

AVALIAÇÃO DA QUALIDADE DO GASTO PÚBLICO E MENSURAÇÃO DA EFICIÊNCIA

ORGANIZADORES: Rogério Boueri • Fabiana Rocha • Fabiana Rodopoulos

Ministério da Fazenda
Secretaria do Tesouro Nacional

Organizadores:

Rogério Boueri

Fabiana Rocha

Fabiana Rodopoulos

AVALIAÇÃO DA
QUALIDADE DO GASTO PÚBLICO
E MENSURAÇÃO DA EFICIÊNCIA

Brasília, 2015

Ministro da Fazenda

Joaquim Vieira Ferreira Levy

Secretário do Tesouro Nacional

Marcelo Barbosa Saintive

Subsecretários

Gildenora Batista Dantas Milhomem

Lísio Fábio de Brasil Camargo

Marcus Pereira Aucélio

Otavio Ladeira de Medeiros

Paulo Fontoura Valle

Pricilla Maria Santana

Organizadores

Rogério Boueri

Fabiana Rocha

Fabiana Rodopoulos (Secretaria do Tesouro Nacional)

Coordenação Editorial: Secretaria do Tesouro Nacional

Revisão de Texto: Beth Nardelli e Fernanda Gomes (Njobs Comunicação)

Diagramação: Daniela Rodrigues (Njobs Comunicação)

Boueri, Rogério; Rocha, Fabiana; Rodopoulos, Fabiana.

Avaliação da Qualidade do Gasto Público e Mensuração da Eficiência/ Rogério Boueri, Fabiana Rocha, Fabiana Rodopoulos (Organizadores) - Brasília: Secretaria do Tesouro Nacional, 2015.

463 p.

1. Visão Macroeconômica do Gasto Público no Brasil. 2. Dívida Pública: Gestão Eficiente. 3. Experiência Internacional na Avaliação do Gasto Público. 4. Conceitos e Mensuração de Eficiência. 5. Modelos Paramétricos e Não Paramétricos. 6. Software Livre R. 7. Federalismo e Descentralização. 8. Tópicos Avançados. I. Título.

Copyright © Tesouro Nacional, 2015

Todos os direitos reservados à Secretaria do Tesouro Nacional – Brasília-DF. Nenhuma parte desta publicação poderá ser reproduzida ou transmitida sem prévia autorização por escrito da instituição. Para permissão de fotocópia ou reimpressão de qualquer parte deste livro, envie, por favor, uma solicitação para:

Secretaria do Tesouro Nacional

Coordenação-geral de Desenvolvimento Institucional (Codin)

Esplanada dos Ministérios, Ministério da Fazenda (MF)

Bloco P, ed. anexo do MF, ala A, Térreo.

CEP: 70.048-900 – Brasília, DF – Brasil.

Telefone: (55) 61 3412-3973 – Fax: (55) 61 3412-1623

email: geifo.codin.df.stn@fazenda.gov.br

Este livro foi elaborado por vários autores, dentre os quais servidores do Tesouro Nacional, USP, FGV, Universidade Católica de Brasília, UFPB e do Banco Mundial. As opiniões, interpretações e conclusões expressas neste livro não refletem necessariamente as opiniões dessas instituições. O Tesouro Nacional se isenta da responsabilidade sobre a exatidão dos dados incluídos no trabalho.

Sumário

Apresentação da Secretaria do Tesouro Nacional	9
Agradecimentos	11
Sumário Executivo	13

PARTE I - UM RETRATO DO GASTO PÚBLICO NO BRASIL: UMA VISÃO MACROECONÔMICA

Capítulo 1 - Um retrato do gasto público no Brasil: por que se buscar a eficiência

Alex Pereira Benício • Fabiana M. A. Rodopoulos • Felipe Palmeira Bardella

1.1 Introdução	19
1.2 Para que governo?	20
1.3 A importância da qualidade do ajuste fiscal	21
1.4 Análise dos gastos públicos no Brasil	24
1.5 Uma visão de longo prazo das finanças públicas: desafios e oportunidades para os gastos públicos brasileiros	41
Referências	48
Apêndice	50

Capítulo 2 - Dívida pública: contribuições de uma gestão eficiente para a estabilidade econômica

Fabiano Silvio Colbano • Mauricio Dias Leister

2.1 Introdução	53
2.2 O papel da dívida pública em um ambiente não ricardiano	54
2.3 A gestão da dívida pública e a suavização da carga tributária	60
2.4 A importância da gestão de riscos	64
2.5 A centralização da gestão da dívida pública	69
2.6 Evidências empíricas da contribuição da gestão da dívida brasileira ao equilíbrio fiscal	71
2.7 Considerações finais	78
Referências	79

Capítulo 3 - Composição ótima do gasto público para o crescimento econômico

Ana Carolina Giuberti - Fabiana Rocha

3.1 Introdução	81
3.2 A literatura teórica	82
3.3 A literatura empírica	87
3.4 Gasto público, eficiência e crescimento econômico	96
3.5 Considerações finais	100
Referências	102

PARTE II - A EXPERIÊNCIA INTERNACIONAL NA AVALIAÇÃO DO GASTO PÚBLICO

Capítulo 4 - Revisões de Despesas na OCDE

Marc Robinson

4.1 Introdução	107
4.2 Holanda	118
4.3 Canadá	121
4.4 Austrália	126
4.5 Dinamarca	129
4.6 Reino Unido	132
4.7 França	136
4.8 Considerações finais	140
Referências	155

Capítulo 5 - Revisão das Despesas Públicas

Ian Lienert

5.1 Introdução	161
5.2 O que é uma PER, seu escopo e estrutura de análise?	162
5.3 Questões tipicamente abordadas nas PERs e ferramentas	165
5.4 De que modo uma PER difere de uma Revisão de Despesa?	184
5.5 Estudos de Caso de PERs: México, Peru, Rússia e Turquia	186
5.6 Lições apreendidas de experiências da PER	200
5.7 Considerações finais	204
Referências	206

PARTE III - O QUE É EFICIÊNCIA?

Capítulo 6 - Conceitos sobre eficiência

Enlison Mattos • Rafael Terra

6.1 Introdução	211
6.2 Conceitos básicos de eficiência na produção de bens	213
6.3 Conceitos básicos de eficiência na produção pública de bens	219
6.4 Considerações finais	232
Referências	233

Capítulo 7 - Fundamentos microeconômicos da mensuração de eficiência

Enlison Mattos • Rafael Terra

7.1 Elementos básicos da teoria da produção	235
7.2 Definições e medidas de eficiência econômica	248
7.3 Considerações finais	263
Referências	265

PARTE IV - COMO MEDIR A EFICIÊNCIA?

Capítulo 8 - Modelos não paramétricos: Análise Envoltória de Dados (DEA)

Rogério Boueri

8.1 Introdução	269
8.2 Índices de produtividade	270
8.3 Modelo CCR com ótica nos insumos	271
8.4 Modelo CCR com ótica nos produtos	276
8.5 Modelos duais	280
8.6 Conjuntos de referência	283
8.7 Modelos com retornos variáveis de escala	285
8.8 Painel de dados em DEA	288
8.9 Tecnologias alternativas de avaliação	290
8.10 Utilizando o painel de dados para a decomposição da variação da produtividade	291
8.11 Variáveis ambientais e variáveis não discricionárias	298
8.12 Modelo de dois estágios	301
Referências	305

Capítulo 9 - Modelos paramétricos: Fronteira Estocástica

Rogério Boueri

9.1 Introdução	307
9.2 Modelo teórico	309
9.3 Estimação da Fronteira Estocástica – fronteira de produção	312
9.4 Estimação da Fronteira Estocástica – fronteira de custo	316
9.5 Dados em painel: eficiência temporalmente constante	318
9.6 Dados em painel: variação temporal da eficiência	321
9.7 Variáveis ambientais	323
9.8 Função translogarítmica inversa (translog)	325
9.9 Decomposição da evolução da eficiência com a FE	329
Referências	331

Capítulo 10 - R: Um *software* livre para mensuração da eficiência

Luis Felipe Vital Nunes Pereira

10.1 Introdução: o desafio computacional da mensuração da eficiência	333
10.2 R: Um <i>software</i> livre	335
10.3 Aplicações de DEA com o pacote Benchmarking	345
10.4 Aplicações de fronteira estocástica com o pacote Frontier	367
Referências	375

Capítulo 11 - Federalismo e Descentralização

Janete Duarte • Sérgio Ricardo de Brito Gadelha • Plínio Portela de Oliveira •

Luis Felipe V. N. Pereira

11.1 Introdução	377
11.2 As transferências intergovernamentais	379
11.3 Distribuição das responsabilidades de gastos	383
11.4 Financiamento da saúde e da educação no Brasil	385
11.5 O papel dos consórcios na eficiência dos serviços descentralizados	389
Referências	391

PARTE V - TÓPICOS AVANÇADOS EM MENSURAÇÃO DE EFICIÊNCIA

Capítulo 12 - Eficiência na provisão de educação e saúde: resenha e aplicações para os municípios brasileiros

*Fabiana Rocha • Janete Duarte • Sérgio Ricardo de Brito Gadelha • José de Anchieta Semedo
Neves • Plínio Portela de Oliveira • Luis Felipe Vital Nunes Pereira*

12.1 Introdução	395
12.2 Revisão da literatura	396
12.3 Tópicos avançados na mensuração da eficiência: a abordagem da metafronteira	404
Referências	418

Capítulo 13 - Detecção de *outliers* em modelos não paramétricos: o método *Jackstrap* ampliado

Maria da Conceição Sampaio de Sousa • Borko D. Stosic

13.1 Introdução	421
13.2 O método <i>Jackstrap</i>	424
13.3 Detecção e inclusão de <i>outliers</i>	424
13.4 Considerações finais	429
Referências	431

Capítulo 14 - Aplicação do Método *Jackstrap* na Atenção Básica à Saúde

Maria da Conceição Sampaio de Sousa • Rodrigo Hitoshi Dias

14.1 Introdução	433
14.2 Insumos e produtos	434
14.3 Resultados	436
14.4 Escores robustos de eficiência	442
14.5 Considerações finais	448
Referências	449

Anexo - Bases de dados disponíveis em áreas finalísticas para a avaliação de políticas públicas	451
Nota sobre os autores	459
Sobre o projeto gráfico	463



Apresentação da Secretaria do Tesouro Nacional

É com muita satisfação que a Secretaria do Tesouro Nacional apresenta esta obra, resultado de um gratificante trabalho desta instituição com algumas das principais universidades do país e com o Banco Mundial.

Este livro tem como objetivo fomentar o debate sobre qualidade do gasto público, bem como disseminar técnicas de mensuração de eficiência do gasto. É uma publicação inovadora por apresentar, em uma única obra, aspectos conceituais, empíricos e a experiência internacional sobre mensuração da eficiência do gasto público.

Como poderá ser observado nos capítulos iniciais, a política fiscal brasileira tem um compromisso com um modelo que combina crescimento econômico com erradicação da pobreza e redução das desigualdades sociais. Esse modelo implica uma demanda crescente por gastos em saúde, educação e benefícios assistenciais, ao mesmo tempo em que exige investimentos em infraestrutura para permitir o crescimento econômico sustentável.

Sem prejuízo da necessidade de manter o compromisso de responsabilidade fiscal para garantir a estabilidade macroeconômica, a busca da qualidade dos serviços públicos é outro conceito que a administração pública vem incorporando ao longo dos anos. Nesse contexto, faz-se necessária a introdução e disseminação da visão de eficiência no setor público brasileiro. Assim, não basta saber quanto o governo gasta com a provisão de bens e serviços, é necessário avaliar e ponderar os custos e os benefícios desses gastos para determinar a respectiva utilidade para o contribuinte. O foco passa para o aumento da consciência a respeito do valor dos recursos públicos, procurando incorporar o valor da eficiência na lógica de funcionamento da administração pública.

Para medir a qualidade do gasto público são necessários dados sobre esse gasto, técnicas de mensuração e gestores bem capacitados. Esta publicação cobre todos esses pontos e poderá ser utilizada por gestores públicos em suas atividades diárias. Adicionalmente, os capítulos aplicados foram elaborados em *software* livre, de forma a evitar que eventuais restrições de *software* impeçam a adoção das técnicas aqui apresentadas.

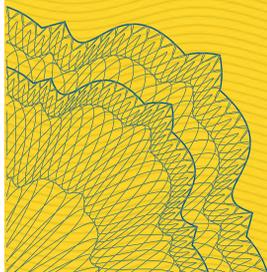
A Secretaria do Tesouro Nacional também dedica especial atenção à formação dos futuros gestores, alunos de graduação e pós-graduação em administração, economia e áreas correlatas. Por esse motivo, esta publicação tem o formato de um livro-texto, com teorias, técnicas e exercícios, o que possibilita sua adoção em sala de aula.

A busca pela melhoria da qualidade do gasto público é missão institucional da Secretaria do Tesouro Nacional e um desafio permanente daqueles que receberam da sociedade brasileira a incumbência da gestão de recursos públicos. Nesse contexto, esta publicação, pelo caráter pioneiro, emerge como uma contribuição significativa e oportuna às discussões necessárias para o desenvolvimento de uma cultura de eficiência do gasto, essencial para o desenvolvimento econômico nacional.

Marcelo Barbosa Saintive

Secretário do Tesouro Nacional

Agradecimentos



Este livro é resultado de um projeto que envolveu a Secretaria do Tesouro Nacional, o Banco Mundial e a academia na busca para fomentar o debate sobre a qualidade e a eficiência do gasto público. A publicação combina a experiência na execução da política fiscal, a pesquisa acadêmica sobre qualidade do gasto público e a experiência internacional, de forma a fornecer ao leitor uma visão ampla sobre como medir a qualidade e a eficiência do gasto público.

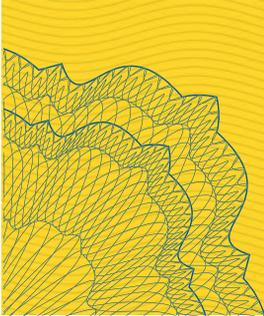
Os autores foram cuidadosamente selecionados e são gestores e/ou pesquisadores de prestígio em suas áreas de atuação. As teorias e técnicas apresentadas cobrem as visões e linhas de pensamento das principais universidades do Brasil, dentre elas a Universidade de São Paulo (USP), a Universidade de Brasília (UnB), a Universidade Católica de Brasília (UCB), a Fundação Getúlio Vargas (FGV/SP), a Universidade Federal da Paraíba (UFPB) e a Universidade Rural de Pernambuco (UFRPE), além da experiência internacional do Banco Mundial.

Os organizadores agradecem ao subsecretário de Assuntos Corporativos da Secretaria do Tesouro Nacional (STN), Liscio Camargo, e ao subsecretário de Planejamento e Estatísticas Fiscais quando da elaboração do livro, Cleber Oliveira, pelo empenho e dedicação que tornaram possível a publicação deste livro e por acreditarem que a solução para um gasto público mais eficiente passa pela capacitação de atuais e futuros gestores. Além disso, agradecemos ao Otaviano Canuto e ao Pablo Fajnzylber, do Banco Mundial, por confiarem no projeto e colocarem a equipe do Banco à disposição. Agradecemos ainda à Edith Kikone pela coordenação dos capítulos feitos pelo Banco Mundial.

Os organizadores também agradecem à Embaixada Britânica, pela parceria no âmbito do Projeto Integração Institucional para Melhoria da Eficiência do Gasto Público, e ao Programa das Nações Unidas para o Desenvolvimento (PNUD), pelo suporte administrativo prestado nos trâmites relativos à contratação de prestadores de serviços diversos. Por fim, gostaríamos de agradecer aos diversos profissionais que colaboraram de alguma forma para o sucesso deste projeto, em especial, à Karla Rocha, da gerência de estatísticas de finanças públicas, , à Helise Gomes e aos colegas da gerência de informação da Coordenação-Geral de Desenvolvimento Institucional. Por fim, à equipe da gerência de estudos econômico-fiscais da Secretaria do Tesouro Nacional, em especial à Janete Duarte e Artur Santos, sem os quais a concretização deste livro não teria sido possível.

Bons estudos.

Os organizadores.



Sumário Executivo

A economia brasileira alcançou, nas últimas duas décadas, conquistas importantes que permitiram melhorias de bem-estar sem precedentes à população. A primeira dessas conquistas foi a estabilização monetária atingida a partir de meados da década de 1990 e que contribuiu decisivamente para a ampliação da credibilidade e da confiança na economia nacional.

O outro grande sucesso logrado pelo nosso sistema econômico foi o ingresso em um processo de redução das desigualdades sociais e da pobreza. Embora esse processo tenha sido beneficiado por elementos externos, como a elevação dos preços das commodities e a alta disponibilidade de liquidez nos mercados financeiros internacionais, é inegável que as políticas instituídas e aprofundadas na década passada tiveram um impacto benéfico poderoso no combate às desigualdades e à miséria.

Não obstante a importância dessas conquistas, novos passos devem ser tomados de maneira a manter a trajetória de desenvolvimento sustentado tão ansiosamente desejada pela população brasileira. E o próximo passo nessa direção está relacionado à qualidade e à eficiência do gasto público.

Se por um lado a capacidade de tributação tem limites a partir dos quais começa a prejudicar a atividade produtiva, por outro, há expectativas concretas da população em relação à provisão serviços públicos. Anseios pela melhoria dos serviços refletem os avanços já obtidos e não tendem a desaparecer.

Forma-se então um dilema, por um lado, a capacidade do governo extrair recursos da sociedade para o financiamento dos serviços

públicos é limitada, por outro, as demandas por tais serviços tendem a crescer. Como conciliar essas duas realidades diante da inexorável restrição orçamentária?

A chave está na racionalização e priorização do gasto público de forma a se atingir os objetivos dos cidadãos em relação aos serviços públicos com a utilização dos recursos disponíveis, sem que haja necessidade de crescimento desmesurado do gasto. Essa racionalização do gasto é, sem dúvida, o grande desafio que ora se apresenta ao desenvolvimento econômico brasileiro.

Faz-se necessário que os diversos projetos e programas governamentais sejam planejados, executados e avaliados no âmbito de uma cultura de racionalidade econômico-social. É primordial que o apreço aos recursos públicos se manifeste não só no que concerne a transparência de como são utilizados, mas também aos efeitos positivos que o seu dispêndio trará à sociedade como um todo, em vista do custo de retirá-los da própria sociedade através da tributação.

Esse volume pretende ser um instrumento para o aprofundamento dessa cultura de racionalização e priorização da despesa pública.

Para tanto, ele aborda primeiramente, a experiência brasileira, apontando suas discrepâncias e os mecanismos para lidar com elas. A seguir trata de experiências internacionais, especialmente aquelas de países com grande tradição na análise as despesas públicas. Por fim, são estudadas as técnicas necessárias para lidar com o assunto, que é por natureza complexo e controverso. Tais metodologias servem tanto para identificar dificuldades e diagnosticar problemas, quanto para apontar soluções, ou pelo menos, balizar decisões bem informadas.

Especificamente, o livro é composto de 14 capítulos, divididos em cinco partes, além de um anexo. Os três capítulos que compõem a **Parte I** trazem uma visão macroeconômica do gasto público brasileiro e de como a busca pela eficiência pode ampliar os recursos destinados às reais necessidades da população. O capítulo 1 retrata a evolução do gasto público, apontando algumas de suas virtudes, mas também peculiaridades e idiosincrasias, de modo a guiar a busca pela eficiência e qualidade. No capítulo 2, a gestão da dívida pública é discutida também com o intuito de torná-la mais racional e eficiente, diminuindo assim o ônus econômico e social

que o seu serviço inevitavelmente impõe às finanças públicas da nação. O capítulo 3 discute a composição do gasto público. Embora reconheça que não há composição inequivocamente ótima entre investimentos e gastos de custeio, esse capítulo busca balizar e apontar a direção que se deve perseguir.

Um pouco da experiência internacional sobre avaliação e revisão das despesas públicas é descrita da **Parte II** do volume. No capítulo 4, são apresentadas as práticas de revisão da despesa pública em seis países da OCDE. Nesse capítulo, o autor ressalta que o *spending review* é “ferramenta usada não somente para reduzir os gastos agregados, mas também capaz de aumentar o espaço disponível para priorizar novas despesas consistentes com a manutenção de uma política fiscal sólida”. O capítulo 5 apresenta exercício semelhante realizado pelo Banco Mundial em relação a quatro nações emergentes (México, Peru, Rússia e Turquia).

A **Parte III** discute o conceito de eficiência sob um ponto de vista teórico e microeconômico. Enquanto o capítulo 6 explicita as definições teóricas de eficiência, o capítulo 7 analisa, ainda teoricamente, as formas de medi-la. Esses dois capítulos têm importância especialmente quando se consideram dois objetivos. Em primeiro lugar, eles servem para a estruturação do modelo básico de avaliação microeconômica da eficiência das ações governamentais, ainda que sob o ponto de vista extremamente conceitual. Em segundo lugar, eles lançam as bases para as ferramentas práticas de mensuração da eficiência descritas na próxima parte do volume.

Portanto, na **Parte IV** são desenvolvidas as principais ferramentas empíricas para aferição de eficiência do gasto público. Essa parte é composta por quatro capítulos, sendo que dois deles, os capítulos 8 e 9, descrevem as metodologias de “Análise Envoltória de Dados” e de “Fronteira Estocástica”, respectivamente. Essas técnicas são utilizadas na análise comparativa de eficiência e vem ganhando cada vez mais espaço na comparação de unidades produtoras de serviços públicos. O capítulo 10 apresenta o software R, no qual essas metodologias podem ser programadas. Por ser um software livre e com inúmeras bibliotecas contendo rotinas para a aplicação das mais variadas técnicas estatísticas e numéricas, esse programa tem se tornado a ferramenta favorita dos praticantes de análise de eficiência pelo mundo afora. O

capítulo 11 fecha essa parte do livro descrevendo como a estrutura federativa brasileira pode favorecer ou desfavorecer a eficiência dos gastos nas áreas de responsabilidade compartilhada entre os diversos entes federativos, como é o caso da saúde e da educação.

A parte final do livro (**Parte V**) apresenta três capítulos nos quais ocorre aplicações e refinamento das técnicas básicas expostas na parte anterior. Mais uma vez o foco recai sobre os gastos com saúde e educação, cujas eficiências relativas são analisadas no capítulo 12. Técnicas para detecção de pontos extremos, ou *outliers*, são apresentadas nos capítulos 13 e 14. É válido notar que, quando se trabalha na avaliação do gasto público em âmbito municipal tais técnicas são imprescindíveis, uma vez que, a probabilidade de que haja incorreções nos dados é muito grande.

Além desses catorze capítulos, o volume também contém um anexo no qual são apresentadas as principais bases de dados disponíveis para análise da qualidade e da eficiência do gasto público. Esse anexo somado ao tom descritivo da maioria dos capítulos atesta o caráter didático do livro, que poderá ser utilizado não só por estudantes universitários, mas também pelos oficiais responsáveis pelo planejamento, execução e avaliação do gasto público nas diversas esferas de governo.

Boa leitura!

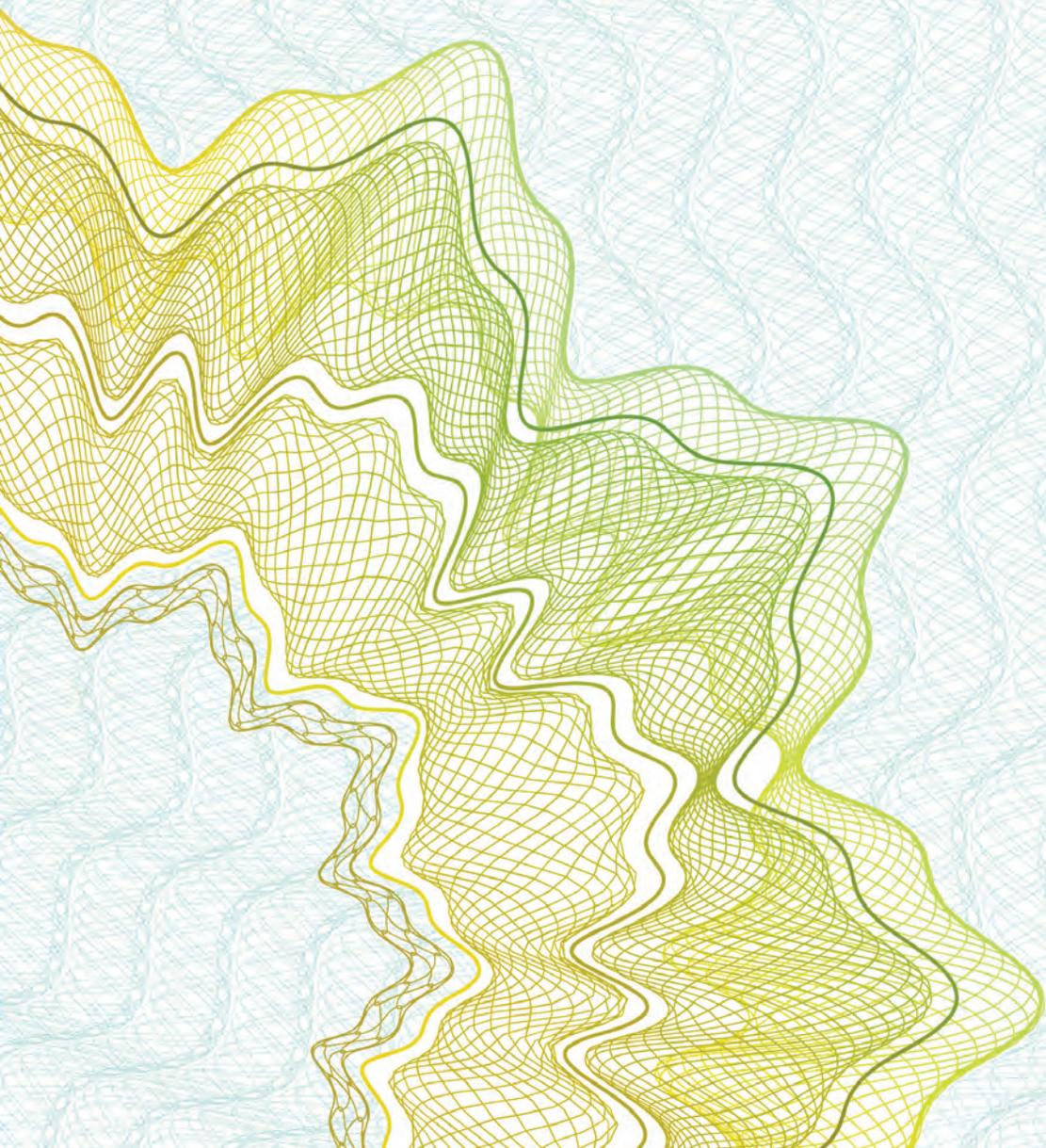
Atenciosamente,

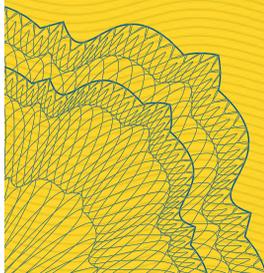
Joaquim Vieira Ferreira Levy

Ministro de Estado da Fazenda

PARTE I

UM RETRATO DO GASTO PÚBLICO NO BRASIL:
UMA VISÃO MACROECONÔMICA





Capítulo 1

Um retrato do gasto público no Brasil: por que se buscar a eficiência

Alex Pereira Benício
Fabiana M. A. Rodopoulos
Felipe Palmeira Bardella

1.1 Introdução

Nos últimos anos, a política fiscal tem ganhado importância nas discussões de política econômica, motivada principalmente pelos desafios postos a partir da crise financeira de 2008. Os tópicos dessa agenda de discussões enfatizam o dilema entre o estímulo fiscal em um ambiente de forte desaceleração econômica global e a consolidação fiscal em um contexto de grave crise das finanças públicas das principais economias desenvolvidas.

