

Os Métodos Estatísticos aplicados à Economia segundo Cinco Abordagens

Bruno César Araújo (Org.), Alexandre Messa Silva, Alan Silva, Danilo Coelho, Dea Fioravante, Fernando Freitas, Júnia Conceição, Luis Kubota, Luiz Dias Bahia, Marco Aurélio Mendonça, Patrick Alves, Rogério Freitas

(DISET – IPEA)

Introdução

Nos dizeres de Intriligator, Bodkin e Hsiao (1996), “pelo menos 80% do material presente nos livros-texto de econometria enfoca puramente em técnicas econométricas.” Na literatura econométrica, exceção feita a alguns periódicos especializados, o mesmo ocorre.

Contudo, tão importante quanto estudar métodos e técnicas econométricas é estudar sua **metodologia** e aplicações. Neste artigo, a metodologia da econometria é definida como o meta-estudo de como suas técnicas se situam no contexto epistemológico e contribuem para o avanço do conhecimento científico.

Não existe apenas uma visão da metodologia da econometria, tampouco de seu conceito e sua finalidade. Por isto, buscaremos interpretar o uso dos métodos econométricos de acordo com cinco paradigmas diferentes: (i) a chamada abordagem “clássica” ou da comissão Cowles, nos anos 1950-60; (ii) a abordagem “inglesa” ou da London School of Economics, dos anos 1970-80; (iii) a abordagem VAR dos anos 1980; (iv) a abordagem da calibragem dos anos 1990 e (v) a microeconometria e os experimentos quase-naturais, que ganham crescente notoriedade, sobretudo na avaliação de políticas públicas e nos estudos do mercado de trabalho. Nossa intenção não é avaliar as abordagens em busca da “melhor” ou “mais correta”, apenas evidenciar o quanto o uso dos métodos econométricos e a própria visão do que é econometria e seus fins variam de acordo com cada uma delas.

Analisaremos como cada abordagem trata a questão do **antes** (relação teoria-modelo, escolha de especificações, análise exploratória dos dados e temas correlatos), **durante** (técnicas econométricas, *data mining* e outros tópicos) e **depois** da aplicação das técnicas econométricas (interpretação dos modelos, apresentação e implicações para políticas públicas). Naturalmente, estas questões estão interrelacionadas e esta sistemática de exposição é meramente ilustrativa. Sempre que possível, a exposição contará com exemplos e temáticas correlatas aos estudos com microdados de produção, tecnologia e inovação.

Assim, o restante do artigo é composto por 6 seções. Na próxima seção, discutiremos o conceito de econometria e seus usos. Na terceira seção, exporemos brevemente as cinco abordagens citadas acima e seus pressupostos. As três seções seguintes tratam do tratamento dado pelas abordagens ao antes, durante e depois da estimação, enquanto os comentários finais se situam na sétima e última seção.

Econometria: o que é e para que serve?

Um conceito relativo

Para definir econometria, a maioria dos autores recorre à famosa passagem do então editor da *Econometrica*, Ragnar Frisch, na edição inaugural daquela publicação: a econometria é “a relação entre teoria econômica, matemática e estatística” ou mesmo a “unificação entre as abordagens teórico-quantitativas e empírico-quantitativas aos problemas econômicos” (FRISCH, 1993, p.1, tradução livre). Ou, de forma análoga, a econometria é constituída da aplicação de métodos matemáticos e estatísticos à análise de conjuntos de dados econômicos, com o objetivo de caracterizar empiricamente teorias econômicas. Pode-se defini-la como um método de análise que agrega estatística, matemática e teoria econômica.

No entanto, a econometria não pode ser compreendida como somente a aplicação adequada de técnicas estatísticas a problemas econômicos. Em realidade, o tripé teoria econômica, estatística e matemática é balanceado de acordo com a ênfase que cada metodólogo dá a cada uma destas pernas e com a ordem de precedência (por exemplo, devemos proceder da teoria aos dados ou o contrário?), e isto envolve considerações epistemológicas da maior importância. Ao longo do artigo, discutiremos oportunamente como o conceito de econometria é definido de acordo com as diferentes abordagens.

Os usos da econometria

De acordo com Hoover (2006), a econometria tem quatro funções:

1. **Testar implicações de uma teoria** – esta função se relaciona com a epistemologia dos chamados verificacionistas ou do falsificacionismo atribuído a Karl Popper. Brevemente, ambos convergem ao defenderem que a avaliação das teorias científicas deve ser feita por meio da avaliação empírica das implicações das mesmas, a partir de hipóteses testáveis empiricamente.¹ Assim, através da correta tradução da teoria em proposições testáveis a partir de modelagem econométrica, a econometria pode servir para testar implicações de uma teoria.

Na prática, porém, esta função da econometria não é trivialmente implementada. No quadro 1, analisamos a “saga” da avaliação empírica do paradigma de vantagens comparativas e discutimos como é difícil elaborar testes

¹ Verificacionistas e falsificacionistas divergem, contudo, quanto às finalidades e interpretações dos testes empíricos. Enquanto os verificacionistas buscam o suporte empírico das predições teóricas, por meio de sua confirmação, os falsificacionistas buscam falsear estas predições, uma vez que **afirmar o conseqüente** pode ser falacioso (a proposição “A implica B, B é verdadeiro, logo A é verdadeiro” pode não ser verdade porque pode haver situações em que B é falso), enquanto **negar o conseqüente** não (“A implica B, B é falso, logo A é falso”). Deste modo, para Popper, não existem teorias das quais se pode derivar proposições testáveis que sejam “comprovadas” ou “verdadeiras”, e sim aquelas que nunca foram criticamente falseadas. O debate sobre qual seria o “nível crítico” de violação da teoria é muito complexo e foge ao escopo do trabalho. Para uma discussão mais profunda sobre estas correntes epistemológicas e seus desdobramentos, vide Blaug (1999), de onde tomamos emprestados os vocábulos verificacionistas e falsificacionistas.

empíricos para suas implicações que sejam deriváveis do paradigma. Não por acaso, o paradigma ainda resiste a despeito de diversas avaliações empíricas desfavoráveis.

2. **Calibrar parâmetros de um modelo** – por vezes, a econometria é utilizada para mensurar parâmetros de um modelo já definido (por exemplo, calcular a elasticidade-renda de uma função consumo). No limite, a econometria pode ser a mensuração de uma relação proposta pela teoria econômica.
3. **Previsão** – a econometria pode ser usada para prever uma relação ou mesmo os valores de uma variável de interesse. Esta previsão pode ser baseada em teoria ou simplesmente pode derivar de uma relação estatística, no entanto, segundo Hoover (2006), a teoria econômica é o que diferencia as regularidades acidentais das genuínas, e permite qualificar e explicar uma previsão.²
4. **Caracterização de relações e fenômenos** – a econometria pode permitir uma visualização dos dados de forma a revelar relações e fenômenos às vezes não previstos pela teoria, possibilitando o estabelecimento de fatos estilizados e *insights* para os economistas teóricos.

Vale notar, novamente, que a validade científica ou não dos usos acima citados depende da abordagem adotada. Por exemplo, as duas primeiras funções estão relacionadas às abordagens “da teoria para os dados”, enquanto as duas últimas, às abordagens “dos dados para a teoria”. Os adeptos da abordagem clássica enfatizam a segunda função da econometria (econometria como mensuração), enquanto os adeptos da abordagem inglesa são mais afeitos à quarta (econometria como observação).

² Para os seguidores fiéis do método hipotético-dedutivo – para quem as teorias científicas consistem em uma rede de leis científicas, das quais conclusões empíricas são dedutíveis empiricamente – a previsão é o simétrico da explicação: se uma conclusão empírica da teoria ainda vai ser observada, temos uma previsão; em verdade, explicação é a previsão aplicada ao passado (Hoover, 2006).

Quadro 1: avaliação empírica dos modelos de vantagens comparativas

À medida que os modelos de vantagens comparativas foram se tornando populares no pós-guerra, eles foram sujeitas intensivamente a testes empíricos. Não raro, a teoria foi contrariada em diversos artigos. Apesar disso, estas teorias ainda são bastante populares e formam o corpo do *mainstream* da economia internacional. Por que isto acontece?

A resposta talvez resida na dificuldade em se formular testes definitivos para estas teorias. Consideremos, por exemplo, o caso Heckscher-Ohlin. Se, por um lado, ele é derivado a partir de hipóteses muito restritivas, difíceis de se verificar na prática, por outro lado o relaxamento de alguma hipótese pode degenerar o modelo. Sendo a escolha do teste adequado fator decisivo para a verificação/falsificação empírica de uma teoria, no caso do comércio internacional o teste definitivo teria que levar em conta que os modelos completos sempre respondem a três questionamentos: (i) quais os bens são transacionados; (ii) com quem é transacionado; (iii) o quanto é transacionado (DEARDOFF, 1984, p. 469). É praticamente impossível formular testes tão compreensivos.

Por isso, Leamer e Levinsohn (1995, p.1341) já na primeira página de seu artigo de revisão da literatura empírica advertem: em comércio internacional, “estime, não teste, não leve a teoria ao pé da letra”. No fundo, as teorias de comércio exterior são condições de arbitragem, e condições de arbitragem quase nunca são atingidas com perfeição. Entretanto, no parágrafo seguinte já temos um segundo conselho: “não trate a teoria de forma tão frouxa”, ou seja, devem-se buscar sempre a derivação de testes e a elaboração de base dados mais consistentes o possível com a teoria.

Visto isso, é natural que estas teorias sejam relativamente robustas a décadas de testes empíricos, os quais nem sempre bem sucedidos. Se por um lado, estes testes não conseguem “testar” perfeitamente as teorias; por outro, não se devem buscar alternativas de modelagem em comércio internacional que se adaptem melhor aos dados apenas com base no empirismo.

É fácil observar que um teste mais rigoroso do modelo de Ricardo seria fatal para a teoria na sua forma pura, pois suas previsões são muito fortes (alto grau de especialização, irrelevância das dotações relativas para o comércio, ausência de questões distributivas [KRUGMAN E OBSTFELD, 2001, p. 34]). Uma alternativa seria abandonar o modelo de Ricardo como opção teórica em nome de um modelo mais sofisticado, como Heckscher-Ohlin. Porém, se dermos a Ricardo uma chance e concentrarmos atenção no apenas no princípio das vantagens comparativas referentes ao trabalho, poderemos enxergar aspectos interessantes.

Diversos autores nos anos 1950-60 adotaram esta estratégia. Um exemplo representativo desta literatura é o trabalho de Balassa (1963), que demonstrou existir uma relação positiva entre produtividade relativa e exportações relativas no comércio bilateral entre EUA e Grã-Bretanha.

Apesar de apresentar produtividade absoluta maior em praticamente todos os setores industriais, os EUA tinham uma exportação relativa maior nos setores em que havia maior produtividade relativa, conforme a previsão do modelo ricardiano. Por sua vez, um teste direto do modelo de Heckscher-Ohlin teria que se basear nas diferenças de preços em autarquia, ou seja, antes do comércio entre os países iniciar (DEARDOFF, 1984, p. 471). O problema é que estes preços de autarquia são inobserváveis. Desta forma, os autores buscaram maneiras indiretas de prever os padrões de comércio, e partiram da constatação de que, na realidade, o modelo Heckscher-Ohlin descreve a troca de bens como uma forma indireta de troca de fatores de produção. Desta forma, o primeiro autor a testar o modelo de Heckscher-Ohlin buscou observar a direção do padrão de comércio baseado na ligação entre dotações relativas de fatores embutidos nos produtos transacionados e o padrão de exportação/importação.

