

UMA VISÃO MAIS TÉCNICA DO ALGORÍTMO DE SIMULAÇÕES DE FUSÃO

*Jorge Fagundes**, *Fabio Kanczuk***

1. Introdução

Este artigo apresenta uma visão relativamente técnica dos procedimentos envolvidos em uma simulação de fusão. De modo geral, nosso intuito foi o de mostrar algumas das principais dificuldades e problemas de cada etapa do processo, além de expor algumas das soluções e metodologias tipicamente utilizadas.

Conforme discutido em um outro artigo por nós elaborado¹, o algoritmo de simulações envolve duas etapas: (i) a análise prévia (*front end*), que consiste na estimação dos parâmetros da demanda e na especificação de sua forma funcional, e (ii) a análise posterior (*back-end*) no qual os parâmetros estimados são combinados com dados de participação de mercado e preços praticados para calcular os efeitos do ato de concentração sobre os preços.

Seguindo essa mesma divisão, este artigo segue o seguinte roteiro. A seção 2 trata da análise *front end*, descrevendo brevemente os dados e formas funcionais utilizadas e discutindo o problema do viés de simultaneidade presente na estimação econométrica, bem como algumas de suas soluções. A seção 3 trata da análise *back-end* e apresenta o sistema de equações empregado quando se utiliza o modelo Bertrand com bens diferenciados. A seção 4 é uma digressão sobre definição de mercado relevante e análise da elasticidade crítica. Finalmente, a seção 5 sugere algumas extensões futuras para a metodologia de simulações, expondo uma conclusão.

2. *Front end: Estimação econométrica da demanda*

A análise prévia (*front end*) de uma simulação de fusão é composta da escolha da forma funcional para a demanda e da estimação econométrica

* Doutor em economia/UFRJ e sócio da Fagundes Consultoria Econômica.

** Doutor em economia/UCLA e professor da USP.

¹ Ver Fagundes e Kanczuk (2004) “O USO DE MODELOS DE SIMULAÇÃO EM ANÁLISES DE FUSÕES”. Revista do Ibrac.

propriamente dita. As subseções a seguir apresentam, inicialmente, os dados geralmente disponíveis para a estimação e as formas funcionais mais tipicamente utilizadas. A seguir, discutem os problemas de natureza econométrica, além das possíveis soluções para remediá-los.

2.1. *Dados e formas funcionais*

Idealmente, um sistema de demanda seria estimado a partir de um rico banco de dados, capaz de produzir estimativas confiáveis de todas as elasticidades próprias e cruzadas relevantes, para todos os produtos do mercado em questão. Na prática, evidentemente, os dados disponíveis nunca são ideais, tanto no Brasil como nos outros países.

Normalmente, no Brasil, as estimações econométricas da demanda são baseadas em dados da Nielsen A.C.. Estes dados são apresentados com periodicidade bimestral, para sete regiões geográficas definidas pela própria Nielsen, e descrevem a evolução dos preços e quantidades vendidas para um conjunto de bens substitutos. Esse banco de dados é complementado com informações tais como índices de inflação – utilizados para deflacionar os preços -, e medidas de riqueza, tais como PIB, ou Consumo das Famílias. Por serem fornecidas por órgãos oficiais, tais como o IBGE, estas informações costumam ser facilmente obtidas.

Mesmo nas melhores situações, em que há dados para uma série de produtos substitutos, para várias regiões geográficas, e num horizonte de tempo relativamente extenso, a quantidade de informação contida nos dados tende a ser uma restrição crucial para o processo de estimação econométrica. Neste sentido, é fundamental lembrar-se do tradicional dilema entre incerteza e viés estatístico: restrições teóricas ao modelo reduzem o número de parâmetros a serem estimados, aumentando o grau de precisão estatística, mesmo que implicando em algum viés. Esse é o conceito que deve nortear a decisão sobre qual a forma funcional a ser utilizada.

Formas funcionais que exibem muitos parâmetros permitem uma enorme gama de padrões de substituição e complementariedade entre os bens, mas também demandam uma enorme quantidade de informação dos dados. Formas funcionais que dependem de poucos parâmetros aplicam restrições aos possíveis padrões de substituição entre os bens, mas demandam menos informações dos dados. A escolha da forma funcional busca equilibrar as vantagens e desvantagens deste *trade-off*.

Comumente, as quatro formas funcionais de demanda empregadas são: (i) demanda linear, (ii) demanda de elasticidade constante (log-linear), (iii) logit, (iv) AIDS ou sistema de demanda quase ideal (Almost Ideal Demand System). A formalização matemática destas quatro formas estão a seguir.

(i) Log-Linear

A forma funcional “log-linear” também é conhecida como elasticidades constantes. As equações do sistema de demanda têm o formato

$$\log(q_i) = \sum_j \varepsilon_{ij} \log(p_j)$$

Em que q_i e p_j denotam, respectivamente, quantidades do bem i e preços do bem j . E os parâmetros ε_{ij} representam a elasticidade (constante) do bem i com relação ao bem j .

(ii) Linear

No caso da forma funcional linear, as equações de demanda seguem o formato

$$q_i = q_i^0 + \sum_j m_{ij} p_j$$

Em que, novamente, q e p denotam, respectivamente, quantidades e preços.

