, pode ser sintetizada pela Figura 1, abaixo. Empregadores, em grande parte, mas também outras entidades e os próprios beneficiários contratam planos de saúde, pelos quais pagam contraprestações às operadoras. Estas contratam prestadores de assistência à saúde, que realizam os serviços contratados pelos beneficiários, quando estes os demandam.

Figura 1 – Mapa da Cadeia Produtiva da Saúde Suplementar



Fonte: Azevedo et al. (2016, p. 30), reproduzido em (CADE, 2018, p.11)

⁵ Percentagens calculadas a partir de dados coletados na sala de situação da ANS (<http://www.ans.gov.br/perfil-do-setor/dados-e-indicadores-do-setor/sala-de-situacao>) em 02/10/2019, referentes a julho de 2019.

⁶ Percentagens calculadas a partir de dados coletados na sala de situação da ANS (<http://www.ans.gov.br/perfil-do-setor/dados-e-indicadores-do-setor/sala-de-situacao>) em 02/10/2019, referentes a julho de 2019. Há beneficiários cuja forma de contratação não foi identificada pela ANS, razão pela qual a soma das porcentagens não soma 100%.

⁷ Na codificação do plano de contas que as operadoras devem enviar para a ANS, a partir de 2019, as despesas assistenciais são discriminadas de acordo com a forma de remuneração do prestador e sua relação com a operadora. Até o terceiro trimestre de 2019, 6,84% das despesas assistenciais foram realizadas em rede própria, 0,62% no SUS (ressarcimento) e 3,30% por reembolso, todas as demais foram realizadas em rede contratada. Os dados das demonstrações financeiras foram coletados no portal brasileiro de dados abertos, em 04/03/2020.

⁸ Percentagens calculadas a partir dos dados das demonstrações financeiras do primeiro trimestre de 2019, constante do portal brasileiro de dados abertos, acessado em 22/09/2019.

Ainda conforme as demonstrações financeiras do terceiro trimestre de 2019, mais de 84% dos serviços são pagos por procedimento, de forma que os prestadores somente são remunerados se realizarem algum serviço. Sendo assim, não basta fazer parte da rede assistencial de um plano, é preciso ser suficientemente atraente aos beneficiários daquele plano para que estes o escolham ao precisarem realizar algum serviço.

Note-se, ainda, que não é comum as operadoras orientarem os beneficiários na escolha dos prestadores. Na maior parte dos planos, os clientes têm acesso a toda rede assistencial sem que precisem ser encaminhados por algum serviço de acolhida inicial. Programas de atenção primária e médicos de família associados aos planos que orientem os beneficiários ainda são raros, a despeito dos esforços do órgão regulador para estimular modelos assistenciais baseados na coordenação do cuidado. Isso significa que a escolha na prestação do serviço dentro da rede assistencial não é intermediada pela operadora, dependendo apenas das variáveis que os beneficiários conhecem e consideram importantes para avaliação de um prestador.

Como a operadora tem pouca influência sobre os serviços que serão utilizados uma vez contratada a rede assistencial e comercializados os planos de saúde, a negociação com os prestadores de serviços influencia consideravelmente seus resultados. Os preços negociados pelos serviços podem significar despesas assistenciais menores ou maiores. Contudo, além dos preços, é preciso considerar como a inclusão de determinado prestador à rede assistencial afeta a propensão à aquisição do plano pelos potenciais contratantes. A operadora faz, a princípio, uma estimativa dos custos e benefícios da inclusão de cada prestador à sua rede no momento do estabelecimento dos contratos com a rede assistencial.

Para muitos dos prestadores de serviços de saúde brasileiros, seus resultados dependem de seu desempenho na saúde suplementar. De acordo com dados da Associação Nacional de Hospitais Privados (Anahp), 30% das despesas com saúde no Brasil, em 2018, foram realizados na saúde suplementar⁹. É razoável supor que a importância do setor suplementar seja ainda maior para os prestadores hospitalares, pois pagamentos diretos de consumidores a prestadores de serviços tendem a ser para procedimentos ambulatoriais,

⁹ Dados constantes do Painel – Saúde em números – 2019, da Associação Nacional de Hospitais Privados, que pode ser obtido em <https://ondemand.anahp.com.br/curso/painel-2019>.

menos custosos e com custos mais previsíveis que os hospitalares. Segundo a Anahp, 90% da receita de seus associados em 2019 provieram de planos de saúde¹⁰.

O desempenho dos prestadores de serviços de saúde que atuam na saúde suplementar depende, portanto, de três fatores:

1. Vinculação a planos de saúde, sendo que quanto mais beneficiários houver nesses planos e maior for a demanda desses beneficiários por cuidado médico, maior tende a ser receita;
2. Forma de remuneração contratada e preços dos serviços;
3. Atratividade do prestador para os beneficiários dos planos aos quais está vinculado.

O objetivo inicial de um prestador que comercialize serviços na saúde suplementar é firmar contratos com operadoras de planos de saúde. Quanto mais beneficiários abranger esses contratos, maior a demanda potencial por seus serviços. A forma de remuneração influencia no risco da atividade. Quando o pagamento é por procedimento, quaisquer eventos não previstos, intercorrências ou procedimentos decorrentes de erros também são remunerados, o que pode não acontecer em outras formas de remuneração. Os preços determinam o lucro do prestador em cada procedimento e seus resultados econômico-financeiros. A atratividade é um dos principais fatores a influenciar a demanda por atendimentos, especialmente quando se considera que os custos marginais de utilização de um prestador para um beneficiário em plano sem mecanismo financeiro de regulação é nulo.

Dado que o poder de mercado de operadoras de planos, definido como a capacidade de estabelecer preços acima de seu custo marginal, depende da demanda por planos de saúde em geral e por seus planos de saúde, em particular, e de seu poder de compra, definido como a capacidade de obter de um fornecedor termos mais favoráveis do que aqueles oferecidos aos outros compradores (OCDE, 1981, citado em GOLDBERG, 2006), e que este também depende da demanda por seus planos e da quantidade de contratos efetivamente comercializados, a análise da concorrência nem sempre pode ser realizada utilizando-se os

¹⁰ ANAPH. Observatório 2019 – Publicação anual – edição 11. Publicação pode ser obtida em <https://ondemand.anahp.com.br/curso/publicacao-observatorio-2019>. Acesso em 02/10/2019.

modelos, parâmetros e ferramentas utilizados para análise de setores nos quais o tamanho e características das redes de prestação de serviços e de clientes não tenha papel tão fundamental. Paralelamente, o poder de mercado de um prestador de serviços depende da demanda pelos serviços comercializados em geral e por seus serviços, em particular, mas esta não pode ser estimada das formas usuais, já que os custos de utilização da rede de serviços são nulos para os beneficiários. Assim, também no setor de prestação de serviços de saúde, a análise de concorrência nem sempre pode utilizar os modelos e as ferramentas usuais.

O Cade, nos atos de concentração que envolvem os setores de prestação de serviços de saúde e de saúde suplementar, utiliza arcabouço para análise baseado nos principais aspectos tradicionalmente apontados na literatura concorrencial como indicadores de existência de poder de mercado e de possibilidade de geração de efeitos negativos. São definidos mercados relevantes, analisadas as participações de mercado, a existência de barreiras à entrada e de rivalidade entre os ofertantes (CADE, 2018). A participação de mercado das operadoras é calculada com base na quantidade de beneficiários no mercado relevante, e a participação de hospitais gerais, com base no número de leitos disponíveis à saúde suplementar. De acordo com Cade (2018), dois tópicos foram objeto de muitos debates e análises do Conselho: participações societárias cruzadas e integrações verticais.

As participações cruzadas, ainda que minoritárias, podem alinhar interesses de agentes anteriormente concorrentes, fazer com que um agente influencie a conduta de outro de forma a diminuir a concorrência entre os dois ou, ainda, garantir o acesso a variáveis sensíveis do outro agente. Os casos que chamaram a atenção do Conselho envolveram grandes ofertantes de planos de saúde (Amil), de serviços hospitalares (Rede D'Or) e de serviços de diagnóstico (Grupo Fleury) (CADE, 2018).

A integração vertical foi discutida em 40% dos 155 atos de concentração analisados em (CADE, 2018). Integrações entre diferentes tipos de prestadores de serviços e entre operadoras e diversos tipos de prestadores foram analisadas, fazendo Cade (2018) concluir que há uma tendência de verticalização nos mercados da saúde suplementar. A assimetria de informação que caracteriza as relações entre prestadores de serviços, operadoras de planos

e beneficiários¹¹ e os potenciais ganhos decorrentes da diminuição dessas assimetrias e do alinhamento de interesses são apontados como motivações para esses atos. Problemas concorrenciais, contudo, podem surgir decorrentes do fechamento de mercados tanto a jusante (operação de planos de saúde) quanto a montante (prestação de serviços).

Para a compreensão da dinâmica na saúde suplementar e a fundamentação da análise de casos complexos, é preciso estabelecer como as relações entre prestadores de serviços de saúde, operadoras de planos de saúde, contratantes de planos e beneficiários influenciam o poder de mercado dos ofertantes no setor de prestação de serviços e de planos de saúde e como diferentes configurações podem resultar em mercados mais ou menos concorrenciais. Modelos econômicos podem ser utilizados para isso. O modelo de preços hedônicos (ROSEN, 1974) permite estimar os preços implícitos de atributos de determinado bem e, satisfeitas certas condições, estabelecer a disposição a pagar dos consumidores em relação a esses atributos. Gaynor e Town (2012) desenvolveram um modelo especificamente para o setor de saúde, no qual consideram as relações discutidas acima e como estimar o poder de mercado de operadoras e planos de saúde e os efeitos de atos de concentração nesses mercados.

Neste trabalho, a literatura concernente à disposição a pagar será revista com o fim de identificar modelos e métodos que podem ser usados para compreensão da dinâmica entre os elos da cadeia na saúde suplementar. A fim de verificar a adequação desses modelos para estudos empíricos sobre a saúde suplementar brasileira, serão realizadas simulações e, com base nestas, será proposta uma estratégia empírica para a estimação de modelos que possam orientar a análise concorrencial.

¹¹ A prestação de serviços médicos caracteriza-se pela assimetria de informação entre ofertante e consumidor. Mesmo após a realização do serviço, o aumento do bem-estar do consumidor decorrente do consumo do serviço é de difícil mensuração por ele mesmo (credence goods). Também as operadoras têm dificuldade para aferição da necessidade de realização de serviços de saúde em seus beneficiários. Se a forma de remuneração não propiciar o alinhamento de interesses, é possível que sejam geradas despesas assistenciais maiores do que em situações nas quais os interesses de prestadores e operadoras estão alinhados (integração vertical, por exemplo). Já os pacientes têm maior conhecimento sobre o seu estado de saúde no momento de contratação de um plano do que as operadoras e também maior conhecimento sobre sua propensão à utilização dos serviços dado que os custos marginais são nulos. A operadora, por sua vez, tem maior conhecimento sobre sua capacidade de cobrir os serviços contratados e manter-se solvente.

2. Revisão da literatura

Em 1974, Sherwin Rosen publicou artigo que deu início a uma ampla literatura sobre preços hedônicos. Estes podem ser definidos como os preços implícitos de atributos de produtos diferenciados. Os preços hedônicos são revelados a partir das quantidades dos atributos associadas a cada produto e dos preços observados desses produtos. Um exemplo bastante utilizado de aplicação de preços hedônicos é o setor imobiliário. O preço de um imóvel é determinado pela quantidade dos diferentes cômodos (banheiros, quartos, etc.), pela localização e por eventuais outros atributos. A partir dos preços dos imóveis, infere-se qual o preço que os consumidores estão dispostos a pagar por um banheiro a mais, um quarto ou qualquer outro atributo, inclusive características externas ao imóvel (GIBBONS, MOURATO e RESENDE, 2014).

Embora o artigo de Rosen (1974) não tenha sido o primeiro estudo sobre o tema¹², inovou ao derivar, a partir dos preços hedônicos, a disposição a pagar (willingness to pay - wtp) dos consumidores. Para isso, Rosen (1974) propôs um método em dois estágios. No primeiro, os preços hedônicos são estimados a partir da regressão do preço do produto nas suas características. Pressupõe-se que o equilíbrio seja competitivo e que os determinantes para a escolha dos consumidores sejam apenas os atributos dos produtos e características dos próprios consumidores. Isso implica que, se duas firmas ofertam um produto com o mesmo conjunto de características, mas preços diferentes, os consumidores somente consideram a oferta do produto com menor preço. A reputação de um produto ou sua marca, podem, contudo, ser consideradas atributos, bem como a qualidade. Oczkowski (2015), por exemplo, inclui como atributo de vinhos a reputação das vinícolas que os produzem e suas classificações em um *ranking* de qualidade.

A partir das estimativas do primeiro estágio, utilizam-se as variações nos preços hedônicos para a recuperação das funções de disposição a pagar (wtp) dos indivíduos. Computam-se os preços marginais estimados pela derivação do preço em relação a cada atributo, e os preços marginais estimados são utilizados no segundo estágio para a estimação da oferta e da demanda de cada atributo.

¹² LANCASTER, Kelvin. A new approach to consumer theory. *Journal of Political Economy*, abril de 1966, é um exemplo de trabalho anterior ao de Rosen e por ele próprio citado.