Precisamente, estamos falando do trabalho de Leontief (1954), que buscou evidências do modelo Heckscher-Ohlin no comércio bilateral entre EUA e o resto do mundo em 1947. Leontief (1954) procedeu da seguinte forma: com mesma tecnologia e equalização dos preços dos fatores entre os países, temos que a matriz insumo-produto dos EUA se aplica também às suas importações. Então, o conteúdo de capital embutido num vetor de quantidades \mathbf{x} é igual a $\sum_i x_i a_{ki}$. Por sua vez, o conteúdo de trabalho nestes produtos é $\sum_i x_i a_{li}$. Logo, a relação capital-trabalho embutida nestes produtos será $\sum_i x_i a_{ki} / \sum_i x_i a_{li}$.

À época, os EUA eram o país mais capital-intensivo do mundo, portanto, com base em Heckscher-Ohlin, suas exportações deveriam ser capital-intensivas e suas importações, trabalho intensivas. O problema é que Leontief (1954) achou o resultado exatamente oposto, pelo menos para a relação capital-trabalho. Este resultado ficou conhecido como "paradoxo de Leontief". O próprio Leontief (1954) explicou o paradoxo afirmando que o trabalho nos EUA era mais produtivo. Se o trabalho nos EUA fosse duas vezes mais produtivo que no resto do mundo, o paradoxo estaria resolvido.

Segundo Helpman (1998, p. 5), Leontief (1954) de alguma forma explicou um paradoxo introduzindo outro, pois como o trabalho nos EUA poderiam ser duas vezes mais produtivo mesmo após o controle pelo uso do capital? Por sua vez, Krugman & Obstfeld (2001, p. 85-6) notam que os próprios números de Leontief apontam para a direção de Heckscher-Ohlin no que tange à escolaridade da mão-de-obra envolvida no comércio internacional e à proporção de engenheiros e cientistas. Isto sugere que as exportações fator de produção não computado dos EUA eram intensivas em mão-de-obra qualificada, ou seja, talvez existisse um outro nos cálculos de Leontief (1954): o capital humano. Como complemento, "uma explicação plausível [para o paradoxo de Leontief] seria a de que os Estados Unidos têm uma vantagem especial ao produzir novos produtos ou bens fabricados com tecnologias inovadoras" (KRUGMAN E OBSTFELD, 2001, p. 86).

Mais geralmente, duas relações ajudam a generalizar o teorema de Heckscher-Ohlin (HELPMAN, 1998, p. 6-8). A primeira relação é da produção. Suponhamos uma matriz \mathbf{A} formada pelos coeficientes tecnológicos a_{ij} , que representam a quantidade do insumo i usado na produção de j . Cada coluna desta matriz descreve a tecnologia empregada em cada produto. Esta notação permite incluir, por exemplo, vários tipos de capital, tipos de trabalho etc. Ignoremos os casos não-pares e suponhamos ainda que esta matriz é quadrada e inversível (não-singular). Vale notar que os coeficientes a_{ij} não necessariamente são fixos e dependem da tecnologia e dos preços fatoriais, porém na estrutura Heckscher-Ohlin os preços dos fatores se equalizam e as tecnologias são as mesmas, logo a matriz \mathbf{A} é a mesma para os países. Sejam \mathbf{V}^z e \mathbf{X}^z respectivamente os vetores coluna dos insumos e dos produtos do país z . Então, pelas condições de pleno emprego, temos que $\mathbf{A}\mathbf{X}^z = \mathbf{V}^z$, para todos os países.

Do lado do consumo, suponhamos preferências são homotéticas e idênticas para todos os países, logo, sem impedimentos ao comércio, a composição do consumo será a mesma em todos os países e proporcional à produção mundial. Isto quer dizer que, sendo $\mathbf{X}^* = \sum_z \mathbf{X}^z$, temos que $\mathbf{C}^z = s_z \mathbf{X}^*$, para todos os países. Se o comércio for balanceado, s_z será a proporção do país z na renda mundial.

As duas relações acima (e suas hipóteses) permitem-nos relacionar o balanço de comércio às dotações de fatores. Basta lembrar que, sendo \mathbf{A} inversível, $\mathbf{X}^z = \mathbf{A}^{-1}\mathbf{V}^z$. Como o vetor das exportações líquidas \mathbf{T}^z é a diferença entre a produção e consumo, vemos que:

$$\begin{aligned} \mathbf{T}^z &= \mathbf{X}^z - \mathbf{C}^z \\ \mathbf{T}^z &= \mathbf{A}^{-1}\mathbf{V}^z - s_z \mathbf{X}^* \\ \frac{1}{s_k} \mathbf{T}^z &= \mathbf{A}^{-1} \left(\frac{1}{s_z} \mathbf{V}^z \right) - \mathbf{X}^* \end{aligned}$$

O pesquisador pode utilizar esta especificação linear para trabalhar com o enfoque setorial (variando as colunas da matriz), o enfoque nos países (variando o z) ou nos dois (construindo um painel). Este tipo de análise pode ser interessante, mas é preciso obter informações sobre produção mundial, balanço de comércio, proporções dos países no consumo mundial, matriz tecnológica e dotações dos fatores nos países. Como o foco do modelo Heckscher-Ohlin é na relação entre dotações e fluxos de comércio, podemos simplificar a equação (1.9). Foi o que fez Vanek (1968 *apud* HELPMAN, 1998, p. 8). Vanek supôs tecnologias \mathbf{A} idênticas para todos os países, de forma que o conteúdo de fatores das exportações líquidas de um país é $\mathbf{F}^z = \mathbf{A}\mathbf{T}^z$. Usando a definição $\mathbf{T}^z = \mathbf{X}^z - \mathbf{C}^z$ em conjunto com as relações de produção e consumo anteriormente definidas, temos que:

$$F^z = V^z - s_z V^*$$

Nesta equação, V^* é o vetor coluna de agregação mundial dos fatores de produção, $V^* = \sum_z V_z$. A equação (1.10) sugere que, se ordenarmos o vetor V^z de forma crescente em relação à abundância $V_1^z / V_1^* < V_2^z / V_2^* < \dots < V_m^z / V_m^*$, o lado direito da equação tenderá a ser negativo para os índices mais baixos e positivo para os mais altos. Assim, o país terá exportações líquidas negativas dos fatores em que ele é relativamente escasso e exportações líquidas positivas dos fatores em que é abundante, que é a previsão do modelo Heckscher-Ohlin.

Com a equação de Vanek, Leamer (1980 *apud* LEAMER E LEVINSOHN, 1995, p. 1364) mostra que o paradoxo de Leontief em verdade não é um paradoxo: é possível que um país tenha exportações trabalho-intensivas sendo capital-intensivo. Isto porque a equação de Vanek é uma equação que relaciona as dotações relativas dos fatores às *exportações líquidas*, e Leontief (1954) calculou as intensidades fatoriais das exportações e importações *separadamente*. Isto só seria correto se, para o ano de 1947 (ano de estudo de Leontief), o balanço comercial americano estivesse em equilíbrio, mas em 1947 ele se apresentava superavitário. No caso de dois fatores (capital, k e trabalho, l), temos que as exportações líquidas fatoriais são $F_k = Exp_k - Imp_k = K - sK^*$ e $F_l = Exp_l - Imp_l = L - sL^*$, em que os sobrescritos referentes ao país foram omitidos por simplicidade. Com os cálculos separados, é possível que as intensidades de exportação e de importação sejam tais que $Exp_k / Exp_l < Imp_k / Imp_l$ e, simultaneamente, $K/L > K^*/L^*$.

Bowen, Leamer & Sveikauskas (1987) construíram uma versão modificada da equação $F^z = V^z - s_z V^*$ bastante intuitiva e de fácil uso. Dividindo esta equação por $s_z V_i^*$, temos a seguinte relação linear:

$$\frac{(F_i^z / s_z)}{V_i^*} = \frac{(V_i^z / s_z)}{V_i^*} - 1, \text{ para todo } i \text{ e } z.$$

Do lado esquerdo da equação acima, temos uma normalização das exportações líquidas do fator em relação à dotação mundial deste fator, ajustada pela proporção mundial de gastos de seu país. Mas é do lado direito que está a grande inovação dos autores: ele representa o quanto a dotação do fator i no país z está acima ou abaixo da média mundial, com o ajuste pelo tamanho do país. De acordo com a equação acima, se o país for relativamente bem dotado, o lado direito será positivo e o lado esquerdo, que representa as exportações líquidas, também. Se o contrário acontecer, o país z será importador líquido deste fator.

Pode-se testar o teorema de Heckscher-Ohlin de duas formas com a equação acima, ambas dependendo de cálculos separados dos lados esquerdo e direito. A primeira forma é comparar os sinais das expressões dos respectivos lados, enquanto a segunda consiste em realizar um teste de correlação de *rankings* dos dois lados. Procedendo das duas formas para 12 fatores de produção (sendo 7 tipos de trabalho) e 27 países para o ano de 1967, Bowen, Leamer & Sveikauskas (1987) encontraram direção errada das exportações líquidas no teste de sinais em mais de 30% das vezes para um terço dos fatores (entre eles o capital). Para o caso do teste de *rankings*, as violações ocorrem em 50% dos casos.

Trefler (1995) revisita o trabalho de Bowen, Leamer & Sveikauskas (1987), para dados de 1983, 9 fatores e 33 países. O "mistério do comércio perdido" de que trata o título do artigo é que a equação de Vanek não apenas prevê o sentido do comércio, mas também o volume de comércio. Os resultados são basicamente os mesmos de Bowen, Leamer & Sveikauskas (1987), visto que o teste de *rankings* apresenta correlação de 0.28 e o teste de sinais é bem-sucedido em aproximadamente 50% dos casos. Entretanto, Trefler (1995) busca algum padrão sistemático para as falhas de previsão:

1. As medidas de exportações líquidas dos fatores F_i são muito menores que as diferenças nas dotações de fatores $V_i - sV^*$, tão pequenas que, mesmo que estas medidas tenham o mesmo sinal, parece não haver comércio de fatores (este é o "mistério do comércio perdido");
2. Os desvios de previsão estão associados com o tamanho do país;
3. Quando existe correção pela PPC, os países mais pobres apresentam sistematicamente valores menores para $F_i - (V_i - sV^*)$. Isto sugere, quando este valor é positivo, o saldo de exportações é menor que o previsto pelas dotações relativas, enquanto que quando ele é negativo, as importações têm valor maior que o previsto. Para os países mais ricos, ocorre o oposto;
4. Os países mais pobres são abundantes em quase todos os fatores, ou seja, apresentam mais insumos para os quais $V_i - sV^* > 0$, enquanto o oposto ocorre para os países ricos.

À guisa de conclusão, a mensagem geral dos textos empíricos, mesmo dos que seguiram com maior rigor uma derivação do teste a partir da teoria, é de que diferenças relativas na produtividade e nas dotações dos fatores têm algum papel na explicação do comércio internacional, ainda que as teorias não se adequem perfeitamente à evidência empírica. Testes como o de Bowen, Leamer & Sveikauskas (1987) e Trefler (1995) não apresentam ajuste perfeito da teoria aos dados, porém, como nota Helpman (1998, p. 10), a teoria não é testada contra uma alternativa específica, portanto, fica difícil estabelecer qual seria um "nível de violação" suficiente para abandonar o paradigma.

Afinal, modelos estatísticos são aplicáveis à economia?

Ainda segundo Hoover (2006), a aplicabilidade de modelos estatísticos e probabilísticos aos dados econômicos nem sempre foi ponto pacífico na ciência econômica. O que garante que nos modelos econométricos os erros do modelo se comportarão de acordo com distribuições conhecidas, ou mesmo que apresentarão propriedades assintóticas (só demonstráveis teoricamente) necessárias para a aplicação de certos métodos?