Note-se que os parâmetros q_i^0 não são necessários para realizar as simulações. Observe-se, também, que há uma relação entre os parâmetros ε_{ij} (elasticidades, forma log-linear) e m_{ij} dada por:

$$\varepsilon_{ij} = (\partial q_i / \partial p_j)(p_j / q_i) = m_{ij} (p_j / q_i)$$

(iii) Logit

O sistema de demanda logit é geralmente motivado por um modelo de utilidade estocástica em que os consumidores fazem uma escolha discreta: um produto deve ser escolhido dentro de uma cesta com N alternativas. Quando a parte não observável das características do produto (variável estocástica na utilidade) é distribuída de acordo com uma distribuição de valor extremo, a probabilidade de escolha de um bem i assume a conhecida forma,

$$\pi_i = \exp(\alpha_i - \beta p_i) / \sum_{j=1}^n \exp(\alpha_j - \beta p_j)$$

Para ser utilizado em simulações de fusões, é conveniente expressar esse modelo em termos tradicionais de antitruste (vide Werden, Froeb e Tardiff (1996)). O produto N fica definido como a escolha “nenhuma das alternativas”, isto é, o bem externo (*outside good*). Assume-se que $p = 0$, para que a utilidade do bem externo seja constante. Denota-se a média ponderada (pelos *market shares*) dos bens internos por p^m .

A probabilidade de escolha de um bem interno, condicional na escolha ter sido de um bem interno, correspondem aos *market shares* tradicionais, e são dados por

$$s_i = \exp(\alpha_i - \beta p_i) / \sum_{j=1}^{n-1} \exp(\alpha_j - \beta p_j)$$

E elasticidade da indústria é definida por

$$\varepsilon \equiv (\partial \log \sum_{j=1}^{n-1} \pi_i(\lambda p_j) / \partial \lambda) \cdot p^m = -\beta p^m \pi_n$$

Note-se que há somente dois parâmetros no modelo: ε que controla a substituição entre o bem externo e os bens internos, e β que controla a substituição entre os bens internos.

(iv) AIDS

A forma de demanda AIDS recebeu este nome por satisfazer as propriedades da teoria do consumidor, tais como homogeneidade de grau zero e simetria (vide Deaton e Muelbauer (1992))², Quando se assume que a demanda segue a forma funcional “AIDS”, as equações têm o formato:

$$q_i = [\alpha_i + \sum_j \gamma_{ij} \log(p_j) + \beta_i \log(x/P)] / p_i$$

em que x representa os gastos totais na cesta de bens substitutos do mercado em questão, e P um índice de preços.

Às vezes, opta-se pela forma simplificada dessa especificação, usualmente denotada por SAIDS (Simplified AIDS), em que os parâmetros β_i

²O modelo Logit também satisfaz várias destas propriedades.

são iguais a zero. Os parâmetros α_i não são necessários para as simulações.

Além destas quatro formas funcionais, cabe lembrar que o *Department of Justice* dos Estados Unidos tem recentemente utilizado demandas com formato BLP (Barry, Levinsohn e Pakes (1995)), as quais são discutidas brevemente na seção 2.4. Deve-se pensar nestas formas funcionais como uma generalização da forma Logit, que permitem análises que levam em consideração a heterogeneidade dos consumidores.

2.2. O problema do viés de simultaneidade³

De posse dos dados necessários, e feita a escolha da forma funcional, o econometrista está pronto para a tarefa de estimar os parâmetros que determinam o sistema de demandas para os bens em questão. Para tal, basta escolher um método econométrico e obter os parâmetros que melhor aproximam a curva de demanda aos dados disponíveis.

Neste ponto, é fundamental lembrar que um conhecido problema à aplicação da tradicional metodologia da regressão por mínimos quadrados (Ordinary Least Squares) é ter o resíduo da equação correlacionado aos regressores. Se isso for verdade, as hipóteses necessárias para se proceder com a estimação não são satisfeitas. E, caso se proceda com a estimação, os parâmetros obtidos apresentam um “erro”, chamado de viés de simultaneidade.

A análise quantitativa de antitruste, em especial a estimação das curvas de demanda, relaciona-se diretamente a este problema. A existência de um equilíbrio de mercado, que determina preços e quantidades, pressupõe a existência de uma curva de oferta para complementar a curva de demanda. A existência dessas duas curvas, por sua vez, implica em uma causalidade reversa.

Pelo lado da demanda, maiores preços correspondem a menores quantidades demandadas. Em contraste, pelo lado da oferta, maiores preços apontam para maiores lucros, e uma reação das empresas no sentido de aumentar a produção e quantidade vendida. Esses efeitos contrários precisam ser identificados e controlados. Caso contrário, o analista estará analisando uma mistura de ambos e obtendo uma elasticidade de demanda incorreta.

Para tornar o viés de simultaneidade o mais explícito possível, considere-se uma economia expressa pelas equações:

³ A referência clássica para o problema de viés e a estimação equações simultâneas é Hausman (1983).

$$q = \alpha - \beta p + \varepsilon \text{ (Demanda)}$$

$$q = \gamma p + \eta \text{ (Oferta)}$$

em que

- q e p denotam quantidade e preço do produto em questão,
- α , β , e γ são números positivos, e
- ε e η são variáveis estocásticas independentes, com distribuição normal de média zero e variâncias respectivamente denotadas por σ_ε e σ_η .