Formalmente, o bem pode ser descrito por seus “n” atributos: $z = (z_1, z_2, \dots, z_n)$, onde z é o bem e z_i é seu iésimo atributo. O preço de z é definido a partir dos preços dos atributos: $p(z) = p(z_1, z_2, \dots, z_n)$. Os consumidores consomem uma unidade do bem, no valor z , podendo suas utilidades serem escritas da seguinte forma:

$$U(x, z_1, z_2, \dots, z_n) \quad (1)$$

Sendo x todos os outros bens consumidos, cujos preços são normalizados ($p_x = 1$). A renda do consumidor pode ser escrita em função de p_x ($y = x + p(z)$, sendo y , a renda) e sua utilidade pode ser descrita a partir de sua restrição orçamentária ($u = u(y - p_x, z_1, z_2, \dots, z_n)$). Invertendo essa equação e mantendo constantes todas as variáveis, exceto z_i , encontra-se uma expressão para a disposição a pagar (wtp) por z_i :

$$W_i = W_i(y - p_x, z_i; \mathbf{Z}_{-j}^*, u^*) \quad (2),$$

Sendo W_i a disposição a pagar por z_i , \mathbf{Z}_{-j}^* , o conjunto ótimo das outras características do produto e u^* a utilidade máxima passível de ser obtida, considerando-se a restrição orçamentária.

A utilidade é maximizada quando $W(z^*; u^*; y) = p(z^*)$.

Do lado da oferta, $M(z)$ é a quantidade de unidades produzidas por uma firma que faz z . O custo total de produção é dado por $C(M; z; \beta)$, na qual β é um vetor com parâmetros da função de produção e de preços dos fatores. Cada firma maximiza o lucro ($\pi = Mp(z) - C(M, z_1, \dots, z_n; \beta)$). A função de oferta é dada por:

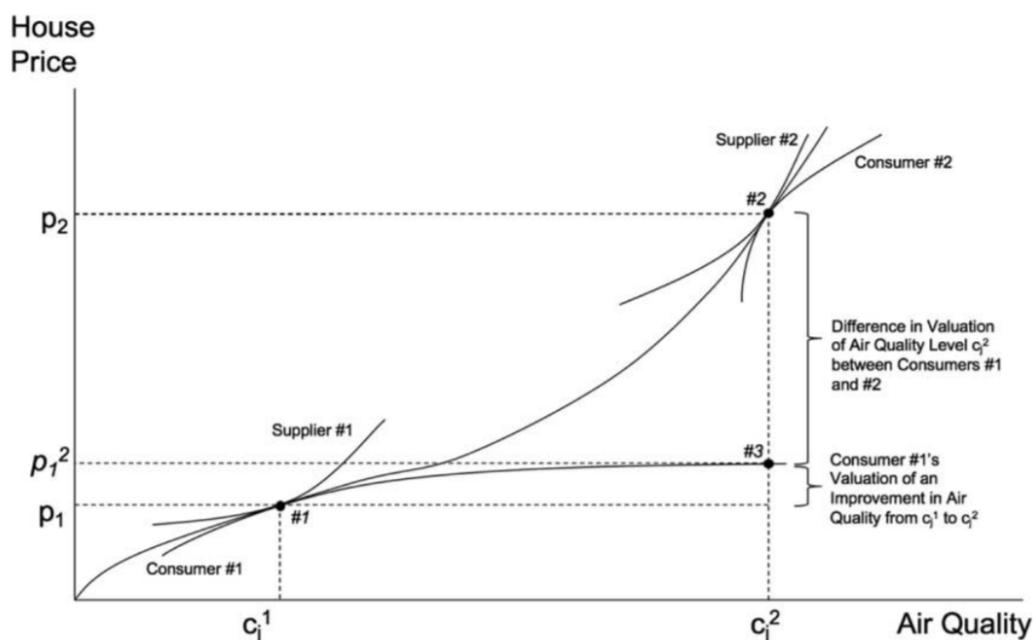
$$\phi(z_1, \dots, z_n; \pi, c) \quad (3),$$

Sendo (z_1, \dots, z_n) o conjunto de atributos e a quantidade desses atributos presentes nos produtos ofertados pela firma, dado um determinado nível de lucros e custos.

As interações no equilíbrio entre consumidores e ofertantes permitem estimar os preços hedônicos dos atributos e construir uma relação gráfica entre os preços do bem e a quantidade no produto de dada característica. Essa representação é comumente chamada de *hedonic price schedule (hps)*. Como exemplo, o gráfico 1, abaixo, apresenta o hps para os preços de casas e a qualidade do ar. Os consumidores escolhem o bem final de acordo com a sua valorização da característica estudada. Em cada ponto do gráfico, o preço marginal da

característica é igual a wtp do consumidor, de modo que esse gráfico pode ser usado para determinar a wtp média na população de interesse. Contudo, é interessante notar que não é possível inferir mudanças na wtp a partir do gráfico, pois cada ponto representa o equilíbrio considerando-se as preferências do consumidor naquele ponto. Considerando-se o gráfico 1, abaixo, se houvesse uma mudança na qualidade do ar, do ponto 1 ao 2, para o consumidor 1, a mudança na wtp seria de p_1 para p_1^2 , mas a análise do gráfico indicaria mudança bem maior (de p_1 para p_2).

Gráfico 1 – HPS de um mercado hedônico da qualidade do ar



Fonte: Extraído de (GREENSTONE, 2015, p 1893).

Como mencionado anteriormente, a inovação de Rosen (1974) foi propor a estimação de um segundo estágio, no qual os preços hedônicos são usados para a recuperação da wtp dos diferentes consumidores.

O modelo econométrico a ser estimado é:

$$p_i(z) = F^i(z_1, \dots, z_n; y_1) \quad (4)$$

$$p_i(z) = G^i(z_1, \dots, z_n; y_2) \quad (5)$$

p_i e z_i são variáveis dependentes conjuntas e y_1 (renda) e y_2 (parâmetros da função de produção e preços de fatores) são variáveis exógenas que deslocam a demanda (4) e a oferta (5), i indexa o atributo.

O procedimento para estimação é:

1. Estima-se $p(z)$ por mínimos quadrados ordinários ou qualquer método usual, sem considerar y_1 e y_2 , ou seja, $p(z)$ como função dos atributos do produto. O resultado é $\hat{p}(z)$;
2. Computa-se $\frac{\delta p(z)}{\delta z_i} = \hat{p}_i(z)$ para cada ofertante e comprador, avaliado na quantidade de características (valores numéricos de z) efetivamente vendidos e comprados;
3. Utiliza-se os preços marginais estimados ($\hat{p}_i(z)$) como variáveis endógenas no segundo estágio de estimação de (4) e (5).

Batik (1987) e Epple (1987) observam que, se a função de preços hedônicos não for linear, o que é o caso mais comum, o consumidor escolhe simultaneamente a quantidade consumida de cada atributo e seu preço marginal. O primeiro aponta que, na estimação de (4), conforme descrita por Rosen (1974), o erro seria composto por um componente relacionado aos gostos do consumidor e um componente aleatório. Se a função de preços hedônicos não for linear, os preços são endógenos e correlacionados ao componente do erro associado aos gostos do consumidor. Os instrumentos propostos por Rosen (1974), quais sejam, os preços marginais ($\hat{p}_i(z)$) produziram resultados viesados. De forma semelhante, Epple (1987) considera que, se o tipo de produto que os consumidores desejarem relacionar-se sistematicamente com y e o tipo de produto vendido com β , no equilíbrio, as características dos compradores relacionam-se com as dos produtores de quem eles compram, e os instrumentos propostos por Rosen (1974) geram resultados não consistentes.

Batik (1987) propõe a utilização de instrumentos que movam a restrição orçamentária do consumidor ou variáveis que movam a função hedônica. Epple (1987) analisa as condições sob as quais a identificação e a estimação de modelos de equilíbrio hedônico são factíveis. Ao longo do tempo, contudo, observa-se que muitos trabalhos evitaram a estimação do segundo estágio do modelo de Rosen (1974), atendo-se a estimativa dos preços hedônicos. Dessa

forma, não foi preciso enfrentar os problemas relacionados à identificação do referido estágio, destinado a estimar a propensão marginal a pagar de consumidores e identificar a oferta e demanda. Paralelamente, algumas propostas foram feitas para a estimação consistente desse estágio. Bishop e Timmins (2011), por exemplo, argumentam que apenas o rearranjo da equação estimada no segundo estágio de (ROSEN, 1974), de forma a isolar a variável endógena (z_i), seria suficiente para a estimação consistente da wtp e das equações de oferta e demanda. Greenstone (2017) cita algumas tentativas de se construir estimativas robustas para o segundo estágio do modelo discutido, mas todas requerem hipóteses bastante fortes não tendo tido grande impacto na literatura.

O arcabouço desenvolvido por Rosen (1974) foi utilizado para a estimação dos preços hedônicos e, por vezes, da demanda em diversos setores: no setor imobiliário, como mencionado acima (FÁVERO, 2011; GIBBONS, MOURATO e RESENDE, 2014), no setor aéreo (SENGUPTA e WIGGINS, 2006), de vinhos (COMBRIS, LECOQ e VISSER, 1997), entre outros. Na saúde, Jensen e Morrisey (1990) estimaram os efeitos marginais de diferentes características de seguros saúde nos preços que as firmas pagam por estes por meio de regressões de preços hedônicos e Basu e Sullivan (2017) analisaram a possibilidade de utilizar esse arcabouço para desenhar uma estrutura de avaliação de tecnologias e serviços de saúde.

Os primeiros estimaram os preços hedônicos para o segmento de seguro saúde norte-americano *fee for service* (FFS) contratado por empregadores. Nesse tipo de seguro saúde, os segurados são reembolsados de suas despesas médicas e têm grande liberdade de escolha. De acordo com os autores, as seguintes características variam nos planos FFS: extensão dos serviços cobertos; provisões aplicáveis de compartilhamento de custos e os limites para despesas diretas (a partir do qual, a seguradora responsabilizar-se-ia completamente pelo pagamento dos serviços). Os autores não estimaram a wtp associada aos atributos desses seguros porque a regulação na maior parte dos estados americanos obriga a inclusão de atributos específicos (coberturas especiais) nos produtos. Em decorrência dessas obrigações, os consumidores podem adquirir produtos com atributos que, em outras condições, não adquiririam. Assim, não seriam atendidas as condições de identificação necessárias ao modelo de Rosen (1974).

A equação estimada por Jensen e Morrisey (1990) é

$$P_i = \beta_0 + B_i\beta_i + G_i\gamma_i + U_{1i}, \quad (6)$$

Na qual i indexa os planos, P_i é o prêmio mensal, B_i é um vetor dos atributos (benefícios providos), G_i é um vetor das características do grupo assegurado e U_{1i} é um erro aleatório.

A principal dificuldade para estimação de (6) é que P_i só é conhecido quando o empregado tem de contribuir para o acesso ao plano. Apenas em 20% dos planos de titulares e em 41% dos planos de dependentes, os beneficiários contribuem. A solução encontrada foi utilizar o modelo de seleção de amostra de Heckman (1980). A contribuição para o plano é modelada a partir de um probit, com os seguintes preditores: existência de acordo sindical; média dos salários na indústria; tamanho da firma; dados demográficos dos trabalhadores e um conjunto de *dummies* para a indústria, região e ano. Assim,

$$C_i = 0 \text{ se } z_{i,\alpha} < U_{2,i} \quad (7)$$

$$C_i = 1 \text{ se } z_{i,\alpha} \geq U_{2,i} \quad (8)$$

C_i é a ocorrência de contribuição e $z_{i,\alpha}$ são as variáveis preditoras. Se U_{1i} e $U_{2,i}$ são distribuídas normalmente, com média 0 e variâncias, respectivamente, σ_1^2 e 1, e têm correlação ρ , então a equação de preços para a subamostra com observações completas é

$$P_i = \beta_0 + B_i\beta_i + G_i\gamma_i + \rho\sigma_i \left[-\phi \left(\frac{z_{i,\alpha}}{\Phi(z_{i,\alpha})} \right) \right] + \epsilon_i \quad (9)$$

$$\epsilon_i = U_{1i} - \rho\sigma_i \left[-\phi \left(\frac{z_{i,\alpha}}{\Phi(z_{i,\alpha})} \right) \right] \quad (10)$$

$$E(\epsilon_i / z_{i,\alpha} \geq U_{2,i}) = 0 \quad (11).$$

ϕ e Φ são, respectivamente, a função de densidade de probabilidade e a função de distribuição acumulada de uma normal.

O procedimento utilizado no artigo de Jensen e Morrisey (1990) foi:

1. Estimação do modelo de seleção por máxima verossimilhança ((7) e (8));
2. Computação de $\hat{\lambda} = -\phi \left(\frac{z_{i,\alpha}}{\Phi(z_{i,\alpha})} \right)$ para cada plano;
3. Estimação dos preços, substituindo $\hat{\lambda} = -\phi \left(\frac{z_{i,\alpha}}{\Phi(z_{i,\alpha})} \right)$.

O resultado é consistente para β , γ e σ_1^2 .

Os autores concluem que os prêmios são afetados de diferentes formas por seus atributos. Provisões para contenção de custos, como o compartilhamento de riscos, são efetivas para a diminuição de prêmios. Paralelamente, planos que cobrem doenças psiquiátricas e dependência química têm prêmios mais altos, mesmo controlando para as outras coberturas e as características do grupo assegurado. Esse resultado, contudo, pode ser decorrente de um viés de auto seleção não controlado no modelo.

Além disso, planos constituídos pelas próprias empresas têm custos 13% maiores que os de planos contratados com as mesmas características. Nesse caso, os resultados também podem ter sido influenciados pela seleção. Quando há planos ofertados por HMOs (Health Maintenance Organizations, organizações que oferecem cobertura em uma rede pré-estabelecida, em geral, verticalizada), os planos FFS são mais caros, o que indica que as pessoas mais saudáveis escolhem os planos da HMOs. É preciso considerar, no entanto, que não se controlou o grau de urbanização da localização das firmas, o que pode influenciar na escolha por diferentes tipos de planos (HMOs estão presentes em áreas mais urbanizadas).