Para muitos economistas de renome, como Keynes, isto simplesmente não é possível. Os dados econômicos são gerados por processos em que não é possível conhecer a distribuição de probabilidades *ex-ante*.

Em outros campos (agricultura, por exemplo), isto também não é garantido, mas a experimentação controlada fundamenta a abordagem probabilística – os trabalhos de Ronald Fisher, o pai da estatística clássica, são um exemplo disso.

Foi Trygve Haavelmo (1944) que teorizou que os dados econômicos obedecem a um **processo de geração dos dados** muito complexo, mas que pode ser aproximado pela análise de regressão, desde que o modelo esteja corretamente especificado e seja derivável de uma teoria verdadeira. No entanto, para este autor a análise precisa ser feita com muito cuidado, pois o processo de geração dos dados deriva de um sistema econômico do qual a modelagem econométrica só é capaz de captar uma parte.

Podemos caracterizar a abordagem clássica como um desenvolvimento a partir das idéias de Haavelmo, como veremos a seguir. Contudo, a idéia de processo de geração de dados é também muito cara a outras abordagens, especialmente à inglesa.

Cinco abordagens à econometria³

Gênesis

A primeira característica distintiva das abordagens econométricas é se ela é “dos dados para a teoria” ou “da teoria para os dados”. Este debate esteve presente nos primórdios da estatística e da metodologia seguida por expoentes como Ronald Fisher ou Gauss (HOOVER, 2006).

A abordagem “dos dados para a teoria” de Fisher na verdade era a aplicação da indução estatística a partir de experimentação controlada. As regressões eram usadas para avaliar se fatores independentes tinham ou não os efeitos hipotéticos, com a garantia de que os experimentos eram aleatórios e controlados e de que poderiam ser redesenhados e refeitos se fosse necessário. Desta forma, a análise de regressão poderia testar hipóteses, e os resultados destes testes constituiriam suporte empírico para teorias.

Por outro lado, a abordagem “da teoria para os dados” é foi a empregada originalmente por Gauss quando da formulação do método de mínimos quadrados ordinários. Hoover (2006) exemplifica as preocupações de Gauss com a modelagem do movimento dos planetas. A teoria era encarada como correta, Gauss estava preocupado

³ Naturalmente, podem existir outras abordagens. Estas cinco foram escolhidas puramente devido à ignorância com respeito a detalhes metodológicos de outras abordagens.

como as teorias se aplicavam aos dados empíricos e com a calibração dos parâmetros na presença de erros de observação aleatórios.

De maneira muito aproximada, podemos inferir que a abordagem clássica e a calibração partem “da teoria para os dados” e as abordagens inglesa, VAR e de experimentos quase-naturais partem “dos dados para a teoria”. Porém, analisemos estas abordagens com mais detalhe.

Abordagem clássica: da teoria aos dados

Como já mencionado, a abordagem clássica é a que se segue diretamente do desenvolvimento das idéias de Haavelmo. Há ênfase no papel da **especificação** correta e da derivação dos testes estatísticos concernentes à teoria, que fornecerão a ligação entre teoria e dados.

Os adeptos da abordagem clássica reconhecem que um modelo nada mais é do que uma representação simplificada de um processo do mundo real. De acordo com Koopmans (1957), um modelo representativo deve ser o mais simples possível e obedecer à regra da parcimônia no que se refere ao espectro de variáveis relevantes para a explicação do fenômeno a que se propõe. Dessa forma, um modelo econômico é constituído de um conjunto de hipóteses que retratam de forma aproximada o comportamento de uma economia. Nas décadas de 1950 e 1960, a preocupação desta geração de econométricos esteve centrada na formulação de modelos econômicos sistêmicos cuja representação econométrica produzisse hipóteses testáveis relevantes à ciência econômica, principalmente quanto à utilização dos resultados na previsão e avaliação dos resultados de políticas econômicas.

Assim, um modelo econométrico dentro desse arcabouço é resultado de um conjunto de equações comportamentais derivadas do modelo, as quais envolvem variáveis observáveis e um termo aleatório ou randômico, que contém todos os fatores que não foram incorporados ao modelo em análise. Além disso, contém afirmações sobre a existência de erros de observação em variáveis do modelo e sobre a especificação da distribuição de probabilidades do termo aleatório. A ênfase sistêmica é o que deu origem às metodologias de estimativas de equações simultâneas, por exemplo. O objetivo dessa formulação é prover uma forma representativa passível de teste empírico, por meio de estimação, teste e checagem do diagnóstico produzido.

Sabe-se, no entanto, que os modelos econômicos muitas vezes simplificam exageradamente uma dada realidade. Alguém poderia argüir que, basicamente, todo e qualquer modelo implica a omissão de variáveis. Entretanto, de acordo com Friedman (1953):

"The relevant question to ask about the 'assumptions' of a theory is not whether they are descriptively 'realistic' for they never are, but whether they are sufficiently good approximations for the purpose at hand. And this question can be answered by only seeing whether the theory works, which means whether it yields sufficiently accurate predictions."

Por isso, a discussão em torno de quão realistas são as hipóteses de um modelo econômico não é uma tarefa de fácil solução. A utilização de métodos estatísticos de estimação proporciona estimativas passíveis de teste. Após essa etapa, procedemos ao teste e checagem de diagnóstico do mesmo. Isto é, testamos as implicações estimadas, além de

verificarmos se o diagnóstico produzido pelas estimativas se ajustam à realidade empírica que nos é apresentada.

Em resumo, podemos caracterizar a abordagem clássica como um método de análise que tem como ponto de partida a teoria econômica. Com base nisso, o pesquisador é capaz de construir um modelo teórico que encerre as principais conclusões teóricas. Em seguida, ao caracterizarmos a distribuição de probabilidades do termo aleatório, produzimos um modelo econométrico que está pronto para ser estimado e testado. A etapa seguinte envolve a utilização de métodos estatísticos de estimação do modelo econométrico, além do teste das implicações produzidas por ele. O conjunto de dados empíricos observáveis, assim como a inferência estatística, é o suporte dessa etapa. Finalmente, caso o modelo econométrico não seja passível de rejeição, pode ser utilizado na previsão do comportamento de variáveis econômicas ou para a elaboração de cenários, quando da formulação de políticas.

Essa abordagem, então, traz à tona uma dicotomia clara: enquanto ao teórico caberia a construção do modelo econômico, o econometrista seria responsável por sua estimação e seu teste. Dessa forma, o trabalho econométrico passava a se constituir basicamente da estimação de regressões estatísticas, sendo que, como consequência, o desenvolvimento da econometria teórica até os anos 1960 se deu, principalmente, em torno da elaboração de estimadores visando à flexibilização da especificação a partir de diferentes condições – eis a gênese dos métodos de “estimação robusta”.

Porém, foi justamente essa flexibilidade que levaria ao descrédito na efetividade daquela dicotomia. Do ponto de vista epistemológico, se por um lado a abordagem clássica sempre sofreu do círculo vicioso do empirismo - como se parte de teorias válidas empiricamente se a observação empírica só pode ser obtida sob a égide de uma teoria verdadeira? – por outro lado o desenvolvimento das estimações robustas tornou a teoria econômica praticamente infalsável, já que nunca estaria claro se a rejeição de um determinado modelo empírico lançaria dúvidas sobre o modelo teórico ou sobre as suposições do método de estimação.

Do ponto de vista meramente estatístico, o uso de sistemas na abordagem clássica fornecia previsões de ajuste muito ruim, não raro piores do que as fornecidas por modelos de apenas uma equação ou mesmo de apenas uma variável. Por sua vez, os economistas teóricos insistiam que a representação econométrica não era adequada à teoria. A popularização dos modelos de otimização dinâmica na presença de expectativas racionais e mercados incompletos levou à hoje famosa crítica de Lucas, segundo a qual as relações estimadas econometricamente não eram estáveis nem estruturais. Por fim, os tomadores de decisão e formuladores de políticas começaram a contestar o uso prático destes métodos econométricos. Em suma, nos dizeres de Pesaran e Smith (1995), a abordagem clássica se mostrava nos anos 1970 estatisticamente *inadequada*, teoricamente *inconsistente* e praticamente *irrelevante*.

A resposta a essa questão foi então, a partir da década de 1970, o divórcio entre teoria e econometria. De um lado, uma parte dos economistas teóricos passou a construir modelos sem qualquer preocupação com a comprovação empírica destes e, de outro, muitos econometristas passariam a desenvolver procedimentos sem preocupação explícita com o embasamento teórico, muitas vezes focalizando exclusivamente no ajuste aos dados.

A abordagem VAR: dos dados, sem teoria

A preocupação da abordagem clássica com equações estruturais diretamente deriváveis da teoria freqüentemente recomendava o uso de sistemas que necessitavam de tantas restrições de identificação que muitas vezes a própria teoria não era capaz de prover (HOOVER, 2006). Esta foi a motivação para o artigo de Christopher Sims chamado “Macroeconomics and Reality” (1980), que lançou uma nova abordagem macroeconômica: a abordagem VAR (*Vector Autoregression*).

A abordagem VAR é aquela em que sistemas irrestritos de equações na forma reduzida são estimados a fim de obter funções de impulso-resposta das variáveis em relação às outras. Na prática, cada equação do sistema VAR regride uma variável em função de defasagens dela mesma e de outras variáveis. Há critérios para a escolha do número de *lags* e um tratamento especial dos resíduos, uma vez que geralmente eles estão inter-relacionados no sistema.⁴

Segundo Hoover (2006), de início não havia preocupação explícita com causalidade nem com simultaneidade. Além disso, ainda que hoje em dia vários adeptos da abordagem VAR se sintam desconfortáveis, a abordagem original de Sims (1980) é declaradamente ateorica. Porém, diante de diversas críticas Sims (1986 *apud* Hoover, 2006) reconheceu que a ordem de causalidade pode afetar as funções impulso-resposta, então propôs a aplicação do chamado SVAR (*structural VAR*), em que um VAR é modificado para lidar com uma ordem causal particular. Em verdade, Pesaran e Smith (1998) notam que a abordagem VAR não precisa ser necessariamente ateorica uma vez que se podem impor restrições nos sistemas derivadas de expectativas racionais e otimização dinâmica.

Contudo, segundo Hoover (2006), a metodologia SVAR precisa buscar na teoria as relações de causalidade, mas quase sempre não consegue uma incorporação formal das mesmas. Neste sentido, a abordagem VAR sofre do mesmo problema da abordagem clássica: quais restrições no sistema precisam ser impostas e, dentre estas, quais são críveis, uma vez que nada garante que os dados econômicos seguem as hipóteses do VAR? Além disso, este autor ainda argumenta que a defesa das relações de causalidade em sistemas SVAR é geralmente informal e casuística.

A abordagem inglesa: dos dados à teoria

Também nos anos 1970 e 1980 surgiu a chamada a abordagem da London School of Economics (LSE), também conhecida como abordagem inglesa, a partir de trabalhos de Denis Sargan e David Hendry. Estes trabalhos foram na contra-mão das proposições da abordagem clássica. Entre as várias críticas imputadas a esta, destacam-se três, a saber (HENDRY, 1987, e SPANOS, 1989):

1. Há diferenças entre modelos econométricos propostos e o processo de geração dos dados produzidos pela realidade. Isto é, a associação entre uma variável teórica e a sua representação estatística não é nada mais do que uma forma de ajustamento do modelo aos dados.

⁴ Não detalharemos a metodologia VAR porque ela é essencialmente macroeconômica, e não se aplica diretamente aos trabalhos com microdados de empresas, no que tange à produção, tecnologia e inovação.