Ao mesmo tempo, a demanda por serviços públicos é crescente e o debate público costumeiramente centra-se no montante de recursos financeiros necessários para atender às necessidades da população, muitas vezes baseado no discurso de que uma quantidade maior de recursos é condição necessária para satisfazer à demanda por mais bens públicos.

Em um contexto de crise e ajustamento fiscal, as discussões sobre o tamanho do ajuste fiscal despertou a necessidade de avaliar outras questões fundamentais, como a qualidade do gasto público e os desafios de longo prazo das finanças públicas.

O objetivo do capítulo é apresentar um panorama dos gastos públicos no Brasil, com intuito de familiarizar o leitor com algumas de suas particularidades e fundamentar o argumento quanto à necessidade de aumentar a eficiência desses gastos. O desafio para esta e para as próximas gerações, portanto, não é apenas conter a expansão do gasto público, mas também avaliar onde ele é pouco produtivo, buscando fazer mais com menos recursos e priorizando a eficiência dos programas públicos.

Este capítulo está organizado da seguinte forma: a seção 1.2 introduz a discussão sobre gasto público, descrevendo as atribuições do governo sob a ótica econômica, decomposto nas funções alocativa, estabilizadora e distributiva. A seção 1.3 apresenta argumentos em favor do aspecto qualitativo de um programa de ajuste fiscal, em detrimento do foco exclusivo na magnitude do esforço fiscal. A seção 1.4 descreve os gastos públicos no Brasil em uma perspectiva internacional e analisa a composição dos dispêndios. Finalmente, a seção 1.5 versa sobre a necessidade de visão de longo prazo para as finanças públicas brasileiras e resume os desafios e oportunidades para esta e para as próximas gerações.

1.2 Para que governo?

De acordo com a classificação tradicional de Musgrave (1959), podemos agrupar as atribuições do governo em três funções: alocativa, estabilizadora e distributiva.

Começando pela última, a ação estatal no exercício da função distributiva refere-se àquelas medidas destinadas a minorar os desequilíbrios de renda e condições de vida entre indivíduos e regiões para níveis socialmente aceitáveis. Como o mercado não é capaz de redistribuir renda de acordo com os ideais de justiça social, compete ao governo interferir na economia, buscando a redistribuição de renda ou alívio da pobreza de acordo com o que a sociedade almeja. Enquadram-se nesse segmento, por exemplo, os gastos com a rede de amparo social e cobertura de benefícios previdenciários por critérios universais, sem preexistência de contribuição por parte do beneficiário. Em um país com muitas desigualdades, como o Brasil, o cumprimento pleno dessa função pode demandar expressivos recursos públicos.

Por sua vez, a função estabilizadora requer do governo ações tempestivas destinadas a controlar os efeitos dos choques econômicos sobre a renda e o consumo. Isso se justifica pela perda de bem-estar social provocada por recessões e superaquecimentos da economia, os quais prejudicam tanto as famílias quanto o funcionamento eficiente dos arranjos produtivos. Para esse fim, a demanda agregada é o canal utilizado para estimular a atividade econômica em momentos de depressão, ou controlá-la em períodos de aceleração inflacionária. Incluem-se neste rol aquelas medidas

destinadas a controlar o gasto público, a tributação, o crédito e outras de natureza regulatória que afetem o nível de demanda agregada.

Finalmente, a função alocativa do governo engloba o fornecimento de bens e serviços que o setor privado é incapaz de fornecer em níveis satisfatórios na vigência exclusivamente dos mecanismos de mercado. Isso ocorre devido a falhas que o mercado privado apresenta, dentre as quais merece destaque a existência dos chamados bens públicos. Esta denominação aplica-se àqueles bens que possuem os seguintes atributos: não são rivais e nem excludentes, ou seja, o consumo do bem por um indivíduo não reduz a disponibilidade para outro e ninguém pode ser impedido de consumi-lo. A título de exemplo, enquadram-se nesta categoria os gastos com Defesa Nacional e Segurança Pública. Sua importância econômica e social enseja que o setor público atue para garantir que esses bens sejam fornecidos em quantidades satisfatórias. Há ainda os chamados *bens meritórios*, os quais – apesar de poderem ser fornecidos pelo setor privado – recomendam um esforço complementar do Estado, dada a relevância do impacto econômico e social que esses bens têm. São exemplos disso ações na esfera educacional e de saúde pública.

1.3 A importância da qualidade do ajuste fiscal

Desde a publicação da *Teoria Geral* de Keynes, em 1936, a teoria econômica convencional é clara em apontar a política de tributação e – principalmente – dos gastos públicos como importante instrumento para mitigação dos ciclos econômicos. Em situações de crise, quando a atividade econômica opera abaixo do seu nível natural, o gasto público ganha destaque como variável de controle das autoridades econômicas para estimular a economia.

Contudo, experiências de ajustes fiscais com efeitos expansionistas na economia enfraqueceram as prescrições keynesianas de política fiscal. Em estudo bastante conhecido, Alesina e Perotti (1997) investigaram empiricamente padrões de resposta da economia em episódios de grandes ajustes fiscais no período 1960-1994. Os resultados encontrados mostraram que ajustes fiscais bem-sucedidos – definidos no sentido de manutenção do equilíbrio orçamentário após a implementação do programa – estavam associados: (i) à ênfase no corte de gastos, ao invés do aumento de

receitas; (ii) ao corte dos gastos correntes, ao invés dos investimentos. Quanto aos efeitos econômicos, foi observado que os ajustes bem-sucedidos resultavam em expansão do crescimento econômico e queda do desemprego, dentre outros resultados positivos.

Esses resultados foram justificados por uma série de argumentos, baseados em pesquisas específicas de outros autores. Dentre esses argumentos, mencione-se a situação em que cortes de gastos percebidos como permanentes possam gerar nos agentes econômicos a expectativa de redução da carga tributária futura, ampliando o consumo privado e, por conseguinte, a demanda agregada já no momento presente. Outro argumento é que o ajuste fiscal pode reduzir a taxa de juros pelo efeito credibilidade, ao diminuir o prêmio de risco da inflação e de um eventual *default*. Além de incentivar o investimento pela elevação do valor de mercado da riqueza privada, a queda da taxa de juros viabiliza a realização de novos investimentos privados, bem como a compra de bens duráveis.

Alesina e Ardagna (2009) fizeram uma atualização daquele estudo ampliando a base de dados e utilizando técnicas de análise mais sofisticadas. Os resultados encontrados confirmam as conclusões do primeiro estudo, enfatizando que cortes de gastos são mais efetivos para estabilizar a dívida e evitar recessões do que aumento de receitas. Cortes de impostos, por sua vez, são mais efetivos para estimular a economia do que aumento de gastos.

Em 2010, o Fundo Monetário Internacional divulgou estudo em que investigava o efeito de consolidações fiscais sobre a atividade econômica, baseado em análise de dados históricos das economias avançadas e do modelo de equilíbrio geral GIMF (*IMF's Global Integrated Monetary and Fiscal Model*).¹ Os resultados encontrados corroboravam a abordagem keynesiana de que uma consolidação fiscal normalmente reduz o produto e o emprego no curto prazo. Esses efeitos contracionistas seriam agravados quando: (i) o risco de solvência fosse percebido como baixo; (ii) ocorresse simultaneamente em vários países; e (iii) se a política monetária estivesse esgotada, com taxas de juros próximas do limite inferior.

1 International Monetary Fund (2010).

Contudo, de modo similar ao obtido nas pesquisas de Alesina e Perotti (1997) e Alesina e Ardagna (2009), é apontado que ajustes baseados somente em aumentos de receitas teriam efeitos mais severos na retração econômica quando comparados à opção de corte de gastos. Os efeitos negativos do aumento de impostos seriam ainda mais acentuados quando concentrados em tributos indiretos, em vez daqueles diretos. Também de forma similar, é identificado que cortes dos gastos com investimentos são mais recessivos do que cortes no consumo do governo, que, por sua vez, são mais recessivos do que cortes de gastos com transferências. Na verdade, esses cortes (sobre transferências) teriam efeitos expansionistas sobre o produto, totalmente em linha com os resultados observados em estudos anteriores. Essas convergências mostram a importância do controle efetivo das despesas públicas correntes como instrumento de estabilização econômica.

Diversas evidências empíricas dão suporte a esse argumento. Gupta et al. (2002), em um estudo de 39 países de baixa renda, encontram evidência de que a composição dos gastos públicos é um determinante importante para o crescimento econômico de curto prazo. Um ajuste fiscal baseado em cortes nos gastos correntes tende a gerar taxas de crescimento mais altas do que ajustes baseados em aumento de impostos. Em particular, gastos de capital devem ser preservados em um ajustamento fiscal, uma vez que implicam maior crescimento, enquanto reduções de pagamento de despesas obrigatórias não prejudicam o crescimento. A construção de uma estrada, por exemplo, ao induzir a expansão da atividade econômica de uma região, poderia gerar efeito sobre o Produto Interno Bruto superior ao resultado individual do empreendimento, do ponto de vista do setor privado.

Além da composição do ajuste fiscal, o tamanho do ajuste também influencia o impacto sobre a atividade econômica, na medida em que é ampliado espaço para a poupança privada financiar os investimentos privados. Ademais, pesquisa empírica conduzida por McDermott e Wescott (1996) concluiu que quanto maior o ajuste fiscal, maior a probabilidade de que este fosse permanente e tivesse impacto positivo sobre o produto e a renda.

O tamanho do ajuste fiscal afeta diretamente o funcionamento dos serviços públicos prestados à população, de modo que uma interrupção sem critério pode causar sérios transtornos. Assim, é

importante a informação de quais programas têm margem para redução de gastos sem grandes prejuízos à qualidade e cobertura dos serviços prestados. Para isso, o monitoramento do nível de eficiência dos gastos surge como uma ferramenta para os gestores públicos realizarem o controle das despesas de modo mais racional.

1.4 Análise dos gastos públicos no Brasil

Nesta seção é apresentado um panorama do gasto público no Brasil com o propósito de discutir três questões: (i) o tamanho do gasto no Brasil; (ii) a composição dos dispêndios; e (iii) a esfera de governo que executa o gasto.

A primeira questão é desenvolvida sob a ótica de comparação internacional do nível de gastos realizados pela administração pública em países com diferentes graus de desenvolvimento econômico e social. O segundo ponto em análise se apresenta por meio da avaliação das despesas segundo a estrutura funcional. Especial atenção é dada para as despesas realizadas na função saúde e educação, que são o foco deste livro. Por fim, são analisados os gastos realizados por esfera de governo no Brasil, destacando-se as competências atribuídas pela Constituição Federal de 1988 aos governos federal, estaduais e municipais no país.

1.4.1 Gasto público no Brasil – uma comparação internacional

O tamanho do gasto público no Brasil pode ser mais bem compreendido quando situado em um painel internacional, comparando-se com grupos de países classificados segundo o grau de desenvolvimento econômico e padrão cultural/social. Utilizou-se a estrutura de agrupamento de países feita pelo Fundo Monetário Internacional na publicação *Fiscal Monitor*,² com inclusão do grupo dos países nórdicos, que têm um padrão elevado de gastos públicos em razão do estado de bem-estar social que implementaram em suas economias.

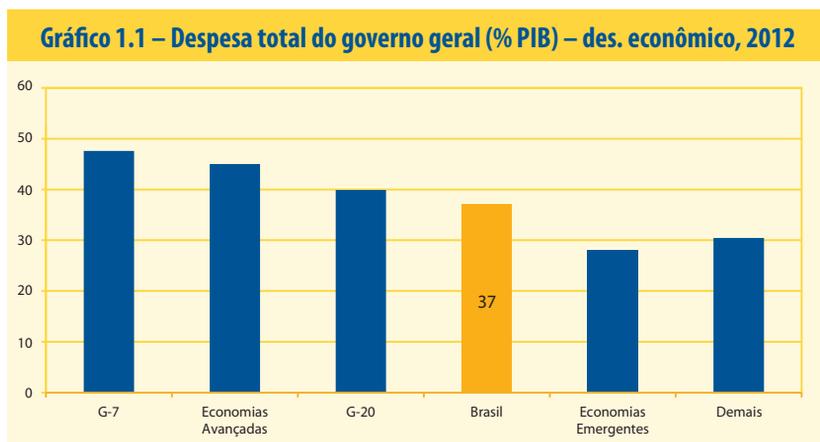
Em termos de base de dados, optou-se por utilizar as estatísticas contidas no *Government Finance Statistics Yearbook* (GFSY) do Fundo Monetário Internacional, que permite a realização de comparação

2 International Monetary Fund (2013).

entre países. No caso do Brasil, os dados são elaborados pela Secretaria do Tesouro Nacional do Ministério da Fazenda (STN/MF), compatíveis com os dados divulgados no Padrão Especial de Disseminação de Dados (PEDD) (*Special Data Dissemination Standard – SDDS*, em inglês). Os dados referem-se, em sua maior parte, ao ano de 2012.³

Relativamente à cobertura de governo, os dados utilizados adotam o conceito de governo geral, que compreende as administrações direta, indireta e as empresas públicas dependentes⁴ de todas as esferas de governo. Dessa forma, é possível fazer comparação entre países com diferentes estruturas político-administrativas, desde estados unitários, até estruturas federativas, como a brasileira, que tem três níveis de governo: federal, estadual e municipal. A despesa total do governo geral inclui despesa primária e juros nominais, mas não inclui gastos com investimento público, pois na metodologia adotada pelo Fundo Monetário Internacional tais gastos são tratados como aquisição de patrimônio público.

Inicialmente, pode-se posicionar o Brasil relativamente aos grupos de países por grau de desenvolvimento econômico, destacando-se os países de economia avançada, os de economia emergente e os grupos G-7 e G-20, que reúnem as maiores economias do mundo.⁵



Fonte: International Monetary Fund (2013).

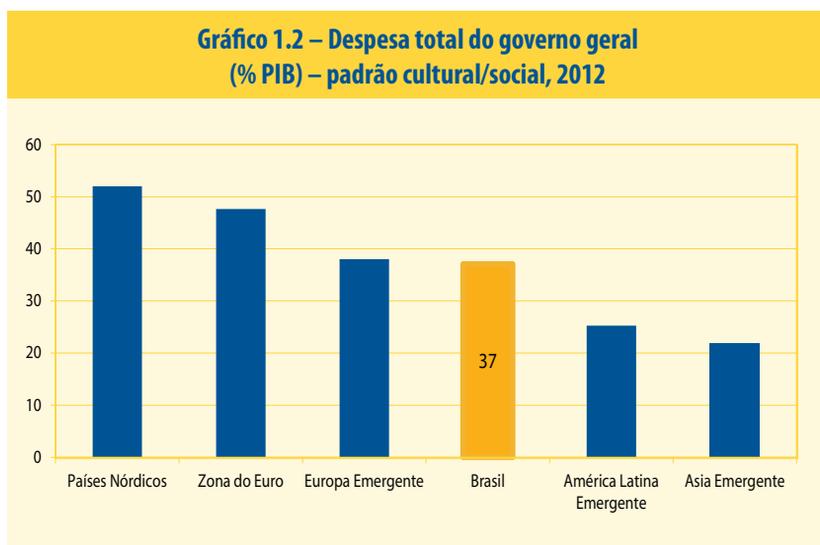
3 Alguns países apresentam dados referentes a períodos anteriores listados no Apêndice 3.

4 Empresas públicas independentes dos orçamentos não estão incluídas no governo geral.

5 Veja a lista de países incluídos em cada grupo no Apêndice 1.

Em 2012, a despesa total do governo geral no Brasil correspondeu a 37% do Produto Interno Bruto (PIB), superior em aproximadamente 5,0 pontos percentuais de PIB ao gasto realizado pelo conjunto dos países de economia emergente (Gráfico 1.1). O padrão brasileiro de gasto público também supera aquele dos países emergentes, bem como os de menor grau de desenvolvimento econômico.

Relativamente às economias avançadas e aos grupos das maiores economias do mundo (G-7 e G-20), os gastos brasileiros são inferiores, especialmente em relação ao padrão dos países europeus, o que fica evidente no Gráfico 1.2. Canadá, Israel e Alemanha apresentam níveis de gasto público de cerca de 40% do PIB, pouco superiores ao padrão brasileiro. Os Estados Unidos, por sua vez, têm gastos similares ao Brasil, em torno de 37% do PIB. Desta feita, o patamar de gasto público no Brasil, ainda que seja inferior ao padrão europeu, é comparável ao de alguns países de economia avançada.

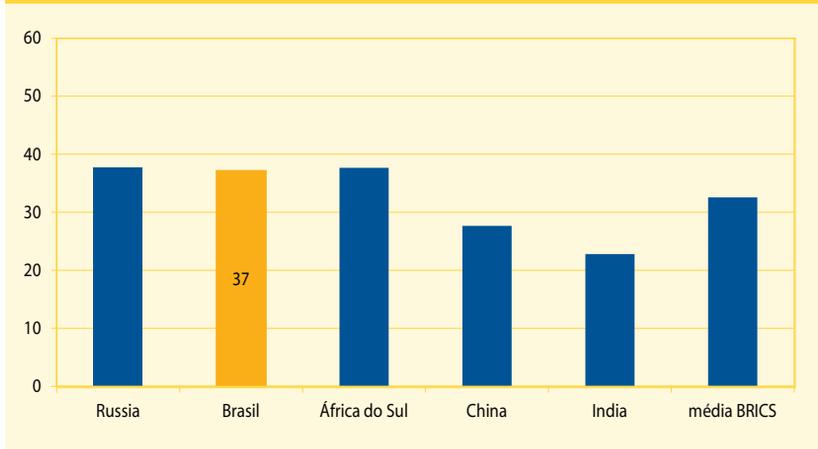


Fonte: International Monetary Fund (2013).

Essa segunda estrutura de agrupamento dos países (Gráfico 1.2) tem como propósito organizar os países segundo similaridades em termos de padrões culturais e sociais. Nesse sentido, os países nórdicos apresentam patamares de gasto público superiores a 50% do PIB de forma a manter o estado de bem-estar social que os caracteriza.

As economias da zona do Euro também se caracterizam por gastos públicos em níveis elevados. No entanto, a Europa tem passado por uma importante crise econômica que pode alterar a estrutura dos seus dispêndios. A crise financeira internacional de 2008 e 2009 obrigou os governos e bancos centrais das economias avançadas a socorrerem bancos e instituições financeiras, de forma a evitar o colapso dos sistemas financeiros naqueles países. Essa política propiciou as condições para a crise fiscal vivenciada a partir de 2011. Os países da zona do Euro tiveram dificuldades de financiar os elevados níveis de gasto dos estados nacionais, e políticas de austeridade fiscal foram discutidas. Um pacto fiscal que limita o déficit público e o endividamento foi aprovado pelo bloco e alguns países realizaram cortes orçamentários expressivos. Dessa forma, o padrão de gasto dos países da zona do euro pode ser alterado após o término desse processo.

Comparativamente às nações emergentes, tanto da América Latina quanto da Ásia, o padrão de gasto público no Brasil é significativamente superior. Dos vizinhos sul-americanos, apenas Colômbia e Costa Rica, com um gasto total em torno de 27% do PIB, aproximam-se do padrão brasileiro. Os demais países da amostra apresentam gastos em patamares mais próximos de 20% do PIB. Por fim, comparando-se com os BRICS (agrupamento que reúne Brasil, Rússia, Índia, China e África do Sul), o total de gasto público no Brasil é levemente superado pela Rússia e pela África do Sul. No entanto, o gasto brasileiro é significativamente superior ao da Índia e ao da China (Gráfico 1.3).

Gráfico 1.3 – Despesa total do governo geral – BRICS (% PIB), 2012

Fonte: International Monetary Fund (2013).

Em resumo, o tamanho do gasto público total no Brasil já apresenta um padrão elevado comparativamente ao padrão internacional. Caso haja condições econômicas para a elevação de gastos públicos no País, esse espaço é certamente reduzido. Nesse sentido, a elevação dos níveis dos serviços públicos ofertados pelo Estado brasileiro deve se dar por meio do aprimoramento da eficiência na aplicação dos recursos públicos.

1.4.2 Análise da composição do gasto público

Uma vez analisado o tamanho do gasto público, a pergunta que se apresenta naturalmente é “com o que o Brasil gasta?”. Na oferta de quais serviços públicos o Estado brasileiro tem empregado mais recursos? Uma das formas de responder tais questões é por meio da análise do gasto público classificado por função de governo. A classificação funcional organiza os gastos públicos por área de atuação governamental, por exemplo, gastos com saúde, educação e previdência social.

No Brasil, tal estrutura de classificação é composta por um conjunto de funções e subfunções que são disciplinadas pela Secretaria de Orçamento Federal (SOF) do Ministério do Planejamento, Orçamento e Gestão (MP). A atual classificação funcional foi instituída pela

Portaria⁶ nº 42, de 14 de abril de 1999, do então Ministério do Orçamento e Gestão (MOG), que atualizou aquela constante da Lei nº 4.320/1964.⁷

A referida portaria estabelece 28 funções de governo que devem ser utilizadas para classificar o gasto público. Tais funções estão dispostas no Quadro 1.1 a seguir.

Quadro 1.1 – Classificação funcional do gasto público	
01 – Legislativa	15 – Urbanismo
02 – Judiciária	16 – Habitação
03 – Essencial à Justiça	17 – Saneamento
04 – Administração	18 – Gestão Ambiental
05 – Defesa Nacional	19 – Ciência e Tecnologia
06 – Segurança Pública	20 – Agricultura
07 – Relações Exteriores	21 – Organização Agrária
08 – Assistência Social	22 – Indústria
09 – Previdência Social	23 – Comércio e Serviços
10 – Saúde	24 – Comunicações
11 – Trabalho	25 – Energia
12 – Educação	26 – Transporte
13 – Cultura	27 – Desporto e Lazer
14 – Direitos da Cidadania	28 – Encargos Especiais

Fonte: Portaria MOG nº 42, de 14 de abril de 1999.

O Gráfico 1.4 apresenta a composição do gasto público no Brasil por função.⁸ Para fins analíticos, além das funções saúde, educação, administração e segurança pública, são apresentadas duas outras classificações que representam agrupamentos das funções relacionadas no Quadro 1.1. Nesse sentido, o item *seguridade social*

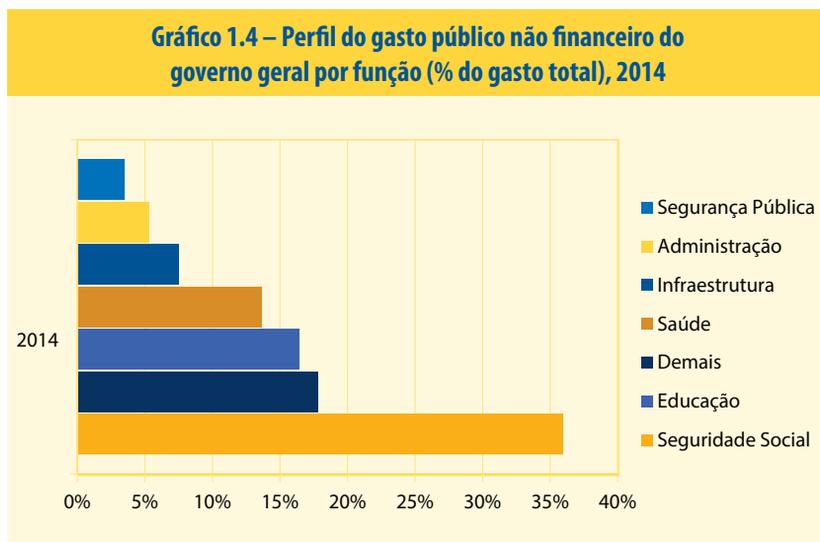
6 A Portaria MOG nº 42, de 14 de abril de 1999, pode ser acessada por meio do endereço eletrônico: <http://www.planejamento.gov.br/secretarias/upload/Legislacao/Portarias/990414_port_42.pdf>.

7 Tratados no inciso I do § 1º do art. 2º e § 2º do art. 8º, ambos da Lei nº 4.320, de 17 de março de 1964.

8 Foi retirada desse cálculo a função *encargos especiais*, pois essa função provoca uma distorção na série, uma vez que representa um volume elevado de gasto decorrente da rolagem da dívida pública.

congrega os gastos nas funções assistência social e seguridade social, enquanto o item *infraestrutura* reúne as funções: urbanismo, habitação, saneamento, comunicações, energia e transporte.

A forma como se dividiram as atividades do governo geral em 2014 pode ser identificada no Gráfico 1.4. Portanto, aproximadamente 63% do total do gasto público não financeiro no Brasil referem-se a despesas realizadas com seguridade social, educação e saúde.



Fonte: Secretaria do Tesouro Nacional. Balanço do Setor Público Nacional.