Por fim, os autores consideram que seus resultados são coerentes com o observado no mercado de seguros de saúde. Ainda que esses produtos sejam bens complexos, podem ser caracterizados como um conjunto de provisões. Quem procura respostas para os custos dessas provisões podem encontrar nas equações de preços hedônicos um meio para sua obtenção.

Basu e Sullivan (2017) discutem o desenvolvimento de uma extensão dos princípios do modelo de preços hedônicos para o desenho de uma estrutura de avaliação de tecnologias e serviços em saúde. Haveria dois problemas para a implementação dessa proposta:

1. Na saúde, raramente, existe um mercado privado que funcione de forma que os preços observados reflitam as oportunidades dos custos de produção;
2. Dificilmente, os consumidores confrontam-se com os custos marginais de realização de um procedimento, pois muitos estão assegurados ou têm direito a subsídios governamentais.

A solução, conforme os autores, é estimar o valor do cuidado à saúde indiretamente, por meio da estimação da função de valor da compra do seguro de saúde. Restariam dificuldades relacionadas à existência de subsídios para a contratação de seguros e os desafios de se calcular os benefícios clínicos e a qualidade do cuidado (atributos dos serviços de saúde).

Outra questão considerada é que as quantidades consumidas de (z_1, \dots, z_n) não refletem apenas os atributos relevantes da assistência à saúde, mas também a probabilidade de se ficar doente, de modo que os valores estimados de (2) e dos preços hedônicos não refletiriam a wtp marginal no equilíbrio de mercado. Ainda assim, Basu e Sullivan (2017) consideram que o arcabouço proposto por Rosen (1974) poderia ser usado para delinear a distribuição do valor desses atributos na população. A avaliação de custo/benefício compararia o aumento nos prêmios totais decorrentes da inclusão da tecnologia no rol de procedimentos cobertos com o seu impacto no orçamento.

Recentemente, Finkelstein, Hendren e Shepard (2019) estimaram a disposição a pagar (wtp) de indivíduos de baixa renda por seguros de saúde. No artigo, não é utilizado modelo de preços hedônicos. Os autores utilizam as regras de elegibilidade para subsídios de planos de saúde contratados por meio da CommCare¹³ em Massachussetts para identificar a disposição a pagar por seguros de saúde da população de baixa renda por meio de regressão descontínua. Embora o artigo não utilize a metodologia que se está discutindo, é interessante observar a estratégia empírica para a estimação da wtp e os resultados obtidos.

Em Massachussetts, são concedidos subsídios para aquisição de seguro de saúde contratados por meio da CommCare para os mais pobres, havendo três níveis de subsídios conforme a renda do indivíduo. As discontinuidades entre os subsídios de uma classe para outra permitem a estimação de uma regressão descontínua (*regression discontinuity design*) para avaliação da disposição a pagar, já que, quando um indivíduo passa de uma faixa de renda para outra, o subsídio reduz de forma descontínua, tornando o preço do seguro mais alto proporcionalmente. Além disso, em um ano em particular (2011), foram ofertados três seguros por meio da CommCare. Dois desses seguros eram bastante semelhantes entre si,

¹³ CommCare é uma espécie de bolsa em que são negociados planos de saúde, subsidiados, para a população pobre de Massachussetts. Foi estabelecida em 2006, no âmbito da reforma da assistência à saúde.

mas o terceiro tinha menor rede assistencial, pior reputação e menor preço. Isso permitiu também se comparar a disposição a pagar da população elegível ao subsídio pelo nível de cobertura.

Partindo de um modelo no qual se pressupõe que todos preferem contratos mais generosos e com base nas estimativas realizadas, chegou-se à conclusão de que, ao longo dos dados utilizados, o custo médio dos indivíduos é maior que sua disposição a pagar. Em outras palavras, não havendo subsídios, a maior parte dos consumidores de baixa renda de Massachusetts não contrataria seguro de saúde. Comparando-se a disposição a pagar por planos com maior cobertura, também se observa que a maior parte prefere os planos com baixa cobertura e, mesmo se não houvesse seleção adversa (indivíduos com maior propensão à utilização dos serviços tendem a adquirir os seguros com maior cobertura, aumentando os prêmios destes), 70% dos indivíduos não comprariam os seguros com maior cobertura. Os autores atribuem esses resultados a um certo otimismo da população estudada de que não enfrentará intercorrências na saúde e à existência de assistência de saúde não paga (indivíduos que são atendidos, mas não pagam a conta) ou gratuita, o que diminuiria a disposição a pagar por seguros das classes mais baixas nos Estados Unidos.

É difícil encontrar configuração institucional semelhante a de Massachusetts que permita replicar o trabalho de Finkelstein, Hendren e Shepard (2019). Seus resultados, contudo, induzem reflexões importantes para o estudo do consumo de seguros de saúde, pois os modelos de aquisição de seguros têm como pressuposto que a disposição a pagar por um seguro é igual ao custo esperado da assistência do indivíduo. Se assim não for, não há aquisição de seguros sem o estabelecimento de subsídios. Para países nos quais a assistência à saúde baseia-se em políticas de seguro universal, se for observada situação semelhante à encontrada no referido artigo, torna-se necessário o estabelecimento de subsídios.

Dos artigos analisados, Jensen e Morrissey (1990) é o que mais se aproxima do objetivo deste trabalho, que é avaliar a dinâmica entre os elos da cadeia da saúde suplementar no Brasil. Como analisaram planos FFS, que funcionam por meio do reembolso de despesas assistenciais, a rede de prestadores não é um atributo do plano. Se o arcabouço proposto por Rosen (1974) fosse utilizado para a análise no setor de saúde no Brasil, seria imprescindível considerar a rede, pois, como mencionado anteriormente, a maior parte dos planos no Brasil provê assistência principalmente por meio de sua rede credenciada. Quanto maior a rede e

mais qualificada, melhor tende a ser a percepção do consumidor sobre o plano. Mesmo as seguradoras de saúde, cujos reembolsos são mais permissivos e generosos, têm a maior parte de suas despesas assistenciais realizadas em prestadores contratados.

Conhecer os preços hedônicos dos atributos de planos de saúde é um primeiro passo para o entendimento de como os preços dos planos são formados e qual a importância da rede, mas é a derivação da disposição a pagar dos consumidores pela inclusão de um prestador que permite estimar com maior rigor o poder de mercado de prestadores de serviços de saúde e os efeitos de atos de concentração no setor de prestação de serviços ou entre prestadores e operadoras no mercado à jusante.

Nesse sentido, Town e Vistnes (2001) foram os primeiros a estimar a disposição a pagar considerando as especificidades de um *option demand market*, ou seja, mercados nos quais o ofertante vende a possibilidade de acesso à uma rede de fornecedores para consumidores que não têm certeza se precisarão utilizar essa rede. Os mercados de planos de saúde são exemplos típicos de mercados desse tipo e foram utilizados pelos autores para estimação empírica de seu modelo.

A principal inovação de (TOWN e VISTNES, 2001) foi definir a disposição a pagar de acordo com o momento, haja vista que, em *option demand markets*, o consumidor adquire o produto para, somente depois, se necessário, acessar a rede contratada. Sendo assim, a disposição a pagar anterior à contratação do produto (disposição a pagar *ex-ante*) depende da probabilidade de o consumidor necessitar dos serviços prestados pela rede contratada, e, se for o caso, da utilidade associada à utilização desses serviços. A disposição a pagar *ex ante* é relevante quando se analisam mercados desse tipo porque o poder de mercado dos prestadores incluídos na rede de produtos se associa diretamente a ela.

Para calculá-la, contudo, é necessário calcular a *interim wtp*, nos termos de (CAPPS, DRANOVE e SATTERTHWAITTE, 2003), que é a disposição a pagar do consumidor quando já sabe a sua necessidade de utilização da rede (no caso de planos de saúde, quando tem uma doença ou condição diagnosticada). A *wtp ex-post* é a disposição a pagar derivada considerando-se que o consumidor já utilizou a rede assistencial, em outros termos, consumiu o produto que adquiriu. Esta é a *wtp* derivada por Rosen (1974) e outros.

A partir desses trabalhos iniciais, derivou-se ampla literatura que visa não apenas estimar a wtp *ex-ante*, mas utilizá-la para simular os efeitos de fusões e aquisições e estimar o poder de mercado de hospitais e de planos de saúde. A seguir, são revisados três artigos que, a partir desse conceito, constroem modelos que permitem estimar o poder de mercado de prestadores de serviços de saúde e de operadoras de planos de saúde, bem como realizar simulações para operações de fusões e aquisições em ambos os mercados (prestadores de serviços e operadoras).

A disposição a pagar de consumidores de planos de saúde pela possibilidade de acesso a um prestador é apresentada por Gaynor e Town (2012) como um meio, semi estrutural, de estimação do poder de mercado de prestadores hospitalares. A partir de um modelo em três estágios no qual, no primeiro, operadoras de planos de saúde e hospitais negociam para determinar a inclusão de um hospital na rede assistencial e os valores a serem pagos; no segundo, os pacientes escolhem os planos; no terceiro, os pacientes têm choque na saúde e escolhem o prestador com base nesse choque (doença ou outra condição que os façam precisar de atendimento médico-hospitalar) e na rede de seu plano, os autores associam à wtp com os preços negociados entre prestadores e operadoras. Vale a pena deter-se um pouco mais neste modelo.

Gaynor e Town (2012) consideram que quatro características são importantes na competição entre hospitais:

1. Pacientes com planos de saúde usam a rede assistencial do plano, que é, em geral, um subconjunto da rede assistencial do lugar;
2. Pacientes não pagam diretamente pelo cuidado;
3. A escolha do seguro é feita antes da necessidade do cuidado (*option demand market*);
4. Hospitais negociam com os seguradores para entrar na rede e o valor reembolsado.

No setor de planos de saúde, as operadoras competem com base em preços e na abrangência e qualidade da rede assistencial, nesse modelo.

A modelagem começa pelo terceiro estágio, no qual o paciente tem o choque na saúde e escolhe o prestador ao qual recorrerá. Existem H planos (h é o plano de cada paciente), N é a quantidade de hospitais, sendo J_h , o conjunto de hospitais na rede do plano ao qual está vinculado o paciente i . No começo do período, o paciente sabe seu choque na saúde, que é indexado por m ($m = 1, \dots, M$, sendo que cada m corresponde a uma doença, e 0 significa sem doença). A probabilidade *ex-ante* de cada choque é $(\rho_{i,1}, \rho_{i,2}, \dots, \rho_{i,m})$. Cada diagnóstico tem um peso (w_1, w_2, \dots, w_m) que descreve a sua importância relativa em termos de custo e de impacto na utilidade do consumidor.

Dado o diagnóstico m , o consumidor escolhe o hospital no qual será tratado. A sua utilidade depende das características do hospital, das suas próprias características, do tempo de viagem ao hospital e de um erro idiossincrático:

$$u_{i,j,m} = w_m [f(x_j, z_i, d_{i,j}, m_i; \theta) + e_{i,j}] \quad (12)$$

Sendo, x_j um vetor com as características do hospital, z_i , vetor com as características do paciente, incluindo idade, sexo e raça, $d_{i,j}$ é a distância do paciente i ao hospital j , m_i é a doença, θ é o vetor de parâmetros a ser estimado, e $e_{i,j}$ é o erro idiossincrático.

Na prática, $u_{i,j}$ é parametrizada para ser uma função linear. O paciente também pode escolher um hospital fora da rede ($J(0)$)¹⁴. Modelando-se a probabilidade do indivíduo i , com a doença m , buscar cuidado no hospital j , como um *logit*, tem-se:

$$s_{i,j,m} = \frac{\exp(f(x_j, z_i, d_{i,j}, m_i; \theta))}{1 + \sum_{k \in J_h} \exp(f(x_j, z_i, d_{i,j}, m_i; \theta))} \quad (13)$$

A utilidade esperada *ex-ante* de um paciente considerando a rede J_h é:

$$w_i(J_h) = \sum_{m=1, \dots, M} (\rho_{i,m} w_m (\ln(1 + \sum_{j \in J_i} \exp(f(x_j, z_i, d_{i,j}, m_i; \theta))))). \quad (14)$$

No segundo estágio, os pacientes escolhem o plano de saúde. Há um conjunto de planos ($h = 1, \dots, H$) disponível para cada indivíduo. Cada plano tem um conjunto de hospitais na

¹⁴ Nas aplicações do modelo, usualmente são consideradas como alternativas fora da rede assistencial do plano prestadores que estão na mesma localidade do beneficiário ou que estão no mesmo mercado relevante dos prestadores analisados. No Brasil, além dos prestadores privados, poder-se-ia considerar os prestadores do sistema público de saúde como uma alternativa fora do plano. Assim, mesmo se a rede assistencial do plano abrangesse todos os prestadores privados locais que prestam serviços para planos de saúde, haveria uma alternativa fora da rede se existir prestador público.

sua rede (J_h) e as operadoras escolhem simultaneamente os prêmios de seus planos. Os atributos dos planos são dados por um conjunto de características ch . A utilidade esperada para o consumidor i do plano j é dada por:

$$v_{i,h} = w_i(J_h; \theta) + \gamma ch + \xi_h + \alpha^P \ln(\text{income}_i - \text{Prem}_h) + e_{i,h} \quad (15)$$

ξ_h captura as diferenças não medidas pelo economista na atratividade do plano h , γ e α^P são parâmetros do modelo, income_i é a renda do indivíduo, Prem_h é o valor do prêmio e $e_{i,h}$ é o erro. A probabilidade de que um indivíduo i adquiriria o plano h é:

$$y_{i,h}(J, \xi, \text{Prem}, \theta, \gamma, \alpha) = \frac{\exp(\tilde{v}_{i,h})}{\sum_h \exp(\tilde{v}_{i,h})} \quad (16)$$

J é uma matrix $H \times H$ com 0 e 1 que denota a rede de todos os planos. A participação de mercado do plano h é dada pela soma das probabilidades de se escolher o plano na população:

$$y_h(J, \xi, \text{Prem}) = \sum_i y_{i,h}(J_1, \dots, J_H, \text{Prem}). \quad (17)$$

O lucro auferido por um plano, considerada uma dada rede, é:

$$\pi_h(J, \xi, \text{Prem}) = (\text{Prem}_h - a_{ch} - \sum_j p_{j,h} q_{j,h}) y_h(J, \xi, \text{Prem}) - \text{Fixed}_h \quad (18)$$

a_{ch} é uma variável que designa os custos variáveis não assistenciais, $p_{j,h}$ são os preços dos tratamentos negociados com o hospital j , $q_{j,h}$, a quantidade de beneficiários que buscaram atendimento em j , e Fixed_h são os custos fixos do plano h . A estrutura de custos é conhecida por todos, e as operadoras escolhem os prêmios para maximizar seus lucros. A condição de primeira ordem para maximização do lucro, considerando as redes fixas, é

$$y_h(J, \xi, \text{Prem}) + (\text{Prem}_h - a_{ch} - \sum_j p_{j,h} q_{j,h}(J, \xi_h, \text{Prem})) \frac{dy_h(J, \xi, \text{Prem})}{d\text{Prem}} = 0 \quad (19)$$

Somente existe um equilíbrio, no qual todos os prêmios satisfazem a equação (19).