Na forma reduzida, o mesmo sistema é expresso por,

$$q = \frac{\gamma\alpha + \gamma\varepsilon + \beta\eta}{\beta + \gamma}$$

$$p = \frac{\alpha + \varepsilon - \eta}{\beta + \gamma}$$

A tradicional regressão por mínimos quadrados ordinários da equação de demanda estimará, sem viés, os parâmetros da função de expectativa condicional, a qual é dada por,

$$E[q | p] = \alpha^* - \beta^* p$$

Como resultado dessa estimação para o parâmetro β^* obtém-se

$$\beta^* = \sigma_{pq} / \sigma_p^2$$

em que σ denota covariância.

Definindo θ por

$$\theta = \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\varepsilon^2}$$

obtém-se,

$$\beta^* = \theta\beta - (1 - \theta)\gamma$$

Das duas últimas expressões, podemos discutir alguns casos:

- Quando somente há choques de oferta, i.e., quando $\sigma_e = 0$, tem-se que $\theta = 1$, e a elasticidade é estimada corretamente: $\beta^* = \beta$
- Caso contrário, a estimação está viesada. A elasticidade estimada é inferior (em módulo) a elasticidade de fato
- Quando somente há choques de demanda ($\sigma_\eta = 0$), a estimação resulta na inclinação da curva de oferta: $\beta^* = -\gamma$
- Neste último caso, ou, mais geralmente, quando os choques de demanda dominam os choques de oferta, até o sinal da elasticidade é obtido incorretamente.

2.3. Estimação através de instrumentos

Para se obter estimações consistentes quando há problemas de viés de simultaneidade utiliza-se a metodologia denominada de regressões por variáveis instrumentais.

Para compreender como esta metodologia funciona, deve-se pensar que os regressores da equação de demanda – i.e., os preços - têm dois componentes. Um desses componentes é correlacionado com o resíduo da equação (os choques de demanda), causando os problemas acima discutidos. O outro componente, por hipótese, não tem correlação com esse termo. Se tivermos informações sobre como isolar este segundo componente, poderíamos focar nas variações dos preços que são exclusivamente devidas a ele. Com isso, estaríamos eliminando os efeitos do primeiro componente e evitando o viés na estimação.

A informação sobre os movimentos dos preços que não são correlacionadas com os choques de demanda é obtida através de variáveis instrumentais, ou instrumentos. Regressões por variáveis instrumentais utilizam essa informação adicional, de forma a isolar os movimentos nos preços, permitindo estimações não viesadas da demanda.

Para concretizar esta idéia, considere novamente a equação de demanda

$$q = \alpha - \beta p + \varepsilon$$

Suponha, adicionalmente, que existam instrumentos “válidos”, denotados por z_i , que estão disponíveis para nosso uso. Para que um instrumento z seja “válido”, ele deve satisfazer duas condições:

- Relevância: A correlação entre z e p não é igual a zero
- Exogeneidade: Correlação entre Z e ε é igual a zero

Se os instrumentos z_i satisfazem essas condições, o parâmetro β pode ser estimado através de uma metodologia conhecida como mínimos quadrados em dois estágios (*Two Stages Least Squares*). Como o nome sugere, essa estimação se dá em duas etapas. A primeira etapa decompõe a variação dos preços em duas componentes, a parte problemática (que é correlacionada com o resíduo da equação de demanda) e a outra parte, livre de problemas. Isto se dá pela regressão,

$$p = \sum_i \pi_i z_i + \xi$$

em que p são os coeficientes estimados e ξ é o resíduo. Este resíduo é a componente correlacionada com os choques de demanda. A outra componente, que pode ser prevista pelo instrumento, e não é correlacionada com os choques de demanda, é dada por,

$$p^* = \sum_i \pi_i z_i$$

No segundo estágio, faz-se uma simples regressão relacionando esta componente com a quantidade:

$$q = \alpha - \beta p^* + \varepsilon$$

Caso os instrumentos utilizados sejam válidos, a estimação da elasticidade β é não-viesada e consistente. Evidentemente, se os instrumentos não são válidos, a metodologia de variáveis instrumentais implica em novos problemas, e em resultados que não correspondem aos desejados. É, assim, importante verificar se os instrumentos satisfazem as duas propriedades anteriormente mencionadas: relevância e exogeneidade.

(i) Relevância

O papel da relevância dos instrumentos é semelhante ao papel do tamanho da amostra. Quanto mais relevante o instrumento, mais acurada será a estimação – o mesmo é verdade para maiores amostras. Instrumentos que explicam uma pequena parte da variação dos regressores são chamados de “instrumentos fracos” (*weak instruments*), e são objeto de muita de pesquisa acadêmica recente. De forma simplificada, para verificar a relevância de um instrumento, testa-se a hipótese de que todos os regressores do primeiro estágio da estimação são iguais a zero (teste F). Tipicamente, quando só há um regressor endógeno, um valor de F superior a 10 indica que os instrumentos são bastante relevantes.

(ii) Exogeneidade

Quando o número de instrumentos é igual ao número de regressores endógenos é impossível testar para endogeneidade. Já no caso em que há mais instrumentos que regressores endógenos, é possível se testar as restrições de superidentificação (*overidentifying restrictions*). Para tal, calcula-se o resíduo da regressão com variáveis instrumentais, utilizando-se os próprios regressores, ao invés de suas previsões. A seguir, corre-se uma regressão deste resíduo nos instrumentos. A estatística $J = mF$, em que F é o teste de todos os coeficientes desta regressão serem zero, é distribuída de acordo com χ^2_{m-k} , em que m e k são, respectivamente, o número de instrumentos e o número de regressores endógenos. Ainda uma outra forma de testar-se para exogeneidade é aplicar-se o teste de Hausman para parâmetros obtidos quando a estimação é feita com instrumentos sabidamente exógenos e parâmetros obtidos quando a estimação utiliza não somente estes instrumentos mas também outros. Como é usual, a variância da diferença entre estas duas estimações é igual à diferença das variâncias de cada uma delas.