As despesas com seguridade social na função previdência social representam 36% do gasto e concentram as despesas do Regime Geral de Previdência Social (RGPS) e dos Regimes Próprios de Previdência dos Servidores Públicos (RPPS) nas três esferas de governo.

O RGPS é gerenciado na esfera federal pelo Ministério da Previdência Social (MPS), dispõe de um orçamento anual da ordem de R\$ 400 bilhões em 2014 e paga uma quantidade aproximada de 28 milhões de benefícios previdenciários e acidentários. Os gastos de assistência social, por sua vez, contemplam as despesas ao amparo da Lei Orgânica de Assistência Social (LOAS), e os gastos com a Renda Mensal Vitalícia (RMV).

O aumento da composição da seguridade social está associado a fatores institucionais – como a decisão de política econômica de conceder aumentos de salário mínimo acima da inflação no período analisado, impactando previdência⁹ e programas assistenciais (LOAS/RMV); e também a fatores estruturais, como o envelhecimento natural da população e o aumento da expectativa de vida.

No que se refere ao RPPS da União, os gastos relativos aos benefícios previdenciários dos inativos e pensionistas aumentaram em virtude da política de governo de valorização do servidor público. Dois fatores explicam o incremento: (i) o crescimento vegetativo da folha salarial; (ii) os ganhos reais concedidos pelo governo, fruto da recomposição salarial do funcionalismo federal em geral e da reestruturação de algumas carreiras.

Em relação à assistência social, com objetivo de estabelecer políticas públicas para combate e erradicação da pobreza, a partir de 2003, foram ampliados programas assistenciais, como o Bolsa Família, cuja finalidade é a transferência de renda para as famílias em situação de pobreza. Depois da seguridade social, as funções educação e saúde concentram os maiores gastos da administração pública brasileira: juntas, as duas funções representaram aproximadamente 31% de toda a despesa pública no ano de 2014.

Quadro 1.2 – Gastos nas funções educação e saúde (% do gasto total)				
	2000	2005	2011	2014
Função Educação	14,0%	13,4%	15,5%	17,8%
Função Saúde	11,5%	14,0%	14,1%	13,6%

Fonte: Secretaria do Tesouro Nacional. Balanço do Setor Público Nacional.

O incremento de 2,1 pontos percentuais nos gastos em saúde entre 2000 e 2014 deve-se, entre outros fatores, à promulgação da Emenda Constitucional nº 29, de 2000, que viabilizou o crescimento dos aportes de estados e municípios para a saúde. No caso federal, entretanto, a regra de definir o volume de recursos a serem aplicados

9 Cerca de 2/3 do total de beneficiários recebem até um salário mínimo.

no ano por meio da correção, pela variação nominal do PIB, do valor apurado no ano anterior, teve o efeito de manter a participação em termos da sua parcela no PIB aos patamares do ano de 2000.¹⁰

O crescimento recente dos gastos da área da educação reflete a Emenda Constitucional nº 53, promulgada no fim de 2006 – que, dentre outras medidas, substituiu o Fundo de Manutenção e Desenvolvimento do Ensino Fundamental (Fundef) pelo Fundo de Manutenção e Desenvolvimento da Educação Básica e de Valorização dos Profissionais da Educação (Fundeb), expandindo os valores da complementação da União aos recursos estaduais e municipais aplicados na educação básica, bem como a implantação do Plano Nacional de Educação – no qual se destaca o intenso processo de reestruturação e expansão das instituições federais de ensino superior e tecnológico.

Por fim, os gastos com infraestrutura no Brasil representam 7% do total, em média. Nesse ponto, chega-se à questão de como são executados esses gastos dentro da estrutura federativa do Estado brasileiro. Estados, Distrito Federal, municípios e União atuam de forma equilibrada no provimento dos serviços públicos? Ou determinado gasto é majoritariamente realizado por alguma esfera de governo?

1.4.3 Análise do gasto público no Brasil com educação e saúde

Diferentemente da previdência e da assistência social, nas quais os gastos públicos são majoritariamente realizados pela esfera federal, os gastos com educação e com saúde são executados pelas três esferas de governo de forma coordenada. Assim, estados, Distrito Federal e municípios assumem, com a União, a responsabilidade de prover serviços públicos nessas áreas. O objetivo desta seção é apresentar a composição dos gastos em saúde e educação no Brasil com intuito de disseminar a visão do aumento da eficiência do setor público em um sistema de governo federativo. O detalhamento sobre descentralização e eficiência será tratado no capítulo 11.

10 Ribeiro; Piola; Servo (2007).

Box 1.1 – Recursos mínimos obrigatórios para saúde e educação

Com a Constituição de 1988, a prestação de serviços de saúde passou a se dar de forma compartilhada entre os três níveis de governo, cabendo à União determinar as diretrizes da política de saúde e aos estados e municípios sua execução.

No âmbito da União, enquanto não existe lei complementar que defina a base de cálculo e os percentuais aplicáveis, vigora a exigência de se aplicar, a cada ano, pelo menos o mesmo valor efetivamente empenhado em ações e serviços públicos de saúde no ano imediatamente anterior (ou o mínimo constitucionalmente exigido para esse ano, em caso de infração à regra constitucional), corrigido pela variação nominal do Produto Interno Bruto (PIB) do ano de elaboração da proposta orçamentária.

Para os estados e municípios, o montante mínimo de recursos aplicados em saúde deveria corresponder a um percentual da receita de impostos e transferências constitucionais e legais. Esse percentual foi gradualmente ampliado e, desde 2004, corresponde a 12% para os estados e 15% para os municípios.

Para a área de educação, a organização de sistemas de ensino é responsabilidade conjunta de todas as esferas de governo (União, estados, Distrito Federal e municípios), que devem cumpri-la em regime de colaboração. Os recursos para o financiamento dos sistemas de ensino provêm de recursos tributários dos orçamentos de todos os entes da Federação e também de uma contribuição social especial, o salário-educação.

No caso da União, o montante mínimo que deve ser destinado à manutenção e ao desenvolvimento do ensino corresponde a 18% da receita de impostos (líquida de transferências pagas). Já os estados, o Distrito Federal e os municípios devem destinar a essa finalidade pelo menos 25% da receita de impostos e transferências líquidas. Esses montantes devem necessariamente ser aplicados: nos respectivos sistemas públicos de ensino; em escolas comunitárias, confessionais e filantrópicas, nos termos da lei; ou ainda em atividades universitárias de pesquisa e extensão. As receitas do salário-educação (que não é imposto, e sim contribuição social) não são incluídas entre os recursos mínimos obrigatórios de nenhum ente da Federação; trata-se de uma fonte adicional de recursos.

(continua)

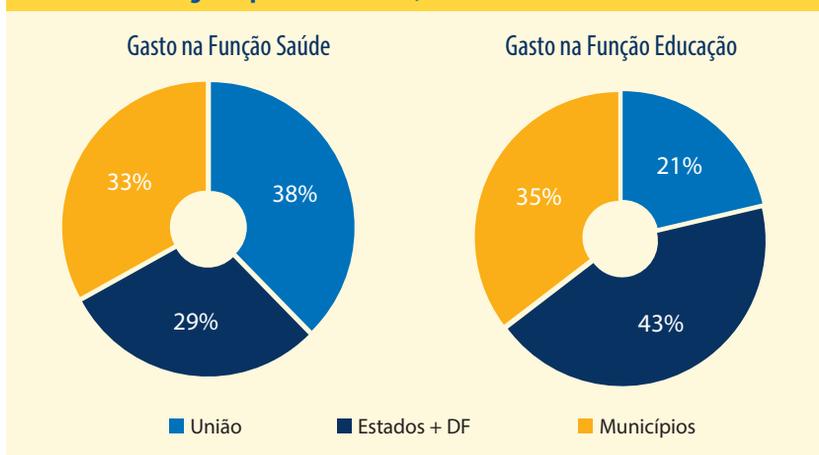
*(continuação)***Box 1.1 – Recursos mínimos obrigatórios para saúde e educação**

Cabe ressaltar que, desde 2011, os recursos mínimos obrigatórios vinculados à educação não são atingidos pela Desvinculação de Receitas da União (DRU).

A DRU constitui elemento essencial para a gestão eficiente dos recursos federais, uma vez que permite a desvinculação de 20% das receitas de impostos e contribuições, possibilitando ao governo maior flexibilidade na aplicação dos recursos. Iniciou-se em 1994, com a criação do Fundo Social de Emergência, aprovado por meio de emenda constitucional. Em 1997, esse dispositivo foi substituído pelo Fundo de Estabilização Fiscal. Em 2000, foi instituída a denominada *Desvinculação de Receitas da União* (DRU), que vigorou até 2003. A DRU foi renovada em 2003, 2007, 2009 e 2011, atualmente com data prevista de término em 2015.

Observa-se que a União responde em maior medida pela saúde (38% do total dos gastos nesta função), enquanto os estados respondem principalmente pela educação (43% do total). Já os gastos dos municípios nessas duas funções estão distribuídos de forma semelhante: 33% na saúde e 35% na educação.

Gráfico 1.5 – Participação de cada esfera de governo no gasto público no Brasil, média de 2000 a 2014



Fonte: Secretaria do Tesouro Nacional. Balanço do Setor Público Nacional.

A forma de organização do sistema educacional no Brasil está disciplinada pela própria Constituição Federal de 1988. Em seu artigo 211, a Carta Magna determina que União, estados, Distrito Federal e municípios atuem de forma colaborativa no provimento de serviços educacionais no território brasileiro. A União deve fornecer assistência aos estados e municípios de forma a assegurar padrão mínimo de qualidade e igualdade de oportunidades. Os estados e o Distrito Federal devem atuar prioritariamente no ensino fundamental e médio, e os municípios, no ensino fundamental e na educação infantil.

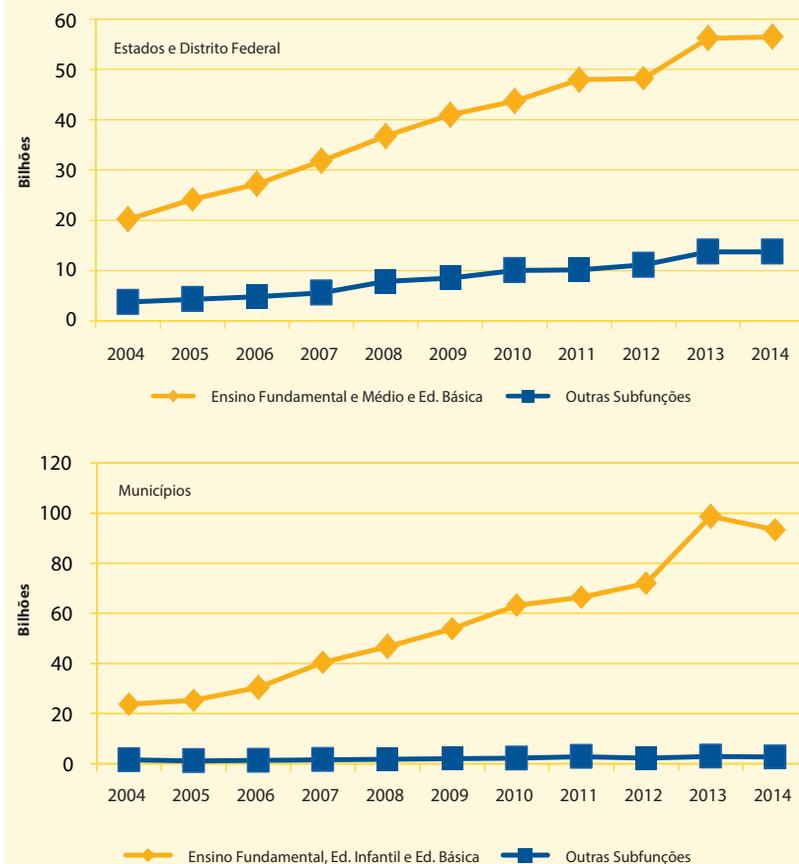
São oito as subfunções da função educação definidas pela Portaria nº 42/2004:

Quadro 1.3 – Classificação funcional do gasto público em educação	
12 – Educação	361 – Ensino Fundamental
	362 – Ensino Médio
	363 – Ensino Profissional
	364 – Ensino Superior
	365 – Educação Infantil
	366 – Educação de Jovens e Adultos
	367 – Educação Especial
	368 – Educação Básica

Fonte: Portaria MOG nº 42, de 14 de abril de 1999.

Os estados e o Distrito Federal estão cumprindo seu papel constitucional de priorizarem o ensino fundamental e médio, pelo menos no que diz respeito à alocação de recursos. O Gráfico 1.6 mostra que o volume de recursos aplicados nessas duas subfunções dobrou no período de 2004 a 2014. Entre as outras subfunções, destaca-se o gasto das administrações estaduais com ensino superior por meio da manutenção de universidades estaduais em quase todas as unidades da Federação.

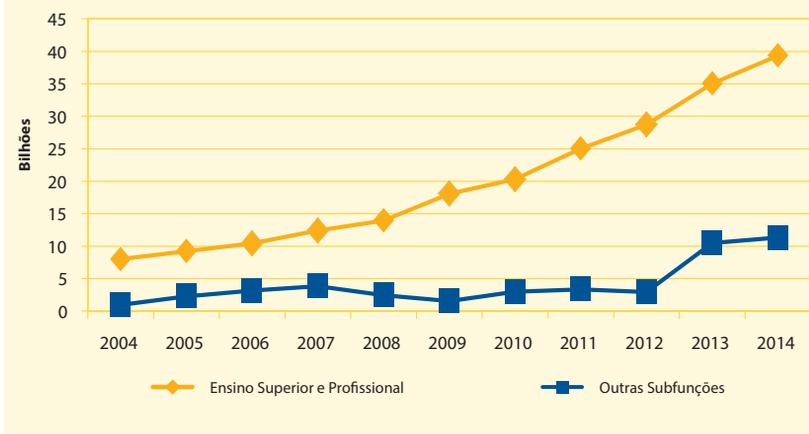
Gráfico 1.6 – Gasto na função educação por subfunção (R\$ bilhões) – estados, Distrito Federal e municípios



Fonte: Secretaria do Tesouro Nacional. Balanço do Setor Público Nacional.

No caso dos municípios, a aplicação de recursos nas subfunções ensino fundamental e educação infantil representa a quase totalidade dos gastos realizados nessa esfera com educação.

Por fim, os gastos do governo federal (Gráfico 1.7) estão concentrados em provimento de ensino superior e ensino profissionalizante por meio das universidades e das escolas técnicas federais.

Gráfico 1.7 – Gastos da União por subfunção (R\$ bilhões)

Fonte: Secretaria do Tesouro Nacional. Balanço do Setor Público Nacional.

Assim como na educação, o sistema de saúde pública no Brasil também está disciplinado pela própria Constituição Federal de 1988. No entanto, neste caso, a Carta Magna estabelece um sistema único de saúde financiado e operacionalizado por todas as três esferas de governo. O sistema é organizado de forma descentralizada, com direção única em cada esfera de governo, e financiado com recursos orçamentários da União, dos estados, do Distrito Federal e dos municípios.

São seis as subfunções da função saúde definidas pela Portaria nº 42/2004:

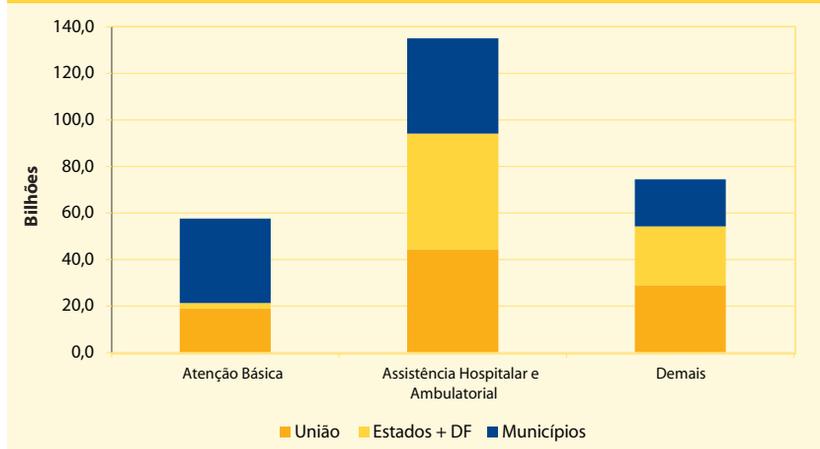
Quadro 1.4 – Classificação funcional do gasto público em Saúde

10 – Saúde	301 – Atenção Básica
	302 – Assistência Hospitalar e Ambulatorial
	303 – Suporte Profilático e Terapêutico
	304 – Vigilância Sanitária
	305 – Vigilância Epidemiológica
	306 – Alimentação e Nutrição

Fonte: Portaria MOG nº 42, de 14 de abril de 1999.

O Gráfico 1.8 a seguir mostra que os maiores gastos são realizados nas subfunções *atenção básica* e *assistência hospitalar e ambulatorial*. Na primeira, os municípios têm maior participação no total de recursos aplicados. Já as despesas com assistência hospitalar e ambulatorial são executadas de forma homogênea por estados, Distrito Federal e União.

Gráfico 1.8 – Gastos públicos no Brasil por subfunção de saúde (R\$ bilhões)



Fonte: Secretaria do Tesouro Nacional. Balanço do Setor Público Nacional.

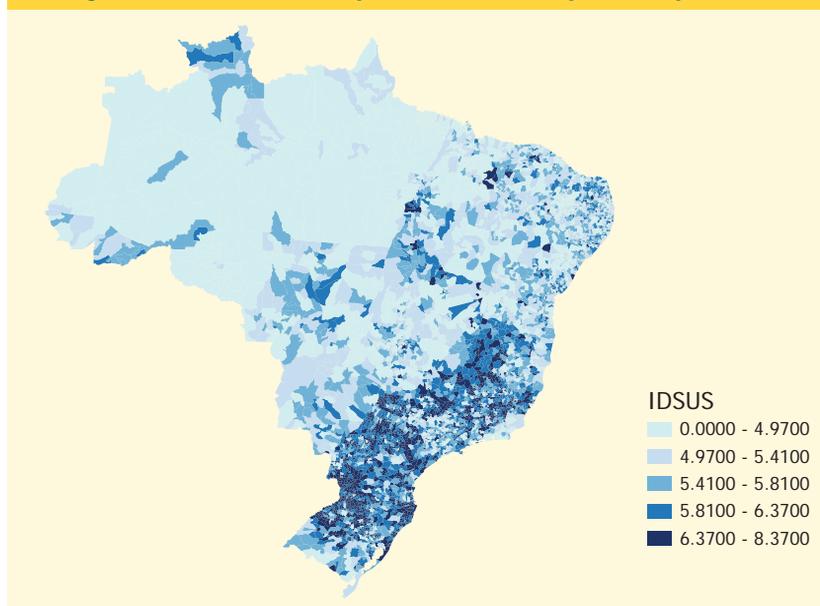
Não obstante o montante dos dispêndios públicos em saúde e educação, o Brasil apresenta grande heterogeneidade interna no que tange ao desempenho dos produtos oferecidos.

Na área da saúde, isso pode ser observado quando avaliamos o Índice de Desempenho do Sistema Único de Saúde (Idsus) de 2011 (Figura 1.1). Referido índice foi criado em 2011 pelo Ministério da Saúde, que avaliou entre 2008 e 2010 os diferentes níveis de atenção (básica, especializada ambulatorial e hospitalar e de urgência e emergência), verificando a infraestrutura de saúde para atender às pessoas e se os serviços ofertados têm capacidade de dar as melhores respostas aos problemas de saúde da população. O índice avalia, com pontuação de 0 a 10, municípios, regiões, estados e país com base em informações de acesso, que mostram como está a oferta de ações e serviços de saúde, e de efetividade, que medem o desempenho do sistema, ou seja, o grau com que os serviços e ações de saúde estão atingindo os resultados esperados.

Importante ressaltar a necessidade de cautela nas comparações individualizadas por municípios, em função da grande diversidade (demográfica, cultural, socioeconômica, geográfica etc.) dos territórios do nosso País. Para uma análise comparativa das notas do Idsus, foram criados grupos homogêneos. Apenas dentro deles, por apresentarem características similares entre si, é possível traçar um paralelo.

O índice apontou que 93,8% dos municípios tiveram nota abaixo da média, estabelecida como 7. Do total dos municípios brasileiros, a maior parte ficou abaixo do regular: 2,4% (132 municípios) tiveram notas variando de 0 a 3,9; 18,3% (1.018) ganharam de 4 a 4,9; 47% (2.616) receberam de 5 a 5,9; 26,1% (1.450) de 6 a 6,9; 6,1% (341) de 7 a 7,9. Apenas seis municípios ficaram com nota acima de 8.

Figura 1.1 – Índice de Desempenho do SUS (Idsus) por município, 2011



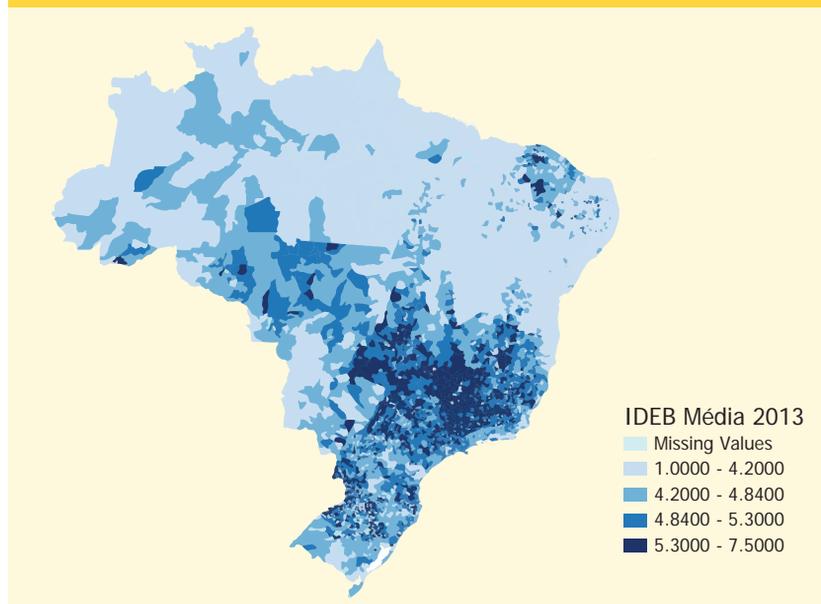
Elaboração dos autores. Fonte: Ministério da Saúde.

Nocasa da educação, o Índice de Desenvolvimento da Educação Básica (Ideb) foi criado pelo Ministério da Educação em 2005, e considera tanto informações de desempenho em exames padronizados quanto as de fluxo escolar, tendo como resultado a combinação de dois outros indicadores: (i) pontuação média dos estudantes em exames padronizados ao final

de determinada etapa da educação básica (5º e 9º anos do ensino fundamental e 3º ano do ensino médio); e (ii) taxa média de aprovação dos estudantes da correspondente etapa de ensino. O indicador é calculado a partir dos dados sobre aprovação escolar, obtidos no Censo Escolar, e médias de desempenho nas avaliações do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), o Sistema Nacional de Avaliação da Educação Básica (Saeb) para as unidades da federação e para o país, e a Prova Brasil – para os municípios.

A Figura 1.2 apresenta a média do Ideb para 5º e 9º anos. Do total de municípios do país, 5.490 tiveram avaliação auferida pelo Ideb 2013 (5º e 9º anos da rede pública) e, destes, 34% apresentaram notas menores que 4 (1.845 municípios). Na ponta oposta, a da excelência, apenas 20% dos municípios apresentaram notas superiores a 5,3 (1.093 municípios). Destes, 65% ficam no Sudeste (706 municípios), 23% no Sul (252 municípios), 9% no Centro-Oeste (99 municípios), 3% no Nordeste (33 municípios) e 0,3% no Norte (3 municípios).

**Figura 1.2 – Índice de Desenvolvimento da Educação Básica (Ideb),
média de 5º e 9º anos, 2013**



Elaboração dos autores. Fonte: Instituto Nacional de Estudos e Pesquisa Educacionais Anísio Teixeira (Inep).

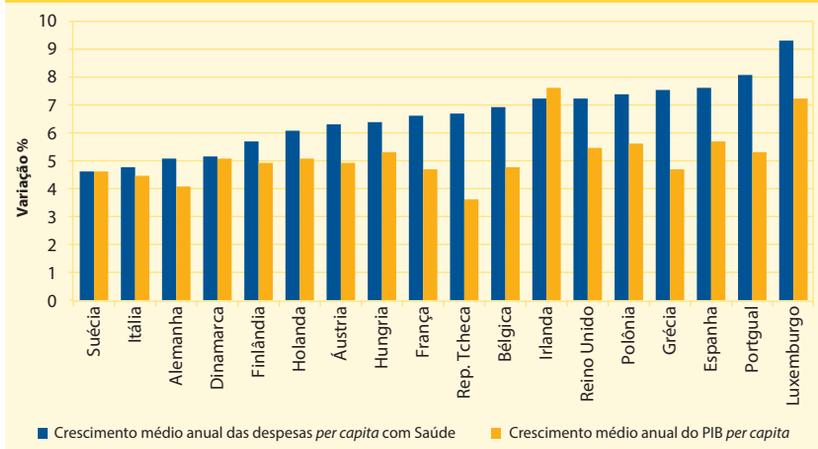
1.5 Uma visão de longo prazo das finanças públicas: desafios e oportunidades para os gastos públicos brasileiros

Tomando a perspectiva de longo prazo, a literatura das finanças públicas apresenta diversas teorias para explicar o crescimento historicamente observado do tamanho dos gastos públicos. Dentre eles, a chamada *Lei de Wagner* atribui essa tendência ao aumento da demanda por bens típicos do setor público por mecanismos intrínsecos ao desenvolvimento econômico. Outra explicação está associada ao aumento dos custos relativos do setor público, em decorrência da tendência de ganhos salariais do setor de serviços em escala superior aos seus ganhos de produtividade, à medida que evolui o processo de desenvolvimento econômico.

A esses elementos agregam-se novos fatores no período contemporâneo, relacionados à dinâmica demográfica. Em síntese, o aumento da longevidade da população terá impactos substanciais sobre as finanças públicas nas próximas décadas. Como seria de se esperar, o item de despesa mais afetado será o gasto previdenciário, que tomará proporção cada vez maior da despesa pública, acompanhando a transformação na pirâmide etária.