No primeiro estágio, operadoras de planos de saúde e hospitais negociam os preços dos tratamentos. Há um preço base por paciente i , e o preço pago por tratamento é a multiplicação desse preço base pelo peso do diagnóstico (w_m). São JH pares negociando os preços $p_{j,h}$. O custo marginal dos hospitais é c_j .

Para analisar a negociação, foca-se na solução de barganha de Nash, especificando os valores relacionados a acordos e desacordos. O valor de acordo para um hospital são as receitas líquidas provindas do fluxo de pacientes do plano. O valor do desacordo é uma receita fixa, dada por r_h . Para a operadora, o valor do acordo são as receitas brutas decorrentes de ter a rede J_h , diminuídas das despesas com o atendimento do paciente nos hospitais que constituem a rede.

Mantendo as redes fixas dos outros planos, a receita bruta de uma operadora é dada por:

$$F_h(J_h) = Prem_h(J_h, J_{-h}) y_h(J_h, J_{-h}, \xi, Prem_h^*), \quad (20)$$

Sendo J_{-h} a rede dos outros planos e $Prem_h^*$, um vetor de prêmios que resolve a condição de primeira ordem da maximização do lucro.

Definindo-se $J_h - j$ como a rede de h sem o hospital j :

$$F_h(J_h - j) = Prem_h^*(J_h - j, J_{-h}) y_h(J_h - j, J_{-h}, \xi, Prem_h^*). \quad (21)$$

O valor do não acordo para a operadora é

$$F_h(J_h - j) - \sum_k p_{k,j} q_{k,h}^{J-j}, \quad (22)$$

Sendo $q_{k,h}^{J-j}$ o fluxo de pacientes para a rede dada por $J_h - j$. Se o excedente líquido do acordo hospital-plano não for maior que zero, não há acordo.

Em resumo, os ganhos de hospitais e operadoras quando há acordo e quando não há são:

Hospitais:

$$\text{Acordo: } H_{agree} = (p_{jh} - c_j) q_{jh}(J_h)$$

$$\text{Desacordo: } H_{disagree} = r_j$$

Operadoras

$$\text{Acordo: } M_{agree} = F_h(J_h) - p_{jh} q_{jh} - \sum_{l \neq j} p_{lh} q_{lh} - cm_{jh}$$

$$\text{Desacordo: } M_{disagree} = F_h(J_h - j) - \sum_{l \neq j} p_{lh} q_{lh} (J_h - j).$$

No equilíbrio de Nash, os lucros de hospitais e operadoras são maximizados, dados os outros preços:

$$\max_p [H_{agree} - H_{disagree}]^\beta [M_{agree} - M_{disagree}]^{1-\beta}, \quad (23)$$

Sendo $\beta \in [0,1]$ a habilidade de barganha relativa.

Diferenciando e resolvendo a equação de equilíbrio para p_{jh} , encontra-se:

$$p_{jh} = (1 - \beta) \left(c_j + \frac{r_h}{q_{jh}} \right) + \frac{\beta}{q_{jh}} (F_h(J_h) - F_h(J_h - j) - cm_{jh}) + \beta \sum_{l \neq j} p_{lh} d_{jln}. \quad (24)$$

A solução do equilíbrio de Nash (eq. (24)) prevê que o preço cobrado por um hospital é crescente nos seus custos, em sua habilidade de barganha, nos preços dos competidores e no valor líquido que o hospital traz para a rede. As operadoras têm poucas ferramentas para afetar seu poder de barganha, sendo uma delas a orientação dos beneficiários no acesso à rede assistencial.

A partir do modelo descrito acima, pode-se desenvolver estratégias empíricas para analisar o impacto de casos de fusão no setor hospitalar, avaliando-se tanto o impacto nos preços negociados pelos hospitais com as operadoras quanto no preço final dos planos de saúde pagos pelos beneficiários.

Uma estratégia é partir da agregação da utilidade esperada ex-ante do consumidor para calcular a disposição a pagar dos beneficiários pela inclusão de um hospital na rede de um plano. Assumindo-se que a disposição a pagar relaciona-se com o valor que o hospital gera para a operadora, pode-se associá-la aos preços por ele negociados. Sendo assim, é possível estabelecer relações funcionais que permitam verificar o impacto nos preços de uma fusão no setor hospitalar.

Considerando o modelo apresentado, define-se $V_h(J_h)$ como a agregação da utilidade esperada *ex-ante* do consumidor (eq. (14)) para um conjunto relevante de beneficiários:

$$V_h(J_h) = \sum_{i \in I_h} W_i(J_h), \quad (25)$$

Sendo I_h o conjunto relevante de beneficiários. A disposição a pagar pela inclusão do prestador j na rede é dada por

$$WTP_j = V_h(J_h) - V_h(J_h - j). \quad (26)$$

A eq. (26) pode ser estimada com os dados de utilização do hospital pelos beneficiários.

Considerando a eq. (24), sabe-se que o valor que um hospital gera para uma operadora é

$$F_h(J_h) - F_h(J_{h-j}) \quad (27).$$

Os preços negociados pelo hospital são função da eq. (27) dividido pelo número esperado de pacientes daquela operadora.

1. Assume-se que a WTP_j (eq. (26)) é uma boa aproximação da eq. (27) ou que a relação funcional entre o preço e a eq. (27) é bem aproximada pela relação funcional entre o preço (ou lucro) e a WTP_j . Assim, pode-se estimar o poder de mercado de um hospital ou estimar os efeitos de uma fusão no setor. Em síntese, os passos para isso são:
2. Estima-se um logit multinomial referente à escolha do hospital, com dados de atendimento (internação) de pacientes, considerando suas características. Obtém-se uma estimativa para θ (eq. 12), parâmetro da utilidade do consumidor;
3. Tendo sido estimado θ e conhecendo-se os pesos dos diagnósticos (w_1, w_2, \dots, w_m), pode-se calcular diretamente a WTP_j , utilizando-se a eq. (14);
4. Com as estimativas do modelo, os valores de $W_i(J_h) - W_i(J_{h-j})$ e a WTP_j são calculados, considerando-se a probabilidade de cada doença para indivíduo, a partir da distribuição empírica dos dados;
5. Para se calcular o impacto no preço de uma fusão entre o hospital J e o hospital K , é preciso apenas calcular os novos valores da $WTP_{j+k} = V_h(J_h) - V_h(J_h - (J - K))$ e usar essas estimativas para encontrar os preços pós-fusão.

O modelo não prevê uma forma funcional específica para a relação entre a WTP e os preços (ou lucros), e os resíduos de uma regressão da WTP nos preços capturam todos os fatores que os afetam e que não foram incluídos na regressão, não havendo interpretação estrutural. Assim, é preciso fazer hipóteses para o correto tratamento dos resíduos.

HO e LEE (2017) constroem um modelo baseado na estimação da wtp e na derivação dos processos de barganha entre prestadores e operadoras e entre operadoras e

contratantes de planos de saúde. A wtp é definida como uma medida da utilidade esperada do consumidor relacionada à rede hospitalar de um plano. Essa medida é usada na estimação da demanda por planos. Esse modelo baseia-se na seguinte estrutura:

1a. Os empregadores¹⁵ e as operadoras negociam os prêmios ($\phi \equiv \{\phi_j\}_{j \in M}$, onde j é a operadora e M , o conjunto de operadoras e ϕ_j é o prêmio pago por domicílio);

1b. Simultaneamente todas as operadoras (j) e os hospitais (i) negociam os preços dos serviços hospitalares ($p \equiv \{p_{ij}\}_{ij \in G}$, sendo p_{ij} o preço pago ao hospital i pela operadora j para tratar os seus beneficiários e G é a rede atual de hospitais e operadoras;

2. Dada as redes hospitalares e os prêmios, os domicílios escolhem a operadora, determinando a demanda pela operadora j ($D_j(G, \phi)$);

3. Depois de aderir a um plano, cada indivíduo fica doente com uma probabilidade e quem fica doente visita algum hospital da rede. Isso determina a quantidade de indivíduos que visita o hospital i por meio da operadora j ($D_{ij}^H(G, \phi)$).

Assume-se que a negociação entre empregadores e operadoras tem o formato de barganhas simultâneas bilaterais de Nash, nas quais o empregador maximiza o bem-estar dos empregados líquido dos prêmios pagos. Os prêmios negociados (ϕ_j) para cada operadora j satisfazem:

$$\phi_j = \operatorname{argmax}_{\phi} [\pi_j^M(G, p, \{\phi, \phi_{-j}\})]^{\tau\phi} \times [W(M, \{\phi, \phi_{-j}\}) - W(M \setminus j, \phi_{-j})]^{(1-\tau\phi)}$$

$\forall j \in M$, (28)

Sendo M o conjunto de operadoras e ϕ_{-j} , o conjunto de prêmios excluído o prêmio da operadora j . O primeiro termo do lado direito da equação são os ganhos com o negócio (GFT_j^M) da operadora e o segundo, os ganhos com o negócio do empregador (GFT_j^E). Ambos os ganhos devem ser iguais ou maiores que zero para que haja negociação. O parâmetro de barganha do prêmio de Nash é representado por $\tau\phi \in [0,1]$. Se $\tau\phi = 1$, as operadoras,

¹⁵ O modelo considera o mercado de planos empresariais norte-americano, que compreende a maior parte dos planos vendidos naquele país. Note-se que, diferentemente do Brasil, nos Estados Unidos, os empregadores, em geral, contratam diversos planos com características diferentes, e os empregados podem escolher a qual dos planos vão aderir (podendo ou não pagar parte ou a totalidade dos prêmios).

simultaneamente, estabelecem preços que maximizam seus lucros e, se $\tau\phi = 0$, os empregadores pagam apenas o suficiente para que as operadoras cubram seus custos.

A mesma hipótese, de negociações simultâneas bilaterais de Nash, é feita para as negociações entre hospitais e operadoras em termo dos preços ($p_{ij} \in p$). Cada preço negociado maximiza o produto bilateral de Nash:

$$p_{ij} = \underset{p}{\operatorname{argmax}} [\pi_j^M(G, \{p, p_{-ij}\}, \phi) - \pi_j^M(G \setminus ij, p_{-ij}, \phi)]^{\tau j} \\ \times [\pi_i^H(G, \{p, p_{-ij}\}, \phi) - \pi_i^H(G \setminus ij, p_{-ij}, \phi)]^{(1-\tau j)} \quad \forall ij \in G \quad (29)$$

Sendo π_j^M , o lucro da operadora j , p_{-ij} , o conjunto de preços excluído o hospital i , π_i^H é o lucro do hospital i (considerado o conjunto de hospitais H) e τj é o parâmetro de barganha do preço dos serviços hospitalares de Nash. O primeiro termo da equação são os ganhos com o negócio da operadora, e o segundo, os ganhos com o negócio do hospital.

A partir dessas duas equações, podem-se derivar as condições de primeira ordem e estimar o impacto da competição nos prêmios e nos preços hospitalares. As especificações da demanda por serviços hospitalares e da demanda por planos serão apresentadas na modelagem empírica.

Ho e Lee (2017) usam o modelo para simular o impacto da redução na competição entre os três maiores ofertantes (Blue Shield of California – BS, Blue Cross – BC e Kaiser Permanente) de planos de saúde contratados por meio do sistema de Aposentadoria dos Empregados Públicos da Califórnia (CalPERS). Foram usados dados referentes a 2004, que abrangem informações sobre adesões aos planos, eventos e admissões hospitalares. Para os beneficiários da BS e da BC, os autores observaram também as escolhas de hospitais, o diagnóstico e os preços totais pagos pelas operadoras aos hospitais dada uma admissão.