2.4. Em busca de instrumentos

Conforme mencionado, para que uma estimação com variáveis instrumentais produza resultados satisfatórios, é necessário que ela esteja baseada em instrumentos válidos. Se não existem instrumentos válidos, de nada adianta a técnica da estimação em dois estágios. Em algumas áreas de pesquisa é possível realizar experimentos artificiais que criam grupos de controle e instrumentos especialmente para se testar uma hipótese. A mesma facilidade

não ocorre em Análises de Antitruste. Tipicamente, a busca por instrumentos é a tarefa mais difícil para o economista, dependendo de criatividade e intuição econômica.

Esta difícil situação produziu uma variedade de ferramentas alternativas, que permitem algum balizamento na hora de se montar o modelo. Antes de enumerá-las, vale dizer que não há uma ferramenta superior ou ferramenta perfeita. Todas envolvem hipóteses, apresentando vantagens e desvantagens. Cada aplicação particular sugere uma forma distinta de estimação, não se devendo descartar a possibilidade de se combinar diferentes ferramentas ou até utilizar alguma ferramenta distinta.

Isto posto, as ferramentas tipicamente utilizadas são:

	<i>Ferramenta</i>	<i>Idéia</i>
i	Nada	Casos particulares, regulação
ii	Cost-Shifters + Restrições	Variáveis de custo de produção
iii	Hedônicos	Outras variáveis da oferta
iv	Painel (Hausman)	Argumento de componentes de variância
v	BLP	Modelo completo, com curva de oferta

(i) Nada

A primeira alternativa é simplesmente assumir que o erro nos preços não é correlacionado com os choques de demanda. Neste caso, utiliza-se o tradicional método dos mínimos quadrados (OLS) para a estimação de demanda.

Há algumas situações em que essa hipótese é bem razoável. Uma delas é quando os preços são regulados e não podem ser alterados. Neste caso, evidentemente, eles não podem ser correlacionados com choques de demanda.

Em outras situações, em que os ofertantes acumulam estoques por não terem oportunidade de alterar os preços, o mesmo argumento pode ser razoável. Nestes casos, é importante que a frequência das observações dos preços seja relativamente elevada.

(ii) “Cost Shifters” e Restrições

Esta alternativa é provavelmente a mais típica. Normalmente há algumas séries de dados sobre os custos de produção, sejam eles os preços de bens intermediários ou salários e custo de capital.

Como se sabe, a curva de oferta pode ser escrita, de forma extremamente geral, por,

$$p = c + \mu$$

Em que c refere-se aos custos marginais e μ ao *mark-up*. Sendo assim, variáveis que capturam variações nos custos marginais também estão relacionadas com os preços. Em geral, considera-se que as mesmas não se correlacionam com os choques de demanda, apesar dos efeitos de equilíbrio geral.

Mesmo nos casos mais fáceis, tipicamente não há uma quantidade de *cost-shifters* suficiente para se estimar uma demanda com padrões de substituição bastante gerais. Isto é, caso a equação de demanda contenha os preços de todos os bens substitutos e complementares ao produto em questão, haverá um grande número de variáveis endógenas. Neste caso, seria necessário que houvesse um *cost-shifter* para cada um desses preços.

Como normalmente não há tantos *cost-shifters*, é comum restringir o formato da demanda, para um caso mais particular. A formalização através da forma funcional Logit apresenta essa vantagem, e por isso mesmo, é tão comumente utilizada. Suas restrições a substitutabilidade de demanda permitem reduzir o número de *cost-shifters* necessários para identificação, em troca de fortes restrições aos padrões de substituição entre os bens.

(iii) Hedônicos

A terceira alternativa utiliza teoria econômica para obter variáveis relacionadas a outra parte da oferta, o *mark-up* (μ). Modelos de Bertrand com bens diferenciados e a própria intuição econômica sugerem que o *mark-up* praticado pelas empresas depende da proximidade do bem em questão aos outros bens, no espaço dos produtos. Isto é, um produtor tende a escolher um *mark-up* maior caso não haja nenhum produto que seja um substituto próximo do seu.

Sendo assim, variáveis que explicam as características dos substitutos, e até o número de substitutos, tendem a influenciar o *mark-up* de um produto. Estas são correlacionadas com o preço mas não com choques de demanda.

(iv) Painel (Hausman)

A metodologia proposta por Hausman (1997) foi explorar a dimensão painel dos dados, o que é tipicamente disponível, e fazer um argumento

sobre os componentes da variância. Ele assume que a forma reduzida dos preços de um produto j na cidade n no tempo t é dada por:

$$p_{jnt} = \alpha_{jn} + \delta_j c_{jt} + w_{jnt}$$

Em que:

- O termo α_{jn} captura qualquer coisa que dependa da cidade e do produto, mas não varia no tempo. Hausman o motiva como diferenciais de salários e de custos de transportes entre diferentes regiões.
- Os choques de custo para cada produto, c_{jt} são comuns em todas as cidades, por hipótese. Por não ter subscrito n , estes custos entram em cada cidade através de efeitos comuns a todo país.
- Os resíduos w_{jnt} são perturbações com média zero. Por hipótese, estes choques são independentes para diferentes cidades.