No caso da Saúde, a situação é ainda mais grave, na medida em que ao fator demográfico acrescenta-se a tendência de elevação do custo relativo desse serviço. Levantamento feito por Erixon e Van Der Marel (2011) em países europeus evidencia o crescimento observado de gastos *per capita* com saúde acima da renda, ao longo das últimas quatro décadas (1960-2005). Para o caso brasileiro, por se tratar de uma economia com nível de desenvolvimento socioeconômico inferior ao europeu, há que se mencionar ainda a necessidade de expansão da infraestrutura instalada para a prestação dos serviços públicos.

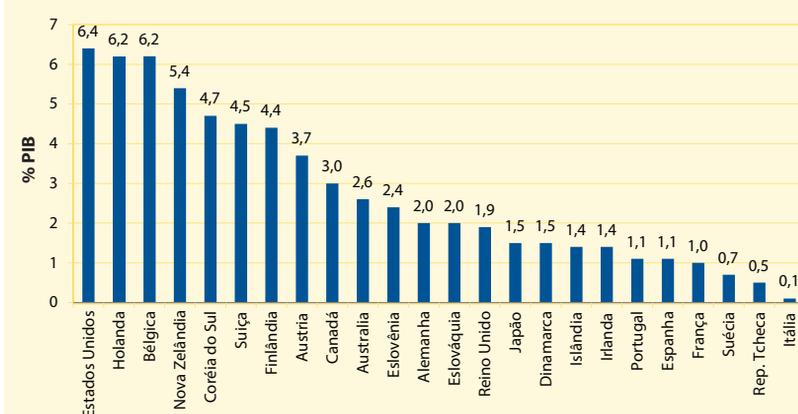
Gráfico 1.9 – Crescimento médio anual *per capita* do PIB e das despesas com saúde, países selecionados, 1960-2005, %



Fonte: Erixon; Van Der Marel (2011).

Em termos prospectivos, é esperado que perdue a dinâmica observada desses gastos nas próximas décadas, devido aos fatores estruturais mencionados. O gráfico a seguir apresenta projeções do FMI para os gastos relacionados ao envelhecimento da população em economias avançadas. É estimado que no período 2014-2030 tais despesas aumentem em 3,8 pontos percentuais do PIB, sendo 74% desse acréscimo no âmbito da saúde e o restante no da previdência. Trata-se de magnitudes expressivas, visto serem gastos permanentes.

Gráfico 1.10 – Países avançados: crescimento estimado das despesas relacionadas ao envelhecimento, % do PIB, 2014-2030

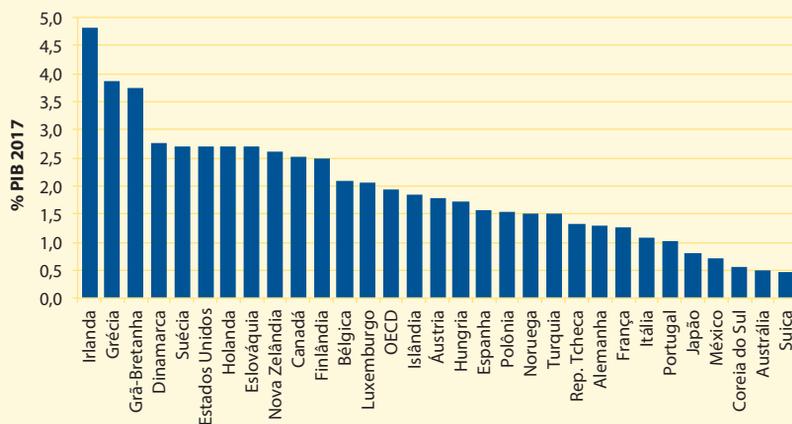


Fonte: International Monetary Fund (2014). Methodological and Statistical Appendix, Statistical Table 24a.

Depreende-se desses cenários a necessidade da realização de reformas estruturais nos sistemas públicos destinados a garantir a viabilidade financeira da prestação desses serviços. Contudo, paralelamente a tais reformas, deverão ser realizados esforços de racionalização dos gastos públicos, por meio dos ganhos de eficiência nesses setores. No tocante a esse aspecto, vale mencionar estudo da Organização para a Cooperação e Desenvolvimento Econômico (OCDE),¹¹ que estimou a economia potencial de gastos dos seus países membros se fossem adotadas medidas que harmonizassem o nível de eficiência entre diferentes unidades prestadoras de saúde e educação. No caso da saúde, os resultados estão reportados a seguir, e mostram que tais ganhos variam de 0,5 a 4,5 pontos percentuais do PIB, conforme o país, alcançando quase 2 pontos percentuais na média. Também se trata de uma magnitude expressiva, e que serve para ilustrar o quanto a análise quantitativa da eficiência do gasto pode ser uma aliada no enfrentamento dos desafios de longo prazo das finanças públicas.

11 Sutherland; Hollter; Rossana (2012)

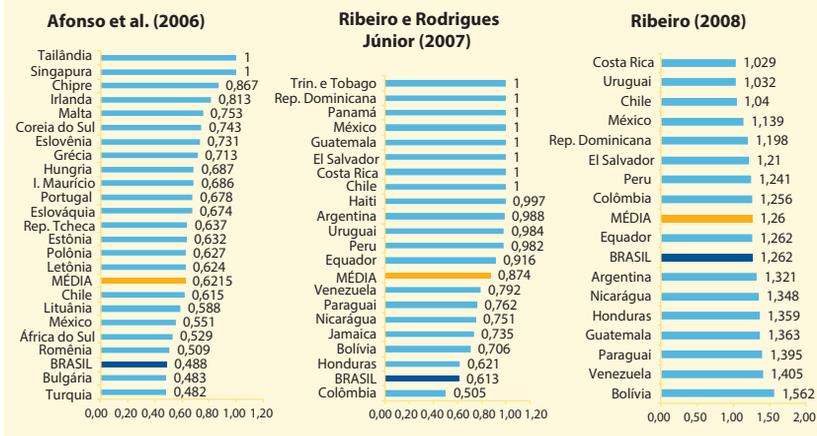
Gráfico 1.11 – Economia potencial com maior eficiência no gasto público com Saúde, países da OCDE, % PIB 2017



Fonte: Sutherland; Holter; Rossana (2012, p. 19).

Nesse sentido, vários estudos mostram o Brasil em uma posição desfavorável frente a um painel internacional, considerando diferentes categorias de países. O gráfico a seguir sintetiza alguns desses estudos para o Brasil.

Gráfico 1.12 – Nível de eficiência do gasto público em países selecionados, conforme três diferentes estudos



Fonte: Afonso; Schuknecht; Tanzi (2006).

Afonso, Schuknecht e Tanzi (2006) traçam um comparativo do Brasil com 24 países em desenvolvimento de diferentes continentes e realidades econômicas e sociais. Nesse estudo, o nível de eficiência do gasto público brasileiro aparece na antepenúltima posição. Ribeiro e Rodrigues Júnior (2007) e Ribeiro (2008) estimam e comparam o nível de eficiência do gasto do setor público brasileiro ao de diversos países da América Latina, que apresentam aspectos econômicos e culturais mais semelhantes aos nossos. Em ambos os estudos, o nível de eficiência do Brasil aparece abaixo da média. Em linha com esses resultados, estudo publicado pelo FMI (2015b, p. 20) estima que a redução de ineficiências dos gastos públicos brasileiros em educação, saúde, assistência social e investimentos públicos permitiria uma economia potencial superior a 3% do PIB.

Em que pese a grande heterogeneidade dos gastos em termos nacionais, o aspecto negativo desse resultado é que – em termos agregados – os recursos públicos no Brasil estão sendo pouco produtivos para os padrões internacionais. Contudo, se analisarmos por outro ângulo, é possível argumentar que, com um dado esforço de racionalização do gasto, também seriam grandes as economias de recursos. Para isso, a utilização de técnicas de mensuração da eficiência das unidades fornecedoras de serviços públicos – especialmente em saúde e educação – pode constituir-se em uma ferramenta muito útil para elaboração de políticas de aprimoramento do gasto público. A economia obtida com a racionalização dos gastos públicos será fundamental para o país enfrentar os desafios postos pela transição demográfica em curso e seus impactos nas contas públicas.

Os dados oficiais indicam que ainda somos uma nação relativamente jovem, o que implica – teoricamente – menor demanda por serviços públicos voltados aos idosos. Contudo, os avanços econômicos e sociais nas últimas décadas ensejaram rápidas transformações no perfil demográfico do país, as quais seguem em ritmo acelerado. Isso pode ser evidenciado no índice de envelhecimento da população brasileira, produzido pelo Instituto Brasileiro de Geografia e Estatística (IBGE). Como observado no Gráfico 1.13, em 2000, para cada grupo de 100 crianças de 0 a 14 anos, havia 18,7 idosos de 65 anos ou mais de idade. Em 2040, a população idosa superaria a de crianças em 13% e, em 2060, a relação poderá ser de 100 para 206,2. Essa mudança no perfil demográfico irá pressionar ainda mais as finanças públicas.

Gráfico 1.13 – Evolução do índice envelhecimento da população – Brasil – 2000/2060



Fonte: Instituto Brasileiro de Geografia e Estatística (2013).

Essa pressão sobre as finanças públicas ocorrerá não somente pela via do gasto previdenciário, mas também pelo acréscimo nos gastos públicos com saúde. Uma sinalização desse fenômeno pode ser encontrada pela evolução dos dados de internações por grupos de doença, associados ao perfil etário correspondente, no período 1999 a 2007. Apesar de referir-se a um período curto quando se trata de transformações demográficas, a proporção das internações por doenças mais associadas a adultos e idosos (neoplasma, doenças do sistema circulatório e digestivo) apresentaram crescimento de 4,12 pontos percentuais no período. Esse dado ilustra a velocidade da mudança no perfil demográfico da população brasileira e seu impacto na demanda pelos serviços públicos.

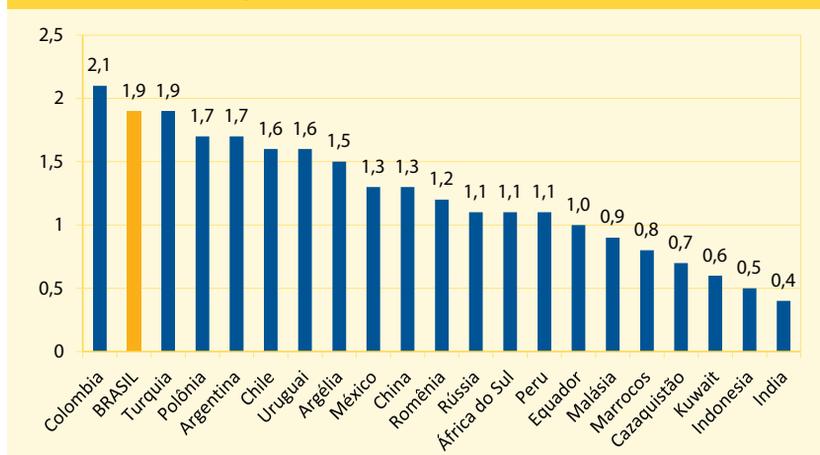
Quadro 1.5 – Internações de acordo com o grupo de doença, em % do total

	1999	2007	Varição em p.p.
Doenças associadas a adultos e idosos	20,54	24,66	4,12
Neoplasma	2,98	5,65	2,67
Doenças do sistema circulatório	9,21	10,22	1,01
Doenças do sistema digestivo	8,35	8,79	0,44
Demais grupos de doenças	79,46	75,34	-4,12

Fonte: World Bank (2011, p. 92).

Em termos prospectivos, estudo recente do FMI¹² apontou que, nos próximos 15 anos, os gastos com saúde no Brasil irão elevar-se em cerca de 1,9 ponto percentual do PIB, o segundo maior impacto dentre os países emergentes pesquisados. Esses dados revelam os desafios fiscais que o país irá enfrentar nos próximos anos com a mudança em curso no perfil demográfico da sua população.

Gráfico 1.14 – Países emergentes: aumento esperado no gasto com saúde, % do PIB, 2015-2030



Fonte: International Monetary Fund (2015a), Methodological and Statistical Appendix, Statistical Table A24..

O crescimento natural dos gastos associados ao envelhecimento da população poderá ser parcialmente compensado por economias geradas pela adoção de melhores práticas de gestão pública, que logrem os mesmos resultados com uma quantidade inferior de recursos públicos aplicados. Portanto, mais do que uma oportunidade, o aumento da eficiência do gasto público no Brasil é uma necessidade. O primeiro passo nessa direção é a mensuração do desempenho das unidades produtoras desses serviços públicos, cujas técnicas e procedimentos afins serão discutidos ao longo deste livro.

12 International Monetary Fund (2015a), Methodological and Statistical Appendix, Statistical A24.

REFERÊNCIAS

AFONSO, A.; SCHUKNECHT, L.; TANZI, V. *Public sector efficiency: evidence for new EU members states and emerging markets*. European Central Bank, 2006. (Working Paper n. 581).

ALESINA, A.; PEROTTI, R. Fiscal adjustments in OECD countries: composition and macroeconomic effects. *IMF Staff Papers*, v. 44, n. 2, Jun. 1997.

ALESINA, A.; ARDAGNA, S. *Large changes in fiscal policy: taxes versus spending*. Cambridge: NBER, 2009. (NBER Working Paper n. 15.438). Disponível em: <<http://www.nber.org/papers/w15438>>.

CECCHETTI, S.; MOHANTY, M.; ZAMPOLLI, F. The future of public debt: prospects and implications. *BIS Working Papers*, n. 300, Mar. 2010.

ERIXON, F.; VAN DER MAREL, E. What is driving the rise in health care expenditures? An inquiry into the nature and causes of the cost disease. *ECIPE Working Paper*, n. 05, 2011.

GUPTA, S. et al. *Política Fiscal, composición del gasto y crecimiento en los países de bajo ingreso*. Departamento de Finanzas Públicas. Fondo Monetário Internacional, 2002.

INTERNATIONAL MONETARY FUND. *World economic outlook – recovery, risk, and rebalancing, chapter 3, will it hurt?* Macroeconomic effects of fiscal consolidation. Washington-DC, Oct. 2010.

_____. IMF Fiscal Monitor. World Economic and Financial Surveys. *Shifting gears: tackling challenges on the road to fiscal adjustment*. Washington, Apr. 2011.

_____. *Government finance statistics yearbook*. Washington-DC, 2013. (v. XXXVII).

_____. IMF Fiscal Monitor. World Economic and Financial Surveys. *Back to Work - How Fiscal Policy Can Help*. Washington: International Monetary Fund, Oct. 2014.

_____. IMF Fiscal Monitor. World Economic and Financial Surveys. *Now Is the Time- Fiscal Policies for Sustainable Growth*. Washington: International Monetary Fund, Apr. 2015a.

_____. IMF Staff Discussion Note. *Fiscal Policy in Latin America: Lessons and Legacies of the Global Financial Crisis*. SDN/15/06, Apr. 2015b.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). Projeção da população por sexo e idade - 2000/2060. Série Relatórios Metodológicos, Rio de Janeiro, Volume 40, 2013. Base de dados: “Indicadores Implícitos na Projeção”, disponível para download em: http://www.ibge.gov.br/home/estatistica/populacao/projecao_da_populacao/2013/default_tab.shtm.

MOTTA, R. S.; MOREIRA, A. Eficiência na gestão municipal do Brasil. *Texto de Discussão IPEA*, n. 1301, 2007.

MCDERMOTT, J.; WESCOTT, R. F. An empirical analysis of fiscal adjustments. *IMF Staff Papers*, v. 43, n. 4, p. 725-753, Dec. 1996.

MUSBRAVE, R. A. *The Theory of public finance*. New York: MacGraw-Hill, 1959.

RIBEIRO, M. B.; RODRIGUES JR., W. Eficiência do gasto público na América Latina. In: INSTITUTO DE PESQUISA ECONÔMICA APLICADA (IPEA). *Boletim de Desenvolvimento Fiscal*. Brasília, 2006.

RIBEIRO, M. B. *Desempenho e eficiência do gasto público: uma análise comparativa do Brasil em relação a um conjunto de países da América Latina*. Monografia Premiada com o 3º Lugar no XIII Prêmio Tesouro Nacional. Qualidade do Gasto Público. Brasília: ESAF, 2008. Disponível em: <http://www.tesouro.fazenda.gov.br/Premio_TN/XIIIpremio/qualidade/3qualidadeXIIIPNTN/Desempenho_Eficiencia_gasto_Publico.pdf>. Acesso em: 6 Jul. 2010.

RIBEIRO, J. A. C.; PIOLA, S.; SERVO, L. M. As novas configurações de antigos problemas: financiamento e gasto com ações e serviços públicos de saúde no Brasil. *Divulgação em Saúde para Debate*, Rio de Janeiro, n. 37, p. 21-43, 2007.

SECRETARIA DO TESOURO NACIONAL. *Balço do Setor Público Nacional*, 2011.

SUTHERLAND, D.; HOLLTER, P.; ROSSANA, M. Fiscal consolidation: how much, how fast and by what means? *OECD Economic Policy Papers*, n. 01, Apr. 2012.

WORLD BANK (2011). *Becoming old in an older Brazil – implications of population aging on growth, poverty, public finance and service delivery*. Conference Edition. Apr. 2011.

APÊNDICE

1. Agrupamento segundo o grau de desenvolvimento econômico

G-7: Canadá, França, Alemanha, Itália, Japão, Reino Unido, Estados Unidos.

Economias avançadas: Austrália, Áustria, Bélgica, Canadá, República Checa, Dinamarca, Estônia, Finlândia, França, Alemanha, Grécia, Islândia, Irlanda, Israel, Itália, Japão, Coreia do Sul, Holanda, Nova Zelândia, Noruega, Portugal, Cingapura, Eslováquia, Eslovênia, Espanha, Suécia, Suíça, Reino Unido, Estados Unidos.

G-20: Argentina, Austrália, Brasil, Canadá, China, França, Alemanha, Índia, Itália, Japão, Coreia, Rússia, África do Sul, Turquia, Reino Unido, Estados Unidos.

Economias emergentes: Argentina, Brasil, Bulgária, Chile, China, Colômbia, Hungria, Índia, Jordânia, Cazaquistão, Letônia, Lituânia, Marrocos, Peru, Polônia, Romênia, Rússia, África do Sul, Tailândia, Turquia, Ucrânia.

Demais: Afeganistão, Armênia, Bolívia, Cabo Verde, República do Congo, Geórgia, Honduras, Lesoto, Maldivas, Moldávia, Mongólia, Tadjiquistão, Iêmen, Bósnia e Herzegovina, Hong Kong, Costa Rica, El Salvador, Jamaica, Ilhas Maurício, San Marino, Gaza, Albânia, Barbados, Belarus, Butão, Macao, Croácia, Egito, Macedônia, Paraguai, Sérvia, Ilhas Seychelles, São Cristóvão e Nevis, Suazilândia, Tunísia.

2. Agrupamento segundo o padrão cultural/social

Países nórdicos: Dinamarca, Finlândia, Noruega, Suécia.

Zona Euro: Áustria, Bélgica, Chipre, Finlândia, França, Alemanha, Grécia, Irlanda, Itália, Luxemburgo, Malta, Holanda, Portugal, Eslováquia, Eslovênia, Espanha.

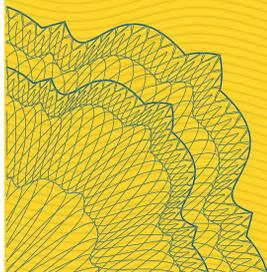
Ásia emergente: China, Índia, Indonésia, Malásia, Paquistão, Filipinas, Tailândia.

Europa emergente: Bulgária, Hungria, Letônia, Lituânia, Polônia, Romênia, Rússia Turquia, Ucrânia.

América Latina emergente: Argentina, Brasil, Chile, Colômbia, México, Peru.

3. Relação dos países que apresentam dados fiscais anteriores a 2011

Suíça (2010), China (2009), Índia (2008), Jordânia (2009), Cazaquistão (2010), Marrocos (2010), Peru (2010), Bolívia (2007), Cabo Verde (2009), República do Congo (2005), Lesoto (2008), Maldivas (2010), Tadjiquistão (2004), Iêmen (2009), Bósnia e Herzegovina (2010), Hong Kong (2010), Costa Rica (2010), Jamaica (2005), San Marino (2009), Gaza (2010), Barbados (2004), Butão (2009), Macedônia (2008), Ilhas Seychelles (2010), Suazilândia (2003).



Capítulo 2

Dívida pública: contribuições de uma gestão eficiente para a estabilidade econômica

Fabiano Silvio Colbano

Maurício Dias Leister

2.1 Introdução

Este capítulo tem por finalidade demonstrar a relevância da gestão da dívida pública para o objeto de estudo deste livro: a eficiência do gasto público. Ao fim da leitura do capítulo, espera-se que esteja claro que uma gestão eficiente da dívida pública permite redução do seu custo de financiamento, abrindo espaço no médio prazo para ampliação de outras despesas consideradas relevantes. Ainda mais, a boa gestão do endividamento é capaz de reduzir o grau de vulnerabilidade das finanças públicas, bem como suavizar o efeito dos ciclos econômicos.

Tal qual um agente privado, que objetiva suavizar seu consumo intertemporalmente e avalia os riscos sobre sua renda futura, é importante que o governo se preocupe com a evolução do seu passivo líquido e também com as variações não esperadas das suas receitas e despesas. Além disso, objetivando maximizar o bem-estar social, há diversas razões que justificam o gerenciamento cuidadoso da dívida pública, por meio da determinação de sua estrutura ótima: (i) diminuição do peso das distorções decorrentes de uma estrutura tributária distorciva; (ii) maior credibilidade das políticas fiscal e monetária; (iii) suavização de problemas de restrição de crédito ao setor privado; (iv) determinação de uma alocação de equilíbrio dos riscos no caso de mercados privados incompletos de ativos; e (v) redistribuição de riscos intra e intergerações.

Também se espera que fique evidente o quanto a gestão da dívida pública tem evoluído ao longo desta última década. A esse respeito, a Secretaria do Tesouro Nacional, ente responsável pela gestão da dívida soberana brasileira, logrou transferir para a sociedade o bônus que o

bom cumprimento dessa tarefa traz consigo. A consistente evolução da classificação de risco torna concretas as conquistas desse processo.

Para atender aos objetivos acima, este capítulo é dividido da seguinte forma. A segunda seção argumenta, do ponto de vista teórico, em favor da relevância econômica da gestão da dívida, inclusive como esta contribui para a suavização de ciclos econômicos. A questão de como a dívida pública pode interferir na volatilidade tributária e no orçamento governamental é o tema da terceira seção. É na quarta seção que a contribuição da gestão de risco para o gestor da dívida pública é abordada. A apresentação de evidências empíricas de como a gestão da dívida pública brasileira contribuiu para redução nos seus custos de financiamento é vista na quinta seção.

2.2 O papel da dívida pública em um ambiente não ricardiano

Nesta seção será abordado o tema da gestão da dívida pública de um ponto de vista teórico-econômico. Nesse sentido, a questão a ser respondida é: de que forma a gestão da dívida pública afeta o ambiente macroeconômico?

A dívida pública tem importantes efeitos sobre a economia real. Sob determinadas hipóteses a cerca das preferências dos agentes privados (a *completude*, a eficiência dos mercados financeiros e a tributação), a trajetória dos gastos públicos será neutra do ponto de vista da alocação real de recursos na economia. Contudo, ainda que estas restrições não se verifiquem no mundo real, estudá-las é importante por duas razões. Primeiramente, o resultado de neutralidade da decisão governamental de se financiar via arrecadação de impostos ou dívida, conhecido como Equivalência Ricardiana (BARRO, 1974), tornou-se um referencial teórico fundamental para as finanças públicas. Em segundo lugar, essa análise constitui-se um primeiro passo para melhor compreensão dos efeitos econômicos do gerenciamento da dívida pública.

Concretamente, conforme argumentam Brennan e Buchanan (1986) e Feldstein (1976), as condições necessárias para que o Teorema da Equivalência Ricardiana seja válido são as seguintes:

1. A existência de mercados de capitais perfeitos, para que os agentes possam emprestar e pedir emprestado fundos à mesma taxa de juro que o Estado;

2. Os agentes não têm qualquer incerteza acerca do nível dos seus rendimentos presentes e futuros;
3. Os agentes, enquanto contribuintes presentes e futuros, comportam-se como se tivessem horizontes infinitos;
4. Os agentes antecipam perfeitamente as responsabilidades fiscais futuras implícitas na dívida pública; e
5. Todos os impostos são *lump-sum*.

Os resultados derivados na Equivalência Ricardiana também se apoiam em duas hipóteses: expectativas racionais¹ e o modelo do ciclo de vida.² Com esses pressupostos, qualquer aumento do déficit público, que redunde em elevação da dívida pública, leva os agentes a aumentarem suas poupanças. Isso porque, dadas as expectativas racionais, eles antecipam a necessidade de aumento futuro nos impostos para que o governo possa cumprir com os pagamentos de maiores encargos financeiros da dívida. Nesse caso, os efeitos econômicos do aumento de gastos do governo seriam inteiramente revertidos pelo aumento da poupança privada para fazer frente ao aumento futuro esperado dos impostos. Também, o aumento da poupança privada reduziria o consumo privado, anulando o efeito do aumento dos gastos do governo sobre o crescimento econômico.

Da mesma forma, um aumento do gasto público que fosse financiado via elevação de impostos não afetaria o crescimento econômico, isto porque a elevação dos impostos diminuiria a renda disponível do setor privado. Consequentemente, também haveria redução do consumo privado, que compensaria o aumento do gasto público.

A taxa de juros também não é afetada pelo aumento da dívida pública, pois como a poupança privada aumenta no mesmo montante que o déficit público, a taxa de juro mantém-se inalterada. Assim, o

1 Assume-se que há um comportamento humano racional, maximizador de utilidades esperadas e apto a processar de maneira ótima todas as informações disponíveis, de forma que os indivíduos não cometem erros sistemáticos ao prever o futuro e fazem eficiente uso de todas as informações disponíveis. Para mais detalhes ver Muth (1961) e Lucas e Sargent (1978).