Como em (GAYNOR e TOWN, 2012), o procedimento empírico começa pela caracterização da demanda individual por hospitais. É utilizado um modelo de escolha discreta que permite a consideração de diferenças de preferências entre os consumidores. Os consumidores são divididos em dez grupos, de acordo com a idade e o sexo. Cada categoria de consumidores κ requer admissão a um hospital com probabilidade γ_κ^α . Condicionada na admissão, o indivíduo recebe um diagnóstico $l \in L$, sendo L o conjunto de diagnósticos possíveis (cardíaco, câncer, neurológico, digestivo, parto e outros), com probabilidade $\gamma_{\kappa,l}$.

Um indivíduo k do tipo $\kappa(k)$ com diagnóstico l deriva a seguinte utilidade da visita ao hospital i :

$$u_{k,i,l,m}^H = \delta_i + z_i v_{k,l} \beta^z + d_{i,k} \beta_m^d + \varepsilon_{k,i,l,m}^H \quad (30)$$

Sendo δ_i os efeitos fixos do hospital, z_i , características observadas dos hospitais, $v_{k,l}$, características dos consumidores, $d_{i,k}$, é a distância entre a localização do consumidor k e o hospital i , m designa o mercado no qual o consumidor está. Os β são os parâmetros associados a essas variáveis e $\varepsilon_{k,i,l,m}^H$, é o termo de erro indissociado, por hipótese independente e identicamente distribuído. A partir dessa equação, pode-se estimar a probabilidade de um indivíduo k , que mora no mercado m , tem o plano da operadora j e o diagnóstico l , visite o hospital i . Essa estimação é feita por máxima verossimilhança.

Com base no modelo de demanda de hospitais estimada, é construída uma medida da utilidade esperada *ex ante* relacionada à rede hospitalar de uma operadora. Essa medida é a disposição a pagar (wtp) que será utilizada, posteriormente, na estimação da demanda por planos. A wtp de um indivíduo k relacionada à rede hospitalar oferecida pelo plano j é

$$WTP_{k,j,m}(G_{j,m}) = \gamma_{\kappa(k)}^a \sum_{l \in L} \gamma_{\kappa(k),l} \log \left(\sum_{h \in G_{j,m}} \exp(\hat{\delta}_i + z_h v_{k,l} \hat{\beta}^z + d_{h,k} \hat{\beta}_m^d) \right) \quad (31)$$

Pressupõe-se que, na escolha dos planos, os domicílios levem em consideração as preferências de todos os seus membros e que as escolhas possíveis sejam optar por planos de uma das três operadoras. Os planos da Kaiser Permanent são considerados a opção externa, ou seja, não são incluídos diretamente na equação que descreve a utilidade dos domicílios ou famílias, de forma que a utilidade de um domicílio ou família f do plano $j \in \{BS, BC\}$, no mercado m é dada por:

$$u_{f,j,m}^M = \delta_{j,m} + \alpha_f^\phi (0,2 \times \phi_j \Phi_{\lambda(f)} + \sum_{\forall \kappa} \alpha_\kappa^W \sum_{k \in f: \kappa(k)=\kappa} WTP_{k,j,m} + \varepsilon_{f,j,m}^M) \quad (32)$$

Sendo $\delta_{j,m}$ o efeito-fixo da operadora controlado pela rede e outras características, ϕ_j é o prêmio pago pela família, que é ponderado pela contribuição do consumidor de 20% (o restante é pago pelo empregador), $\lambda(f)$ é o tipo de família, que pode ser individual, casal ou família (casal com filhos), $\Phi_{\lambda(f)}$ é um vetor de prêmios associado a cada tipo de família ([1,2, 2.6]) e $\varepsilon_{f,j,m}^M$ é o termo de erro. Nos mercados em que planos da Kaiser estão disponíveis, a utilidade associada a seus planos é derivada por meio da equação

$$u_{f,Kaiser,m}^M = \alpha_f^\phi (0,2 \times \phi_{Kaiser} \Phi_{\lambda(f)}) + \alpha_m^K d_f^K + \varepsilon_{f,Kaiser,m}, \quad (33)$$

Sendo d_f^K a distância entre o consumidor e o hospital mais próximo da rede da Kaiser.

A probabilidade de uma dada família f escolher a operadora j é

$$\hat{\sigma}_{f,j,m}(\phi, G) = \frac{\exp(\hat{u}_{f,j,m}^M)}{\sum_{n \in M_{z(f)}} \exp(\hat{u}_{f,n,m}^M)} \quad j \in M_{z(f)}, \quad (34)$$

$M_{z(f)}$ denota o conjunto de operadoras na região em que mora a família f . As probabilidades são estimadas por máxima verossimilhança.

Definidas as utilidades, passa-se a construção dos objetos restantes necessários à estimação da demanda por hospitais e por planos e os preços dos serviços hospitalares. Como mencionado anteriormente, os dados utilizados no artigo abrangem apenas informações sobre o total pago por uma operadora a determinado hospital. Para construir uma medida do preço negociado entre cada par de operadora e prestador, o total pago ao hospital é dividido pela soma dos pesos associados a cada diagnóstico das admissões. Formalmente, o procedimento é o seguinte:

1. Considerando A_{ij} o conjunto de admissões observadas entre o hospital i e a operadora j , assume-se que, para qualquer admissão $a \in A_{ij}$, o pagamento total observado é $p_a^o = p_{ij}^* \times DRG_a + \varepsilon_a$, sendo p_{ij}^* o preço por admissão (considerando um peso do diagnóstico de 1). Esse valor é multiplicado pelo peso do diagnóstico associado à admissão a , e ε_a representa o termo com média zero associado a um choque no pagamento.
2. Assim, p_{ij}^* é $\hat{p}_{ij} \equiv (\sum_{\forall a \in A_{ij}} p_a^o) / \sum_{\forall a \in A_{ij}} DRG_a = p_{ij}^* + \varepsilon_{ij}^A$. Sendo $\varepsilon_{ij}^A \equiv \hat{p}_{ij} - p_{ij}^* = (\sum_{\forall a \in A_{ij}} \varepsilon_a) / (\sum_{\forall a \in A_{ij}} DRG_a)$.

Estimados os preços dos serviços hospitalares, é possível calcular o lucro da operadora j que atua no mercado m :

$$\pi_j^M(G, p, \phi) = \sum_m (\phi_j \Phi' D_{j,m}(\cdot) - D_{j,m}^E(\cdot) \eta_j - \sum_{h \in G_{j,m}^M} D_{h,j,m}^H(\cdot) p_{h,j}), \quad (35)$$

O primeiro termo do lado direito designa a receita da operadora, sendo $D_{j,m}(\cdot)$ um vetor com a quantidade de famílias de cada tipo beneficiárias na operadora j . O segundo

representa o total dos custos variáveis não hospitalares, sendo η_j esses custos unitários, e o terceiro termo representa o pagamento aos hospitais, sendo $D_{h,j,m}^H(\cdot)$ a demanda pelos serviços hospitalares dos beneficiários da operadora j , considerando cada hospital h .

A estimação de η_j é feita conjuntamente a de τ , por meio de mínimos quadrados generalizado (GMM) em dois estágios, sob a hipótese de que $E[w^n(\theta) \times Z^n] = 0$ para $n \in \{1,2,3\}$. A média do erro $\{w^n\}_{n \in \{1,2,3\}}$ é por construção 0, e $\{Z^n\}_{n \in \{1,2,3\}}$, que são os instrumentos, é uma constante e a quantidade de hospitais na rede de cada segurador. Assim, sendo $\theta \equiv \{\eta, \tau\}$, pode-se rescrever a condição de primeira ordem associada a eq. (28), usando a definição dos lucros da eq. (35), o que resulta em:

$$w_j^1(\theta) = \tau\phi \times \frac{\partial \pi_j^M}{\partial \phi_j} - (1 - \tau\phi) \times \left(\frac{\pi_j^M \times (\Phi' \widehat{D}_j(\cdot) + 0,8 \sum_{k \in M} \phi_k \Phi' \frac{\partial \widehat{D}_k(\cdot)}{\partial \phi_j})}{GFT_j^E(\cdot)} \right) \quad \forall j \quad (36), \text{ onde}$$

$$\frac{\partial \pi_j^M}{\partial \phi_j} = \Phi \times \widehat{D}_j(\cdot) + \phi_j \left(\Phi' \frac{\partial \widehat{D}_j(\cdot)}{\partial \phi_j} \right) - \frac{\partial \widehat{D}_j^E(\cdot)}{\partial \phi_j} \eta_j - \sum_{h \in G_j^H} \frac{\partial \widehat{D}_{h,j}^H(\cdot)}{\partial \phi_j} \hat{p}_{h,j}. \quad (36A)$$

Todos os termos dessas equações com exceção de θ , que contém η e τ , podem ser computados a partir das equações estimadas anteriormente. O GFT_j^E pode ser derivado a partir da utilidade do consumidor relativa à aquisição de planos eq. (32) e das hipóteses referentes aos choques de demanda. Assim,

$$GFT_j^E \equiv W(M) - W(M \setminus j) = \left(\sum_m \sum_{f \in F} \frac{1}{\alpha_f} \log \frac{\sum_{k \in M} \exp(\tilde{u}_{f,k,m}^M)}{\sum_{k \in M \setminus j} \exp(\tilde{u}_{f,k,m}^M)} \right) - 0,8 \sum_{k \in M} \phi_k \Phi' [\Delta_j^M \widehat{D}_k] \quad (37)$$

Nessa equação, $\tilde{u}_{f,k,m}^M$ é a estimativa de eq. (32) e $\Delta_j^M \widehat{D}_k$ é a mudança na adesão de beneficiários k em todos os tipos de famílias e mercados quando a operadora j é removida do conjunto de escolhas do empregador. O segundo termo do lado direito da equação é multiplicado por 0,8, que é a contribuição do empregador para os prêmios.

São derivadas as condições de primeira ordem do processo de negociação entre hospitais e operadoras. Derivado todo o modelo, além das estimativas, como custos marginais e os parâmetros de barganha de Nash, é simulado o impacto de remover uma operadora do conjunto de escolhas dos consumidores, que, como afirmado anteriormente, é o objetivo do artigo. São feitos exercícios em que são removidos a Kaiser Permanent e a Blue

Cross. Além disso, são simulados os resultados se as operadoras não tivessem que barganhar com os empregadores e a determinação dos prêmios fosse feita como em Nash-Bertrand. Na simulação em que a Kaiser é removida, os prêmios sobem aproximadamente 14% a 17%. Os preços dos serviços hospitalares não sobem. Já no caso da remoção da Blue Cross, há a diminuição dos prêmios e dos pagamentos feitos aos hospitais.

Para além dos resultados das estimativas e das simulações, esse trabalho é importante porque representa um passo além em relação ao artigo de Gaynor e Town (2012) tendo sido construído modelo estrutural, o que permite a interpretação de cada variável estimada e da relação entre elas, inclusive da wtp.

Posteriormente, Ho e Lee (2019) apresentam modelo muito semelhante ao apresentado em (HO e LEE, 2017), mas no qual, ao invés de um processo de barganha Nash em Nash (Nash in Nash)¹⁶ na negociação entre prestadores de serviços e operadoras, consideram um processo *Nash in Nash with threat of replacement* (Nash em Nash com ameaça de substituição). Em síntese, esse também é um processo de negociação simultânea entre pares de operadoras e hospitais, mas as operadoras podem ameaçar excluir um prestador da sua rede e substituí-lo por um hospital que não estava na rede. Isso gera um limite endógeno, que depende da existência de substitutos ao prestador, aos preços resultantes das negociações Nash em Nash.

As redes de prestadores, nesse caso, são determinadas endogenamente. Enquanto na barganha Nash em Nash (HO e LEE, 2017), prevê-se a formação de redes extensas, pois quanto maior a rede, menor a contribuição de cada hospital e menor o valor pago, no modelo com negociações *Nash in Nash with threat of replacement*, tal como descrita acima, a exclusão de um (ou vários) prestador (es) pode aumentar o lucro da operadora. Algumas razões consideradas pelos autores para que as operadoras diminuíssem suas redes hospitalares é *cream-skimming*¹⁷, orientação dos beneficiários para os prestadores com menores preços

¹⁶ Processo de barganha Nash em Nash (Nash in Nash bargain) é quando pares de firmas negociam simultaneamente (por exemplo, pares de prestadores de serviços hospitalares e operadoras). A negociação de cada par resulta em uma solução de Nash dado que todos os outros pares também alcançaram uma solução de Nash.

¹⁷ Cream-skimming é quando um ofertante seleciona os indivíduos que consumirão seus produtos, de forma a somente atender os indivíduos que gerem mais lucro. No presente caso, a hipótese dos autores é de que os consumidores mais propensos a ficar doentes buscariam planos com redes mais extensas. Dessa forma, reduzir a rede seria uma forma de evitar a adesão desses consumidores ao plano.

ou redução os preços negociados com os hospitais. Como a estrutura do modelo é muito parecida a do artigo sumariado anteriormente (HO e LEE (2017)), não será descrita neste documento.

Na análise empírica, a partir dos mesmos dados utilizados no artigo anterior, os autores usam uma mudança na rede assistencial proposta pela Blue Shield para verificar se os resultados previstos pelo modelo são coerentes com o ocorrido. Até 2004, a rede assistencial da Blue Cross abrangia quase todos os hospitais nos mercados em que atuava. A rede da Blue Shield também era bastante extensa e incluía quase todos os hospitais. Contudo, em junho daquele ano, a última requereu ao regulador californiano a exclusão de 38 hospitais de alto custo da sua rede no ano seguinte. A proposta não foi aceita, sendo excluídos apenas 24 hospitais. Os autores simulam os resultados se a proposta da Blue Shield tivesse sido aceita integralmente. Além disso, o ajuste do modelo apresentado é avaliado comparando-se os hospitais que o modelo prevê que seriam excluídos da rede com os hospitais que constavam da proposta de exclusão da operadora. Os resultados do modelo são coerentes com o ocorrido na Califórnia.