2. Hendry (1987) admite que as informações contidas nos bancos de dados podem ir além do previsto por um arcabouço teórico qualquer. Logo, é criticável a posição clássica de tomar como ponto de partida para a modelagem econométrica dado enfoque teórico. Na verdade, é imprescindível deixar espaço para que informações adicionais contidas nos dados aflorem e permitam um espaço científico mais amplo ao pesquisador. Isso, entretanto, não implica abandonar as técnicas utilizadas pela abordagem tradicional. A diferença mais significativa é a maneira como esse modelo (de regressão) é inicialmente formulado. Ou este é advindo da teoria (abordagem tradicional), ou é o mais genérico possível, para daí, extraíndo os "excessos" (via análise estatística), chegarmos a um modelo "enxuto" representativo da realidade.
3. Por último, não podemos tomar a tradução imediata de um modelo econômico em modelo econométrico como um fato sem contestações. Quando introduzimos um termo aleatório no modelo econômico e afirmamos ser ele então um modelo econométrico, estamos qualificando um modelo estatístico cuja base de dados conveniente é experimental. Essa afirmação, segundo Hendry (1987) e Spanos (1989), não é tão óbvia assim, porque não há correspondência entre um modelo estatístico e os dados observáveis que escolhemos para representar as variáveis teóricas nele contidas. Além disso, a identificação entre dados experimentais e aqueles gerados pela natureza não é evidente.

Sendo assim, a modelagem econométrica, segundo a abordagem da LSE, deve ser construída com base em um modelo estatístico cuja especificação seja a mais generalizada possível, retratando um mecanismo que mimetize as informações contidas no conjunto de dados, tendo em vista os modelos estimáveis.

Podemos apresentar um sumário dessa abordagem da seguinte forma: a teoria econômica, como arcabouço lógico de idéias, não pode estar totalmente descolada do processo gerador que caracteriza o comportamento dos dados na natureza. Com base nisso e na interação entre esse processo e o conjunto de dados coletados, construímos um modelo econômico estimável. Esse modelo é produto da construção de um suporte estatístico generalizado, ou seja, o modelo estatístico proposto não admite a simplicidade como ponto de partida, e sim a generalização. Este é um ponto importante, porque qualifica esta abordagem como "do geral para o específico" com respeito à modelagem.

A partir disso, procede-se à sucessiva eliminação de especificações concorrentes por meio do princípio da parcimônia, ou seja, buscando a especificação mais parcimoniosa possível que contenha as mesmas informações encontradas no modelo mais geral⁵. Assim, existe uma clara inter-relação entre os vários estágios decisivos do método, ao mesmo tempo em que se manifesta uma preocupação em seguir um processo analítico que parte da generalização e chega à simplificação.

No campo da macroeconometria, a abordagem inglesa é aquela identificada com os modelos de equilíbrio de longo-prazo expressos por modelos de correção de erros e análises de cointegração. Porém, sua sistemática pode ser aplicada a outros campos de estudo, como veremos em seção posterior.

⁵ Para uma visão mais detalhada deste procedimento, ver Mizon (1984) e Hendry (1988).

Apesar do rigor estatístico envolvido nessa abordagem, determinadas críticas são comumente levantadas a seu respeito. Primeiramente, o modelo geral é especificado a partir do bom senso, da disponibilidade de dados e de análises exploratórias. Dessa forma, enquanto de um lado não há qualquer garantia de que a adequação do modelo ao suposto processo de geração de dados corresponda ao verdadeiro processo (este processo de escolha do modelo pode levar à “convergência a um máximo local”), de outro é possível a existência de uma especificação alternativa não aninhada na especificação geral (HOOVER, 2006).

Em segundo lugar, segundo Faust and Whiteman (1995 e 1997), a grande quantidade de testes realizados ao longo do procedimento tornaria as estatísticas reportadas impassíveis de interpretação. A observação dos autores se deve a evidências de que os algoritmos de *data-mining* envolvidos na transformação do geral para o específico seriam responsáveis por uma fonte de erro a partir das distorções de tamanho dos testes. Basta lembrar que quando um modelo está mal-especificado, os testes estatísticos podem estar errados; porém se a essência da abordagem inglesa é a busca pela melhor especificação, inevitavelmente teremos que testar coeficientes a partir de especificações preliminares. Deste modo, não se pode garantir que este processo convergirá para o modelo “verdadeiro”.

Por fim, cabe salientar que o resultado final do método pode levar a resultados de grande qualidade estatística, mas de escasso significado econômico. Ainda que as teorias de cointegração representem um passo adiante em relação aos modelos puros de séries temporais, Pesaran e Smith (1995) criticam a abordagem inglesa no sentido de que a teoria tem um papel somente na especificação (flexível) da relação de longo-prazo a ser estimada.

A calibragem: da teoria à simulação, da simulação aos dados⁶

Outra reação às críticas à abordagem clássica foi a proposição de uma agenda de pesquisa ainda mais comprometida com a teoria: a quantificação dinâmica de modelos macroeconômicos de equilíbrio geral, sugerida por Finn Kyndland e Edward Prescott (1991). De acordo com esta abordagem, não há “teste” da teoria econômica, uma vez que o processo se inicia com um modelo teórico de crescimento ou de ciclos de negócios (em geral, a partir de agentes representativos com expectativas racionais) que tem seus parâmetros ajustados, a fim de se efetuar uma simulação para posterior comparação dos resultados com os dados reais. Ao contrário da abordagem clássica, os parâmetros nestes modelos não são estimados, e sim tirados das contas nacionais, das “grandes razões macroeconômicas” (como relação capital-trabalho, investimento/PIB etc), de fatos estilizados, do senso comum, enfim, de outras fontes informais (HOOVER, 2006).

Os adeptos da calibração costumam defendê-la como a “verdadeira econometria” na essência do tripé teoria econômica, matemática e estatística (a despeito da ênfase estar nas duas primeiras). A calibragem teria vantagens sobre as abordagens estatísticas porque:

1. É possível partir diretamente dos modelos, que segundo os adeptos da calibragem não estão sujeitos a teste.
2. Ainda que sejam aproximações da realidade, os modelos a serem calibrados podem ser desenhados a fim de mimetizar apenas as relações de interesse, e

⁶ Esta é outra metodologia que não detalharemos por ser essencialmente macroeconométrica.

3. Esta é uma vantagem em relação às de, por exemplo, máxima verossimilhança porque a parametrização de modelos que seguem este princípio leva em conta o ajuste total do modelo, o que inclui dimensões às vezes irrelevantes para a análise.

No entanto, Hoover (2006) enumera duas críticas contundentes a esta abordagem. A primeira é que não existem regras formais para o processo de escolha dos parâmetros para o modelo a ser simulado, como também não existem critérios objetivos para avaliação do ajuste da simulação aos dados reais e, por consequência, para comparação entre modelos rivais.⁷

Por sua vez, Pesaran e Smith (1995) consideram a calibragem um “passo atrás” para a econometria, uma vez que a ênfase está na estimação e não em testes ou avaliação de modelos. Além disso, para os modelos serem computacionalmente operacionalizáveis, eles precisam ser simples e supor agentes representativos com conjuntos de informação homogêneos. Sem embargo, a modelagem econométrica “homogênea” é a regra até alguns econométricos mudarem seu foco para a microeconometria. Esta é justamente a próxima abordagem que analisaremos.

Microeconometria: do desenho de experimentos aos dados, dos dados à teoria⁸

As primeiras bases micro-econômicas disponibilizadas nos anos 1950 revelaram padrões e características que não eram facilmente previstas nos modelos teóricos de consumo e demanda vigente na época. Indivíduos relativamente semelhantes apresentavam escolhas diferentes, distintas remunerações composições de riqueza, demonstrando a fraqueza do conceito de agente representativo (HECKMAN, 2000, p. 265). Revelaram-se importantes fontes de heterogeneidade e diversidade individual, antes escondida sob a informação agregada das séries temporais. As ferramentas econométricas utilizadas até então foram alvo de evoluções e reformulações e algumas vezes abandonadas por novas abordagens. Apesar de uma ampla classe de modelos quantitativos já estabelecidos, ocorreu um esforço no desenvolvimento de toda uma nova variedade de modelos microeconômicos, dada a necessidade de explorar e interpretar novas fontes de informação disponíveis nas novas bases de dados. Este desenvolvimento em métodos quantitativos aplicados à microeconomia procurava um adequado tratamento dos dados no estabelecimento de relações de causa e efeito, e interpretações de fatos econômicos em situações onde os casos contrafactuais não estavam disponíveis. Esta é basicamente a gênese da microeconometria (HECKMAN, 2000, p. 266).

Em verdade, a direção das análises caminhou do macro para o microeconômico, estimulando a formulação de teorias que procuravam explicar o comportamento das unidades observacionais. As análises econométricas baseadas em respostas ao nível microeconômico descobriram uma alta diversidade e heterogeneidade existente na realidade da vida econômica. Essas novas evidências desconstruíram a figura de um

⁷ Alguns economistas, jocosamente, chamam a calibração de “econometria sem variância.”

⁸ Esta será a parte que analisaremos melhor, pois guarda relação direta com os estudos da produção, tecnologia e inovação.

indivíduo representativo correspondendo ao comportamento médio da população, e conseqüentemente modificaram a forma como se pensava e avaliava em políticas públicas.

A microeconometria, ao menos em sua vertente quase-natural, busca lidar com a simulação de experimentos quase-naturais – inclusive, expressões como “tratamento”, “grupos de casos e controles e outras são oriundas de experimentação natural. Neste sentido, a abordagem pode ser encarada como sendo “dos dados para a teoria”, as evidências empíricas servem para avaliar a validade de explicações rivais, embora de uma maneira que tenta mimetizar os experimentos de Fisher, só que desta vez com dados econômicos e com a pseudo-experimentação.

A fim de avaliar o efeito do tratamento, é necessário fazer uma avaliação contrafactual, isto é, responder à pergunta: “O que aconteceria caso aqueles que receberam o tratamento não o tivessem recebido (ou vice-versa)?”. Em outras palavras, “qual o efeito médio do tratamento”.

Caso a distribuição entre os agentes que recebem o tratamento e os que não fosse aleatória dentro da amostra, estimar tal efeito médio seria muito simples: bastaria testar *ex-post* a diferença de médias da variável supostamente impactada pelo tratamento, para o grupo de casos (composto por quem recebeu o tratamento) e de controles (composto por quem não o recebeu). Está implícita a suposição de que se que foi tratado não o tivesse sido, as variáveis de interesse para estes indivíduos teriam valores semelhantes aos dos que não receberam – o que é uma forma de imputação, uma vez que não dispomos destes dados.

Entretanto, o problema não é tão trivial porque em geral os agentes econômicos se auto-selecionam: por exemplo, indivíduos desempregados escolhem participar de programas de requalificação, de forma que é preciso levar isto em conta ao analisarmos os impactos destes programas nos salários destes indivíduos; firmas decidem exportar ou não, de forma que a avaliação dos impactos da exportação sobre a competitividade precisa levar isto em conta.

No entanto, quando os determinantes da auto-seleção são conhecidos e suficientes pode-se recorrer aos chamados experimentos quase-naturais. Em um experimento quase-natural, a suposição básica é de que a auto-seleção se dá a partir de características observáveis, de forma que, condicionalmente a estas características, a distribuição do tratamento é pseudo-aleatória. Em outras palavras: se um par caso-controle apresenta determinantes observáveis da auto-seleção tão semelhantes, então se pode tomar a diferença de médias na variável de interesse como forma de avaliação do efeito do tratamento, pois a única diferença relevante para o problema de pesquisa entre os indivíduos é o fato de um ter recebido o tratamento e o outro não.