A hipótese econômica, implícita nos subscritos, é que os choques de demanda estão restritos a serem independentes entre diferentes cidades (entram somente nos preços através do termo w), enquanto os choques de custos podem ser comuns entre diferentes cidades (entram através de c). Em linguagem mais econômica, choques econômicos que afetam preços são diferentes para cada cidade e não tem um componente nacional comum.

Como as outras, esta hipótese já foi objeto de controvérsias. Uma vez aceita, ela implica que os preços do produto em outras cidades podem ser utilizados como instrumentos. Isto porque eles são correlacionados com o preço do produto (através do componente de custo), mas não são correlacionados com choques de demanda.

(v) BLP

Devido ao forte apelo econômico, a metodologia BLP (Barry, Levinsohn e Pakes (1995)) tem sido intensivamente utilizada pelo Departamento de Justiça nos Estados Unidos nos últimos anos. Pelo que sabemos, ela ainda não foi empregada no Brasil em nenhum caso. De certa forma, esta metodologia constitui uma variação da metodologia (iii), mas com uma especificação mais completa da modelagem.

Para obter a curva de demanda, BLP assume que as preferências dos consumidores têm forma Logit. A existência de heterogeneidade entre os

consumidores, modelada através de coeficientes estocásticos, permite eliminar a forte restrição sobre substitutabilidade.

Em particular, considerando-se o mercado de automóveis, BLP impõem que a utilidade do consumidor i de consumir o produto j é dada por

$$u_{ij} = \alpha \log(y_i - p_j) + x_j \beta + \xi_j + \sum_k \sigma_k x_{jk} v_{ik} + \varepsilon_{ij}$$

Em que y_i representa a renda desse consumidor, v_i são outras características desse consumidor, x_j são atributos do produto (tais como potência, tamanho, consumo de gasolina), α , β e σ são parâmetros a serem estimados, e ξ e ε são perturbações com média zero. ξ tem, como interpretação natural, características do produto que não são observadas pelo econometrista, mas são observadas pelos consumidores. ε tem distribuição de valor extremo, o que implica na forma funcional Logit.

Para determinar a oferta, BLP assumem que os custos marginais de produção do produto j são dados por

$$\ln(mc_j) = w_j \gamma + \omega_j$$

Em que w são atributos observáveis do produto, ω são atributos não observáveis, e γ um parâmetro a ser estimado. A seguir, BLP assume que as empresas escolhem preços de forma a maximizar lucro, jogando o típico jogo de Bertrand com bens diferenciáveis (one-shot Nash).

Para a estimação, realizada através da aplicação dos Métodos Gerais dos Momentos, BLP utilizam tanto a curva de demanda quanto a curva de oferta, i.e., as condições de primeira ordem da maximização de lucro. Como instrumentos, além de alguns cost-shifters disponíveis, eles utilizam características dos outros produtos, os quais medem a proximidade (no espaço das características) entre os produtos considerados

Por considerar a de oferta explicitamente, BLP conseguem obter um maior número de momentos (maior número de equações), o que tende a garantir melhores estimações dos coeficientes. Uma possível crítica seria a necessidade de se fazer uma hipótese com relação à oferta. Contudo, essa hipótese (Bertrand em bens diferenciados) seria feita de qualquer forma durante a fase de simulação de fusões.

3. *Back end: A simulação propriamente dita*

A análise posterior (*back-end*), que é a simulação propriamente dita, é composta das seguintes três etapas: (i) escolha do modelo, (ii) calibração, e (iii) previsão dos efeitos da fusão.

A escolha do modelo resulta de uma combinação de lógica e metodologia científica, Teoria Econômica, e informações sobre a magnitude das elasticidades anteriormente estimadas. A criação de novos modelos de competição é uma área ativa de pesquisa acadêmica, a qual deverá influenciar a prática antitruste num futuro breve. Por hora, o típico é adotar-se modelos de equilíbrio não cooperativo, chamados “One Shot Nash”, visando a obtenção dos “efeitos unilaterais” da fusão. No caso de atos de concentração de produtos diferenciados, por exemplo, geralmente se adota o modelo de Bertrand-Nash. No caso de produtos homogêneos, o modelo Cournot-Nash é mais adequado.

A calibração do modelo corresponde a uma determinação de parâmetros de modo que o modelo escolhido refleta as condições de equilíbrio prevalentes antes do ato de concentração. Tipicamente, ajusta-se o custo marginal de tal forma que a margem seja igual à prevista pelo modelo comportamental dos agentes na situação pré-fusão, dados os respectivos preços e participações de mercado de cada produto.

Para a previsão dos efeitos da fusão, supõe-se a manutenção dos custos de produção, e calculam-se os novos preços a partir de expressões que relacionam a margem das firmas à elasticidade-preço da demanda. Alternativamente, computa-se a redução de custos necessária para evitar a elevação de preços.

De modo tentar clarificar essas etapas para o leitor, a seguir considera-se o caso da adoção de um modelo de Bertrand-Nash para bens diferenciados, o qual é sem dúvida o modelo mais comumente empregado para análises de efeitos unilaterais.

Seja uma economia em que cada produto j é produzido por uma única firma i . Cada firma pode produzir vários produtos, e a relação $F(\cdot)$ mapeia os produtos em direção às firmas que os produzem. O custo marginal de produção de cada produto é c_j , por hipótese constante. Assume-se que não há economias de escopo, mas firmas distintas podem ter custos distintos. Preços e quantidades são denotados por p_j e q_j . Assume-se competição de Bertrand, isto é, cada firma escolhe seus preços tomando os outros preços como dados.