Do mesmo modo que o modelo anterior, este pode ser usado para simular o impacto nos preços pagos pelos consumidores, nos preços negociados de serviços hospitalares e no bem-estar dos diferentes grupos de agentes de mudanças no ambiente concorrencial tanto das operadoras de planos de saúde quanto dos hospitais. Além disso, como citado anteriormente, a principal inovação é que a rede é determinada endogenamente, portanto, também as consequências referentes aos tamanhos das redes podem ser estimados/simulados.

Todos os artigos revistos têm em comum a possibilidade de se estimar a disposição a pagar dos consumidores referente a determinada rede de um plano ou mesmo à inclusão de um prestador neste rede. As derivações feitas a partir do arcabouço construído por Rosen (1974) são úteis para se compreender o quanto os consumidores valorizam um atributo do plano e, por exemplo, se diferenciam um leito em um hospital certificado de um leito em um hospital comum, mas não permite extrair conclusões sobre os efeitos de fusões horizontais.

(ROSEN, 1974) contudo, tem duas limitações importantes. A primeira é que trata diferentes itens de um mesmo atributo como semelhantes. Assim, não se diferencia, por

exemplo, a inclusão de um hospital de outro qualquer. É possível contornar essa limitação tratando hospitais, de acordo com suas características, como atributos diferentes (por exemplo, quantidade de hospitais não acreditados e quantidade de hospitais acreditados). A segunda limitação desse modelo é que não é possível inferir a relação entre os preços dos planos e a wtp obtida, como é possível por meio do modelo de (GAYNOR e TOWN, 2012). Ocorrendo uma fusão no setor hospitalar, mas não havendo diminuição dos prestadores hospitalares ou da quantidade de leitos ofertadas, não há alteração nos preços de equilíbrio no arcabouço construído por Rosen (1974). Há que se considerar, também, que estimar consistentemente o segundo estágio desse modelo não é simples.

Os modelos que analisam especificamente a dinâmica no setor de saúde, utilizando o conceito de disposição a pagar *ex-ante*, como (GAYNOR e TOWN, 2012), (HO e LEE, 2017) e (HO e LEE, 2019) permitem estimar a disposição a pagar (wtp) *ex-ante* dos consumidores de planos de saúde por uma determinada rede hospitalar ou prestador e relacionar essa variável aos preços pagos para os prestadores de serviços pelas operadoras. O arcabouço desenvolvido em (GAYNOR e TOWN, 2012) pode ser usado para análise de concentrações horizontais no setor hospitalar ou mesmo de concentrações verticais, assumidas algumas hipóteses. É importante notar que os dados necessários para a estimação de modelos semelhantes a este nem sempre são facilmente obtidos.

Como o leitor percebeu, os artigos de HO e LEE (2017, 2019) têm muitas semelhanças com o de GAYNOR e TOWN (2012), razão pela qual, na introdução a esta revisão, afirmou-se que seriam sumariados dois conjuntos distintos de artigos, os primeiros relacionados ao modelo de ROSEN (1974) e o segundo relacionado à derivação da wtp *ex-ante*, como definida em (TOWN e VISTNES, 2001). As principais contribuições dos artigos que utilizaram esse conceito, após o trabalho inicial de Town e Vistnes (2001), em especial dos artigos revistos nesta seção ((GAYNOR e TOWN, 2012), (HO e LEE, 2017) e (HO e LEE, 2019)) é o desenvolvimento de modelos que analisam não apenas a escolha do beneficiário pelo prestador de saúde, mas também a escolha do plano e o processo de barganha entre prestadores de serviços e operadoras. Se em ((GAYNOR e TOWN, 2012) não é estabelecida uma forma específica para a relação entre o lucro dos prestadores e a wtp, em (HO e LEE, 2017) e (HO e LEE, 2019) são desenvolvidos modelos teóricos, o que permite estabelecer relações entre as variáveis e fazer simulações diversas. Diferentes atos de concentração

podem ser analisados a partir dessa metodologia. Todavia, como em GAYNOR e TOWN (2012), os dados necessários para isso podem ser de difícil obtenção.

3. Simulação do modelo apresentado em GAYNOR e TOWN (2012)

Ainda que na proposta de estimação apresentada por Gaynor e Town(2012) não seja definida forma funcional específica para a relação entre a disposição a pagar dos beneficiários pela inclusão de um prestador na rede assistencial de um plano e a receita líquida adicional para o plano decorrente dessa inclusão, a proposta permite estimar o impacto de fusões e aquisições de hospitais no setor suplementar de forma semiestruturada. Os resultados de estimativas desse tipo seriam um importante complemento à análise já feita pelo Conselho, em casos nos quais há concentração horizontal no setor hospitalar ou nos quais a verticalização pode implicar menor disponibilidade de leitos para os ofertantes não verticalizados de planos de saúde.

Nesta seção, será simulada a aplicação do procedimento descrito em GAYNOR e TOWN (2012) para a estimativa da disposição a pagar *ex-ante* dos beneficiários pelos serviços dos prestadores. O motivo pela qual se propõe a simulação é não haver dados suficientes nas bases de dados públicas para a realização dessas estimações. Contudo, com exceção das probabilidades de ocorrência de doenças e de seus pesos para os beneficiários, essas variáveis são coletadas pelo órgão regulador, de forma que, mostrando-se viável a proposta, esta poderá ser implementada posteriormente, em casos específicos, podendo os dados necessários serem requeridos dos interessados e de terceiros.

3.1. Dados

Para a simulação do procedimento em tela, foi construída uma base de dados composta de 1100 observações. Os beneficiários foram numerados de 1 a 1100. O sexo de cada beneficiário foi atribuído aleatoriamente, conforme uma distribuição de Bernoulli, com probabilidade = 0,50. As idades foram atribuídas também aleatoriamente, conforme uma normal, cuja média é 45 anos, e o desvio-padrão é 23,61. Esses parâmetros foram baseados na distribuição etária dos beneficiários que realizaram algum procedimento custeado pelas operadoras de planos de saúde no estado de São Paulo, em 2017, conforme os dados

divulgados no portal brasileiro de dados, coletados no âmbito do padrão da Troca de Informações na Saúde Suplementar (TISS)¹⁸. Se o valor atribuído aleatoriamente para a idade era menor que zero, ele foi convertido a zero e, se era maior que 100, foi convertido a 100.

Os 1100 beneficiários hipotéticos podem ter tido três diferentes tipos de choque de saúde: 1; 2 e 3, cujas probabilidades de ocorrência na amostra são: 5%; 15% e 80%. A partir dessas probabilidades, os choques de saúde foram atribuídos aos beneficiários.

Ao constatar um choque em sua saúde, os beneficiários podem recorrer a um prestador da rede assistencial de seu plano ou a um prestador não vinculado (particular ou público). Nesta simulação, três prestadores compõem a rede assistencial do plano (a, b e c) e a opção do recurso ao prestador fora da rede assistencial e denominada por y. As probabilidades de um beneficiário recorrer a um ou outro prestador relacionam-se ao tipo de choque de saúde. Na tabela 1, abaixo, são apresentadas as probabilidades conforme o choque de saúde.

Tabela 1 – Probabilidades Simuladas de Choques na Saúde

		Prestadores			
		A	B	C	Y
Choques de saúde	1	70%	20%	0	10%
	2	40%	30%	20%	10%
	3	35%	35%	20%	10%

Elaboração dos autores.

As distâncias entre as residências dos beneficiários e a localização dos prestadores aos quais recorreram foram atribuídas aleatoriamente, conforme distribuições normais. Supôs-se que os beneficiários, na média, percorrem distâncias crescentes de acordo com o prestador, como pode ser observado na tabela 2, abaixo. Se as distâncias atribuídas aleatoriamente eram menores que 0, foram convertidas a zero.

Tabela 2 – Média e Desvio-Padrão das Probabilidades Simuladas de Distribuição

Prestador	Média da distribuição	Desvio-Padrão
A	20	5
B	15	3
C	10	2
Y	10	2

Elaboração dos autores.

¹⁸ <http://dados.gov.br/dataset/procedimentos-hospitalares-por-uf>. Acesso em 22/10/2009.

O prestador A é acreditado. Os prestadores fora da rede assistencial das operadoras podem ser acreditados. Uma distribuição de Bernoulli, com probabilidade de 50%, foi assumida para a acreditação dos prestadores fora da rede assistencial do plano.

Os dados finais são apresentados nas tabelas 3 a 6, abaixo, conforme o hospital em que o beneficiário foi atendido. Pode-se observar que a característica que mais apresentou variação conforme o hospital é a distância percorrida (tabela 6), o que era esperado, dada a forma como foi construída a amostra. A distribuição das doenças, de acordo com os prestadores, também apresentou variações (tabela 3).

Tabela 3 – Quantidade de beneficiários atendidos em um hospital por tipo de doença					Tabela 4 – Quantidade de beneficiários atendidos em um hospital por sexo			
		Doenças					Sexo	
		1	2	3			Feminino	Masculino
Hospital	A	25	77	331	Hospital	A	200	233
	B	15	38	279		B	164	168
	C	0	32	178		C	104	106
	Y	3	33	89		Y	63	62
Tabela 5 – Idade Média dos beneficiários, conforme o hospital de atendimento					Tabela 6 – Distância média percorrida, conforme o hospital de atendimento			
		Idade Média					Distância média	
A		47			A		19,2	
B		42,6			B		15	
C		43,9			C		10,1	
Y		44,1			Y		9,85	

Elaboração dos autores.

3.2. Estimação

A estimação das probabilidades de os beneficiários recorrerem a determinado prestador pode ser feita por meio de uma regressão logística multinomial, como sugerem Gaynor e Town (2012). Os coeficientes de uma regressão deste tipo são pouco informativos e mesmo a direção do efeito não é totalmente determinada por eles. A probabilidade da escolha de um prestador é dada por:

$$P(m = j | x) = \frac{\exp(x\beta_j)}{1 + \sum_{h=1}^J \exp(x\beta_h)} \quad j = 1, \dots, J \quad .19 \quad (38)$$

A utilidade esperada ex-ante de um paciente considerando a rede J_h é:

$$w_i(J_h) = \sum_{m=1, \dots, M} (\rho_{i,m} w_m (\ln(1 + \sum_{j \in J_i} \exp(f(x_j, z_i, d_{i,j}, m_i; \theta))))). \quad (14)$$

Note-se que o denominador da eq. (38), pode ser usado diretamente para o cálculo da utilidade ex-ante de um beneficiário associada à rede assistencial J_h , por meio da eq. (14).

A partir da base de dados criada, foram usadas as variáveis idade e sexo dos beneficiários, acreditação do hospital, distância percorrida e tipo da doença para as estimativas das probabilidades de recurso a determinado prestador.

Os modelos foram estimados no *R: A language and environment for statistical computing*, com a utilização do pacote *nnet* (VENABLES e RIPLEY, 2002). Uma das variáveis a serem explicadas é normalizada, e os resultados das estimativas são apresentados em relação a esta variável. Assim, os resultados apresentados são de $\frac{P(m=j|x)}{P(m=1|x)}$. Usualmente, esta forma de apresentação é melhor porque tem uma interpretação direta. Os coeficientes indicam o quanto a probabilidade da escolha de j aumentaria (ou diminuiria) em relação a probabilidade da escolha de referência (denominador) se determinada variável fosse alterada. No presente caso, utilizou-se os hospitais fora da rede assistencial (y) como variável base. Como o somatório de todas as probabilidades têm de ser 1, $p(m = y | x) = \frac{1}{1 + \sum_{h=1}^J \exp(x\beta_h)}$. Assim,

$$p(m = j | x) = \frac{\exp(x\beta_j)}{1 + \sum_{h=1}^J \exp(x\beta_h)}, \text{ como visto na eq. (38).}$$

3.3. Resultados

Os resultados da equação completa mostram que os únicos coeficientes significativos são a idade na equação para o prestador C e a distância para os prestadores A e B. Mesmo o tipo de doença, cujas distribuições diferiam de forma que um beneficiário tende a buscar mais

¹⁹ Note-se que essa equação é a equação (13) ($s_{i,j,m} = \frac{\exp(f(x_j, z_i, d_{i,j}, m_i; \theta))}{1 + \sum_{k \in J_h} \exp(f(x_j, z_i, d_{i,j}, m_i; \theta))}$), sendo x todas as variáveis a serem estimadas.

ou menos um prestador de acordo com a doença, não foram significativos, como pode ser observado na figura 3, abaixo.

Figura 3 – Resultado da estimação da probabilidade de utilização dos serviços de um prestador

=====			
Dependent variable:			
	a	b	c
	(1)	(2)	(3)

idade	0.001 (0.007)	0.003 (0.006)	0.005 (0.006)
sexo	-0.244 (0.338)	-0.227 (0.284)	0.184 (0.256)
acred	2.252*** (0.386)	-12.133*** (0.002)	-17.231*** (0.00000)
distancia	0.683*** (0.067)	0.780*** (0.063)	0.056 (0.056)
doenca1	-0.448 (0.921)	-0.289 (0.798)	-13.787*** (0.00000)
doenca2	-1.804*** (0.517)	-1.410*** (0.379)	-0.884*** (0.308)
Constant	-8.806*** (0.936)	-7.829*** (0.780)	0.301 (0.609)

Akaike Inf. Crit.	1,170.299	1,170.299	1,170.299
=====			
Note:	*p<0.1; **p<0.05; ***p<0.01		

Elaboração dos autores.