Existem várias formas de manifestação da auto-seleção: heterogeneidade entre os agentes, participação não-aleatória em programas ou até mesmo informações faltantes para indivíduos ou atrito de painel.⁹ Trata-se a auto-seleção de diversas maneiras: o controle estatístico de médias condicionais a partir observáveis em equações lineares, algoritmos de

⁹ Talvez o caso mais famoso da relação entre informações faltantes e auto-seleção foi o motivador da análise hoje conhecida como *Heckit*: dados faltantes sobre horas trabalhadas e salários significam que os salários não estão identificados para trabalhadores desempregados, embora estes possuam uma função de oferta de trabalho. Como mensurar efeito sobre os salários ou sobre a oferta de mão de obra uma vez que os salários estão disponíveis somente para os trabalhadores efetivamente empregados? Esta informação é muito relevante e inferências sobre a força de trabalho como um todo não podem se basear somente nos dados para aqueles que participam do mercado de trabalho, a participação ou não também é uma informação relevante para a modelagem.

matching a partir de seleção em observáveis, modelos em que a equação de participação é levada em conta (modelagem *Heckit*) e outras. Não está no escopo deste artigo discutir estas técnicas em detalhe, o leitor interessado pode buscá-la em manuais de econometria como o de Wooldridge (2002).

Na visão de Heckman (2000), uma agenda de pesquisa econométrica deveria focar em três aspectos:

1. Abordagem econométrica estrutural: de escopo mais amplo. Aqui, há a complexidade de se estimar modelos estruturais com heterogeneidade nos interceptos, e os parâmetros econômicos e comparações de bem-estar são bem definidos.
2. Abordagem do efeito de tratamento: mais indicada para análises de programas com cobertura parcial (grupos de tratamento e controle). Mais interessante para questões relacionadas a políticas ou programas de cobertura parcial.
3. Abordagem do efeito de tratamento na margem: podem ser implementados em associação com modelos não paramétricos de equilíbrio geral.

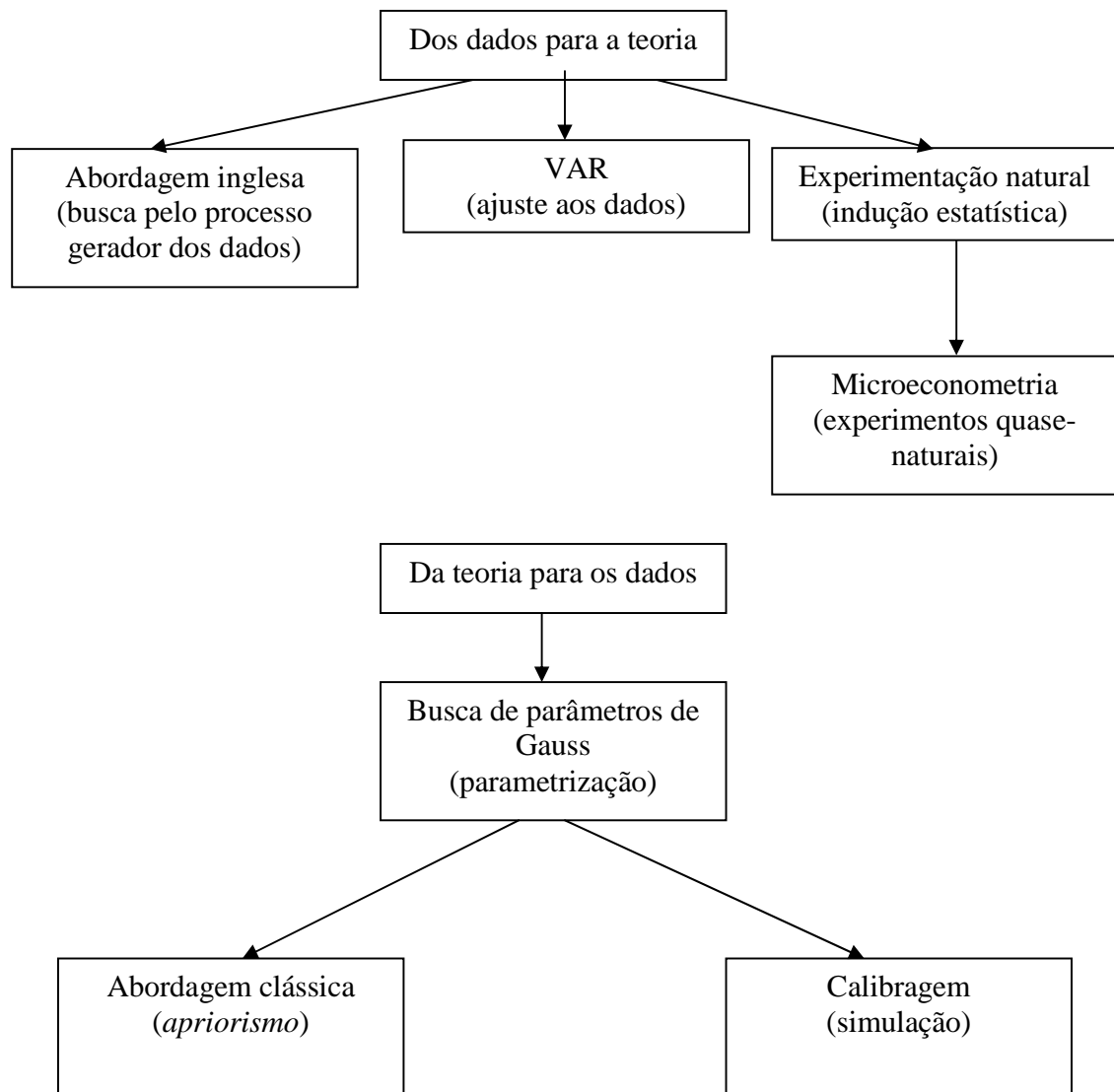
Em todos os casos, é indispensável compreender as fontes de identificação dos modelos, relaxar formas funcionais arbitrárias bem como as hipóteses de distribuição das variáveis de interesse ou modeladas. A área em que houve mais avanços foi a segunda, mas mesmo assim há de se reconhecer que modelos de simulação para novas propostas de políticas públicas são bastante limitados.

Por fim, Heckman (2000) reconhece que evidências construídas a partir de análises microeconômicas já influenciaram o desenvolvimento de teorias macroeconômicas que tem abandonado a figura do agente representativo. Métodos microeconômicos foram desenvolvidos com o objetivo de avaliar um conjunto de políticas sociais, entretanto o próprio autor (HECKMAN, 2000, p.307) reconhece que o alcance dos métodos quantitativos em microeconomia possui suas limitações. Por exemplo, uma política de incremento da escolaridade de toda a população terá como consequência a diminuição do retorno para a educação (embora seja boa), e a metodologia de tratamento de efeitos pode se ineficiente para a avaliação de programas de cobertura universal (devido à dificuldade de se estabelecer grupos de comparação).

Genealogia das abordagens

De forma ilustrativa, eis um esquema da genealogia das abordagens econométricas. Como no começo da seção, a divisão básica foi o emprego de estratégias “dos dados para a teoria” ou “das teorias para os dados”. Questões cronológicas não foram levadas em consideração.

Diagrama 1: Genealogia das abordagens econométricas



Elaboração própria.

Antes da estimação

Nesta seção, discutiremos o que os especialistas têm a dizer sobre como proceder antes de começar a estimar modelos econométricos. Ainda que existam várias recomendações “transversais”, algumas recomendações não estão desprovidas de contexto, isto é, cada autor elabora seus conselhos a partir de um paradigma. Sempre que evidentes, estes paradigmas permearão a discussão.

A relação teoria-modelo

O problema que permeia qualquer abordagem econométrica, exceto a abordagem original do VAR (abordagem atórica) e da calibragem (abordagem puramente teórica), é o “buraco negro” entre a teoria econômica e a prática econométrica. Para complicar, livros textos respeitáveis repetem problemas-padrão e soluções-padrão que fazem com que os estudantes fiquem presos a este universo “padronizado” e não se tornem aptos a pensar sobre as surpresas do mundo real da econometria aplicada. No contexto real não existem problemas-padrão, no entanto as soluções-padrão persistem e os econométricos aplicados muitas vezes são obrigados a fazer transformações *ad hoc* em soluções padrão.

Em linhas gerais, a relação teoria-modelo e a precedência ou não da teoria em relação aos dados já foram discutidas na seção anterior. Vimos que a abordagem clássica e a calibragem dão bastante ênfase à teoria. Os adeptos da abordagem clássica se preocupam se os modelos econométricos são deriváveis da mesma, e as interpretações dos resultados também a ela estão sujeitas. No entanto, mesmo os adeptos da abordagem clássica têm que reconhecer que, ainda que a teoria econômica deve ser o ponto de partida e a força guia da especificação do modelo econométrico, o processo de especificação de modelos não é de forma alguma trivial. Não há um consenso sobre os passos corretos para se chegar a uma especificação correta, mesmo quando suportado por uma teoria, que é abstrata. É um processo de inovação no qual ainda não há conclusões sobre o que é correto e incorreto e, antes de tudo, sobre como julgar o “correto” e o “incorreto”.

Esta concepção mais realista com respeito à relação teoria-modelo, mesmo para as abordagens mais comprometidas com a teoria, vai de encontro às proposições dos manuais de econometria, que tomam a especificação correta do modelo como um pressuposto:

“The regression model is correctly specified. Alternatively there is no specification bias or error in the model used in empirical analysis.”(GREENE, 2003, 5ª ed., p. 78).

Ou, mais ironicamente:

“As you wander through the thicket of models, you may come to question the meaning of the Econometric Scripture that presumes the model is given to you at birth by a wise and beneficent Holy Spirit”(LEAMER 1996 *apud* KENNEDY, 2002, p. 579).

Outro ponto digno de nota é mesmo as abordagens menos comprometidas com a estimação de equações estruturais não subestimam o papel da teoria para a especificação, pelo contrário. A teoria econômica é o ponto de partida para especificação. A diferença, notadamente da abordagem inglesa e da microeconometria a este respeito, é que, como mencionamos, admite-se que as informações contidas nos bancos de dados podem ir além do previsto por um arcabouço teórico. É interessante estabelecer um “diálogo” entre teoria e dados, deixar espaço para que informações adicionais contidas nos dados aflorem e permitam um espaço científico mais amplo ao pesquisador. Mas isto também não quer dizer que estes pesquisadores não tenham uma estratégia definida para estabelecer este diálogo. No entanto, vamos deixar a exposição destas estratégias para a seção seguinte.

Análise exploratória dos dados

Independentemente da abordagem escolhida, torna-se pré-requisito fundamental um estudo detalhado dos dados utilizados antes da aplicação mecânica de modelos prontos que podem ser mal especificados e inadequados, dado o contexto e o comportamento dos dados em questão. Belsley e Welch (1988), enfatizam bem a importância de se conhecer o contexto antes de partir para especificação e estimação dos resultados:

“Don’t try to model without understanding the nonstatistical aspects of the real-life system you are trying to subject to statistical analysis. Statistical analysis done in ignorance of the subject matter is just that – ignorant statistical analysis.”(BELSLEY E WELCH, 1988, p.447)

Inspecionar os dados implica em saber como fazê-lo, como procurar o que se deve encontrar e interpretar o que foi encontrado, como se comprometer a quebrar a lacuna entre o teórico e o aplicado. Infelizmente, mesmo com uma vasta literatura sobre análise de dados, especificação, etc, não é possível determinar as regras de bolso para se fazer isto corretamente.