2 Defende que os agentes econômicos determinam a distribuição da sua renda entre consumo e poupança, sempre com vistas à manutenção de um padrão estável de consumo ao longo da vida, apesar das variações correntes da renda disponível. Em outras palavras, os agentes determinam seus níveis de consumo de acordo com sua expectativa de renda permanente. Quaisquer diferenças entre a renda corrente e a renda permanente afetarão seus níveis de poupança. Para mais detalhes ver Modigliani e Brumberg (1954).

déficit não provoca qualquer redução no ritmo de acumulação do estoque de capital, nem deterioração das contas externas. Enfim, a dívida pública não afeta a riqueza do setor privado.

Para a economia como um todo, a diferença entre o governo financiar seu aumento de gasto via dívida pública ou via elevação dos impostos é que, no primeiro caso, ele provoca elevação da poupança privada e redução da poupança pública, mantendo-se a poupança doméstica total inalterada, enquanto, no segundo, não há mudança na poupança dos agentes. Em todo caso, o crescimento econômico também não seria afetado.

A conclusão, apresentada por Hermann (2002, p. 7), é que

o déficit público não traria, portanto, qualquer benefício em termos de crescimento econômico, tendo, ao contrário, um impacto negativo sobre o bem-estar da sociedade, representado pelo ônus da dívida a ser paga pelas gerações futuras. Daí a recomendação de uma política fiscal de permanente equilíbrio orçamentário.

Ao manter constante o nível corrente de gastos, restringe-se o conjunto de políticas por meio das quais o governo pode afetar a economia. Disto, segue que a proposição ricardiana é condição suficiente para a neutralidade do gerenciamento da dívida pública, já que os agentes podem antecipar perfeitamente o estado futuro da economia. Ou seja, a gestão da dívida torna-se irrelevante.³

No entanto, é possível derivar condições mais gerais sobre a estrutura da economia sob as quais o gerenciamento da dívida é neutro mesmo em um ambiente não ricardiano.⁴ Nesse caso, pode-se pensar a neutralidade da dívida pública como uma extensão da Equivalência Ricardiana. A principal diferença entre ambas é que, por definição, a escolha dos instrumentos de dívida não afeta a arrecadação de

3 Aqui cabe um esclarecimento: para Buchanan (1976), a paternidade do teorema proposto por Barro (1974) deve ser atribuída a Ricardo, e, portanto, segundo esse entendimento, é possível afirmar que Ricardo era cético em relação às hipóteses de uma teoria que ganhou publicidade mais de cem anos depois. Para Buchanan, *op. cit.*, depois de Ricardo enunciar claramente a equivalência, ele nega a sua validade, devido à presença daquilo que hoje é conhecido como *ilusão fiscal*.

4 Como se pode ver, a equivalência ricardiana somente é válida sob condições bastante restritivas, tornando-a pouco realista. As críticas a essas hipóteses estão difundidas em vários estudos. As condições 1 e 2 são criticadas por Tobin (1980). Para Ricardo (1820, 1821), as condições 3 e 4 não parecem ser críveis. E a condição 5 é criticada por Tobin (1980) e Brennan e Buchanan (1986).

impostos (e os gastos) correntes, mas apenas no futuro e em diferentes estados da natureza.⁵ Dessa forma, mudanças na composição da dívida alteram seu custo de carregamento futuro para vários estados da natureza, afetando conseqüentemente o passivo total do governo.

Missale (1999, p. 11) apresenta o *Teorema da Neutralidade da Dívida*, formado por um conjunto de condições suficientes para que as operações financeiras do governo não alterem o equilíbrio da economia em um ambiente não ricardiano:

Dados os impostos correntes e uma trajetória exógena para os gastos públicos, o gerenciamento da dívida pública não afetará a alocação real de recursos se:

- i. os agentes privados são racionais; ou as gerações atuais de indivíduos percebem um aumento futuro dos impostos decorrentes das ações do governo ou eles estão ligados às gerações futuras através de transferências com motivações altruístas;
- ii. a proporção de impostos futuros paga por cada indivíduo independe do estado da natureza;
- iii. os mercados de capitais são perfeitos. Em particular, ou não há restrições às vendas a descoberto de ativos governamentais ou existem substitutos privados para estes ativos;
- iv. o uso de dívida não cria valor, ou seja, não há possibilidade de esquemas Ponzi com dívida;
- v. a estrutura tributária não causa distorções na economia (impostos são do tipo *lump-sum*).

Uma análise mais detalhada das condições impostas pelo teorema explica por que o gerenciamento da dívida pública afeta

5 Para um dado evento futuro, o custo de carregamento da dívida dependerá da sua composição, a qual afeta, conseqüentemente, o volume necessário de impostos para se manter determinada trajetória de dívida. A composição ótima da dívida pública é definida com base nas preferências sociais entre custo e risco. Assim, o gestor de dívida pública define o perfil desejado para o passivo de longo prazo, ou seja, a composição ótima, de modo que o financiamento ocorra da forma menos onerosa possível sem, no entanto, ocasionar elevada exposição a riscos. A questão da composição será tratada com mais detalhes ao longo do capítulo.

a atividade econômica e aponta para os efeitos em termos de bem-estar decorrentes de medidas específicas de política econômica. Particularmente, a escolha dos instrumentos de dívida tem importantes implicações sobre a redistribuição de riscos na economia e sobre a eficiência do sistema tributário e dos mercados financeiros.

Em um mundo com incerteza, o *trade-off* custo-risco para uma dada composição da dívida tem importantes implicações para a estrutura tributária.⁶ Como a tributação futura depende da composição da dívida, mudanças em sua composição alteram a tributação nos momentos seguintes. Se a estrutura tributária é distorcida, tais mudanças impactam diretamente a oferta total de trabalho e o investimento, com efeitos sobre o crescimento econômico e o bem-estar social. Nesse caso, a composição da dívida pode contribuir para minimizar os impactos das distorções geradas pela estrutura tributária e, com isso, aumentar o nível de bem-estar social.

Para alcançar uma estrutura tributária eficiente, a composição da dívida é relevante para o governo, já que ela determina em que medida as distorções podem ser suavizadas. Com um mercado completo de ativos contingentes aos diversos estados da natureza, esse trabalho é facilitado, e o governo pode programar as políticas ótimas. No entanto, mais importante do que ter um mercado completo é a capacidade do governo de emitir títulos com indexadores coerentes aos fatores de riscos econômicos mais relevantes, fato que contribuiria para a diversificação de riscos na economia intra e intergerações.

O problema da inconsistência temporal das políticas fiscal e monetária é outra razão para que a gestão da dívida seja relevante. Na presença de uma estrutura tributária ineficiente, há incentivos futuros para que o governo altere a tributação ótima definida inicialmente. Isto acontece porque, uma vez emitida determinada composição de dívida, as regras ótimas de políticas fiscal e monetária se alteram. A composição e a maturidade da dívida pública são relevantes para determinar o tamanho do incentivo governamental

6 Correspondendo à estrutura tributária ótima, existe uma única composição da dívida pública associada a esta estrutura. Portanto, em um ambiente dinâmico e estocástico, a teoria de tributação ótima é também uma teoria para o gerenciamento ótimo da dívida pública, como mostrado em Lucas e Stokey (1983).

para se desviar do plano tributário original. Portanto, é possível elevar *ex-ante* a credibilidade das políticas macroeconômicas simplesmente com a escolha da composição adequada de dívida. Essa situação também poderia ser contornada no caso de o governo ser capaz de se comprometer com a tributação ótima inicialmente escolhida, que implicaria a impossibilidade de alteração *ex-post* das políticas escolhidas anteriormente.

Um exemplo de como a composição da dívida pode aumentar a credibilidade da política monetária está a seguir. Uma dívida pública com um percentual alto de títulos prefixados cria um incentivo para que o governo incorra em uma inflação mais alta no futuro, que poderia reduzir o valor real da dívida e provocar perdas para os investidores. Nesse caso, se os agentes são racionais, eles cobrarão um prêmio mais alto de inflação sobre os papéis prefixados, elevando o custo desses papéis para o governo e, assim, reduzindo os incentivos do governo a gerar uma inflação mais alta. Esses incentivos desaparecem caso a dívida tenha uma parcela maior de papéis indexados à inflação ou à taxa de câmbio. Dessa forma, o governo pode usar a gestão da dívida pública como uma indicação de que há comprometimento com a política monetária previamente estabelecida.⁷

Goldfajn e Paula (1999) trabalham com um contexto mais próximo da realidade e relaxam a hipótese da equivalência ricardiana, e também descartam a observância de outras duas hipóteses: inexistência de distorções tributárias, e existência de mercados completos e informação simétrica. Com isso, o resultado a que chegam é um ambiente onde o perfil de vencimentos da dívida torna-se um importante instrumento para amenizar a variação dos impostos. Ou seja, quando tais hipóteses não são consideradas, a composição da dívida pública se torna um importante instrumento para a suavização de impostos, para a partilha do risco e para a construção de credibilidade do governo.

7 Importante salientar que a gestão da dívida pública para fins de sinalização de comprometimento com a política monetária não pode ser levada ao extremo. Isso porque objetivos diversos para a gestão da dívida pública (suavização de carga tributária e sinalização para a política monetária, por exemplo) podem gerar diretrizes contraditórias entre si. Portanto, entende-se que o gestor de dívida deve primeiramente perseguir as diretrizes da boa gestão da dívida e, naquilo que não for conflitante com tais princípios, atuar no sentido de reforçar a sinalização da política monetária.

Compreendido, então, que a gestão da dívida pública é tarefa relevante para a economia de um país, é o momento de apresentar quais seriam os princípios, os objetivos e a composição que a dívida pública deve respeitar. Esse é o tema da próxima seção.

2.3 A gestão da dívida pública e a suavização da carga tributária

A gestão eficiente da dívida pública deve contribuir para um melhor resultado fiscal. Para entendimento dessa questão, é oportuno que se comente alguns aspectos teóricos referentes à gestão de dívida soberana. Primeiramente, um gestor de dívida pública deve escolher quais títulos públicos emitirá, questão fundamental na determinação tanto dos tipos de riscos aos quais a dívida estará exposta, o que se reflete no nível total de risco embutido no estoque de dívida, quanto dos respectivos custos. Com isso, aspectos relacionados a indexadores e estrutura de vencimentos estão no foco do gestor da dívida.

Quanto aos indexadores, os mais usuais são: prefixados, remunerados por índices de preços, indexados à taxa flutuante e à taxa de câmbio. Cada indexador embute um grau de risco específico.⁸ A determinação de qual é mais ou menos arriscado depende da particularidade de cada devedor. Em outras palavras, o risco depende da natureza dos ativos e das receitas do emissor,⁹ pois são estes que garantirão o pagamento dos juros e principal dos títulos emitidos. Portanto, no caso das dívidas soberanas, é desejável, entre outros aspectos, que os gastos com o serviço da dívida contratada sejam pró-cíclicos em relação às receitas arrecadadas pelo governo e correlacionados com os ativos estatais.

Deve-se ter em mente outros dois aspectos importantes para a escolha dos títulos a serem emitidos: custo e prazo. Assim, os títulos menos arriscados do ponto de vista do emissor (em geral mais arriscados para o investidor), serão acompanhados por maiores exigências em

8 A seção 2.4 discute a questão da gestão de riscos da dívida pública.

9 Por exemplo, um país onde o governo detenha controle sobre grandes reservas petrolíferas, cuja extração é destinada majoritariamente à exportação, pode não ter o risco cambial como sua principal fonte de risco. Isso porque as receitas que esse governo arrecada estão fortemente atreladas às divisas estrangeiras. Por outro lado, um país com reduzida atividade em sua balança comercial e diminuto volume de reservas internacionais pode encarar os títulos cambiais como extremamente arriscados, pois tanto suas receitas quanto seus ativos não possuem correlação com as divisas estrangeiras.

relação aos custos a eles associados. Analogamente, os prazos (tempo a decorrer até o vencimento) tendem a ser maiores quanto mais arriscado for o título para o emissor (menos arriscado para o investidor).¹⁰

Aqui deve estar claro que uma composição de dívida que se mostre mais vulnerável, à luz das particularidades do emissor, reflete-se em maior custo de financiamento ao longo do tempo, mesmo que em determinado momento a emissão de dívida mais arriscada seja menos custosa. Portanto, a busca permanente da menor vulnerabilidade deve se traduzir em menores pressões sobre as finanças públicas, com comprometimento cada vez menor de recursos orçamentários, necessários para fazer frente aos juros dessa dívida. A busca de uma composição ótima de indexadores da dívida pública atua justamente na conquista do binômio menor vulnerabilidade/menor custo de emissão.

Feito esse breve relato, é possível compreender que a escolha dos títulos que comporão a dívida pública de um país, bem como a sua representatividade no estoque, não é fruto de uma regra padronizada. É uma decisão que envolve análises profundas de riscos, custos e prazos de acordo com as particularidades de cada governo. E, para tal, é fundamental a criação de capacidade técnica, passando por regras claras de governança, eficiente estrutura institucional, centralização da gestão¹¹ e formação e investimento em equipes qualificadas.

Além do princípio supracitado de perseguição da composição ótima com a finalidade de redução da vulnerabilidade das finanças públicas, a perseguição do princípio de suavização da carga tributária (*tax-smoothing*) e da redução dos efeitos dos choques macroeconômicos sobre os contribuintes e sobre o orçamento do governo também devem ser preocupações do gestor da dívida pública.

É desejável que uma carga tributária seja constante no tempo (*suavização de impostos*), o que garante a consistência intertemporal na decisão dos agentes econômicos. Assim, o governo deve estruturar a dívida pública com o objetivo de estabilizar o imposto no tempo, considerando os diferentes estados da natureza.

10 Daqui em diante, caso não se faça menção em contrário, o risco será tratado pela ótica do emissor.

11 Sobre a centralização da gestão, ver seção 2.5.

A suavização da volatilidade tributária significa que, se os gastos do governo se elevam, por uma razão qualquer, durante determinado período de tempo, em vez de aumentar os impostos, o governo se financia por meio da emissão de títulos públicos. Analogamente, quando os gastos se reduzem e, mantida a carga tributária, os superávits orçamentários daí advindos devem ser utilizados para redução do endividamento público.¹²

Uma carga tributária volátil também redundando em outros custos além dos envolvidos na arrecadação de impostos, tais como as distorções provocadas nas decisões econômicas e as incertezas geradas aos negócios. Considerando todos esses custos fica claro que, sob a ótica dessa teoria, os déficits públicos são importantes porque minimizam os problemas associados às variações tributárias, dado que são usados para manter a carga tributária constante ao longo do tempo, apesar de flutuações tanto nos gastos quanto na receita do governo.

Choques sobre o orçamento governamental também podem ser minimizados pela boa gestão da dívida pública. Na medida em que uma gestão eficiente da dívida minimiza os efeitos dos choques sobre o orçamento do governo, ela contribui para a manutenção dos gastos públicos essenciais. Por conseguinte, a maior estabilidade do orçamento pode elevar a qualidade desses gastos.¹³

Uma boa administração do perfil da dívida em termos de indexadores, prazos e concentração de vencimentos garante que se caminhe cada vez mais para um binômio custo-risco que redunde em um suprimento eficiente das necessidades de financiamento do governo ao menor custo de financiamento no longo prazo, respeitando-se a manutenção de níveis prudentes de risco.

Vale aqui discutir brevemente os tipos de indexadores da dívida pública sob a ótica dos impactos de cada um no orçamento

12 A razão para tal está explicitada no trabalho de Barro (1979), que difundiu a teoria conhecida como *tax-smoothing*. Nele, o autor defende que o governo, no intuito de minimizar os custos de distorção envolvidos na arrecadação de impostos – esses custos incluem transferência de recursos dos indivíduos para o governo e custos de coleta de impostos –, utilize-se da dívida pública para tornar regular a carga tributária ao longo do tempo.

13 Na prática, o gestor da dívida deve ter atenção especial não somente para os riscos do serviço da dívida, mas também para os relacionados ao estoque, pois muitas vezes o aumento inesperado deste altera a percepção dos investidores sobre a sustentabilidade da dívida, obrigando o gestor a elevar a carga tributária para sinalizar comprometimento com o seu pagamento.

do governo. Essa questão é explorada por Togo (2007) e aqui reproduziremos os argumentos desse trabalho de forma resumida. Em relação à parcela da dívida prefixada, sabe-se que ela pode proteger a posição fiscal em momentos de choque negativo de oferta agregada. Isso porque tal choque gera redução do produto e elevação do nível de preços. Assim, é de se esperar que as receitas governamentais se reduzam e os gastos aumentem. Portanto, enquanto o serviço da dívida indexada à inflação tende a aumentar com o nível de preços, os títulos prefixados atuam no sentido contrário, reduzindo os juros da dívida em termos reais. Dessa forma, nesse cenário os títulos prefixados fornecem proteção e previsibilidade ao balanço fiscal do governo.

Em relação aos títulos indexados a índices de preços, a proteção do orçamento governamental ocorre em cenário de choque negativo de demanda. Dado que a queda na demanda conduz a uma contração nos níveis de preços e no produto, as receitas governamentais tendem a cair, pois são positivamente correlacionadas com a atividade econômica, e os gastos do governo tendem a se elevar, devido às medidas anticíclicas. Paralelamente, o serviço da dívida indexada à inflação se reduz, acompanhando o nível de preços e oferecendo proteção à posição fiscal.

Já os títulos com taxas de juros flutuantes se mostram favoráveis ao orçamento governamental quando o ambiente é de choque negativo de demanda agregada ou choque positivo de oferta, dado que a taxa de juros de curto prazo tende a ser reduzida, gerando também redução nos gastos com juros nominais desses títulos. No caso de um choque positivo de demanda ou de um choque negativo de oferta agregada, a taxa de juros de curto prazo tende a aumentar em resposta ao aumento no nível de preços, elevando os gastos com juros nominais. Dessa forma, como argumenta Alves (2009, p. 12),

em comparação com os prefixados, os títulos com taxas de juros flutuantes tendem a se comportar bem diante de choques de demanda, mas não em eventos de choques de oferta. A comparação entre títulos indexados à inflação e aqueles com taxas de juros flutuantes depende da medida em que a taxa de juros de curto prazo mudará em comparação com a taxa de inflação. Portanto, depende do que acontecerá com a taxa real de juros.

Por fim a parcela da dívida constituída por papéis indexados a moeda estrangeira pode trazer maior volatilidade. Isso porque, para as contas públicas, uma depreciação da moeda doméstica pode representar aumento do valor do serviço desses títulos da dívida pública. Importante destacar que o grau de vulnerabilidade de cada país em relação à dívida em moeda estrangeira depende de um conjunto de fatores, entre eles a quantidade de reservas internacionais e o grau de abertura econômica.

Enfim, o que deve ficar claro é que uma das funções da composição da dívida é amortecer a variabilidade dos gastos com o seu serviço. Assim, como afirmam Goldfajn e de Paula (1999), a escolha da composição da dívida deve levar em conta as categorias que minimizem o risco de flutuações no custo da dívida, bem como na carga tributária.¹⁴ Sob esse aspecto, o tema gestão de riscos da dívida pública é tratado na seção seguinte.

2.4 A importância da gestão de riscos

Como visto nas seções anteriores, a estrutura da dívida é um fator importante a ser monitorado para o alcance de uma gestão eficiente dos gastos públicos, uma vez que auxilia na redução dos riscos associados a flutuações no seu custo. Nesse sentido, a composição e o perfil de dívida que o governo deve perseguir é assunto de perene interesse dos gestores públicos, tendo o gestor de riscos da dívida pública o papel importante de apontar os dilemas envolvidos em termos de custos e riscos para as diferentes estratégias de dívida de longo prazo.

Quanto a esse tema, esta seção apresentará os objetivos, as diretrizes, os indicadores e os instrumentos usados na gestão de riscos da dívida pública brasileira.¹⁵

14 Aqui vale uma qualificação. É importante ter em mente que não se deve privilegiar unicamente a minimização de riscos a qualquer custo. Isso porque a redução de risco marginal tende a ser cada vez mais custosa à medida que se atingem graus de risco menores. Portanto, a partir de determinado ponto, o custo de se diminuir uma unidade de risco pode ser extremamente alto, não valendo a pena do ponto de vista de gestão estratégica da dívida pública.

15 Para cumprir tal missão, esta seção se apoiou basicamente em dois capítulos do livro *Dívida Pública: a experiência brasileira*, organizado por Silva, de Carvalho e de Medeiros (2009). Os capítulos especificamente são os apresentados por Alves e Silva (2009) e por Silva, Cabral e Baghdassarian (2009).

Segundo Silva, Cabral e Baghdassarian (2009, p. 175),

o gestor de riscos identifica possíveis riscos para a implementação da estratégia de dívida e refina as estimativas (alvos) para a composição e o perfil da dívida no futuro [...]. É também seu papel o monitoramento da execução da estratégia e, quando necessário, a sugestão de medidas corretivas na condução da estratégia de emissões.

Os riscos mais relevantes para a dívida pública são: risco de mercado, risco de refinanciamento, risco orçamentário e risco de demanda. O risco de mercado captura a possibilidade de elevação no estoque da dívida decorrente de alterações nas condições de mercado que afetem os custos dos títulos públicos, tais como as variações nas taxas de juros de curto prazo, de câmbio e de inflação, ou na estrutura a termo da taxa de juros. No caso específico da Dívida Pública Federal (DPF) brasileira, a presença de títulos remunerados pela taxa Selic ou por índices de preços e de títulos denominados em moeda externa faz com que alterações nessas variáveis afetem o custo da DPF (SECRETARIA DO TESOURO NACIONAL, 2012). No mercado financeiro, esse tipo de risco está relacionado à volatilidade dos preços dos ativos. Porém, quando se trata de dívida pública, esse risco refere-se a mudanças no valor da carteira (estoque da dívida).

A análise inicial dos riscos do estoque de dívida pública envolve um conjunto abrangente de indicadores de risco, tais como composição, prazo médio, *duration* de repactuação e perfil de maturação (medido como um percentual da dívida que vence no curto prazo, por exemplo).¹⁶

O risco de refinanciamento da dívida pública, por sua vez, é definido como o risco de mudanças abruptas no perfil de pagamentos da dívida no momento do seu refinanciamento (SILVA; CABRAL; BAGHDASSARIAN, 2009). No extremo, esse risco pode redundar em incapacidade de um governo de refinar parte ou a totalidade da dívida vincenda em determinado momento.

16 No entanto, há outros tipos relevantes de medidas que, inclusive, abrangem modelos financeiros mais sofisticados e com simulações estocásticas. A primeira dessas medidas é o *Cost-at-Risk* (CaR), que busca simular o estoque da dívida a partir de choques em seus indexadores durante o período de um ano. Outras medidas que envolvem simulações estocásticas são o *Cash-Flow-at-Risk* (CfaR) e o *Budget-at-Risk* (BaR).

A Secretaria do Tesouro Nacional utiliza três indicadores para mensurar esse tipo de risco: prazo médio, perfil de maturação da dívida (em especial o percentual da dívida vencendo no curto prazo) e *Cash-Flow-at-Risk* (CfaR). Cada um desses indicadores apura o risco de refinanciamento sob uma ótica diferente. Assim, o prazo médio explicita um ponto de equilíbrio de todos os vencimentos de dívida. Como é uma média, o acompanhamento da evolução desse indicador ao longo do tempo pode ajudar a prevenir encurtamentos sistemáticos da dívida, o que poderia trazer problemas para os gestores (SILVA; CABRAL; BAGHDASSARIAN, 2009).

O indicador de percentual da dívida vencendo em 12 meses é complementar ao prazo médio, pois tem foco no curto prazo. Como afirmam Silva, Cabral e Baghdassarian (2009, p. 181), “enquanto o prazo médio mede possíveis reduções sistemáticas nos prazos dos fluxos de pagamentos, o percentual da dívida vencendo em 12 meses está mais focado nas necessidades de caixa para honrar os pagamentos em um ano”.

Finalmente, o *Cash-Flow-at-Risk* (CfaR) mede, “a um dado nível de significância, o valor máximo do fluxo de caixa (pagamentos) em datas ou períodos específicos no futuro”. Vale lembrar que esse risco não está associado a títulos prefixados, pois, dada a natureza desses papéis, já se sabe de antemão qual o valor exato do fluxo de caixa. Porém, a regra é diferente para títulos a taxas flutuantes ou indexados à inflação/câmbio, dado que não é possível saber antecipadamente qual será o valor exato dos desembolsos exigidos em moeda doméstica. Para estes, o uso do CfaR é bastante útil na mensuração do risco de refinanciamento.

O risco orçamentário, por sua vez, consiste no risco de que o serviço da dívida dentro do ano fiscal supere o valor originalmente aprovado pelo Congresso no orçamento. O indicador que mensura esse risco é o *Budget-at-Risk* (BaR). Segundo Silva, Cabral e Baghdassarian (2009, p. 183),

o BaR é bastante similar ao CfaR no sentido de que ambos medem a incerteza de fluxos de caixa. A diferença essencial entre os dois é que o BaR é focado no período fixo de um ano (ano fiscal), enquanto o CfaR é mais flexível e pode ser computado para qualquer data ou período específico.

Outra diferença é que o BaR possui um valor de referência exogenamente determinado, o limite orçamentário, ao passo que o CfaR provê, para um dado nível de significância, o valor máximo esperado para o fluxo de caixa em determinada data.

O constante monitoramento do risco orçamentário é importante porque o gestor, ao perceber a possibilidade de exceder o orçamento, antecipa ou evita uma missão potencialmente desgastante de apresentar ao Congresso um requerimento de créditos suplementares para honrar os compromissos da dívida.

Por fim, o risco do lado da demanda representa o risco de mudanças bruscas na demanda por títulos públicos. O guia condutor principal de mudanças repentinas na demanda desses papéis é a taxa de juros, embora, obviamente, possa haver múltiplas razões para tal. Nesse sentido, parcela expressiva dos demandantes de títulos públicos, em suas estratégias de investimentos, obedece a limites de exposição ao risco de taxa de juros.

Isso implica que o gestor de dívida tem de estar ciente de que há restrições, usualmente de curto prazo, para a transferência de risco de taxa de juros ao mercado. Uma das métricas mais utilizadas pelos participantes do mercado financeiro em termos de sensibilidade de taxa de juros é o *Value-at-Risk (VaR)*. Nas palavras de Silva, Cabral e Baghdassarian (2009, p. 184),

durante momentos de volatilidade, agravando tal situação, o *VaR* pode atingir níveis elevados e levar a operações de *stop-loss* por parte dos investidores. O efeito de tais mudanças na demanda pode ser desastroso à implantação de uma estratégia de dívida.