A equação estimada prevê corretamente 79% das observações, como pode ser visto na tabela 7, abaixo, mas a previsão de y é particularmente ruim. Para essa alternativa, o modelo prevê corretamente apenas 18% das observações. A porcentagem de observações previstas corretamente não é consideravelmente alterada quando são mantidas apenas as variáveis cujos coeficientes foram significativos (idade e sexo).

Tabela 7 – Quantidade de observações na amostra x quantidade de observações previstas

		Dados			
		A	B	C	Y
Previsto	A	389	0	0	21
	B	23	291	43	14
	C	4	40	166	68
	Y	17	1	1	22

Elaboração dos autores.

3.4. Cálculo da disposição a pagar

Para calcular as utilidades *ex-ante* de cada beneficiário é preciso conhecer a probabilidade de ocorrência de cada doença (ρ) e seu peso (w) na utilidade. Nesta simulação, com o objetivo de simplificar, foi estabelecido que as probabilidades de ocorrência de cada choque de saúde e os seus pesos são os mesmos para todos os beneficiários. As probabilidades de ocorrência dos choques de saúde são as mesmas definidas anteriormente, quando a base de dados foi montada, sendo 5% 15% e 80% para os tipos 1, 2 e 3, respectivamente. Considerando que, usualmente, as doenças mais prevalentes têm menores custos de tratamento e as mais raras, maiores, e que também em termos de bem-estar, as condições mais comuns tendem a ter menor impacto, foram atribuídos os seguintes pesos as condições, 1, 2 e 3, respectivamente, 1000, 100 e 10.

Para o cálculo da wtp de um beneficiário, é preciso calcular as utilidades associadas à possibilidade de utilização de cada prestador. Considerando-se os resultados apresentados na figura 3, as utilidades de um beneficiário i associadas às possibilidades de utilizar os prestadores A, B e C são:

$$U_{Ai} = -8,806 + 0,001 \times idade(i) - 0,244 \times sexo(i) + 2,252 \times creditação \\ + 0,683 \times distancia(i) - 0,448 \times doenca1(i) - 1,804 \times doenca2(i)$$

$$U_{Bi} = -7,829 + 0,003 \times idade(i) - 0,227 \times sexo(i) - 12,133 \times creditação \\ + 0,780 \times distancia(i) - 0,289 \times doenca1(i) - 1,410 \times doenca2(i)$$

$$U_{Ci} = 0,301 - 0,005 \times idade(i) + 0,184 \times sexo(i) - 17,231 \times creditação \\ + 0,056 \times distancia(i) - 13,787 \times doenca1(i) - 0,884 \times doenca2(i)$$

Tendo em conta que, nesta simulação, todos os beneficiários têm as mesmas probabilidades de terem as doenças 1, 2 ou 3 (5%, 15% e 80%, respectivamente) e que os pesos dessas condições são 1000, 100 e 10, $\sum_{m=1, \dots, M} \rho_{i,m} w_m = 73$ (ver eq. (14)) para todos os beneficiários. Assim, a utilidade esperada do beneficiário i associada à rede do plano composta por A, B e C, tendo em conta a eq. (14), é

$$U(i) = 73 \times (\log(1 + \exp(U_{Ai}) + \exp(U_{Bi}) + \exp(U_{Ci})))$$

Somando-se a $U(i)$ de todos os beneficiários do plano que contém os três hospitais, que, por definição é a wtp de cada beneficiário, encontra-se a wtp agregada associada à rede do plano. Nessa simulação, o valor da wtp agregada é 330930. Este número, isoladamente, tem pouco significado. No próximo item, serão discutidas possibilidades de utilização da wtp na análise de atos de concentração, mas uma possibilidade direta de utilização é a comparação entre a wtp tendo em vista diferentes redes assistenciais. Por exemplo, se for assumida que as utilidades associadas aos prestadores A e B não se alteram com a exclusão do hospital C da rede assistencial do plano, a wtp se o hospital C estiver fora da rede assistencial seria 314361, uma redução de 5%.

3.5. Utilização da wtp na análise de atos de concentração

Nos artigos estudados, desde (TOWN e VISTNES, 2001) assume-se ou, no caso dos artigos mais recentes, como (HO e LEE, 2017) e (HO e LEE, 2019), deriva-se uma relação entre a wtp e o lucro dos prestadores de serviços de saúde e também entre a wtp e os preços dos prestadores. É a partir dessas relações que a wtp costuma ser utilizada na análise de atos de concentração. Mas há também outras possibilidades.

Em atos de concentração vertical, se a autoridade antitruste considerar que o bem-estar dos consumidores é relevante na análise, pode usar a wtp diretamente para aferir o impacto da operação no bem-estar dos beneficiários. Imagine, por exemplo, que se está propondo a aquisição de um prestador por uma operadora. Presume-se que, a partir da operação, todos os beneficiários da operadora terão acesso àquele prestador e que os demais beneficiários não mais terão acesso²⁰. Calcula-se a wtp de antes da operação e a de após a operação e

²⁰ No Brasil, a Lei 9.656, de 1998, proíbe o estabelecimento de contratos de exclusividade entre prestadores e operadoras. Ainda assim, é possível que, a partir de um ato de verticalização, o prestador verticalizado passe a atender mais beneficiários da operadora a qual está vinculado e, por conseguinte, retem menos leitos ou outros recursos para o atendimento de beneficiários de outras operadoras. É provável que os contratos do prestador

comparam-se os resultados. Se a posterior for igual ou maior, o ato não afetaria o bem-estar dos consumidores. Pode-se, ainda, continuar a análise, considerando-se os demais estágios dos modelos de (GAYNOR e TOWN, 2012), (HO e LEE, 2017) e (HO e LEE, 2019), simulando-se o impacto nos preços dos serviços dos prestadores e dos planos de saúde e o novo equilíbrio resultante.

Em atos de concentração horizontal em mercados de prestação de serviços de saúde, a análise da relação da *wtp* com os lucros e preços dos prestadores pode ser considerada para avaliação do ato. Capps, Dranove e Satterthwaite (2003) sugerem que uma forma de avaliar se os prestadores são concorrentes próximos, ou seja, fazem parte do mesmo mercado relevante, é calcular a *wtp* dos consumidores antes da fusão e posteriormente à fusão. Supondo-se que, após a operação, os prestadores antes concorrentes passariam a negociar conjuntamente a adesão às redes dos planos, a exclusão do prestador da rede significaria a exclusão dos dois prestadores que antes eram concorrentes. De acordo com os autores, a perda na *wtp* decorrente da exclusão dos dois conjuntamente pode ser maior que a soma da perda na *wtp* associada a cada prestador isoladamente, o que indicaria que os prestadores envolvidos na operação são substitutos próximos.

A intuição por trás dessa possibilidade é simples. Considere um plano cuja rede é composta pelos prestadores A, B e C, sendo que A e B prestam os mesmos serviços e são localizados perto um do outro. Se A sair da rede, a *wtp* dos beneficiários diminui, mas não tanto, já que B, muito similar a A, continua na rede. Se B sair, o mesmo ocorre. Contudo, se A e B se fundirem, de forma a negociarem conjuntamente com a operadora, se forem excluídos da rede, a *wtp* dos beneficiários dos planos será consideravelmente diminuída, já que restará apenas C, que não é um substituto tão próximo de A e B. Assim, o todo é (ou pode ser) maior do que a simples soma das suas partes, ou seja, a diferença da soma das *wtp* relacionadas a A e B e a *wtp* relacionada ao prestador resultante da fusão de A e B é uma medida da substitubilidade dos prestadores que pode ser usada, por exemplo, para definir mercado relevante.

fossem revistos. Aqui, para simplificar, considerar-se-á que todos os contratos com outras operadoras seriam rompidos.

No modelo que foi simulado, se for considerado que as utilidades dos beneficiários não se alteram com a nova configuração de mercado, a soma da wtp associada a A com a wtp associada a B é 243.883. Já a wtp associada ao prestador resultante da fusão de A e B é 277.438. Esta é, portanto, maior que a soma em 33.554. Os prestadores, considerando-se essa estimativa, são concorrentes, e os preços aumentariam após a fusão. Contudo, é importante considerar se os parâmetros das utilidades dos consumidores após a fusão continuam os mesmos. O mais prudente pode ser estimar os parâmetros da utilidade dos consumidores novamente, considerando que os hospitais fundiram-se e, a partir daí, calcular a nova wtp.

Também é importante observar que, enquanto Capps, Dranove e Satterthwaite (2003) defendem esse procedimento como um meio de definir mercados relevantes, Ho e Lee (2017, 2019) assumem um mercado relevante previamente e estimam a wtp considerando os ofertantes daquele mercado. Nas estimativas concernentes a um plano específico, as alternativas fora da rede são consideradas como a alternativa externa na estimação. O mesmo acontece em (GAYNOR e TOWN, 2012) (ver nota de rodapé 13), que têm como pressuposto que todos os prestadores estão em um mesmo mercado relevante.

Se o objetivo for apenas o cálculo da wtp, não há, a princípio, problema em considerar prestadores em outros mercados relevantes ou não se definir o mercado relevante anteriormente à estimação. Nesse caso, todavia, será preciso cuidado com a definição do que é a alternativa externa na estimação dos parâmetros da utilidade do consumidor, além de ser necessário ter a cautela de se reestimar os parâmetros da função de utilidade, como observado anteriormente caso se pretenda utilizar a wtp para definição de substitutos.

A diferença entre a soma das wtp associada aos prestadores que pretendem se fundir e a wtp do prestador resultante da fusão associa-se também ao aumento do poder de mercado dos prestadores. Considerando-se que a wtp é uma boa aproximação do valor que um hospital (ou outro prestador) gera para uma operadora (eq. (27)) e que os preços negociados são função desse valor, a diferença da wtp gerada por A e B, conjuntamente, comparada a soma da wtp gerada por A e da wtp gerada por B é uma medida do ganho de poder de mercado decorrente da operação.

Como visto anteriormente, em (GAYNOR e TOWN, 2012) não se estabelece uma forma específica para a relação entre os preços dos serviços dos prestadores de serviços e a wtp, mas se sabe que existe uma relação direta entre as duas variáveis. Assim, se, no caso de determinada operação, conseguir-se estabelecer uma forma consistente para a relação entre os preços e a wtp, sendo possível regressar aqueles na wtp, será possível, também, estimar o impacto nos preços de mudanças na configuração do mercado. Assumindo-se que os custos se manteriam constantes, seria possível simular, de outro modo, as mudanças no poder de mercado.

Essas são as formas mais simples de utilização da wtp na análise de atos de concentração. Havendo necessidade e dados, os demais estágios dos modelos discutidos ((GAYNOR e TOWN, 2012), (HO e LEE, 2017) e (HO e LEE, 2019)) poderiam também ser estimados. Nesse caso, diversas simulações seriam possíveis, podendo-se estimar o impacto nos preços dos planos de saúde, o poder de mercado das operadoras haja vista as mudanças esperadas, os resultados das barganhas entre hospitais e operadoras, em suma, o novo equilíbrio tanto no mercado de prestação de serviços quanto de planos de saúde.

4. Conclusão

Neste trabalho foram apresentadas duas formas diferentes de estimação da disposição a pagar de consumidores. O modelo de Rosen (1974), bastante utilizado em diversos mercados, permite estimar, assumidas certas hipóteses, os preços hedônicos, ou seja, o quanto um consumidor estaria disposto a pagar por determinada característica do produto/serviço. Uma hipótese desse modelo é de que o consumidor não diferencia a qualidade para uma mesma característica. Por exemplo, um leito de hospital teria o mesmo preço hedônico, independentemente do hospital no qual está. É possível, contudo, diferenciar os hospitais na estimação (leitos de hospitais acreditados x leito de hospitais não acreditados, por exemplo).

O modelo de Gaynor eTown (2012), bem como os de Ho e Lee (2017, 2019) foram formulados em um arcabouço de competição entre hospitais. Nesse caso, os hospitais são vistos como produtos/serviços diferenciados. Além disso, considera-se que a maior parte dos pacientes que acessam os hospitais têm planos de saúde, de forma que a escolha ocorre entre

os hospitais das redes assistenciais dos planos. A disposição a pagar dos beneficiários pela inclusão de um prestador na rede assistencial do seu plano estaria relacionada ao valor negociado entre o prestador e o plano de saúde. Sendo assim, a partir da primeira poder-se-ia medir o impacto nos preços de fusões e aquisições ainda que não se tenha definido uma relação funcional entre a disposição a pagar e os preços, como no caso de (GAYNOR e TOWN, 2012).

Como as estimações com base nos procedimentos realizados nesses artigos exigem base de dados ampla, foi feita uma simulação da proposta de Gaynor e Town (2012) e estudados os dados públicos disponíveis para verificar a viabilidade de sua aplicação na saúde suplementar. O regulador coleta a maior parte dos dados necessários para essas estimações, contudo, como não estão identificados os prestadores e beneficiários, não é possível usar os dados públicos para a aplicação desse modelo. Na estimativa da disposição a pagar, haveria ainda a dificuldade em relação à obtenção de dados e o cálculo da probabilidade de ocorrência de doenças e de seus pesos para os beneficiários. Para a estimativa da relação entre os preços e a disposição a pagar, também seriam necessários dados sobre os primeiros, que não são coletados sistematicamente pelo regulador, mas que podem ser obtidos pelo Cade.