O lucro de cada firma i é dado por:

$$\pi_i = \sum_{j, F(j)=i} (p_j - c_j) q_j$$

em que p (sem subscrito) denota um vetor com todos os preços. Para cada produto k , a condição de primeira ordem para maximização do lucro é:

$$0 = \frac{\partial \pi_i}{\partial p_k} = q_k + \sum_{j, F(j)=i} (p_j - c_j) \frac{\partial q_j}{\partial p_k}$$

Para a obtenção de $\partial q_j / \partial p_k$ utilizam-se, em cada etapa, estimações econométricas de elasticidade de demanda.

Este sistema de equações é então utilizado em duas etapas:

- (i) obtenção dos custos marginais (calibração); e
- (ii) obtenção dos novos preços para a nova estrutura de mercado, isto é, para uma nova função.

4. Digressão: Mercados relevantes e o teste da elasticidade crítica

Como já extensamente argumentado anteriormente, o uso de simulações não substitui a análise estrutural tradicional, mas a complementa, apresentando vantagens e desvantagens. Uma das vantagens principais é dispensar a definição *prévia* de mercado relevante, uma tarefa que é muitas vezes é complicada, e que torna o uso de simulações ainda mais importante.

Isto não significa que uma análise de simulações é contrária ao conceito de mercado relevante, mas que o mercado relevante fica endógena e implicitamente determinado em meio aos procedimentos empregados em uma simulação de fusão. De qualquer forma, a não determinação *prévia* de mercado relevante tende a confundir muitos analistas, levantando críticas que fazem pouco sentido econômico, mas cuja força retórica pode causar grandes confusões.

De forma a esclarecer este ponto, desfazendo as confusões argumentos puramente retóricos, Jonathan Baker (1997), então diretor do Bureau de Economia do FTC (Federal Trade Commission) dos EUA, opinou:

“Parte da promessa do uso de métodos empíricos em análise de fusão é que eles tornam a definição de mercado menos importante. De fato, caso possa ser mos-

trado diretamente que uma fusão prejudica a competição, não há necessidade de gastar esforços na definição de mercado... Muitos entendem que o Clayton Act 7 requer que a Corte identifique os mercados de produto e geográfico que seriam prejudicados com a fusão. Uma possível adaptação – que harmonize as vantagens da análise econométrica com os requerimentos legais – poderia ser denominada de definição de mercado “*res ipsa loquitur*”. Quando um piano cai na calçada, a lei não pergunta se alguém foi negligente, mas vai direto à questão de quem. De forma semelhante, caso se possa demonstrar diretamente que uma fusão provavelmente prejudicará a competição, existe algum mercado relevante que seria prejudicado, mas não é necessário especificar suas fronteiras”

Também com o mesmo objetivo, mas de forma bem mais irônica, Charles James (2002)⁴, então *attorney general* da Divisão de Antitruste do Departamento de Justiça americano, afirmou:

“...sob a influência de sodium penthatol, a maioria dos economistas concederia que definição de mercado não é particularmente importante em uma análise de efeitos unilaterais”

O objetivo dessa seção é discutir uma metodologia quantitativa para a determinação de mercado relevante, conhecida como o “Teste da Elasticidade Crítica”. Essa metodologia é um meio termo entre rigor analítico das simulações e a linguagem mais comum à grande parte dos analistas de antitruste.

Nosso intuito e nossa opinião, cabe logo esclarecer, não são diferentes dos expressos nessas citações. Continuamos acreditando que simulações de fusões “completas” são a melhor forma de se analisar quantitativamente uma fusão. Se o teste da elasticidade crítica apresenta a vantagem de usar uma linguagem um pouco mais fácil de ser compreendida, ele não tem a mesma lógica impecável das simulações, podendo provocar resultados enganosos, quando aplicado sem o devido cuidado.

⁴ Vide Woo (2004), capítulo 4.

4.1. Breve histórico, o teste da elasticidade crítica e a análise de perda crítica

A maioria das investigações sobre fusões conduzidas pelo *Federal Trade Commission* e pelo *Department of Justice* dos Estados Unidos envolveram algum questionamento sobre a proximidade dos produtos em questão, e sua suficiência para que estes produtos fossem incluídos num mesmo mercado relevante. Por muitos anos, essa questão de como determinar as fronteiras do mercado foi respondida sem nenhum rigor. As cortes geralmente declaravam que a resposta estava relacionada às magnitudes das elasticidades cruzadas e a possibilidade dos consumidores substituírem um produto por outro, mas não havia uma sugestão de quão grandes deviam ser estas elasticidades e a possibilidade de substituição.

Esse ambiente de pouca precisão e clareza começou a ser alterado quando da publicação das *merger guidelines* em 1982, e da articulação do paradigma do monopolista hipotético. Esse paradigma, como se sabe, responde à questão de como determinar as fronteiras do mercado. O mercado relevante não deve ser mais extenso que um grupo de produtos e de áreas geográficas tal que um único ofertante desses produtos nessas áreas encontre lucrativo impor uma elevação nos preços que seja pequena, mas significativa e não transitória.

Esse teste foi inicialmente criticado por não ser operacional e muito teórico. Mas, após um período de tempo relativamente curto, tornou-se óbvio que a implementação do teste do monopolista hipotético poderia ser realizada através de uma análise de *break even* relativamente simples, conhecida como “Teste da Elasticidade Crítica” ou “Análise de perda crítica”. De fato, apesar de utilizarem uma linguagem diferente, esses testes nada mais são que uma forma tecnicamente rigorosa de se determinar as fronteiras de um mercado relevante.