Logo, o acompanhamento desse indicador por parte do gestor de riscos da dívida pública é justificado e necessário. Os autores ainda lembram que “esse acompanhamento é especialmente importante naqueles países que estão em processo de aumentar o prazo de seus títulos e a participação de títulos prefixados”, como é o caso brasileiro.

Outro instrumento utilizado por muitos gestores de dívida pública é a definição e o contínuo acompanhamento da composição ótima de longo prazo, também conhecida como

benchmark. Tal instrumento consiste em “um conjunto de indicadores de dívida relevantes, tais como composição, duração, perfil da dívida, etc. A ideia é que ele consista em uma meta de longo prazo, representando as preferências da sociedade” (SILVA; CABRAL; BAGHDASSARIAN, 2009, p. 186).

Para Alves e Silva (2009, p. 152), o passo seguinte é

a elaboração de uma estratégia de transição – que se baseia nos objetivos indicados pelo *benchmark*, respeitadas as condições iniciais (isto é, a composição corrente da dívida e sua estrutura de vencimentos), e procura tratar a questão de quão rápida deveria ser a convergência para o perfil de dívida desejado no futuro.

Obviamente essa etapa deve vir acompanhada de perene administração de risco e monitoramento de sua execução.

Nesse sentido, desde 2001 a Secretaria do Tesouro Nacional explicita os objetivos e as metas anuais para o perfil da dívida no Plano Anual de Financiamento (PAF). Como explicam Silva e Medeiros (2009), “mais que uma ferramenta de planejamento, o PAF consolidou-se como um instrumento de ampliação da transparência e da previsibilidade no gerenciamento da dívida pública”. Assim, o PAF fornece análise detalhada do programa de ações da STN, expondo diretrizes e metas a serem observadas pelos gestores da dívida pública federal em cada ano.¹⁷

Entretanto, vale adiantar um aspecto elencado pelos autores (SILVA, MEDEIROS, 2009, p. 115):

tomando por base as estratégias traçadas, o PAF apresenta os valores mínimos e máximos, projetados para o final do período, de cada um dos indicadores (da dívida pública federal) considerados relevantes, expressos na forma de limites indicativos. O documento também apresenta capítulo específico com indicadores dos riscos financeiro e de refinanciamento.

17 Para maiores detalhes sobre o PAF, ver Silva e Medeiros (2009).

Quanto à composição ótima de longo prazo, é importante salientar que esta seja elaborada sob uma visão de gestão integrada de ativos e passivos (ALM). Segundo Silva, Cabral e Baghdassarian (2009, p. 190), “faz todo o sentido para o governo, ao avaliar seus riscos, não gerenciar seus passivos sem levar em conta a estrutura e as características de seus ativos”. Os autores ainda afirmam que o maior ativo de um governo é a capacidade, ou o direito, de cobrar impostos. Assim, “as características dos superávits primários futuros torna-se um fator-chave para a determinação da estrutura ótima da dívida”.

Compreendido até aqui a relevância da gestão da dívida para a economia, quais objetivos e diretrizes ela deve perseguir e o que se espera da boa gestão de riscos, é momento de averiguar como esses aspectos têm evoluído na economia brasileira, no que se refere à DPF. Esse assunto é apresentado na próxima seção.

2.5 A centralização da gestão da dívida pública

As seções anteriores descreveram as funções, os cuidados e os dilemas presentes na atuação dos gestores de dívida pública. Esta seção ilustra como um desenho institucional adequado favorece a gestão eficiente da dívida pública e, em última medida, o próprio gasto público.

Borrensen e Cosío-Pascal (2002) elencam três modelos básicos de estruturas organizacionais de administração da dívida pública em que os agentes principais podem ser: (i) o Banco Central; (ii) o Ministério das Finanças (ou Fazenda/Economia); ou (iii) um Departamento de Dívida Pública (na literatura internacional, *Debt Management Office* – DMO).

Os autores afirmam que o arranjo onde o Banco Central é o agente principal na gestão da dívida acontece principalmente quando o gerenciamento de dívida está fortemente integrado com operações monetárias, ou quando o Banco Central é o responsável pelo desenvolvimento e bom funcionamento do mercado de títulos públicos. Nesse caso, a autoridade monetária pode ser responsável, dentre outras funções e dentro de parâmetros bem especificados, pela estratégia e pela administração de curto prazo da dívida pública.

No segundo modelo o Ministério das Finanças é responsável, entre outras atribuições, pelo planejamento tático e estratégico da dívida pública. As funções de administração da dívida são geralmente

realizadas por setores dentro do Ministério, por exemplo, um Departamento ou Secretaria do Tesouro. Esse é o modelo brasileiro.

Por fim, a criação de um DMO como agente principal representa maior separação institucional entre as políticas fiscal, monetária e de gestão da dívida pública. Nessa estrutura, a agência pauta sua atuação nas diretrizes de longo prazo definidas pelo Ministério das Finanças e conta com maior liberdade e autonomia para a gestão técnica da dívida. Borrensen e Cosío-Pascal (2002) lembram ainda que em geral, neste tipo de arranjo, o Banco Central está bem mais distante das questões relacionadas à dívida.

Apesar disso, há boas razões para a adoção do modelo em que a gestão da dívida é responsabilidade do Ministério das Finanças, tal qual é feito no Brasil. A principal delas refere-se à importância de se manter relacionamento estreito com setores do governo ligados à execução orçamentário-financeira e à condução da política fiscal. Desta forma, aproveitam-se as sinergias inerentes à coordenação entre essas áreas. De fato, é possível manter esse arranjo sem que isso redunde em perda de autonomia tanto para o planejamento quanto para a execução, em níveis tático e estratégico.

Nyawata (2012) defende a centralização da gestão da dívida pública no Ministério das Finanças em detrimento do Banco Central. Isso porque conflitos potenciais podem surgir de acordo com os arranjos institucional, legal e administrativo pertinentes às decisões sobre perfis de maturidade, restrições sobre os volumes e a falta de acordo formal entre o Tesouro Nacional e o Banco Central. Para o autor, essa defesa fica ainda mais forte quando se consideram três aspectos: (i) visão integrada do financiamento do setor público; (ii) política pública de fomento ao mercado financeiro; e (iii) externalidades positivas que títulos públicos exercem sobre os demais instrumentos financeiros e o restante da economia.

Para Nyawata (2012), se a autoridade monetária for responsável pela totalidade ou parte da emissão de títulos públicos, pode-se gerar conflito com os objetivos do governo na gestão da dívida pública. E, no caso de as duas instituições emitirem títulos públicos, abre-se espaço para a segmentação de mercado, com disputa por investidores entre o Banco Central e o gestor da dívida pública, especialmente se os vencimentos dos títulos forem semelhantes.

Além disso, a existência de dois emissores de dívida soberana pode prejudicar a liquidez, bem como potencial corrida para títulos do Banco Central em detrimento dos títulos do Tesouro Nacional em momentos de maior instabilidade financeira. Desta forma, Nyawata (2012) defende que a autoridade monetária não emita títulos próprios e faça uso dos papéis de dívida emitidos pelo governo na condução da política monetária.

Outro desenho possível, não contemplado no trabalho de Borrensen e Cosío-Pascal (2002), é permitir que cada ente de governo crie e administre sua própria dívida de acordo com sua necessidade e conveniência. Por exemplo, cada um dos ministérios de determinado governo emita e administre títulos públicos para o financiamento de suas necessidades. O ponto fraco desse desenho é justamente a incapacidade de se ter uma visão integrada e eficiente do financiamento do setor público, o que pode gerar disputa por investidores entre os diferentes entes públicos. Uma consequência previsível desse cenário é o maior custo endividamento para o governo, bem como a dificuldade de se manter uma percepção coesa do risco soberano brasileiro por parte de investidores internos e externos.

2.6 Evidências empíricas da contribuição da gestão da dívida brasileira ao equilíbrio fiscal

A gestão da DPF conquistou na última década importantes melhoras. A redução da vulnerabilidade foi o avanço mais expressivo em termos de gestão de dívida pública, e esta foi alcançada, entre outros fatores, por meio da contínua redução do volume de dívida externa, da redução expressiva da dívida interna indexada ao câmbio, do alongamento da maturidade, do aumento da participação de títulos prefixados e indexados à inflação, e da busca de maior diversificação na base de investidores.

Particularmente no caso brasileiro, as receitas estão fortemente ligadas à dinâmica do mercado doméstico, de forma que os recursos oriundos da atividade exportadora e das reservas internacionais acumuladas são capazes de oferecer proteção cambial somente até determinado montante. Nesse contexto, os títulos prefixados carregam o menor grau de risco para o emissor, porque além de

ser denominado em moeda doméstica, no ato da emissão do título já se sabe previamente o seu custo total e o valor dos juros e do principal a serem pagos em cada momento. Embora eles aumentem a previsibilidade dos custos de financiamento e dos fluxos de pagamento, uma vez contratados não variam conforme a dinâmica de arrecadação mais à frente, pois estão fixados. Obviamente, como dito anteriormente, a este menor grau de risco para o emissor corresponde um custo maior de emissão e um prazo de vencimento potencialmente menor, reflexo do *trade-off* entre custo e risco.

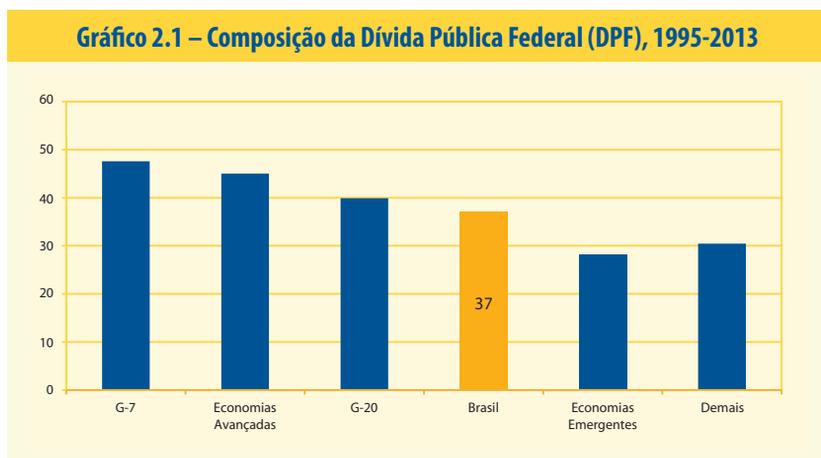
A segunda categoria com menor risco para o governo brasileiro são os títulos indexados a índices de preços. Também são denominados em moeda local, assim como a maior parte dos recursos arrecadados, e a mecânica desse tipo de papel é a seguinte: o investidor é remunerado por uma taxa de juros real prefixada mais o valor correspondente à inflação no período. Considerando a existência de correlação positiva entre as receitas governamentais e a inflação, eles costumam oferecer boa proteção às finanças públicas. Nesse caso em particular, apesar do risco mais reduzido, o prazo de vencimento desse título tende a ser maior inclusive que os com taxas flutuantes e cambiais. Já seu custo de emissão tende a ser superior ao de ambos.

Já os títulos indexados às taxas de juros flutuantes, apesar de denominados em reais, embutem um risco maior, dado que a correlação entre as taxas de juros de curto prazo e a arrecadação governamental é inversa. Por fim, o título mais arriscado, no caso brasileiro, é o atrelado ao câmbio. A razão está no fato de o seu custo ser determinado pelo comportamento da variação de uma moeda estrangeira, que não está sob estrito controle do governo brasileiro, podendo gerar muita vulnerabilidade às finanças públicas. Contudo, esse é o tipo de dívida com menor exigência de custo pelo investidor, pois é o título, entre todos os citados, com maior correlação com as taxas de juros internacionais, dada a sua denominação em moeda estrangeira. Além disso, tais títulos possuem prazos de vencimentos bem maiores, se comparados com os demais.

Os frutos naturais desse processo foram a redução nos riscos de mercado e de refinanciamento da dívida, com efeitos benéficos sobre o custo da dívida, promovendo maior previsibilidade ao orçamento e gerando economia de recursos para o governo.

A chancela desse processo, por parte das finanças internacionais, se deu com a conquista da classificação de risco conhecida como *grau de investimento* (*investment grade*) em 2008, concedida pelas agências de classificação de risco Fitch e Standard & Poor's.

Alguns números podem contextualizar esse ponto. O Gráfico 2.1 ilustra bem a melhora na composição da DPF nas últimas décadas, por meio da redução da parcela indexada ao câmbio ou à taxa flutuante e aumento da parcela prefixada ou indexada a índices de preços. No início de 2001, os títulos atrelados ao câmbio e à taxa flutuante representavam 83,4% do estoque da DPF, ao passo que em dezembro de 2013 essa parcela se reduziu para 23,4%. No mesmo período, a participação dos títulos prefixados e os indexados a índices de preços aumentou de 16,6% do total da DPF para 76,5%.



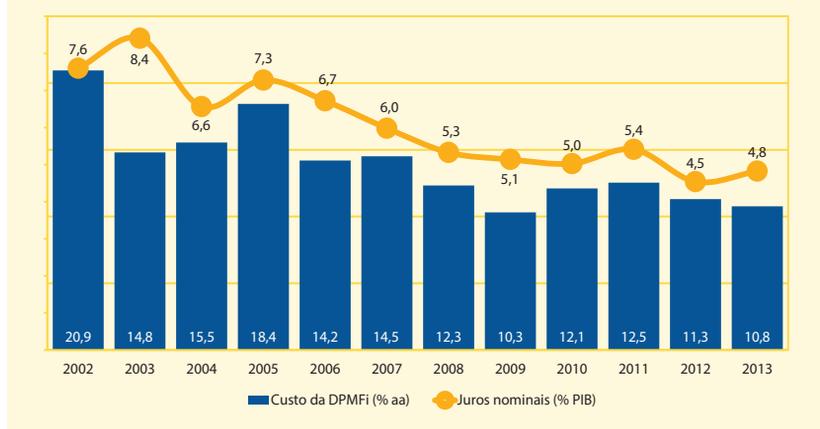
Fonte: Secretaria do Tesouro Nacional.

Em termos de custo médio mensal, em 2002 o custo da Dívida Pública Mobiliária Federal Interna (DPMFi)¹⁸ era de 20,9% ao ano, ao passo que em dezembro de 2013 caiu para 10,8%, como podemos ver no Gráfico 2.2. Tamanha redução no custo da dívida pública

¹⁸ A DPF é composta pela Dívida Pública Mobiliária Federal interna (DPMFi) e pela Dívida Pública Federal externa (DPFe). A DPMFi é a parcela mais expressiva da DPF, representando cerca de 95% da DPF. Embora a série histórica mais longa de custo da dívida, iniciada em 1997, refira-se apenas à DPMFi, entendemos que, dada a sua grande representatividade dentro da DPF, não há prejuízo na análise da evolução do custo da dívida total utilizando a DPMFi como referência.

contribuiu para que a apropriação de juros nominais na Dívida Líquida do Setor Público Consolidado (DLSP) caísse de 7,6% do PIB em 2002 para 4,8% do PIB em 2013.

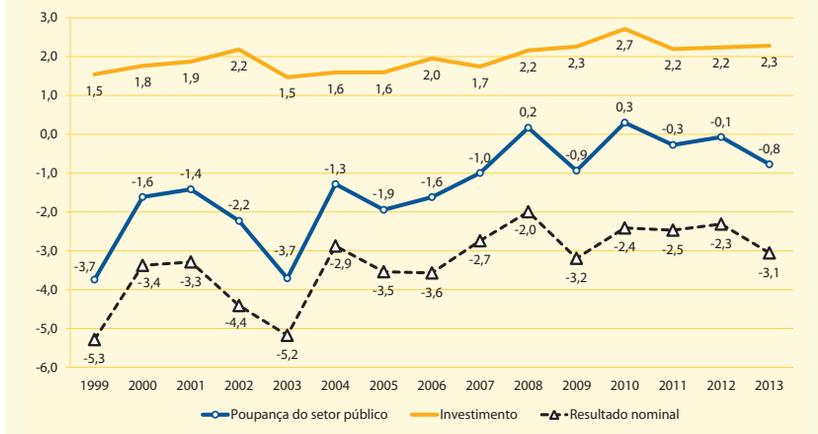
Gráfico 2.2 – Custo da DPMFi e gastos com juros nominais do Setor Público Consolidado, 2002-2013 (% do PIB)



Fonte: Secretaria do Tesouro Nacional (Custo da DPMFi) e Banco Central do Brasil (Juros Nominais).

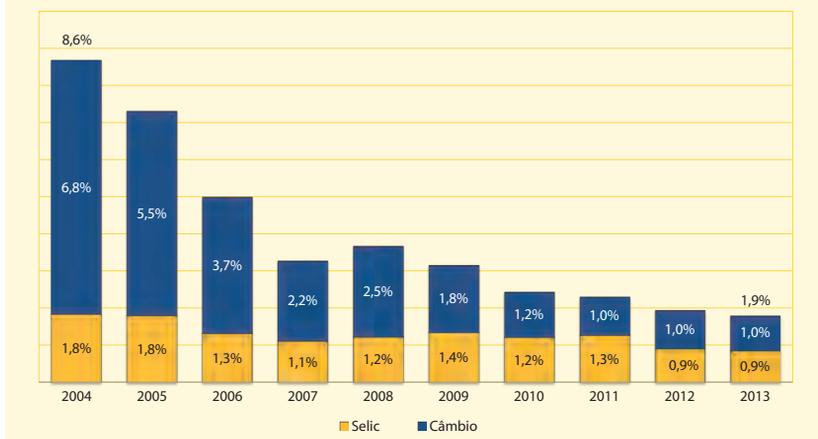
Como consequência da redução na conta de juros nominais e da realização de consistentes superávits primários anuais pelo setor público, o resultado nominal também melhorou: de um déficit de 4,5% do PIB em 2002, o setor público passou a apresentar um déficit de apenas 3,1% do PIB em 2013. Com isto, a poupança pública¹⁹ melhorou consideravelmente nesse período, permitindo o aumento do investimento do setor público, como se pode ver no Gráfico 2.3.

19 A poupança pública ($S_G = T - G$) é calculada pela diferença entre receitas correntes e despesas correntes (que inclui juros). Como ela não inclui o investimento público (I_G), que é uma despesa primária, se quisermos calcular a poupança pública a partir do resultado nominal (RN), devemos fazer: $S_G = RN + I_G = RP - J + I_G$, em que RP é o resultado primário do setor público e J é o juro nominal. Por exemplo, para um déficit nominal de 2% do PIB e um investimento governamental de 3% do PIB, teríamos uma poupança pública de 1% do PIB.

Gráfico 2.3 – Evolução da poupança do setor público (% do PIB), 1999-2013


Elaboração dos autores. Fontes: Dados disponibilizados por MF/SPE e Ipea (Investimento) e Banco Central do Brasil (Resultado Nominal).

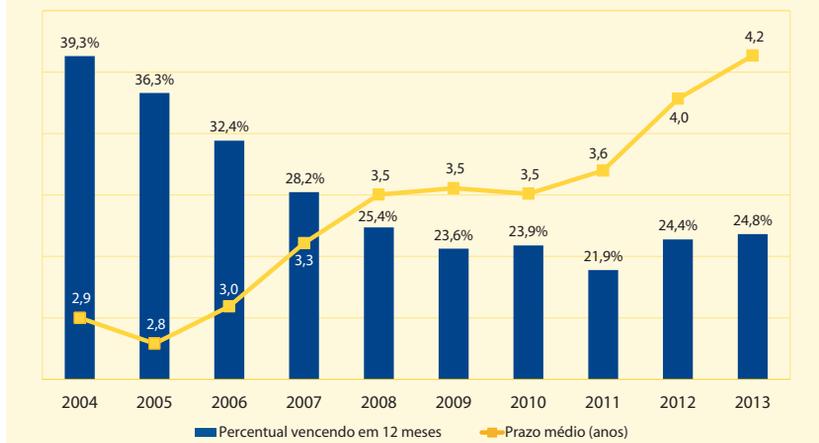
O Gráfico 2.4 mostra a evolução do risco de aumento no estoque da DPF em situações de grandes e persistentes turbulências. Define-se estresse, nesse exercício, por um choque de 3 desvios padrão sobre a média da taxa de juros Selic real e da desvalorização cambial real acumuladas em 12 meses, aplicado sobre as parcelas do estoque da DPF remuneradas por taxas de juros flutuantes ou pela variação cambial. Aplica-se o cenário de estresse para o período de um ano sobre a dívida em Selic e, instantaneamente, para a correção da dívida cambial. De 2004 a 2013, o impacto de um choque de proporções extremas nos juros e no câmbio diminuiu de 8,6% do PIB ao final de 2004 para apenas 1,9% do PIB ao final de 2013. Ou seja, devido à mudança no perfil da dívida neste período, houve uma redução de 77,9% na sensibilidade do estoque da DPF à choques de taxa de juros e câmbio.

Gráfico 2.4 – Impacto de um cenário de estresse sobre o estoque da DPF, 2004-2013


Fonte: Secretaria do Tesouro Nacional.

Outro aspecto significativo para o gestor da dívida diz respeito à estrutura de vencimentos, que determina o risco de refinanciamento da dívida. Em termos de maturidade, também se verificou melhoras: em 2004, a DPF exibia um prazo médio de 2,9 anos, elevando-se para 4,1 anos em 2013. Já o percentual da DPF que vence em 12 meses caiu consideravelmente, de 39,3% para 24,8%, como podemos ver no Gráfico 2.5, a seguir.

Gráfico 2.5 – Indicadores de maturidade da DPF, 2004-2013



Fonte: Secretaria do Tesouro Nacional. Relatório Mensal da Dívida Pública Federal.

Todos esses indicadores apresentados permitem concluir que a gestão da DPF nos últimos anos levou a um perfil de endividamento que tem trazido menores riscos às contas públicas, menores custos e conseqüente melhora na classificação de risco atribuída pelas agências de *rating*. Em paralelo, houve redução do gasto com juros nominais do setor público, contribuindo para a consolidação fiscal brasileira.

2.7 Considerações finais

Ao longo deste capítulo ficou claro que a gestão da DPF é relevante no que diz respeito à conquista de um melhor perfil do endividamento público brasileiro, abrindo espaço para que os gastos com juros sejam reduzidos e, conseqüentemente, liberando recursos no médio prazo para outros projetos governamentais.

Ainda mais, reconhecendo que a tese da equivalência ricardiana pressupõe hipóteses muito restritas, foi mostrado que, em um ambiente onde estas são relaxadas, a dívida pública é capaz de suavizar a carga tributária, gerando maior bem-estar para a sociedade. Também ficou evidente que a boa gestão da dívida pública, preocupada em buscar uma composição ótima e com o acompanhamento de instrumentos de riscos relacionados ao seu gerenciamento, é capaz de trazer maior estabilidade ao orçamento governamental.

Concluiu-se que a centralização da gestão da dívida pública em um único departamento é benéfica, pois as decisões implícitas à gestão da dívida envolvem análises profundas de riscos, custos e prazos de acordo com as particularidades de cada governo. Assim, é necessário o fomento da capacidade técnica, eficiente estrutura institucional e centralização da gestão.

REFERÊNCIAS

- ALVES, L. F. *Composição ótima da dívida pública brasileira: uma estratégia de longo prazo*. Brasília: ESAF. Monografia premiada com o terceiro lugar no XIV Prêmio Tesouro Nacional – 2009: Homenagem a Euclides da Cunha. Política Fiscal e Dívida Pública. Brasília (DF), 2009.
- ALVES, L. F.; SILVA, A. C. Planejamento estratégico da dívida pública federal. In: SILVA, A. C; CARVALHO, L. O. de; MEDEIROS, O. L. de (Org.). *Dívida pública: a experiência brasileira*. Brasília: Secretaria do Tesouro Nacional; Banco Mundial, 2009. p. 149-171.
- BARRO, R. J. Are government bonds net wealth? *Journal of Political Economy*, v. 81, p. 1095-1117, 1974.
- _____. On the determination of the public debt. *Journal of Political Economy*, v. 87, p. 940-971, 1979.
- BORRESEN, P.; COSIO-PASCAL, E. *El papel y la organización de una oficina de la deuda*. Nova York; Genebra: Nações Unidas, 2002.
- BRENNAN, H. G.; BUCHANAN, J. M. The logic of the ricardian equivalence theorem. In: BUCHANAN, James M.; ROWLEY, Charles K.; TOLLISON, Robert D. (Ed.). *Deficits*. Oxford: Basil Blackwell, 1986. p. 79-92.
- BUCHANAN, J. M. Barro on the Ricardian equivalence theorem. *Journal of Political Economy*, v. 84, n. 2, p. 337-342, 1976.
- FELDSTEIN, M. Perceived wealth in bonds and social security: a comment. *Journal of Political Economy*, v. 84, n. 2, p. 331-336, 1976.
- GOLDFAJN, I; PAULA, A. Uma nota sobre a composição ótima da dívida pública – reflexões para o caso brasileiro. Rio de Janeiro: Departamento de Economia PUC-Rio, 1999. (Texto para Discussão n. 441).
- HERMANN, J. A macroeconomia da dívida pública: notas sobre o debate teórico e a experiência brasileira recente (1999-2002). *Cadernos Adenauer*, Ano III, n. 4, p. 41-70, nov. 2002.
- LUCAS JR., R. E.; SARGENT, T. J. After keynesian macroeconomics. In: _____. *After the phillips curve: persistence of high inflation and high unemployment*. Boston: Federal Reserve Bank of Boston, 1978.
- LUCAS JR., R. E.; STOKEY, N. L. Optimal fiscal and monetary policy in an economy without capital. *Journal of Monetary Economics*, v. 12, p. 55-94, 1983.

MISSALE, A. The effects of debt management. In: _____. *Public Debt Management*. Oxford University Press, 1999. (Capítulo 2; p. 10-28).

MODIGLIANI, F.; BRUMBERG, R. Utility analysis and the consumption function: an interpretation of cross section data. In: KURIHARA, K. (Org.) *Post-keynesian economics*. New Jersey: Rutgers University Press, 1954. p. 388-436.

MUTH, J. F. Rational expectations and the theory of price movements. *Econometrica*, n. 29, p. 315-335, jul. 1961.

NYAWATA, O. Treasury bills and/or central bank bills for absorbing surplus liquidity: the main considerations. *IMF Working Paper* n. 12/40, 2012.