Segundo Kennedy (2002), dados são números dentro de um contexto, então é imprescindível conhecer o contexto e os conceitos relativos àqueles dados. Na lide com as pesquisas de inovação, a recomendação é extremamente válida. Existe um debate muito intenso sobre métrica da inovação, mas é consenso de que o processo de coleta dos dados e a interpretação dos conceitos de inovação por parte dos respondentes devem ser levados em conta na interpretação dos resultados, para citar apenas dois aspectos das pesquisas de inovação nas empresas. Por exemplo, na Europa as pesquisas de inovação são respondidas por correio, o processo de resposta ao questionário não é assistido e há baixas taxas de resposta. Ainda, a resposta não é aleatória, o que envia a extrapolação dos resultados para o conjunto da indústria nestes países.

No quadro 2 a seguir, há um exemplo de como erros de medida nos dados influenciaram uma série de decisões judiciais em favor de minorias norte-americanas com respeito a discriminação por parte dos bancos na concessão de empréstimos.

Quadro 2. Erros de medida, conclusões enviesadas

No início dos anos 90 vários bancos americanos foram processados por discriminarem negros e hispânicos na oferta de crédito. Day e Liebowitz (1998) contestam as evidências empíricas que motivaram estes processos. Eles apontam inconsistências na base de dados utilizada nestes estudos:

"There are good economic reasons to be skeptical of claims that lenders discriminate against minorities in their approval of mortgage applications. Discriminators who would turn down a good loan harm themselves by turning down a profit opportunity. Further, the current regulatory climate has put great pressure on mortgage lenders to ensure that its employees do not discriminate against minorities.

The support for the belief that banks discriminate is based largely on the data constructed by the Boston Fed. Yet, we have shown this data set to be deeply flawed in a way that is likely to bias the results. Although some other researchers believe that the errors in the data can be repaired and that such repaired data support the conclusion of the Boston Fed, our analysis of this data indicates otherwise. Our reworking of the data provides no evidence for the conclusion that banks systematically discriminate against minority groups. But we find it unreasonable to think that all the errors in this data set can be found with the techniques at hand. It seems imprudent, to us, to base any policy decisions on analyses of these data.

Finally, we must ask whether this type of problem is endemic to other studies of discrimination. If imperfect data tend to cause findings of discrimination where none may occur, then extreme vigilance is required by those conducting such studies to ensure that the data used are pristine. Have researchers taken sufficient care in their creation and use of data? Are the data used sufficiently good proxies for the purposes to which they are put? At this time we can only ask the questions— others will have to provide the answers." (DAY E LIEBOWITZ, 1998, p. 21)

Durante a estimação

Novamente, a relação teoria-modelo – especificação

Como vimos, a especificação econométrica depende da abordagem adotada. Segundo Kennedy, em seu livro “*A Guide to Econometrics*” (1998), existem três estratégias que tratam deste problema. A estratégia mais usual dos trabalhos empíricos inicia o processo de modelagem partindo de um modelo simples para um modelo mais geral, guiado por testes de diagnóstico. Esta estratégia é conhecida como *average economic regression* (AER) e se identifica muito com a abordagem clássica, e, em certa medida, com certas aplicações microeconômicas. Ela parte de um modelo teórico inicial, tomado como correto, para uma primeira especificação da regressão. O modelo é julgado com base nos resultados de testes de diagnóstico. Caso o teste aponte para algum problema, o econometrista é induzido a trocar o estimador original por um mais robusto.¹⁰ Caso o novo estimador não resolva o problema, o pesquisador é levado a procurar uma especificação alternativa para o modelo. Esta busca é geralmente orientada por valores elevados do coeficiente de determinação (R^2) e da estatística t dos coeficientes que o pesquisador acredita, a priori, serem diferentes de zero.

Portanto, a base desta abordagem é manter uma sensibilidade pelo que é mais simples devido à impossibilidade de incluir em um modelo todas as possíveis variáveis explicativas em suas devidas formas funcionais.

A vantagem de se iniciar a análise com um modelo mais simples (abordagem *bottom-up*) estaria no fato que os modelos mais sofisticados seriam mais sensíveis a erros e inconsistências nos dados, à “sujeira” e ao caráter não experimental dos dados. Nos modelos mais simples, a origem das falhas no modelo é mais fácil de ser identificada. Além disso, tendem a estimular o aprofundamento do conhecimento, pois aprender como e porque um modelo possui baixa performance é uma importante informação para o desenvolvimento do modelo.

A crítica a esta abordagem parte daqueles que acham que é preciso cautela na decisão de substituir um estimador no primeiro sinal de problema apontado por algum teste de diagnóstico (veja seção 3, sobre a infalibilidade da teoria que este tipo de estratégia pode resultar). Defendem que o correto é reavaliar cuidadosamente a especificação do modelo, antes de condenar o estimador. Pois, caso a especificação do modelo esteja muito distante do “verdadeiro modelo” a probabilidade de erro tipo I do teste em questão não será válida.¹¹ A utilização do R^2 e do valor “correto” do sinal dos coeficientes estimados como guias para a incorporação de informações *ad hoc* no modelo também é vista com ceticismo (a este respeito, veja a subseção seguinte sobre *data mining*).

Uma segunda abordagem, chamada *test, test, test (TTT)* e imune a esta crítica, considera a econometria uma ferramenta para testar modelos teóricos rivais. Inicia-se com uma especificação do modelo mais geral do que a expectativa do pesquisador em relação à especificação que será selecionada no final. Várias restrições são testadas com o objetivo de simplificar a especificação inicial do modelo. Em seguida, o modelo é sujeito a uma bateria

¹⁰ Por exemplo, se o teste de Durbin-Watson indicar que os erros possuem correlação serial, o estimador de Mínimos Quadrados Ordinários deve ser substituído por um de Mínimos Quadrados Generalizados.

¹¹ O erro tipo I é definido como a probabilidade de errar ao rejeitar a hipótese nula.

de testes de diagnóstico, com a finalidade de encontrar problemas na especificação do modelo. Diferente da abordagem anterior, um diagnóstico negativo, tal como um pequeno valor do Durbin Watson, é interpretado como um problema de especificação em vez de uma necessidade de um estimador mais sofisticado. Este processo é repetido até que os resultados dos testes de diagnóstico permitam ao pesquisador concluir que o modelo é aceitável. Esta abordagem é mais identificada com a abordagem VAR e com a da LSE.

A maior vantagem da abordagem “geral para específica” (*top down*) é que em casos onde o modelo geral incorpora o verdadeiro modelo gerador dos dados, os erros tipo I dos testes de diagnóstico são aqueles que realmente foram especificados pelo pesquisador, tipicamente 5%. Por outro lado, não se deve esperar que, com frequência, o modelo geral incorpore exatamente o verdadeiro modelo. Portanto, o máximo que pode ser dito é que a abordagem “geral para específica” é menos propensa a gerar testes enviesados.

Alguns dos problemas desta abordagem são citados por autores como Magnus (1999):

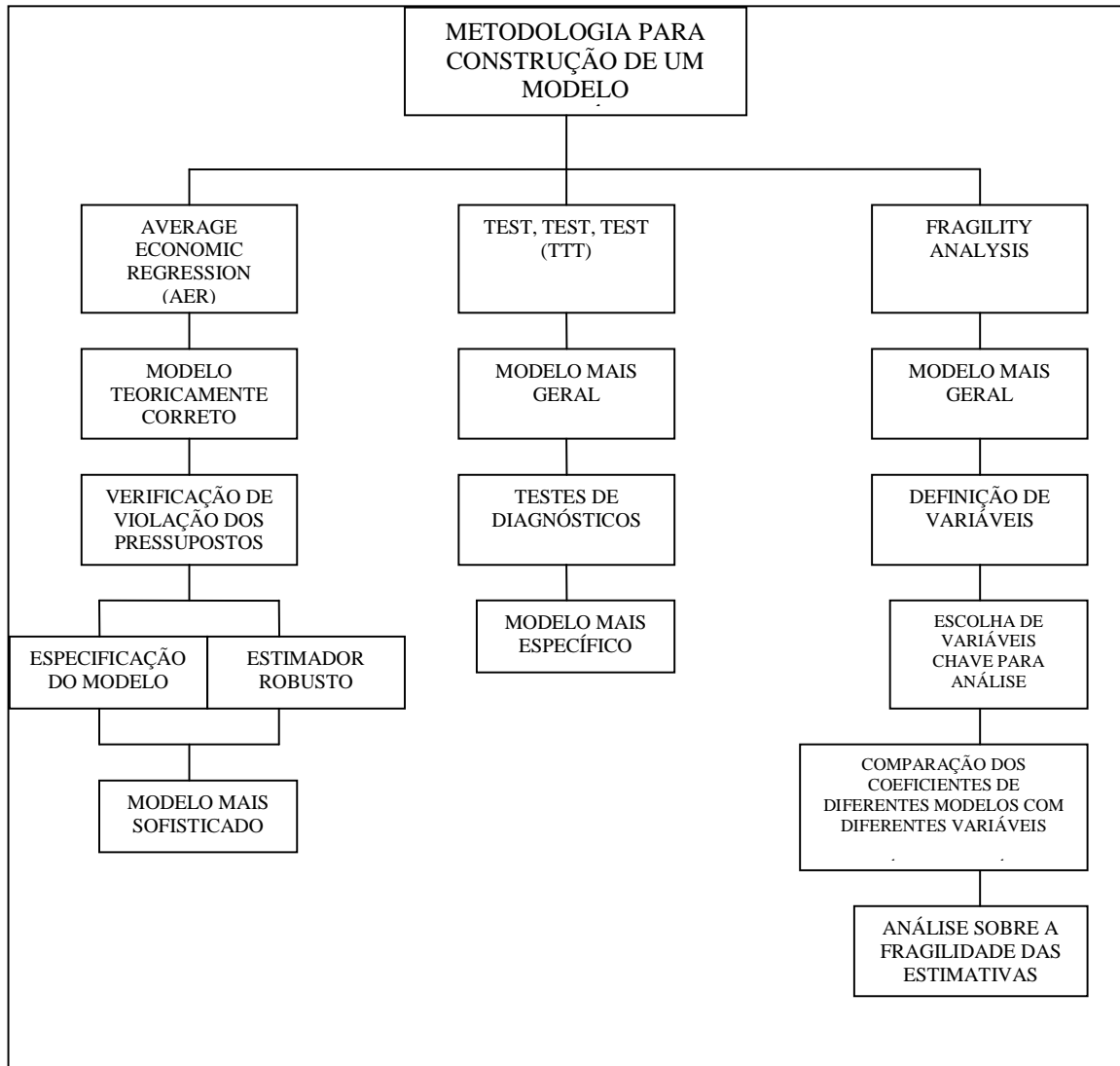
“... it does not work. If you trying to estimate such a large model, which has everything in it that you can think of, you get nonsensical results. Everyone who has done empirical work knows this. The second problem is that you cannot really discover anything new and interesting in this way, because the interesting bit is the constructions of the top model and we are not told how this is done. Therefore no applied economist proceeds in this way. Instead they follow the bottom-up approach. In the bottom-up approach one starts with a simple model and builds up from there. This is, in fact, how scientists in other discipline work.”

Ou seja, alguns econometristas não acreditam na possibilidade de estimar um modelo completamente geral, por isso na prática faz-se necessário adotar uma forma simples, que significa começar a especificação de “baixo para cima”. Uma outra desvantagem de modelos que partem do geral para o específico é que a forte dependência nos testes de diagnóstico aumenta a chance de erros tipo I em cascata, este problema é agravado pelo pequeno grau de liberdade devido à especificação geral do modelo e pelo fato que para vários testes apenas as propriedades assintóticas são conhecidas.