De forma resumida, o teste da elasticidade crítica é composto de três etapas⁵:

(i) Obtenção da margem de lucro dos atuais ofertantes no mercado escolhido (conjunto de produtos e áreas geográficas). A partir desta margem de lucro, através do uso de fórmulas matemáticas simples, baseadas na hipótese de maximização de lucro, obtém-se a qual seria a elasticidade de demanda (pré-

⁵ A análise de perda crítica é equivalente ao teste da elasticidade crítica. Sua única distinção, é utilizar uma métrica distinta: ao invés de elasticidades, utilizam-se reduções nas quantidades vendidas.

fusão) que faria com que um monopolista hipotético desejasse elevar os preços em uma certa porcentagem (5% ou 10%, tipicamente). Esta elasticidade é denominada de “elasticidade crítica”.

(ii) Obtenção da elasticidade de fato observada no mercado escolhido.

(iii) Comparação entre a elasticidade de fato e a elasticidade crítica. Caso a elasticidade de fato seja (significativamente) superior à elasticidade crítica, deve-se expandir o mercado escolhido. Caso a elasticidade de fato seja (significativamente) inferior à crítica, deve-se reduzir o mercado escolhido. Caso seja (aproximadamente) igual, o mercado escolhido pode ser considerado um mercado relevante.

4.2. Limitações e problemas do teste da elasticidade crítica

O fundamental para se perceber as limitações e problemas do teste da elasticidade crítica é observar que, de forma geral, esse teste apresenta as mesmas vantagens e desvantagens da metodologia de simulação de fusão. Por um lado, assim como as simulações, esse teste oferece uma resposta quantitativa e precisa a pergunta realizada. Por outro lado, novamente da mesma forma que as simulações, ele está sujeito a hipóteses que merecem ser analisadas com cuidado e escrutínio.

Para compreender esta observação, note-se que o resultado de um teste de elasticidade crítica é conseqüente de uma estimação de elasticidade, a qual é realizada através de hipóteses sobre as formas funcionais da demanda, além de estar sujeita aos mais diversos problemas econométricos. Ademais, para completar, o teste envolve uma maximização de lucro de um monopolista hipotético, o que é necessariamente função de hipóteses, já tal monopolista não existe no mundo presente.

Essa observação sobre a “equivalência” entre os testes de elasticidade crítica e das simulações de fusão, quanto as suas vantagens e desvantagens relativas, por sua vez, sugere dois comentários:

(i) Em primeiro lugar, essa perspectiva desafia aqueles que são céticos com relação ao uso de simulações, mas não a uma determinação tecnicamente rigorosa do mercado relevante. Isto não faz nenhum sentido lógico.

(ii) Em segundo lugar, essa observação sugere que para que o teste do monopolista hipotético seja empregado corretamente, é necessário analisar suas potenciais dificuldades e hipóteses envolvidas.

Esta percepção levou a um recente interesse dos órgãos de defesa nos Estados Unidos, que nos últimos dois anos vêm solicitando comentários sobre o uso do teste da elasticidade crítica e da análise de perda crítica para determinação de mercado relevante. Com isso, deu-se origem a uma série de discussões e debates entre alguns pesquisadores de renome, como Katz e Shapiro (2003 e 2004), Scheffman e Simons (2003) e Brien e Wickelgreen (2004).

Sobre este ponto, vale notar os argumentos desenvolvidos por Katz e Shapiro (2003), sobre a falha nos argumentos de alguns pareceristas, quando estes buscam expandir ao máximo o mercado relevante, ao obterem margens de lucro elevadas. A lógica dessa inferência está baseada na seguinte intuição. Se a margem de lucro é elevada, uma pequena perda nas vendas seria suficiente para tornar uma elevação de preços não lucrativa. A demanda teria que ser *relativamente* menos elástica para que isso ocorresse, mas isso não seria plausível em um mercado pequeno. Portanto, o mercado é provavelmente grande. Por outro lado, conforme estes autores apontam, a mesma lógica contraria uma intuição igualmente razoável: a margem pode ser elevada porque as empresas perceberam que os consumidores não respondem muito a elevações de preços (a demanda é relativamente inelástica). Sendo assim, se por um lado a elasticidade crítica é baixa, por outro, a elasticidade de fato também deve ser baixa.

Este *conundrum* pode ser resolvido de forma relativamente simples. As margens de lucro obtidas na prática oferecem relativamente poucas informações sobre as elasticidades. Isto se deve a imprecisões estatísticas ou, simplesmente, porque os dados contábeis disponíveis não correspondem aos conceitos econômicos desejados. Sendo assim, as agências de defesa podem incorrer em graves equívocos caso utilizem somente estas margens para determinar as elasticidades. Em complemento, é fundamental buscar outras informações concernentes às respostas dos consumidores a aumentos de preços, sejam essas evidências anedóticas, resultados de experimentos naturais, ou resultados de estimações econométricas da demanda.

Finalmente, ainda vale mencionar as críticas de Werden (2003) sobre a escolha prévia de uma certa porcentagem nos aumento de preços, para a realização do teste. Conforme Werden mostra, através de ilustrações que reproduzem casos verídicos, é possível que para um monopolista hipotético não seja lucrativo elevar os preços em 5%, mas que seja lucrativo elevar os preços em 50%. Evidentemente, as simulações de fusão não estão sujeitas a este problema por não especificarem os aumentos de preços a priori.