RICARDO, D. Funding system. In: SRAFFA, P. (Ed.). *The works and correspondence of David Ricardo: pamphlets and papers 1815-1823*. Cambridge: Cambridge University Press for the Royal Economic Society, 1820. v. IV.

_____. *On the principles of political economy and taxation*. In: SRAFFA, P. (Ed.). *The works and correspondence of David Ricardo: on the principles of political economy and taxation*. 3. ed. Cambridge: Cambridge University Press for the Royal Economic Society, 1821. v. I.

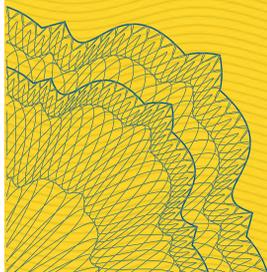
SECRETARIA DO TESOUREIRO NACIONAL. *Dívida pública federal: Plano Anual de Financiamento 2012*. Brasília, 2012. (n. 2).

SILVA, A. C.; CABRAL, R.; BAGHDASSARIAN, W. Gerenciamento de riscos da dívida pública federal. In: SILVA, A. C; CARVALHO, L. O. de; MEDEIROS, O. L. de (Org.). *Dívida pública: a experiência brasileira*. Brasília: Secretaria do Tesouro Nacional; Banco Mundial, 2009. p. 173-218.

SILVA, A.; MEDEIROS, O. L. Conceitos e estatísticas da dívida pública. In: SILVA, A. C; CARVALHO, L. O. de; MEDEIROS, O. L. de (Org.). *Dívida pública: a experiência brasileira*. Brasília: Secretaria do Tesouro Nacional; Banco Mundial, 2009. p. 101-126.

TOBIN, J. Government deficits and capital accumulation. In: TOBIN, J. *Asset accumulation and economic activity*. Oxford: Basil Blackwell, 1980. p. 49-72.

TOGO, E. Coordinating public debt management with fiscal and monetary policies: an analytical framework. *World Bank Policy Research Working Paper*, n. 4.369, 2007.



Capítulo 3

Composição ótima do gasto público para o crescimento econômico

Ana Carolina Giuberti
Fabiana Rocha

3.1 Introdução

O impacto da política fiscal sobre o crescimento de uma economia pode ser medido de duas maneiras. A primeira avalia a produtividade e a eficiência do setor público que são, em última instância, aspectos que afetam o crescimento. Nessa vertente da literatura estão os trabalhos que medem o desempenho do setor público e sua eficiência a partir das relações entre produto e insumo. O restante deste livro, portanto, trata dessas questões.

A segunda forma analisa a relação entre o nível e a composição do gasto público e o crescimento econômico. O foco inicial e que recebeu maior atenção da literatura foi o impacto do tamanho do governo (medido por variáveis fiscais agregadas, como o gasto total) sobre o crescimento. No entanto, os trabalhos empíricos não encontraram evidência de uma relação positiva entre gasto agregado e crescimento da economia, pelo contrário, alguns estudos reportaram uma relação negativa e significativa entre estas variáveis – por exemplo, Barro e Sala-i-Martin (1995) e Fölster e Henrekson (1997) – e outros, uma relação negativa, mas não significativa – por exemplo, Heitger (2001), Bassanini, Scarpett, Hemmings (2001).

Tais resultados levaram a literatura sobre o tema a destacar a importância da análise desagregada, uma vez que nem todos os tipos de gastos têm o mesmo efeito sobre o crescimento: a composição da despesa ganhou, portanto, relevância nesses estudos. Dois fatores justificam sua importância: primeiro, enquanto o tamanho do governo é uma questão de escolha pública, sua composição é objeto de disputa política. A distinção entre gastos públicos *produtivos* e *improdutivos* e o estabelecimento de como um país pode melhorar

seu desempenho econômico, mudando a combinação entre os dois, pode ajudar na acomodação dessa disputa. Segundo, depois de sucessivos ajustes fiscais, a questão difícil é onde fazer cortes fiscais adicionais. Que componente do gasto público deve ser cortado? A resposta deve obviamente depender, entre outras coisas, da contribuição que cada um dos componentes do gasto tem para o crescimento econômico.

Estudos mais recentes no tema passaram a reavaliar a relação entre tamanho/composição fiscal e crescimento sob a ótica da eficiência. Eles diferem, dessa forma, do restante da literatura empírica de crescimento ao explicitamente considerarem que essa relação depende da eficiência do setor público, combinando, dessa forma, as duas maneiras de medir o impacto da política fiscal sobre o crescimento presentes na literatura.

Assim, o objetivo deste capítulo é apresentar e discutir a literatura acerca do impacto do gasto público sobre o crescimento econômico de longo prazo, com foco na composição das despesas. Para tanto, está dividido em cinco seções, além desta introdução. A próxima seção apresenta a literatura teórica sobre o tema. A seção três traz as evidências empíricas tanto para a economia brasileira quanto as apresentadas pela literatura internacional. A seção quatro discute os trabalhos mais recentes que relacionam gasto público, eficiência e crescimento econômico, e a quinta e última seção apresenta as conclusões.

3.2 A literatura teórica

Desde a década de 1960, a relação entre gasto público e crescimento da economia é estudada no *mainstream* econômico por pesquisadores: primeiro, na linha dos modelos neoclássicos de crescimento, cujo expoente maior é o modelo de Solow (1956), seguido, a partir de meados da década de 1980, pelos modelos neoclássicos de crescimento endógeno, quando o debate ganhou impulso, em particular, com a publicação do artigo de Barro (1990).

Antes de apresentar os resultados dessa literatura, é interessante ressaltar as diferenças no processo de crescimento de longo prazo entre os dois modelos e, conseqüentemente, as diferenças no impacto do gasto público nesse processo. O modelo de Solow parte de uma função de produção com retornos constantes de

escala e retornos marginais decrescentes à acumulação dos fatores de produção – capital e trabalho. O principal resultado do modelo, centrado nessas características da função de produção, é que a taxa de equilíbrio de crescimento do produto por trabalhador¹ no longo prazo é determinada pela taxa de progresso técnico, exógena ao modelo.² Fatores que afetam a acumulação de capital, como a taxa de poupança, modificam apenas o patamar do equilíbrio, i.e., o nível de produto e capital por trabalhador, mas não a taxa de crescimento neste ponto. Assim, um aumento na taxa de poupança da economia, que eleva a taxa de investimento e resulta em um nível de produto maior por trabalhador, proporciona uma taxa de crescimento superior à taxa de progresso técnico apenas durante a transição para o novo ponto de equilíbrio. Uma vez alcançado o novo estado estacionário, é a taxa de progresso técnico que determina o crescimento de longo prazo do produto por trabalhador.

A consequência desse modelo para os efeitos da política fiscal é direta: políticas que afetam a taxa de poupança da economia, como variações no gasto público, bem como na dívida pública, ou alterações no sistema tributário, determinam o nível de produto no equilíbrio e afetam a taxa de crescimento durante o período de transição, mas não promovem maior crescimento da economia no *steady-state*.

A inovação presente nos modelos de crescimento endógeno, que os diferencia desses primeiros modelos neoclássicos de crescimento, é a hipótese de que a função de produção apresenta retornos marginais não decrescentes à acumulação de capital. O capital é entendido em sentido amplo nesta literatura, e inclui o capital humano e o estoque de conhecimento da economia. Em geral, os modelos consideram que o capital, ou um subgrupo desse capital, apresenta retornos

1 Os modelos de crescimento, tanto neoclássicos quanto os de crescimento endógeno, reportam seus resultados de longo prazo, quando a economia alcança seu estado estacionário ou de equilíbrio (*steady-state*), em termos da taxa de crescimento do produto por trabalhador ou da taxa de crescimento do produto *per capita*, sob a hipótese de que a força de trabalho é igual à população. No que se segue, o termo por trabalhador/*per capita* será suprimido para evitar constantes repetições no decorrer do texto, mas a taxa de crescimento mencionada será sempre a taxa por trabalhador/*per capita*. Do mesmo modo, as expressões longo prazo, estado estacionário e equilíbrio não serão constantemente mencionadas, mas a análise da taxa de crescimento refere-se à taxa de longo prazo, exceto quando mencionado o contrário.

2 No equilíbrio, a taxa de crescimento do produto agregado é determinada pelo progresso técnico e pela taxa de crescimento da população, assumindo-se que esta é igual à força de trabalho.

constantes de escala. A formulação tradicional considera o produto como uma função linear do estoque de capital, $y=Ak$, onde y e k estão medidos em unidades por trabalhador. Outra estratégia é considerar a existência de externalidades, *spillover* e bens públicos com os retornos constantes de escala nos fatores acumuláveis, que tornam o retorno social à acumulação de capital não decrescente, ainda que o retorno privado seja decrescente. Com isso, a taxa de equilíbrio de crescimento do produto por trabalhador no longo prazo é determinada endogenamente pelos modelos, e os fatores que afetam a acumulação de capital, ou seja, que afetam as decisões de poupar e investir dos agentes, irão produzir efeitos duradouros sobre a taxa de crescimento no estado estacionário. Consequentemente, a política fiscal passa a influenciar tanto o nível de produto quanto a taxa de crescimento da economia no longo prazo. A direção e a intensidade desse efeito variam de acordo com o modelo e as hipóteses adotadas.

Na linha dos modelos neoclássicos de crescimento, Carboni e Medda (2011) estudam o tamanho ótimo do governo e a composição do gasto público que maximizam a taxa de crescimento da economia na transição para o novo estado estacionário e o nível de produto *per capita* no longo prazo. Seguindo Arrow e Kurz (1969), os autores consideram toda a despesa pública como um processo cumulativo destinado a criar capital público produtivo, capital este considerado como insumo na função de produção do setor privado. O resultado encontrado é uma relação não monotônica entre o nível de produto no estado estacionário e o tamanho do governo: se, a princípio, um aumento do gasto público total eleva o nível de produto no equilíbrio, uma vez que esse gasto compõe o capital público, utilizado como insumo na função de produção, a partir de determinado ponto, definido pelos parâmetros do modelo, o aumento passa a ter efeitos deletérios sobre o produto. Do mesmo modo, uma mudança na composição do gasto pode elevar o nível de produto de equilíbrio, se forem elevadas as despesas com maior elasticidade na função de produção. No entanto, esse processo de realocação de despesa para aumentar o nível de produto é limitado, uma vez que o capital público possui retornos decrescentes de escala. De modo semelhante, os autores mostram que os efeitos do tamanho do governo e da composição do gasto sobre o crescimento econômico no período de transição também são não monotônicos;

os efeitos são positivos até o ponto em que o produto marginal do capital público é igual ao custo marginal de financiá-lo, dado pela arrecadação de impostos.

Este *trade-off* entre o benefício de um maior gasto público e o seu custo foi primeiro apontado por Barro (1990), em seu artigo seminal, ao incorporar o setor público no modelo de crescimento endógeno pela inclusão dos serviços públicos (gasto) como insumo na função de produção dos agentes da economia, função esta que apresenta retornos constantes de escala para capital e gasto público juntos, mas retornos decrescentes para o capital privado. Como o gasto público amplia a produtividade do capital no setor privado, é denominado gasto produtivo. O modelo assume ainda que não há externalidades associadas ao uso dos serviços públicos e que os gastos realizados são financiados por um imposto constante sobre a renda, de modo que o orçamento está sempre equilibrado; o governo não incorre em déficits ou superávits. O resultado é que a taxa de crescimento do consumo e, portanto, a taxa de crescimento da economia no estado estacionário, depende da alíquota do imposto de renda e do tamanho do governo, medido pela razão gasto produtivo/produto. Assim, um aumento no tamanho do governo, por meio do aumento do gasto público, apresenta dois efeitos sobre a taxa de crescimento. O primeiro efeito é o aumento da razão gasto produtivo/produto, que por sua vez eleva a produtividade do capital no setor privado e, portanto, o crescimento da economia. O segundo efeito deriva da hipótese de orçamento equilibrado: o aumento no gasto do governo deve ser financiado pelo aumento no imposto de renda, que por sua vez é um imposto distorcivo, o que reduz a taxa de crescimento no longo prazo. Enquanto o primeiro efeito dominar o segundo, o efeito total será o aumento no crescimento econômico de longo prazo. Isso implica que há um tamanho ótimo para o governo além do qual os aumentos nos gastos reduzem a taxa de crescimento de longo prazo, ainda que estes gastos aumentem a produtividade do capital privado, pois são financiados por aumentos nos impostos que distorcem a decisão entre consumo e poupança das famílias.

Barro (1990) considera ainda uma extensão do modelo citado anteriormente, na qual, além da presença do gasto público na função de produção, inclui o consumo do governo como argumento da

função de utilidade de cada família. Mantida a hipótese de orçamento equilibrado, o novo gasto deve ser financiado por um imposto sobre a renda maior. A consequência é que a taxa de crescimento de longo prazo, neste caso, é menor do que a taxa obtida no modelo em que o gasto é todo alocado na função de produção. Por suas características e resultados, esse tipo de gasto público ficou consagrado na literatura como gasto improdutivo. O efeito negativo do consumo do governo sobre o crescimento de longo prazo está diretamente relacionado à forma de financiamento desse gasto, por meio de um imposto distorcivo. O financiamento de forma não distorciva, por um imposto *lump-sum*, tornaria o efeito do consumo sobre o crescimento neutro. Considerando que o modelo não inclui a escolha entre trabalho e lazer dos indivíduos, o imposto sobre consumo teria o mesmo resultado que um imposto tipo *lump-sum* no modelo.

De modo geral, podem-se distinguir quatro tipos de impacto da política fiscal na economia a partir do modelo de Barro. Esses resultados dependem da combinação entre o tipo de gasto realizado e sua forma de financiamento, considerando a restrição do orçamento equilibrado. Assim, gastos produtivos financiados por impostos não distorcivos afetam positivamente o crescimento econômico, de modo inequívoco. Já gastos produtivos financiados por impostos distorcivos possuem efeito positivo sobre o crescimento quando o tamanho do governo for menor do que o tamanho ótimo, e efeitos negativos quando o tamanho do governo ultrapassar esse limite. Por outro lado, gastos com consumo do governo financiados por impostos não distorcivos não afetam a taxa de crescimento, enquanto o consumo do governo financiado por impostos distorcivos deprime o crescimento de longo prazo.

O modelo acima se tornou base para diversos trabalhos na literatura de crescimento endógeno, como o de Devarajan, Swaroop e Zou (1996), que estudam o efeito da composição do gasto público sobre o crescimento de longo prazo da economia. No entanto, diferentemente de Barro (1990), que distingue entre dois tipos de gastos públicos, produtivos e improdutivos – de acordo com a forma em que esses gastos são alocados no modelo, se na função de produção ou na função de utilidade, respectivamente –, Devarajan, Swaroop e Zou (1996) supõem a presença de dois tipos de gastos públicos, produtivos e improdutivos, na função de produção agregada

da economia. Seguindo Barro (1990), os autores consideram que as despesas são financiadas por um imposto constante sobre a renda, de modo que o orçamento está sempre equilibrado. O resultado é que a taxa de crescimento da economia depende da composição entre os dois tipos de gastos. A classificação do gasto em produtivo ou improdutivo, por sua vez, é determinada endogenamente pelo modelo: define-se gasto produtivo aquele cujo aumento na participação orçamentária eleva a taxa de crescimento da economia no estado estacionário. Este resultado depende não apenas da elasticidade do produto em relação aos dois tipos de gastos incluídos no modelo, como também da participação inicial de cada gasto no orçamento. Nesse caso, não é necessário que um dos dois tipos de despesa apresente elasticidade negativa (seja substituto à produção privada), para que seja feita a classificação entre gasto produtivo e improdutivo. Ambos os componentes podem ser complementares à produção privada, no entanto, o que apresentar maior elasticidade é considerado produtivo em relação ao outro. Além disso, para que aumentos em gastos considerados produtivos elevem a taxa de crescimento do produto, é necessário que a sua participação no orçamento não esteja em um patamar muito elevado.³

Nota-se que os resultados apresentados por Carboni e Medda (2011), no que tange aos efeitos da composição do gasto público sobre o crescimento, assemelham-se aos resultados de Devarajan, Swaroop e Zou (1996): em ambos os casos, há uma composição ótima para o gasto público. No entanto, enquanto no primeiro artigo os efeitos são limitados ao crescimento no período de transição para um novo estado estacionário da economia, no segundo, os efeitos sobre a taxa de crescimento do produto são permanentes.

3.3 A literatura empírica

Embora a relação entre gasto público e crescimento esteja definida na literatura teórica, do ponto de vista empírico não há um consenso sobre os efeitos da política fiscal sobre o crescimento de longo prazo. Tal questão fica evidente em diversos trabalhos que trazem uma revisão dessa literatura, por exemplo, os artigos de Easterly e Rebelo

3 Em termos dos parâmetros do modelo, é necessário que a participação no orçamento seja menor do que a razão das elasticidades.

(1993), de Levine e Renelt (1992), o artigo de Nijkamp e Poot (2004) e os trabalhos de Kneller, Bleaney e Gemmel (1998) e Gemmell e Kneller (2001).

Easterly e Rebelo (1993) argumentam que os resultados das variáveis fiscais nos modelos empíricos de crescimento dependem do conjunto de variáveis de controle e de condições iniciais. Usando um conjunto de regressões em *cross-section*, os autores encontram evidências de que o investimento público em transporte e comunicação em países em desenvolvimento leva a um maior crescimento econômico. Para outras categorias de gasto público as evidências não são, de fato, conclusivas, se eles constituem gasto *produtivo*.

Nijkamp e Poot (2004) apresentam fortes evidências neste sentido a partir de uma meta-análise estatística para uma amostra de 123 estudos de caso publicados entre 1983 e 1998. Em geral, considera-se que o consumo do governo, a tributação e os gastos com defesa são prejudiciais ao crescimento econômico, enquanto os gastos com educação e infraestrutura apresentam efeito positivo. No entanto, os autores mostram que as evidências com relação a esses efeitos esperados são fracas, embora a importância comum atribuída aos gastos com infraestrutura e educação seja confirmada, e que os resultados sejam sensíveis à metodologia utilizada. A análise do tópico mais estudado no que tange ao impacto da política fiscal sobre o crescimento do produto – o efeito do tamanho do governo – deixa clara essa questão: estudos de *cross-section* apresentam maior probabilidade de encontrar um efeito negativo do tamanho do governo sobre o crescimento do que modelos de painel. Ademais, modelos de *cross-section* que excluem a renda inicial como variável explicativa apresentam menor probabilidade de encontrar um efeito significativo da política fiscal do que os modelos que incluem essa variável; e regressões convencionais de crescimento parecem ser menos informativas do que outros métodos, dado o grande percentual de resultados inconclusivos.

Gemmell e Kneller (2001) fazem uma interessante revisão dessa literatura dividindo os trabalhos empíricos em três gerações. A primeira geração de estudos é anterior aos modelos de crescimento endógeno e testa uma variedade de hipóteses *ad hoc* acerca do impacto da atividade do governo no crescimento da economia. Os autores

destacam que devido à má especificação dos modelos, à limitação dos dados e a falhas nas técnicas econométricas utilizadas, os resultados encontrados não são robustos nem comparáveis. A segunda geração de trabalhos sobre os impactos da tributação e da despesa pública no crescimento de longo prazo da economia tem por base os modelos neoclássicos de crescimento e/ou os modelos de crescimento endógeno. As metodologias econométricas aplicadas nesses trabalhos tendem a ser mais sofisticadas do que as utilizadas anteriormente, com a estimação, em alguns casos, de modelos por variáveis instrumentais para lidar com o possível problema de endogeneidade e causalidade reversa entre o crescimento do produto e as variáveis fiscais, despesa e tributação. Não obstante, os resultados permanecem não robustos. Apenas para destacar algumas das diferentes conclusões apresentadas pelos trabalhos desta geração, cita-se, de um lado, o artigo de Fölster e Henrekson (1997) que reporta efeitos negativos e significantes da despesa total sobre o crescimento da economia, e de outro, o trabalho de Agell, Lindth e Ohlsson (1999), segundo o qual a despesa total não apresenta efeitos significantes. Do mesmo modo, Barro e Sala-i-Martin (1995) concluem que a despesa em educação possui efeito positivo e significativo sobre o crescimento, enquanto Easterly e Rebelo (1993) não encontram evidências que suportem esta conclusão.

Resultados distintos dos esperados teoricamente também são apresentados por Devarajan Swaroop e Zou (1996). A partir do modelo teórico, exposto anteriormente, os autores estimam os efeitos da composição do gasto público sobre o crescimento de longo prazo para um conjunto de 43 países em desenvolvimento ao longo de 20 anos, e reportam que para todas as categorias de despesas tradicionalmente vistas como produtivas, como investimento em capital e despesas com transporte e comunicação, saúde e educação, os coeficientes estimados foram negativos ou insignificantes. Já a despesa corrente, como categoria ampla, considerada improdutivo no modelo de Barro (1990) e nos demais, apresentou impactos positivos no crescimento econômico de longo prazo. Os autores concluem que os investimentos em capital nos países em desenvolvimento devem ter sido excessivos, de modo que estes se tornaram improdutivos na margem. A contrapartida é que os gastos correntes, por terem sido restringidos, revelaram-se produtivos na margem, demonstrando um caso de má alocação dos recursos nesses países.

Um ponto importante desses modelos de segunda geração, enfatizado por Gemmell e Kneller (2001, p. 107), é que a maioria dos trabalhos neste grupo estuda a tributação e o gasto público de forma isolada, ou seja, inclui apenas uma ou outra variável no modelo estimado, e não leva em conta a restrição orçamentária do governo e a forma de financiamento implícito da despesa. Do ponto de vista teórico, Adam e Bevan (2005) concluem que o resultado de uma política fiscal sobre o crescimento da economia depende da forma de financiamento dessa política. Além do resultado teórico, essa questão tem uma importante implicação empírica para a estimação dos efeitos da tributação e do gasto público como argumentam Kneller, Bleanney e Gemmel (1999). Os autores demonstram que, dada a restrição linear entre tributação e despesa imposta pela restrição orçamentária do governo, o coeficiente estimado para cada variável fiscal não deve ser interpretado como o impacto da variação em uma unidade desta variável sobre o crescimento do produto, mas como o impacto desta variável descontado o efeito da variável fiscal omitida, que constitui o elemento financiador implícito para a mudança na variável analisada. Apenas quando o impacto da variável fiscal omitida é igual a zero, ou seja, ela é neutra, é que o coeficiente estimado de fato reflete o efeito da variável estimada sobre o crescimento. Portanto, deve-se evitar o problema de má especificação do modelo omitindo apenas categorias fiscais consideradas neutras. Essa conclusão, no entanto, é contestada por Adam e Bevan (2005, p. 581). Para eles, embora desejável, a hipótese de neutralidade de uma categoria da receita, ou despesa, ou resultado fiscal não é necessária, nem provável, em termos empíricos. Neste caso, os coeficientes estimados devem ser lidos como o efeito de uma variável fiscal particular sobre o crescimento, descontado o efeito da variável omitida.

Com base na questão acima, Gemmell e Kneller (2011) classificam como de terceira geração⁴ os trabalhos que reconhecem de forma

4 Os autores classificam o artigo de Devarajan, Swaroop, Zou (1996) como de terceira geração por estudar a composição do gasto público, embora este não considere a restrição orçamentária do governo, ao não incluir a tributação nos modelos estimados. Entretanto, ainda que este artigo avance na análise, tanto em termos dos gastos analisados quanto da técnica econométrica, utilizando médias móveis da taxa de crescimento real *per capita* do produto para o período de cinco anos adiante, como forma de controlar a endogeneidade e a causalidade reversa, acredita-se que ele esteja mais próximo dos trabalhos de segunda geração, e, por isso, optou-se por apresentar os seus resultados empíricos com demais trabalhos desse grupo.

explícita ou implícita o papel da restrição orçamentária do governo na estimação dos efeitos da política fiscal, ao incluir simultaneamente no modelo pelo menos duas dentre as três categorias fiscais – tributação, despesa e resultado fiscal (déficit/superávit) – e que adotam métodos avançados recentes de estimação em painel ou séries de tempo, incluindo teste para a possível endogeneidade da política fiscal. Os resultados desses modelos tendem a ser mais robustos, reportando, em geral, efeitos negativos para a tributação, efeitos positivos para algumas categorias de investimento e efeitos negativos ou neutros para o consumo e o gasto com seguridade social. Ademais, os autores ressaltam que a omissão de variáveis fiscais não neutras pode levar a resultados bastante distintos para os parâmetros estimados, o que não deve ser interpretado como falta de robustez, mas sim como reflexo das diferentes categorias omitidas, convergindo assim com a conclusão de Adam e Bevan (2005).

Dentre os trabalhos desta geração, Miller e Russek (1997) reportam que não apenas a forma de financiamento dos gastos públicos afeta o impacto da política fiscal sobre o crescimento, mas também que este efeito é diferente entre países desenvolvidos e em desenvolvimento. Enquanto, para os primeiros, aumentos no gasto público financiados por dívida não alteram a taxa de crescimento e aumentos financiados por elevação nos impostos reduzem o crescimento econômico, para países em desenvolvimento, a expansão das despesas públicas via endividamento tem impacto negativo sobre a economia, mas a expansão financiada por aumento na carga tributária favorece o crescimento. Os autores argumentam que a diferença nos resultados entre os dois grupos de países pode refletir a forma de financiamento da dívida, se por títulos ou por emissão de moeda. Se os países em desenvolvimento utilizam o financiamento por emissão de moeda mais frequentemente, então, essas diferenças sugerem que o financiamento por moeda restringe o crescimento econômico, enquanto o financiamento por títulos, não. Contudo, essa hipótese não é testada explicitamente. Além disso, resultados distintos para as categorias de gasto público também foram encontrados. Aumentos nos gastos com defesa, saúde e seguridade social financiados por dívida tendem a reduzir o crescimento econômico em países em desenvolvimento. Porém, se o maior endividamento for direcionado para a expansão das despesas com educação, o resultado é positivo para a economia. A partir desses

resultados, Miller e Russek (1997) inferem que os gastos com defesa, saúde e seguridade social devem representar uma parcela muito grande do orçamento dos países em desenvolvimento, enquanto os gastos com educação correspondem a uma parcela pequena. Desse modo, pode existir um limite ótimo para a participação de cada gasto do orçamento na promoção do crescimento econômico. Novamente, os autores não testam essa hipótese.