O modelo de Rosen (1974) também poderia ser útil na análise de operações verticais nos setores de prestação de serviços de saúde e planos de saúde e mesmo de operações horizontais no setor de prestação de serviços. Sua vantagem é requerer menos dados. Os modelos ora estudados podem ser um complemento importante a outras análises empreendidas pelo Cade. À medida em que os mercados de prestadores de serviços de saúde e de planos de saúde, em muitos casos já concentrados, concentram-se ainda mais e que a relação entre grandes grupos dos diferentes elos da cadeia da saúde suplementar torna-se mais relevante na organização desses setores, compreender melhor as relações entre rede assistencial e planos de saúde torna-se imprescindível.

Bibliografia

ALMEIDA, Silvia. Poder compensatório e política de defesa da concorrência: referencial geral e aplicação ao mercado de saúde suplementar brasileiro. Tese apresentada à Escola de Economia de São Paulo da Fundação Getúlio Vargas, como requisito para obtenção do título de doutora, 2009.

ANAHP. Observatório 2019, ed. 11.

ANAHP. Painel – Saúde em números, 2019.

ARGARWAL, Nikhil. An empirical model of the medical match. *American Economic Review*, 2015.

BASU, Arniban, SULLIVAN, Sean. Toward a hedonic value framework in health care. *Value in Health*, 2017.

BATIK, Timothy. The estimation of demand parameters in hedonic price models. *Journal of Political Economy*, 1987.

BHATTACHARYA, Jay, ISEN, Adam. On inferring demand for health care in the presence of anchoring, acquiescence, and selection biases. Working Paper nº 13865, National Bureau of Economic Research, 2008.

BISHOP, Kelly, TIMMINS, Christopher. Hedonic prices and implicit markets: estimating marginal willingness to pay for differentiated products without instrumental variables. Working Paper no 17611, National Bureau of Economic Research, 2011.

Conselho Administrativo de Defesa Econômica (Cade). Atos de concentração no mercado de planos de saúde, hospitais e medicina diagnóstica. *Cadernos do Cade*, julho/2018.

CAPPS, Cory, DRANOVE, David, SATTERTHWAITE, Mark. Competition and market power in option demand markets. *RAND Journal of Economics*, 2003.

CHIN, Tuang-Leong, CHAU, K. A critical review of literature on hedonic price model. *International Journal for Housing and its applications*, 2003.

COLLARD-WEXLEY, Allan, GOWRISANKARAN, Gautam e LEE, Robin. Nash-in-nash bargaining: a microfoundation for applied work. National Bureau of Economic Research, working paper 20641, 2014 (revisado em 2017).

COMBRIS, Pierre, LECOQ, Sébastien, VISSER, Michael. Estimation for a hedonic price equation for Bordeaux wine: Does quality matter? *Economic Journal*, 1997.

EPPLÉ, Dennis. Hedonic prices and implicit markets: estimating demand and supply functions for differentiated products. *Journal of Political Economy*, 1987.

ERICSON, Keith, STARC, Amanda. Measuring consumer valuation of limited provided networks. *American Economic Review: Papers & Proceedings*, 2015.

FÁVERO, Luiz. Preços hedônicos no mercado imobiliário comercial de São Paulo: a abordagem da modelagem multinível com classificação cruzada. *Estudos Econômicos*, 2011.

FINKELSTEIN, Amy, HENDREN, Nathaniel, SHEPARD, Mark. Subsidizing health insurance for low-income adults: Evidence from Massachusetts. *American Economic Review*, 2019.

GAYNOR, Martin, HAAS-WILSON, Deborah. Change, consolidation and competition in health care markets. *Journal of Economic Perspectives*, 1999.

GAYNOR, Martin, TOWN, Robert. Competition in health care markets. Working Paper n° 12/282. The Centre for Market and Public Organisation, 2012.

GAYNOR, Martin, HO, Kate, TOWN, Robert. The industrial organization of health-care markets. *Journal of Economic Literature*, 2015.

GIBBONS, S., MOURATO, S., RESENDE, G.M. The Amenity Value of English Nature: A Hedonic Price Approach. *Environmental Resource Economics* 57, 175–196, 2014.

GOLDBERG, Daniel. Poder de compra e política antitruste. Editora Singular, 2006.

GREENSTONE, Michael. The continuing impact of Sherwin Rosen's "Hedonic Prices and Implicit Markets: Product differentiation in pure competition". *Journal of Political Economy*, 2017.

HLAVAC, Marek. Stargazer: Well-Formatted Regression and Summary Statistics Tables. R package version 5.2.1 (<https://CRAN.R-project.org/package=stargazer>).

HENDREN, Nathaniel. Measuring ex-ante welfare in insurance markets. Working Paper n° 24470, National Bureau of Economic Research, 2018.

HO, Katherine. Insurer-provider networks in the medical care markets. *American Economic Review*, 2009.

HO, Kate e LEE, Robin. Insurer competition in health care markets. *Econometrica*, 2017.

HO, Kate e LEE, Robin. Equilibrium provider networks: Bargaining and exclusion in health care markets. *American Economic Review*, 2019.

JENSEN, Gail, MORRISEY, Michael. Group health insurance: a hedonic approach. *The Review of Economics and Statistics*, 1990.

OCDE. Buying power: the exercise of market power by dominant buyers. Report of the Committee of Experts on Restrictive Business Practices, 1981.

OCZKOWSKI, Edward. Hedonic wine price functions with different prices. *Australian Journal of Agricultural and Resource Economics*, 2015.

PONTES, Eduardo, PAIXÃO, Luiz, ABRAMO, Pedro. O mercado imobiliário como revelador das preferências pelos atributos espaciais: uma análise do impacto da criminalidade urbana no preço de apartamentos em Belo Horizonte. *Revista de Economica Contemporânea*, 2011.

ROSEN, Shewin. Hedonic prices and implicit markets: product differentiation in pure competition. *The Journal of Political Economy*, 1974.

SENGUPTA, Arniban, Wiggins, Stevens. Airline pricing, price dispersion and ticket characteristics on and of the internet. Working Paper nº 06-07. Net Institute, 2006.

TOWN, Robert e VISTNES, Gregory. Hospital competition in HMO networks. *Journal of Health Economics*, 2001.

VENABLES, W. N., RIPLEY, B. D. *Modern Applied Statistics with S*. Fourth Edition. Springer, New York. ISBN 0-387-95457-0, 2002. WOOLDRIDGE, Jeffrey. *Econometric Analysis of Cross Section and Panel Data*. The MIT Press, 2002.

Anexo: Código da Simulação

```
options(stringsAsFactors = FALSE, scipen = 999)
```

Instalar o pacote dplyr, se não tiver instalado

```
library(dplyr)
```

Construção da base de dados

```
benef<-seq(1,1100, by=1)

set.seed(57)
sexo<-rbinom(1100, 1, 0.5)

idade<-rnorm(1100, mean= 45, sd=23.61)

for (i in 1:1100)
{if (idade[i]<0)
{idade[i]=0}}

for (i in 1:1100)
{if (idade[i]>100)
{idade[i]=100}}

dados<-data.frame(benef, idade, sexo)

doencas<-c(1,2,3)
cid<-sample(doencas, size=1100, replace=TRUE, prob=c(0.05, 0.15, 0.8))
table(cid)

prest<-c("a","b","c", "y")

adoenc1<-sample(prest, size=65, replace=TRUE, prob=c(0.70, 0.2, 0, 0.1))
adoenc2<-sample(prest, size=159, replace=TRUE, prob=c(0.4, 0.3, 0.2, 0.1))
adoenc3<-sample(prest, size=876, replace=TRUE, prob=c(0.35, 0.35, 0.2, 0.1
))

hospital<-rep("x", times=1100)
hospital[1:65]<-adoenc1
hospital[66:224]<-adoenc2
hospital[225:1100]<-adoenc3

dados<-data.frame(dados, cid)
dados<-arrange(dados, cid)
dados<-cbind(dados, hospital)

dados$acred<-rep(0, times=1100)

for (i in 1:1000)
{if (dados$hospital[i]=="a")
```

```

    dados$sacred[i]=1}

acred_out<-rbinom(100, 1, 0.5)

dados<-arrange(dados, hospital)
dados$sacred[1001:1100]<-acred_out

table(dados$hospital)
dist_a<-rnorm(396, mean= 20, sd=5)
dist_b<-rnorm(376, mean=15, sd=3)
dist_c<-rnorm(227, mean=10, sd=2)
dist_y<-rnorm(101, mean=10, sd=2)

for (i in 1:396)
{if (dist_a[i]<0)
  {dist_a[i]=0}}

for (i in 1:376)
{if (dist_b[i]<0)
  {dist_b[i]=0}}

for (i in 1:227)
{if (dist_c[i]<0)
  {dist_c[i]=0}}

for (i in 1:101)
{if (dist_y[i]<0)
  {dist_y[i]=0}}

dados<-arrange(dados, hospital)
dados$distancia<-rep(0, times=1100)
dados$distancia[1:396]<-dist_a
dados$distancia[397:772]<-dist_b
dados$distancia[773:999]<-dist_c
dados$distancia[1000:1100]<-dist_y

dados$doenca1<-rep(0, times=1100)

for (i in 1:1100)
{if (dados$cid[i]==1)
  {dados$doenca1[i]=1}}

dados$doenca2<-rep(0, times=1100)

for (i in 1:1100)
{if (dados$cid[i]==2)
  {dados$doenca2[i]=1}}

remove(adoenc1, adoenc2, adoenc3, benef, cid, dist_a, dist_b, dist_c,
       doencas, hospital, idade, prest, sexo, dist_y, acred_out)

```

Sumário dos dados finais

```
with(dados, table(hospital, cid))  
with(dados, table(hospital, sexo))  
idade<-dados %>% group_by(hospital) %>%  
  summarise(media=mean(idade))  
distancia<-dados %>% group_by(hospital) %>%  
  summarise(media=mean(distancia))
```

Regressão

Para estimar as regressões e para apresentar os resultados, vamos usar os pacotes nnet e stargazer

```
library(nnet)
```

```
library(stargazer)
```

Regressões estimadas:

```
dados$hospital<-as.factor(dados$hospital)  
dados$hospital<-relevel(dados$hospital, ref="y")  
  
eq1 <- multinom(hospital ~ idade + sexo + acred + distancia + doenca1 + do  
enca2, data = dados)  
  
eq2<-multinom(hospital ~ idade + distancia + doenca1 + doenca2 , data = da  
dos)
```

Resultados:

```
stargazer(eq1, type="text", out="equacao_1.txt")  
stargazer(eq2, type="text", out="equacao_2.txt")
```

Previsões

```
previsao<-table(predict(eq1), dados$hospital)  
previsao2<-table(predict(eq2), dados$hospital)
```

Calcular a WTP

Primeiro passo: calcular a utilidade associada a cada prestador

```
Ba<-rep(0, times=1100)
for (i in 1:1100)
{Ba[i]<-(-8.806+0.001*dados$idade[i]-0.244*dados$sexo[i]+2.252*dados$acred
[i] +
      0.683*dados$distancia[i]+0.448*dados$doenca1[i] - 1.804*dados$d
oenca2[i])}
```

```
Bb<-rep(0, times=1100)
for (i in 1:1100)
{Bb[i]<-(-7.829+0.003*dados$idade[i]-0.227*dados$sexo[i]-12.133*dados$acred
d[i] +
      0.780*dados$distancia[i]-0.289*dados$doenca1[i] - 1.410*dados$d
oenca2[i])}
```

```
Bc<-rep(0, times=1100)
for (i in 1:1100)
{Bc[i]<-(-0.301-0.005*dados$idade[i]+0.184*dados$sexo[i]-17.231*dados$acred
[i] +
      0.056*dados$distancia[i]-13.787*dados$doenca1[i] -0.884*dados$d
oenca2[i])}
```

Segundo Passo: Calcular somatório para cálculo da wtp do consumidor i associado à rede

```
base<-rep(1, times=1100)
base1<-rep(1, times=1100)
for(i in 1:1100)
{base[i]<-log(base1[i]+exp(Ba[i])+exp(Bb[i])+exp(Bc[i]))}
```

```
base2<-rep(1, times=1100)
for(i in 1:1100)
{base2[i]<-log(base1[i]+exp(Ba[i])+exp(Bb[i]))}
```

```
baseA<-rep(1, times=1100)
for(i in 1:1100)
{baseA[i]<-log(base1[i]+exp(Bb[i])+exp(Bc[i]))}
```

```
baseB<-rep(1, times=1100)
for(i in 1:1100)
{baseB[i]<-log(base1[i]+exp(Ba[i])+exp(Bc[i]))}
```

```
baseAB<-rep(1, times=1100)
for(i in 1:1100)
{baseAB[i]<-log(base1[i]+exp(Bc[i]))}
```

Terceiro passo: Considerar a doença (probabilidade de ocorrência e o peso para cálculo da utilidade ex-ante do beneficiário)

```
util<-rep(1, times=1100)
for(i in 1:1100)
{util[i]<-73*base[i]}

util2<-rep(1, times=1100)
for(i in 1:1100)
{util2[i]<-73*base2[i]}

utilA<-rep(1, times=1100)
for(i in 1:1100)
{utilA[i]<-73*baseA[i]}

utilB<-rep(1, times=1100)
for(i in 1:1100)
{utilB[i]<-73*baseB[i]}

utilAB<-rep(1, times=1100)
for(i in 1:1100)
{utilAB[i]<-73*baseAB[i]}
```

Cálculo da wtp

```
greg1<-sum(util)
greg2<-sum(util2)
gregA<-sum(utilA)
gregB<-sum(utilB)
gregAB<-sum(utilAB)

wtp<-greg1-greg2

wtpA<-greg1-gregA
wtpB<-greg1-gregB
wtpAB<-greg1-gregAB

soma<-wtpA+wtpB
dif<-wtpAB - soma
```