Por fim, a metodologia conhecida como *fragility analysis*, inicia o processo de especificação selecionando uma família geral de modelos para levar a cabo uma *extreme bound analysis*, nos quais os coeficientes das variáveis-chaves são estimados utilizando todas as possíveis combinações das outras variáveis do modelo cuja relevância não se tem certeza. Quanto menor for a distância entre o menor e o maior valor das várias estimativas do parâmetro de interesse, mais confiável será a informação produzida. Se esta distância for grande, uma tentativa de estreitá-lo é feita por meio de “*fragility analysis*”, um método Bayesiano, que incorpora informação não-amostrada na estimação.

Um dos principais problemas desta abordagem é a falta de um consenso para determinar o quão amplo deve ser o intervalo das estimativas produzidas que levem à conclusão de que os resultados são frágeis. No entanto, a maior parte dos estudos conclui que os resultados são frágeis. Além disso, muitos pesquisadores não estão familiarizados com a abordagem Bayesiana.

O diagrama a seguir ilustra as fases das três abordagens descritas acima, enquanto o próximo quadro explica os principais testes de especificação mais comumente empregados na abordagem clássica, principalmente.



TESTES DE ESPECIFICAÇÃO E DIAGNÓSTICO:

- 1) **Testes de variável omitida:** são utilizados para testar se uma variável conhecida é uma variável relevante ou não. Exemplo: testes F e t .
- 2) **Reset test:** são usados para testar se o modelo possui alguma variável omitida desconhecida, também são utilizados para testar a forma funcional do modelo.
- 3) **Testes da forma funcional:** existem dois tipos. O primeiro tipo não especifica uma forma funcional específica do modelo e são baseados nos resíduos recursivos e no rainbow teste. No segundo tipo, a forma funcional é posta à prova e é testada utilizando uma forma funcional mais geral tal como a transformação Box-Cox.
- 4) **Testes de mudanças estruturais:** testa a estabilidade dos parâmetros. Exemplo: Chow test, cusum e cusum-of-square tests, e testes de falha de previsão.
- 5) **Testes de outliers:** estes testes, entre os quais se enquadram os testes de normalidade, também são utilizados para testar a especificação do modelo. Exemplos: Jarque-Bera, Sahpiro-Wilk, Cook *outlier test* e uso de medidas de DFITS.
- 6) **Testes de erros não esféricos:** testa se os erros são heterocedásticos ou se possuem correlação serial. Exemplos de testes de correlação serial: Durbin-Watson, Durbin's h , m , Breusch-Godfrey, Godfeld-Quandt. Exemplos de testes de heterocedasticidade: Breusch Pagan e White.
- 7) **Teste de Exogeneidade:** frequentemente referidos como testes de Hausman, testam a correlação contemporânea entre os regressores e os erros. Exemplos: testes de erros de medida e testes de viés de simultaneidade.
- 8) **Testes de transformação de dados:** não possuem uma hipótese específica alternativa e são considerados variantes do teste de Hausman. Exemplos: *grouping test* e *differencing test*.

Data mining

Eis um assunto que divide opiniões, a começar pelas definições. Em princípio, os adeptos da abordagem clássica são veemente críticos com respeito ao *data mining*, entendido por estes como o processo de maximização *ad hoc* do ajuste do modelo econométrico aos dados. Por exemplo, se a teoria econômica provê um modelo formal de crescimento da firma, o pesquisador consegue manipulá-lo a fim de obter uma especificação estimável, e, com o intuito claro de maximizar o ajuste do modelo aos dados, este pesquisador incorpora *ad hoc* variáveis ao modelo econométrico, então ele estará fazendo *data mining* no pior sentido do termo. Neste caso, a teoria é corrompida em nome do ajuste aos dados.

Entretanto, mais modernamente (para todas as abordagens “dos dados para a teoria”), o procedimento do *data mining* é visto como uma boa maneira de se combinar a econometria teórica com a aplicada. Trata-se do ato de modelar os dados com a finalidade de descobrir, empiricamente, um determinado padrão comportamental que pode complementar a teoria econômica. O papel do *data mining* entendido neste sentido para as abordagens VAR, inglesa e mesmo microeconômica é claro: pode ser um instrumento útil para se chegar a um bom modelo, com coeficientes significativos, elevado R^2 , Durbin Watson próximo de 2, etc. A grande vantagem deste método é apontar padrões nos dados que a teoria não explica, sugerindo muitas vezes a necessidade de repensar os modelos teóricos dado as mudanças no contexto real, que podem ser caracterizadas através dos dados. Nas palavras de Heckman:

“While there are serious problems in using the data that suggest a theory to test that theory, even more important problems arise from refusing to learn from the data in revising economic models (HECKMAN, 2000).”

Entretanto, uma das desvantagens do *data mining* é a distorção gerada nos níveis de significância usuais. Na prática desta “mineração”, os níveis de significância tradicionalmente usados (1, 5 e 10%), não são os níveis reais, passam a ser apenas nominais. Portanto, quando um *data miner* encontra coeficientes diferentes de zero, com seus respectivos *t*-estatísticos significativos a um nível de significância de 5%, pode-se dizer que tal significância é apenas nominal. Quando várias regressões são geradas sobre um mesmo corpo amostral, a probabilidade de erro tipo I real é muito maior que a significância de 5%, convencionalmente apresentada. Este aspecto é uma das principais fontes de crítica aos adeptos das abordagens “dos dados para a teoria”.

Depois da estimação

Os testes estatísticos e sua interpretação

Conforme Hair et. al. (2005), quase todas as técnicas econométricas são baseadas na inferência estatística dos valores ou relações de uma população entre variáveis de uma amostra aleatória extraída desta população. Interpretar inferências estatísticas requer que o

pesquisador especifique os níveis de erro estatístico aceitável. A abordagem mais comum é especificar o nível do erro Tipo I, também conhecido como alfa. Este erro é a probabilidade de rejeitar a hipótese nula quando a mesma é verdadeira, ou seja, a chance de o teste exibir significância estatística quando na verdade esta não existe. Ao especificar um nível alfa, o pesquisador estabelece os limites permitidos para erro, especificando a probabilidade de se concluir que a significância existe quando na realidade esta não ocorre.

Quando especifica o nível de erro Tipo I, o pesquisador também determina um erro associado, chamado de erro Tipo II, ou beta. Este erro representa a probabilidade de não rejeitar a hipótese nula quando na realidade esta é falsa. O **poder** do teste de inferência estatística é a probabilidade 1 menos beta. Poder é a probabilidade de rejeitar corretamente a hipótese nula, quando esta deve ser rejeitada.

Apesar de a especificação de alfa estabelecer o nível de significância estatística aceitável, é o nível de poder que determina a probabilidade de “sucesso” em encontrar as diferenças se elas realmente existem. Erros Tipo I e Tipo II são inversamente relacionados, por isso o pesquisador deve fazer um compromisso entre o nível alfa e o poder resultante.

O poder não é apenas uma função de alfa, é determinado por três fatores:

1. Tamanho do efeito (*effect size*). A probabilidade de atingir significância estatística é baseada não apenas em considerações estatísticas, mas também na verdadeira magnitude do efeito de interesse. Por isso, a escolha do nível de significância também deve considerar a escala de medida das variáveis. Na microeconometria, devido à capacidade computacional e à disponibilidade de grandes bases de dados, não raro se obtêm coeficientes diferentes de zero de acordo com os níveis usuais (geralmente 0,05 e 0,01), mas que são números decimais de quatro casas.
2. Alfa. Quando alfa se torna mais restritivo, o poder diminui. Isso significa que quando o pesquisador reduz a chance de encontrar um efeito significativamente incorreto, a probabilidade de corretamente encontrar um efeito também diminui. Diretrizes convencionais sugerem níveis de alfa de 0,05¹² ou 0,01 (embora certamente “Deus ame tanto o 0,06 quanto o 0,05” (ROSNOW E ROSENTHAL, 1989)). Entretanto, o pesquisador deve considerar o impacto dessa decisão sobre o poder, antes de selecionar o nível de alfa.
3. Tamanho da amostra. Em qualquer nível de alfa, tamanhos amostrais aumentados sempre produzem maior poder do teste estatístico. No entanto, aumentar o tamanho da amostra também pode produzir poder “excessivo”. Ou seja, se aumentarmos o tamanho da amostra, efeitos cada vez menores serão considerados estatisticamente significativos, até o ponto em que, para uma amostra muito grande praticamente qualquer efeito será significativo. O pesquisador sempre deve ter em mente que o tamanho da amostra impacta o teste estatístico, tornando-o exageradamente sensível para amostras grandes ou insensível para amostras pequenas.

¹² Dallal (2006) ilustra a influência de Fisher ao estabelecer 0,05 como um nível comumente aceito de significância estatística.

As relações entre alfa, tamanho da amostra, tamanho do efeito e poder são complexas, e Cohen (1977) sugere que estudos devem ser planejados para atingir níveis alfa de pelo menos 0,05, com poder de 80%. Para atingir tais níveis de poder, os três fatores – alfa, tamanho da amostra e tamanho do efeito – devem ser considerados simultaneamente.

A tabela a seguir ilustra o efeito recíproco entre o tamanho da amostra, o nível de significância escolhido e o número de variáveis independentes na detecção de um R^2 relevante, para o caso de uma regressão linear múltipla. Os valores da tabela são o R^2 mínimo que o tamanho da amostra especificado detecta como estatisticamente significativo no nível especificado alfa com uma probabilidade (poder) de 80%.

Ela deve ser interpretada da seguinte forma. Se um pesquisador empregar cinco variáveis independentes, especificar um nível de significância de 0,05 e estiver satisfeito em detectar o R^2 estatisticamente significativo 80% das vezes (o que corresponde a um poder de 80%), uma amostra de 50 respondentes detectará valores R^2 maiores ou iguais a 23%. Se a amostra aumentar para 100 respondentes, então os valores de R^2 de 12% ou mais serão detectados.

Tabela 1 – O R^2 mínimo que pode ser considerado estatisticamente significativo com um poder de 0,80 para diversos números de variáveis independentes e tamanhos de amostras (%)

Tamanho da amostra	Nível de significância alfa = 0,01 Número de variáveis independentes				Nível de significância alfa = 0,05 Número de variáveis independentes			
	2	5	10	20	2	5	10	20
20	45	56	71	-	39	48	64	-
50	23	29	36	49	19	23	29	42
100	13	16	20	26	10	12	15	21
250	5	7	8	11	4	5	6	8
500	3	3	4	6	3	4	5	9
1000	1	2	2	3	1	1	2	2

Fonte: Hair et. al. (2005), p. 148.

Obs: - significa não aplicável

Mais importante ainda, o pesquisador também pode determinar o tamanho da amostra necessário para detectar efeitos para as variáveis independentes individuais, dado o tamanho do efeito esperado (correlação), o nível alfa e o poder desejado.

A importância do conceito de poder estatístico para a análise de resultados de pesquisa é apresentada por Rosnow e Rosenthal (1989). Os nomes a seguir são fictícios. Smith conduziu um experimento com $n = 80$ para mostrar os efeitos de estilo de liderança na produtividade, e concluiu que o estilo A é melhor que o estilo B. Jones, autor do conceito do estilo B replicou o estudo com $n = 20$. Jones encontrou um valor t de 1,06, com 18 graus de liberdade, p-valor $>0,30$, ao passo, que Smith encontrara um valor t de 2,21, com 78 graus de liberdade, e p-valor $<0,05$. Não obstante a discrepância dos valores, a magnitude dos efeitos obtidos nos dois estudos é **idêntica**. Como o tamanho da amostra de

Jones é menor, o seu poder para rejeitar a hipótese a um nível de 0,05 era de 0,18, ao passo que para Smith, com uma amostra muito maior, era de 0,60, mais de três vezes maior.