5. Conclusões: O futuro das simulações de fusão

Neste artigo, foram analisados alguns tópicos da metodologia de simulações de fusões, em especial a estimação econométrica do sistema de demanda, o problema do viés de simultaneidade, e a busca de instrumentos válidos. De forma bem mais breve, discutiu-se a modelagem econômica utilizada, e sua relação com a definição de mercados relevantes. A título de conclusão, arriscamos um prognóstico sobre os futuros desenvolvimentos técnicos dessa metodologia, e sobre sua incorporação à análise antitruste no Brasil e no resto do mundo.

Quanto à estimação econométrica, deve-se notar que já está havendo a incorporação de técnicas mais avançadas por parte dos órgãos de defesa americanos. Estimações de demanda BLP já têm sido utilizadas com sucesso em alguns casos, e seu uso deve rapidamente se espalhar para análises realizadas em outras partes do mundo. Como a estimação BLP original (Berry, Levinsohn e Pakes (1995)) utiliza uma grande quantidade de dados, deve-se esperar que haja alguns avanços nas simplificações que tornam a metodologia mais amplamente utilizável.

Mais além das demandas BLP, a literatura acadêmica tem recentemente progredido em técnicas de estimação para modelos explicitamente dinâmicos (vide, por exemplo, Pakes e McGuire (1994)), os quais já estão automaticamente acoplados a modelos econômicos que consideram a entrada ou saída de firmas na indústria. Sem dúvida, este parece ser o futuro da chamada “Nova Organização Industrial Empírica”. Contudo, acreditamos que aqui a quantidade de informações necessárias para a implementação das novas técnicas deverá ser um limitante crítico para sua utilização em casos de antitruste, pelo menos num futuro breve.

Quanto à utilização de econometria para a definição de mercado relevante, o prognóstico é uma grande interrogação. De um lado, aqueles que usam econometria e simulação parecem acreditar que uma prévia definição de mercado relevante não é necessária, podendo até causar confusões quando empregadas inocentemente. De outro lado, esta é a forma mais fácil de mostrar qual as vantagens da econometria para aqueles que são avessos aos métodos quantitativos. Talvez a divulgação e aumento do uso de testes de elasticidade crítica seja uma etapa necessária, antes de seu desaparecimento.

É na escolha da modelagem, acreditamos, que os próximos anos trarão as mudanças e modificações mais interessantes. Atualmente, o paradigma são os modelos de Bertrand e Cournot estáticos, utilizados para o

cálculo dos “efeitos unilaterais”. Mas a preocupação tanto dos órgãos americanos como da Comunidade Européia nos chamados “efeitos coordenados” pode mudar esse foco. Pesquisa acadêmica recente já indica como separar qual a parcela do *mark-up* que é devida a estes dois efeitos (Nevo (2001)), mas ainda conhecemos pouco sobre quais são os modelos úteis para se prever os efeitos coordenados de uma fusão. Talvez o caminho mais promissor seja a adaptação do modelo de Green e Porter (1986) para análises de antitruste. Uma discussão detalhada sobre estes dois efeitos, e sobre esses modelos é assunto para um outro artigo.

Referências

- Baker, J. B. (1997) “Contemporary Empirical Merger Analysis” *George Mason University Law Review* 5: 347-61
- Barry, S., Levinsohn, J. e Pakes, A. (1995) “Automobile Prices in Market Equilibrium” *Econometrica* 63(4): 841-890
- Deaton, A. e Muellbauer, J. (1980) “An almost Ideal Demand System” *American Economic Review* 70-3, 312-26.
- Green, E and Porter, R. (1984) “Noncooperative Colusion under Imperfect Price Information” *Econometrica* 52(1): 87-100.
- Hausman, J. A. (1983) “Especification and Estimation of Simultaneous Equation Models” Capítulo 7 em *Handbook of Econometrics*, edição de Griliches, Z. e Intriligator, M. D.. Amsterdam, Elsevier.
- Hausman, J. A. (1997) “Valuation of New Goods under Perfect and Imperfect Competition” em *The Economics of New Goods*, Bresnahan, T. e Gordon R. J. Eds, NBER Studies in Income and Wealth 58, The University of Chicago Press
- Katz, M. L. e Shapiro, C. (2003) “Critical Loss: Let’s Tell the Whole Story” *Antitrust Magazine*, Spring 2003
- Katz, M. L. e Shapiro, C. (2004) “Further Thoughts on Critical Loss” *The Antitrust Source*, March 2004
- Nevo, A. (2001) “Measuring Market Power in the Ready-to-Eat Cereal Industry” *Econometrica* 69(2): 307-42
- O’Brien, D. P. e Wickelgren, A. L. (2004) “The State of Critical Loss Analysis: Reply to Scheffman and Simons” *The Antitrust Source*, March 2004
- Scheffman, D. T. e Simons, J. J. (2003) “The State of Critical Loss Analysis: Let’s Make Sure We Understand the Whole Story”, *The Antitrust Source*, November 2003
- Werden, G. J. (2003) “Hospital Mergers and the Hypothetical Monopolist Test” Manuscript

Werden, G. J., Froeb, L. M. e Tardiff, T. J. (1996) “The Use of Logit Model in Applied Industrial Organization” *International Journal of the Economics of Business* 3: 83-105

Wu, L. (2004) *Economics of Antitrust: New Issues, Questions, and Insights* NERA Economic Consulting.