Toda essa discussão está relacionada à nona regra de Kennedy (2002): “não confunda significância estatística com magnitude de significado”. O autor cita uma série de outros pesquisadores que fazem ressalvas à importância que é dada aos testes. Kennedy ressalta que a “santificação” via significância deve ser trocada por buscas de evidência adicional, tanto corroborando quanto – principalmente – contradizendo. Algumas questões que o autor coloca: Se a teoria é correta, há implicações testáveis? Você pode explicar uma gama de descobertas relacionadas? Você pode encontrar um conjunto de evidências consistentes com sua hipótese mas inconsistente com hipóteses alternativas? Sua teoria pode sobrepujar a dos rivais no sentido de explicar outros modelos?

Como traduzir resultados de modelos em implicações teóricas ou sugestões de política?

Não raro, os pesquisadores não conseguem fazer isto com o mesmo sucesso com o qual aplicam os métodos quantitativos e econométricos. Não existe uma regra a ser seguida, mas certos princípios (expostos de maneira simplificada) podem ajudar.

Entender o que está em jogo - assim como é necessário conhecer bem o contexto em que os dados foram gerados e os conceitos com os quais eles se relacionam, é fundamental compreender o contexto em que a estimação é feita. Isto significa compreender os antecedentes históricos do problema a ser tratado, bem como conhecer a literatura teórica e trabalhos empíricos anteriores a respeito do tema. Se o propósito do artigo é testar uma teoria, então há de se ter cuidado ao se derivar as hipóteses que serão submetidas ao teste; se o objetivo é ampliar um modelo ou aplicá-lo para o caso brasileiro, por exemplo, então é preciso contextualizar os resultados com o resto da literatura.

Não confundir correlação com causalidade – este é um debate antigo na filosofia da ciência, não afeito apenas à ciência econômica, e foge ao escopo do trabalho esmiuçá-lo. A única intenção de expor esta proposição é que por vezes temos em trabalhos acadêmicos, esquematicamente, o seguinte: “A ‘explica’ (do ponto de vista econométrico) B, B é bom, logo devemos induzir A”. O “induzir” pode ser substituído por equivalentes como “desenhar políticas públicas para” ou “fomentar”, e outros.

Os problemas deste esquema são que: (i) devemos ser céticos com respeito à causalidade de A para B (aliás, será que sempre que A ocorrer, B também ocorrerá?), e (ii) o que fazer para “induzir” A deve levar em conta o contexto do problema. Não por acaso, este item se relaciona intimamente com o anterior.

Um exemplo extremo é o seguinte: um pesquisador estima, a partir de um modelo log-log, que a elasticidade de exportações com respeito ao pessoal ocupado é de 0,5. Salvo melhor juízo, isto não significa que esta relação seja determinística (problema (i)), e tampouco que o governo deva subsidiar o emprego como política de promoção às exportações (problema (ii)).

Sempre reportar uma análise de sensibilidade – a contextualização dos resultados deve ser feita não somente a partir de outros trabalhos e do contexto econômico,

mas também a partir de especificações, escolhas de *proxies* e técnicas de estimação alternativas, para ver o quão robustos são os resultados. Em certo sentido, isto guarda relação com as idéias de Ed Leamer sobre como proceder com especificação. Uma sugestão dada por Kennedy (2002) é reportar no texto todas as regressões estimadas, e não apenas as boas.

Reprodutibilidade – Koenker (2006) nota como já é parte do folclore da econometria aplicada que mesmo regressões simples apresentadas por um economista não são reprodutíveis por outro (ou até pelo próprio pesquisador, dependendo de seu grau de organização...). À medida que os métodos econométricos evoluíram, o problema parece ter se agravado, uma vez que o pesquisador interessado em reproduzir resultados de outro não apenas depende dos dados utilizados mas também do algoritmo empregado. De fato, é uma regra científica básica que os experimentos e as análises devem ser reprodutíveis por outros pesquisadores. Felizmente, diversos periódicos adotam como regra o fornecimento dos algoritmos e, quando possível, das bases de dados utilizadas no artigo. Porém, econometristas aplicados deveriam adotar esta postura mesmo sem a pressão dos pareceristas.

Síntese

Peter Kennedy (2002) expôs em seu artigo “*Sinning in the basement: what are the rules? The ten commandments of applied econometrics*” os “dez mandamentos” da econometria aplicada. Para este autor, mesmo que não haja uma correspondência direta entre a teoria econométrica e a prática - e, por isso, os econometristas aplicados estão sempre “pecando” -, existem regras para estes pecados. E talvez a primeira regra seja admiti-los, e não necessariamente evitar pecar a qualquer custo. Esta é uma visão compartilhada por vários econometristas aplicados: a melhor política é transparecer o que está sendo feito, ou seja, como a teoria foi adaptada, quais e porque os testes foram aplicados, porque foram escolhidas as *proxies* etc.

A contribuição deste artigo é recomendar não apenas que esta postura seja adotada em relação aos **métodos** empregados e sua mecânica, mas também à **metodologia**. Vimos que os métodos e até mesmo a interpretação dos resultados variam de acordo com a abordagem metodológica.

Quase todos os artigos de econometria apresentam uma seção intitulada “metodologia”, na qual na verdade se expõem os métodos econométricos adotados, e não a abordagem metodológica de fato adotada. A postura de explicitar a escolha metodológica, entendida em seu sentido estrito, talvez enriqueça seminários acadêmicos sobre textos empíricos – nos quais não raro as críticas oriundas de um adepto de uma determinada abordagem dirigidas ao texto do colega, adepto de outra, são genéricas no sentido de que se aplicariam a qualquer artigo que adote a abordagem “rival”.

Referências

- Balassa, B. “An Empirical Demonstration of Classical Comparative Cost Theory”. *Review of Economics and Statistics*, 45, pp.231-8, agosto de 1963.
- Belsey, D. A. e Welch, R. E. “Modelling energy consumption –using and abusing regression diagnostics.” *Journal of Business and Economic Statistics*, 6, 44207, 1988.
- Blaug, M. *Metodologia da Economia*. São Paulo: Ed. Edusp, 1999.
- Bowen, H., E. Leamer & L. Sveikauskas. “Multicountry, Multisector Tests of the Factor Abundance Theory”. *American Economic Review* 77 (5), pp.791-801, 1987.
- Dallal, G. E. *Why $P=0.05$?* Disponível em: <http://www.tufts.edu/~gdallal/p05.htm>. Acessado em 23 de novembro 2006.
- Day, T. E. e Liebowitz, S. J. “Mortgage lending to minorities: where is the bias?” *Economic Inquiry* 36, 1998.
- Deardorff, A. “Testing Trade Theories and Predicting Trade Flows”. In Jones, R. & P. Kenen (eds.), *Handbook of International Economics vol. I*, pp. 467-518. Amsterdam: Ed Elsevier, 1984.
- Faust, J. e Whiteman, C.H., “General-to Specific Procedures for Fitting a Data-Admissible, Theory-Inspired, Congruent, Parsimonious, Encompassing, Weakly-Exogenous, Identified, Structural Model to the DGP: A Translation and Critique”, *Carnegie-Rochester Conference Series on Economic Policy*, Vol. 47, 1997.
- Faust, J. e Whiteman, C.H., Commentary on Grayham E. Mizon ‘Progressive Modeling of Macroeconomic Time Series: The LSE Methodology’. In: K. Hoover (ed.) *Macroeconometrics: Developments, Tensions, and Prospects*, Dordrecht: Kluwer, 1995.
- Friedman, M. *Essays in positive economics*. Chicago: University of Chicago Press, 1953.
- Frisch, R. “Editor’s Note”. *Econometrica*, 1 (1), pp. 1-4.
- Greene, W. *Econometric Analysis*, 5ª Edição, Prentice Hall, 2003.
- Haavelmo, T. “The statistical implications of a system of simultaneous equations”. *Econometrica*, 11 (1), pp. 1-12.
- Hair, J. F. et. al. *Análise multivariada de dados*. Porto Alegre: Editora Bookman, 2005.
- Heckman, J. *Microdata, heterogeneity and the evaluation of public policy*. Nobel Prize Winner’s lecture, 2000.
- Helpman, E. “The Structure of Foreign Trade”. *NBER Working Paper n° 6752*, Cambridge, MA, outubro de 1998.
- Hendry, D. F. “Econometric methodology: a personal perspective.” In: BEWLEY, T. F. *Advances in econometrics*. Cambridge: Cambridge University, 1987.
- Hendry, D. F., Encompassing, *National Institute Economic Review*, (August), 88-92, 1988.

Hoover, K. "The Methodology of Econometrics." In: T.C. Mills and K. Patterson (eds.) *Palgrave Handbooks of Econometrics*, vol 1, Econometric Theory, MacMillan, 2006.

Intriligator, M. D., Bodkin, R. G. e Hsiao, C. *Econometric Models, Techniques and Applications*. Upper Saddle River, NJ, Prentice Hall, 1996.

Kennedy, P. *A guide to econometrics*, 4ª edição, MIT Press, 1998.

Kennedy, P. E. "Sinning in the basement: what are the rules? The ten commandments of applied econometrics." *Journal of Economic Surveys*, v. 16, n. 4, p. 569-589, 2002.

Koenker, R. *Reproducible Econometric Research*. Disponível em <http://www.econ.uiuc.edu/%7Eroger/research/repro/repro.pdf> . Acessado em 1 de dezembro de 2006.

Koopmans, T. C. Three essays on the state of economics science. New York: McGraw-Hill, 1957.

Krugman, P. e M. Obstfeld. *Economia Internacional: Teoria e Política*. São Paulo: Ed. Makron Books, 2001.

Kydland, F. e Prescott, E. "The econometrics of general equilibrium approach to business cycles". *Scandinavian Journal of Economics* 93 (2), pp. 161-78.

Leamer, E. & J. Levinsohn. "International Trade Theory: the Evidence". In Grossman, G. & K. Rogoff (eds.), *Handbook of International Economics vol. III*, pp. 1339-94. Amsterdam: Ed Elsevier, 1995.

Leamer, E. "Questions, theory and data". In Medema, G. e Samuels, W. (eds.) *Foundations of Research in Economics: How do Economists Do Economics?* Aldershot, UK, Edward Helgar, pp. 175-90, 1996.

Leamer, E. "The Leontief Paradox Reconsidered". *Journal of Political Economy* 88, pp. 495-503, 1980.

Leontief, W. "Domestic Production and Foreign Trade: The American Capital Position Re-Examined". *Economia Internazionale* 7, 1954.

Magnus, J. R. "The success of econometrics". *De Economist*, 147, pp. 55-71, 1999.

MIT Press, 2002.

Mizon, G. E. "The Encompassing Approach in Econometrics." In: D.F. Hendry e K.F. Wallis (eds.) *Econometrics and Quantitative Economics*. Oxford: Basil Blackwell, pág. 135-172, 1984.

Pesaran, M. H. e Smith, R. "The Role of Theory in Econometrics." *Journal of econometrics*, 67, pp. 61-79, 1995.

Rosnow, R. e Rosenthal, R. "Statistical procedures and the justification of knowledge in psychological science". *American Psychologist*, 44, 1276-84, 1989.

Sims, C. "Are forecasting models usable for policy analysis" *Federal Reserve Bank of Minneapolis Quarterly Review* 10 (1), 2-15, 1986.

Sims, C. "Macroeconomics and reality". *Econometrica*, 48 (1), 1-48.

Spanos, A. *Statistical foundations of econometric modeling*. Cambridge: Cambridge University Press, 1989.

Trefler, D. "The Mystery of the Missing Trade and Other Mysteries". *American Economic Review* 85, pp. 1029-46, 1995.

Vanek, J. "The Factor Proportions Theory: The N-Factor Case. *Kyklos* 21 (4) pp. 749-756, 1968.

Wooldridge, J. M., *Econometric Analysis of Cross Section and Panel Data*. Cambridge, 2005.