Entretanto, como apresentado anteriormente, o trabalho de Devarajan, Swaroop e Zou (1996) investiga justamente a existência de um limite ótimo para a participação dos diferentes gastos no orçamento do governo no aumento do crescimento. Para as despesas com saúde e defesa, os resultados encontrados corroboram a hipótese levantada por Miller e Russek (1997): os efeitos negativos ou não significantes reportados por Devarajan, Swaroop e Zou (1996) indicam que a parcela desses gastos ultrapassou o limite máximo, tornando-os improdutivos na margem. Por sua vez, a hipótese de que os gastos com educação sejam produtivos não é corroborada.

Os impactos no crescimento econômico tanto do déficit fiscal quanto das diferentes formas de financiamento desse déficit são estudados por Adam e Bevan (2005). A partir de uma amostra de 45 países não membros da Organização para Cooperação e Desenvolvimento Econômico (OCDE), de 1977 a 1999, os autores concluem que o impacto do déficit público depende do valor inicial deste. Para déficits menores do que 1,5% do PIB, uma expansão nos gastos produtivos ou nos gastos não produtivos, ou uma redução na tributação, financiados por aumento no déficit, têm impacto positivo no crescimento. Entretanto, se o déficit inicial for maior do que o limite acima, o impacto passa a ser negativo. Em relação às formas de financiamento do déficit público, os resultados mostram que se o financiamento por títulos de dívida for usado para aumentar o gasto produtivo ou reduzir a carga tributária, seu efeito sobre o crescimento do produto é estatisticamente neutro. Já se o financiamento das políticas acima for por senhoriagem, o impacto sobre o crescimento será positivo, se a senhoriagem ficar abaixo do limite de 1,25% do PIB, e não apresentará efeitos negativos, se estiver acima desse limite. No entanto, se a senhoriagem ultrapassar o limite de 1,25% do PIB e for utilizada para financiar as despesas residuais, não produtivas, o crescimento é reduzido de forma acentuada. Nesse sentido, os resultados de Adam

e Bevan (2005) qualificam a hipótese levantada por Miller e Russek (1997) de que o efeito negativo da expansão dos gastos nos países em desenvolvimento está ligado ao financiamento do déficit público por emissão de moeda: apenas quando o financiamento supera a marca de 1,25% do PIB e é utilizado para financiar gastos não produtivos, observa-se redução no crescimento econômico.

Alguns estudos relacionando política fiscal e crescimento econômico foram desenvolvidos para a economia brasileira. E embora a literatura empírica internacional tenha crescido desde meados da década de 1980, é apenas no período recente, a partir do fim da década de 1990, que se observa um número significativo de trabalhos empíricos, cujo foco principal é avaliar os impactos dos gastos públicos no crescimento da economia brasileira no longo prazo.

Dentre esses estudos destacam-se os trabalhos de Blanco e Herrera (2006), Cândido Júnior (2001), e Silva e Cândido Junior (2009), que utilizam modelos de correção de erros para estimar a relação entre política fiscal e crescimento econômico. Blanco e Herrera (2006) partem de um modelo autorregressivo de defasagens distribuídas (ARDL) com correção de erros e utilizam dados anuais para o período de 1950 a 2000, para estimar o impacto da política fiscal no crescimento da economia brasileira. Os autores consideram a restrição orçamentária do governo ao incluírem a tributação e a dívida pública no modelo estimado, de modo a controlar para o potencial efeito negativo do financiamento público na atividade econômica, seguindo assim na linha dos estudos de terceira geração. Os resultados obtidos apontam que, no longo prazo, os efeitos da tributação e do gasto público foram semelhantes aos observados no curto prazo: os impostos apresentam forte impacto negativo sobre o produto *per capita* da economia; os gastos com consumo e seguridade social não afetam o produto; e os gastos com subsídios possuem efeito negativo.

Cândido Jr. (2001) avalia empiricamente a relação entre gastos públicos (agregados) e crescimento econômico no Brasil entre 1947 e 1995. As estimativas da elasticidade gasto-produto aparecem sempre como negativas, quando se utiliza o conceito mais restrito de gasto público (consumo mais transferências). Quando o gasto público passa a incluir também os investimentos das administrações

públicas, o impacto torna-se positivo. A estimativa do diferencial de produtividade entre os setores público e privado indica que o setor público tem uma produtividade de apenas 60% da do setor privado. As estimativas de curto prazo, no entanto, apontam que os gastos públicos defasados um ano têm impacto positivo sobre o produto, sendo esse efeito revertido no longo prazo.

Silva e Cândido Júnior (2009) estudam a relação de longo prazo a partir de um modelo vetorial autorregressivo (VAR) na forma de modelo de correção de erros para o Brasil, no período de 1970 a 2003, e de mais cinco economias latino-americanas: Argentina (1970-2002), Chile (1970-2002), Colômbia (1970-2002), México (1970-2002) e Venezuela (1970-2002). Em particular, os autores investigam os efeitos do consumo e investimento do governo sobre produto e seus componentes e analisam as consequências para a trajetória do produto de choques nos gastos públicos. As estimações foram realizadas para cada país individualmente, mas de um modo geral, observou-se que os investimentos das administrações públicas possuem efeito positivo no crescimento de longo prazo da economia. Quanto ao consumo do governo, o impacto estimado foi negativo para a Argentina, o Brasil e a Venezuela, porém positivo para as economias chilena e mexicana. Dado que o México apresentou a mais baixa relação consumo do governo/PIB dentre os países analisados e que, no Chile, a participação do consumo do governo no produto apresentou trajetória de queda no período, os autores argumentam que para estes casos é possível que a baixa participação do consumo gere efeitos marginais positivos, como no modelo de Devarajan, Swaroop e Zou (1996).

Souza, Kannebley e Diniz (2010) avaliam a relação entre política fiscal e crescimento econômico a partir de um modelo ARDL e de dados para a receita e para a despesa por função da União, estados e municípios, no período 1980-2008. Seguindo o trabalho de Barro (1990), que classifica as despesas entre produtivas e improdutivas, os autores concluem que os gastos produtivos agregados afetam de forma positiva o crescimento de longo prazo da economia, enquanto que a tributação sobre a renda apresenta efeito negativo. A análise dos gastos produtivos de forma desagregada destaca efeitos positivos principalmente para gasto com infraestrutura. Neste trabalho os autores não incluem o déficit público, de modo a controlar para a restrição orçamentária do governo. Contudo, em versão anterior –

Souza (2007) – essa questão é tratada explicitamente com a inclusão do déficit fiscal e a reestimação dos modelos para a inclusão das diferentes categorias fiscais omitidas: os gastos não produtivos e a tributação não distorciva. Embora a teoria postule que essas categorias sejam neutras em relação ao crescimento econômico, o impacto estimado para ambas foi negativo. A não neutralidade dessas variáveis, entretanto, não afetou o sinal e a significância encontrados para os coeficientes das variáveis gastos produtivos e tributação distorciva.

Alguns trabalhos têm por foco o impacto dos investimentos em infraestrutura sobre o crescimento da economia brasileira no longo prazo, e neste ponto os resultados são congruentes: todos encontram uma relação positiva entre o estoque de capital público em infraestrutura e o crescimento da economia. Ferreira (1996) considera como capital federal de infraestrutura o capital investido em telecomunicações, energia elétrica, portos, setor marítimo e ferrovias, e encontra esse resultado para o período de 1970-1993. Já Ferreira e Malliagos (1998) observam esse mesmo impacto positivo, para o período 1950-1995, considerando o setor rodoviário nos gastos em infraestrutura. Ferreira e Araújo (2006) ampliam o período de análise, de 1960 a 2000, e também concluem que o impacto do estoque de capital é positivo. Outro resultado importante apresentado nesses trabalhos é a constatação da queda contínua dos investimentos públicos em infraestrutura. Como mostram Ferreira e Araújo (2006, p. 2), a participação do investimento público no PIB vem caindo nos últimos 25 anos, e afetou os investimentos em infraestrutura em praticamente todos os setores. Enquanto entre 1969 e 1984 a formação bruta de capital do governo foi em média 4,03%, entre 1985 e 2001 reduziu-se para 1,85%, chegando a 0,9%, entre 2002 e 2005. Essa queda nos investimentos públicos em infraestrutura reflete as crises enfrentadas pela economia brasileira ao longo das últimas décadas – o descontrole da inflação até meados da década de 1990, as crises fiscais e financeiras da União, o endividamento excessivo dos governos estaduais – bem como a rigidez do orçamento brasileiro que torna os investimentos públicos a principal rubrica de ajuste fiscal. No entanto, dada a importância do capital público em infraestrutura no crescimento da economia, a tendência observada de queda nesses investimentos gera uma expectativa de menores taxas de crescimento no futuro.

Assis e Dias (2004) também avaliam somente o impacto do investimento público agregado sobre o crescimento econômico. Concluem que quanto maior a razão capital público/capital privado, maior a taxa de crescimento econômico. Encontram evidência, ainda, de que a carga tributária exerce um efeito negativo sobre o crescimento econômico quando ultrapassa o nível ótimo de participação do governo na economia em termos de tributação, estabelecido como sendo de 25% do produto. Assim, concluem que os efeitos da política fiscal sobre o crescimento econômico são ambíguos.

Rocha e Giuberti (2007) analisam a relação entre composição dos gastos e crescimento econômico dos estados brasileiros para o período de 1986 a 2003. As autoras partem do modelo teórico de Devarajan, Swaroop e Zou (1996) e a partir de um modelo de dados em painel apresentam o seguinte diagnóstico: os gastos com defesa, educação, transporte e comunicação possuem um efeito positivo e significativo sobre o crescimento; os gastos com capital afetam positivamente a taxa de crescimento e este efeito é aparentemente não linear; e as despesas correntes primárias do governo apresentam uma relação aparentemente positiva e não linear. Desse modo, confirmam os resultados padrão da literatura para o gasto produtivo e encontram evidências de que a despesa com consumo do governo não é tão prejudicial à economia como é argumentado pela literatura.

3.4 Gasto público, eficiência e crescimento econômico

Sanchez (2004) considera em seu modelo dois tipos de gasto público: investimento e consumo público produtivo, gasto este referente à manutenção do investimento público. A hipótese do modelo é que esse tipo de consumo pode substituir a acumulação de capital público: a realocação da despesa pública em favor do consumo produtivo leva a uma menor acumulação de capital em infraestrutura, mas também a um uso mais eficiente desse capital, de modo que pode eventualmente elevar a infraestrutura disponível em termos de unidade de eficiência. Desse modo, é a infraestrutura efetiva, determinada pelo estoque nominal de capital e pelo índice de eficiência, que entra na função de produção da economia e possui, portanto, efeito sobre o produto. O índice de eficiência, por sua vez, é função da parcela do gasto destinado ao consumo produtivo, i. e.,

da parcela destinada ao uso e manutenção da infraestrutura física. No entanto, o efeito de uma realocação do gasto em favor do consumo do governo é indeterminado *a priori*, uma vez que essa realocação, embora aumente o índice de eficiência, de um lado, reduz o estoque de capital, do outro. Ademais, o efeito é apenas sobre o nível de produto e o crescimento no período de transição para o novo estado estacionário, dado que no modelo postulado a autora considera retornos decrescentes para o estoque de capital público e privado. O resultado depende, portanto, da análise empírica. Calibrando o modelo para a economia americana, a autora mostra que as trajetórias de queda na despesa com investimento e de aumento na despesa com consumo produtivo do governo, observadas para os EUA, no período de 1952 a 2001, seguiram de perto a composição ótima derivada pelo modelo.

Existem ainda dois trabalhos empíricos que ligam a eficiência do gasto público, entendida da forma feita no restante deste livro, e o crescimento econômico.

O primeiro é o de Angelopoulos, Philippopoulos e Tsiona (2008), que trata somente do gasto agregado, mas que deve ser discutido pois pode facilmente ser expandido para uma avaliação da composição do gasto.

Angelopoulos, Philippopoulos e Tsiona (2008) observam que se por um lado o governo fornece bens e serviços e corrige as falhas de mercado, por outro gera uma série de distorções na economia (os gastos devem ser financiados por impostos que alteram os incentivos existentes).

Na realidade o governo se defronta com um *trade-off* que depende da combinação tamanho-eficiência. Se o tamanho do governo é muito grande e/ou se sua eficiência é muito baixa, os efeitos negativos devem suplantar os positivos. Em outras palavras, a relação entre o tamanho do governo e o crescimento deve ser não monotônica.

Para avaliar empiricamente essa hipótese, eles incorporam duas medidas de eficiência em um modelo de crescimento à la Barro (1990). Com isso o crescimento passa a ser determinado não só pelo tamanho do governo, mas também por uma variável resultante da interação entre o tamanho do governo e as medidas de eficiência.

São utilizadas duas medidas de eficiência. A primeira é o índice construído por Afonso, Schuknecht e Tanzi (2005) para os países da OCDE. Afonso, Schuknecht e Tanzi (2005) criam um índice composto de desempenho do setor público a partir de sete grupos de indicadores que procuram captar as principais funções/áreas de atuação do governo: administração, educação, saúde, infraestrutura, estabilização, redistribuição e alocação (esses três últimos tentam incorporar o que ficou conhecido como funções musgravianas do governo). Cada um desses grupos, por sua vez, é composto por subindicadores. Por exemplo, o indicador de saúde tem como componentes a taxa de mortalidade infantil e a expectativa de vida. O peso de cada grupo de indicadores no índice composto é o mesmo (1/7) e o peso de cada subindicador no grupo de indicadores também é o mesmo e depende do número de subindicadores considerados.

No caso do grupo de indicadores de saúde, como são considerados somente dois indicadores, cada um deles tem 50% de peso. Uma vez construído o indicador de desempenho, ele é dividido pelos recursos usados (o gasto público) para que possa captar a eficiência. Como o índice de desempenho e o gasto são medidos em unidades diferentes, para ser possível estabelecer comparações entre os governos dos diferentes países, o desempenho e os gastos de cada país são expressos relativamente à média do desempenho e dos gastos de todos os países.

O índice de eficiência, em suma, mede a eficiência do setor público em alcançar um conjunto de objetivos e consiste basicamente na razão do indicador composto de desempenho (produto) e o gasto público relacionado a esse indicador (insumo).

Também é obtida uma estimativa da eficiência técnica, usando a abordagem *Free Disposable Hull* (FDH), uma metodologia similar à Análise Envoltória de Dados (DEA⁵) que define fronteiras não paramétricas a partir dos valores extremos do espaço dimensional de insumos e produtos, mas sem adotar a hipótese de convexidade.

Angelopoulos, Philippopoulos e Tsiona (2008) adotam exatamente o mesmo procedimento, porém avaliam somente quatro áreas:

5 Análise Envoltória de Dados é a tradução literal do termo em inglês *Data Envelopment Analysis* (DEA).

administração, educação, infraestrutura e estabilização. Uma vez que não existem informações para muitos países para todas as áreas, optam por sacrificar o número de áreas a fim de aumentar o tamanho da amostra a ser analisada. Eles utilizam, então, uma amostra de 64 países, desenvolvidos e em desenvolvimento, durante o período 1980-2000.

A segunda medida de eficiência é resultante da estimação de um modelo de fronteira estocástica, como será discutido no capítulo 10. O índice composto de desempenho é utilizado como medida de produto e o gasto total do governo como parcela do produto é usado como insumo.

A evidência obtida mostra que, de fato, a relação entre tamanho do governo e o crescimento econômico é não monotônica, indicando que o *mix* tamanho-eficiência é mais importante para o crescimento do que o tamanho do governo propriamente dito.

A extensão direta do trabalho de Angelopoulos, Philippopoulos e Tsiona (2008) envolveria avaliar o impacto da composição e da eficiência sobre o crescimento econômico. Os países podem apresentar eficiências diferentes, por exemplo, nos seus gastos em educação e saúde e, assim, seria fundamental avaliar qual o efeito sobre o crescimento da interação entre eficiência e gastos em diferentes áreas. Em suma, é preciso combinar a discussão da composição dos gastos com uma avaliação mais específica da eficiência dos diferentes tipos de gastos.

Já Gonand (2007) avalia o impacto sobre o crescimento econômico de um aumento na eficiência do gasto público em educação primária e secundária, ainda que em uma abordagem de *accounting*.

Nesse sentido, há uma tentativa de desagregação maior, ainda que não sejam consideradas as eficiências dos outros tipos de gasto.

Como será visto no capítulo 7, a maior eficiência pode ser usada para reduzir insumos (eficiência orientada pelos insumos) ou aumentar produtos (eficiência orientada pelo produto), o que leva à existência de dois canais através dos quais uma melhoria na eficiência poderia promover o crescimento econômico.

No caso de ganhos de eficiência resultantes da redução de insumos, dois mecanismos dinamizam o nível de produção do setor privado:

1. O aumento no emprego privado resultante da mudança de recursos para o setor privado, o que promove a produção desse setor.
2. Diferença de crescimento da produtividade entre o setor público e o setor privado. Os ganhos de produtividade, na média, são estruturalmente mais baixos na educação pública do que na privada. Diante disso, mudar os professores do setor público para o setor privado aumenta o nível de unidades de trabalho eficientes na economia ainda mais e, conseqüentemente, o nível de produto.

Os ganhos de eficiência que aumentam o produto, por sua vez, incrementam a produtividade do trabalho no setor privado e o produto, porque melhoram o estoque de capital humano a gastos públicos inalterados. Tal efeito se materializa gradualmente ao longo do tempo, à medida que coortes mais bem treinadas e relativamente mais produtivas vão progressivamente substituindo as coortes mais antigas, com menor nível de educação e que vão paulatinamente se aposentando.

Os resultados são apresentados para 12 países da OCDE, sendo as medidas de eficiência (orientadas pelo produto ou pelo insumo) obtidas de Sutherland et al. (2007). Estes calculam estimativas de eficiência dos gastos em educação primária e secundária usando a metodologia DEA em que são utilizados como insumos o número de professores por aluno e o *back-ground* socioeconômico dos alunos e, como produtos, o escore médio no Programa Internacional de Avaliação de Alunos (Pisa) e a homogeneidade dos escores nesse programa. A evidência empírica obtida por Gonand (2007) sugere que os ganhos de eficiência terão feitos maiores sobre o produto se eles forem usados para aumentar os resultados em educação, em vez de usados para reduzir insumos. Um aumento de 10% no resultado do produto (o que equivale aproximadamente a aumentar o número médio de anos de escolaridade em um ano, com os mesmos insumos) leva a um aumento no produto, em média, de 3% a 6% no longo prazo. Se os ganhos de eficiência implicam transferência de recursos para o setor privado (redução de insumo), o impacto estimado sobre o produto passa a ser de menos de 1%.

3.5 Considerações finais

Diante da grande quantidade de resultados apresentados pela literatura internacional e brasileira, a conclusão mais apropriada parece ser a de que os efeitos da política fiscal sobre o crescimento do produto dependem das características de cada economia e de cada período. Ainda assim, alguns fatores destacam-se como importantes nessa análise: a composição da despesa, a estrutura tributária e a restrição orçamentária do governo. A literatura teórica prevê distintos efeitos sobre o crescimento do produto para os diferentes tipos de gastos públicos, de acordo com o canal pelo qual esses gastos afetam a economia: se pela função de produção das firmas ou da utilidade das famílias. De modo análogo, os diversos tributos afetam de forma distinta os incentivos dos agentes, o que também reflete no crescimento. O déficit fiscal e o seu financiamento também operam como fatores de influência na economia e, portanto, no seu crescimento. A literatura empírica, embora nem sempre confirme as predições teóricas em termos de sinal e significância de cada uma das categorias fiscais, confirma a existência de efeitos distintos entre as diferentes variáveis que compõem a política fiscal.

REFERÊNCIAS

- ADAM, C. S.; BEVAN, D. L. Fiscal deficits and growth in developing countries. *Journal of Public Economics*, v. 89, p. 571-597, 2005.
- AFONSO, A.; SCHUKNECHT, L.; TANZI, V. Public sector efficiency: an international comparison. *Public Choice*, v. 123, p. 321-347, 2005.
- AGELL, J.; LINDTH, T.; OHLSSON, H. Growth and the public sector: a reply. *European Journal of Political Economy*, v. 15, p.359-366, 1999.
- ANGELOPOULOS, K.; PHILIPPOPOULOS, A.; TSIONAS, E. Does public sector efficiency matter? Revisiting the relation between size and economic growth in a world sample. *Public Choice*, v. 132, p. 245-278, 2008.
- ARROW, K. J.; KURZ, M. Optimal public investment policy and controllability with fixed private savings ratio. *Journal of Economic Theory*, v. 1, p. 141-177, 1969.
- ASSIS, L. C. de; DIAS, J. Política fiscal, nível tecnológico e crescimento econômico no Brasil: teoria e evidência empírica. In: ENCONTRO NACIONAL DE ECONOMIA (ANPEC), 32., 2004, João Pessoa. *Anais...* João Pessoa, 2004.
- BARRO, R. J. Government spending in a simple model of endogenous growth. *Journal of Political Economy*, v. 98, p. 103-125, 1990.
- BARRO, R.; SALA-I-MARTIN, X. *Economic growth*. McGraw-Hill: New York, 1995.
- BASSANINI, A.; SCARPETT, A.; HEMMING, P. Economic growth: the role of policies and institutions: panel data evidence from OECD countries. *OECD Economics Department Working Papers*, Paris, n. 283, 2001.
- BLANCO, F.; HERRERA, S. The quality of fiscal adjustment and the long-run growth impact of fiscal policy in Brazil. In: ENCONTRO NACIONAL DE ECONOMIA (ANPEC), 34., 2006, Salvador. *Anais eletrônicos...* Salvador, 2006. Disponível em: <<http://www.anpec.org.br/encontro2006/artigos/A06A087.pdf>>. Acesso em: 4 nov. 2010.
- CÂNDIDO JR., J. O. *Os gastos públicos no Brasil são produtivos?* Brasília: IPEA, fev. 2001. (Texto para discussão n. 781). Disponível em: <http://www.ipea.gov.br/pub/td/td_2001/td_781.pdf>. Acesso em: 5 nov. 2010.

CARBONI, A.; MEDDA, G. Size and composition of public spending in a neoclassical growth model. *Metroeconomica*, v. 62, n. 1, p. 150-170, 2011.

DEVARAJAN, S.; SWAROOP, V.; ZOU, Z. The composition of public expenditure and economic growth. *Journal of Monetary Economics*, v. 37, p. 313-344, 1996.

EASTERLY, W.; REBELO, S. Fiscal policy and economic growth: an empirical investigation. *Journal of Monetary Economics*, v. 32, p. 417-458, 1993.

FERREIRA, P. C. Investimento em infra-estrutura no Brasil: fatos estilizados e relações de longo prazo. *Pesquisa e Planejamento Econômico*, v. 26, n. 2, p. 231-252, 1996.

FERREIRA, P.C.; ARAÚJO, C. H. V. *Growth and fiscal effects of infrastructure investment in Brazil* [2006]. Disponível em: <<http://www.fgv.br/professor/ferreira/BrazilInfra4.pdf>>. Acesso em: 26 maio 2010.

FERREIRA, P. C.; MALLIAGROS, T. G. Impactos produtivos da infra-estrutura no Brasil – 1950/95. *Pesquisa e Planejamento Econômico*, v. 28, n. 2, p. 315-338, 1998.

FÖLSTER, S.; HENREKSON, M. Growth and the public sector: a critique of the critics. *European Journal of Political Economy*, v. 15, p. 337-358, 1997.

GEMMEL, N.; KNELLER, R. The impact of fiscal policy on long-run growth. In: EUROPEAN COMMISSION. Directorate-General for Economic and Financial Affairs. European Economy. *Current Issues in Economic Growth*, n. 1, 2001. Disponível em: <http://ec.europa.eu/economy_finance/publications/publication7600_en.pdf>. Acesso em: 7 jun. 2011.

GONAND, F. The impact on growth of higher efficiency of public spending on schools. *OECD Economics Department Working Papers*, n. 547, 2007.

HEITGER, B. The scope of government and its impact on economic growth in OECD countries. *Kiel Working Paper*, n. 1.034, 2001.

KNELLER, R.; BLEANNEY, M.; GEMMEL, N. *Growth, public policy and the government budget constraint: evidence from OECD countries*. [S.l.]: School of Economics, University of Nottingham, 1998 (*Discussion Paper* n. 98/14). Disponível em: <<http://www.nottingham.ac.uk/economics/documents/discussion-papers/98-14.pdf>>. Acesso em: 6 jun. 2011.

LEVINE, R.; RENELT, D. A sensitivity analysis of cross-country growth regressions. *American Economic Review*, v. 82, p. 942-963, 1992.

MILLER, S. M.; RUSSEK, F. S. Fiscal structures and economic growth: international evidence. *Economic Inquiry*, v. 35, p. 603-613, jul. 1997.

NIJKAMP, P.; POOT, J. Meta-analysis of the impact of fiscal policies on long-run growth. *European Journal of Political Economy*, v. 20, p. 91-124, maio 2004.

ROCHA, F.; GIUBERTI, A. C. Composição do gasto público e crescimento econômico: uma avaliação macroeconômica da qualidade dos gastos dos estados brasileiros. *Economia Aplicada*, v. 11, n. 4, p. 463-485, 2007.

SANCHEZ, R. M. P. *Characterizing the optimal composition of government expenditures*. 2004. Disponível em: <<http://eprints.ucm.es/7723/1/0409.pdf>>. Acesso em: 20 jul. 2012.

SILVA, A. M. A. da; CÂNDIDO JR., J. *Os impactos macroeconômicos dos gastos públicos na América Latina*. Rio de Janeiro: IPEA, nov. 2009. (Texto para Discussão n. 1434). Disponível em: <http://www.ipea.gov.br/sites/000/2/publicacoes/tds/td_1434.pdf>. Acesso em: 8 abr. 2011.

SOLOW, R. A contribution to the theory of economic growth. *Quarterly Journal of Economics*, v. 70, p. 65-94, 1956.

SOUZA, G. S. de. *Política fiscal e crescimento econômico: evidências para o caso brasileiro*. 2007. 146 f. Dissertação (Mestrado em Economia) – Departamento de Economia da Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto da Universidade de São Paulo (FEA-RP/USP), Universidade de São Paulo, Ribeirão Preto, 2007.

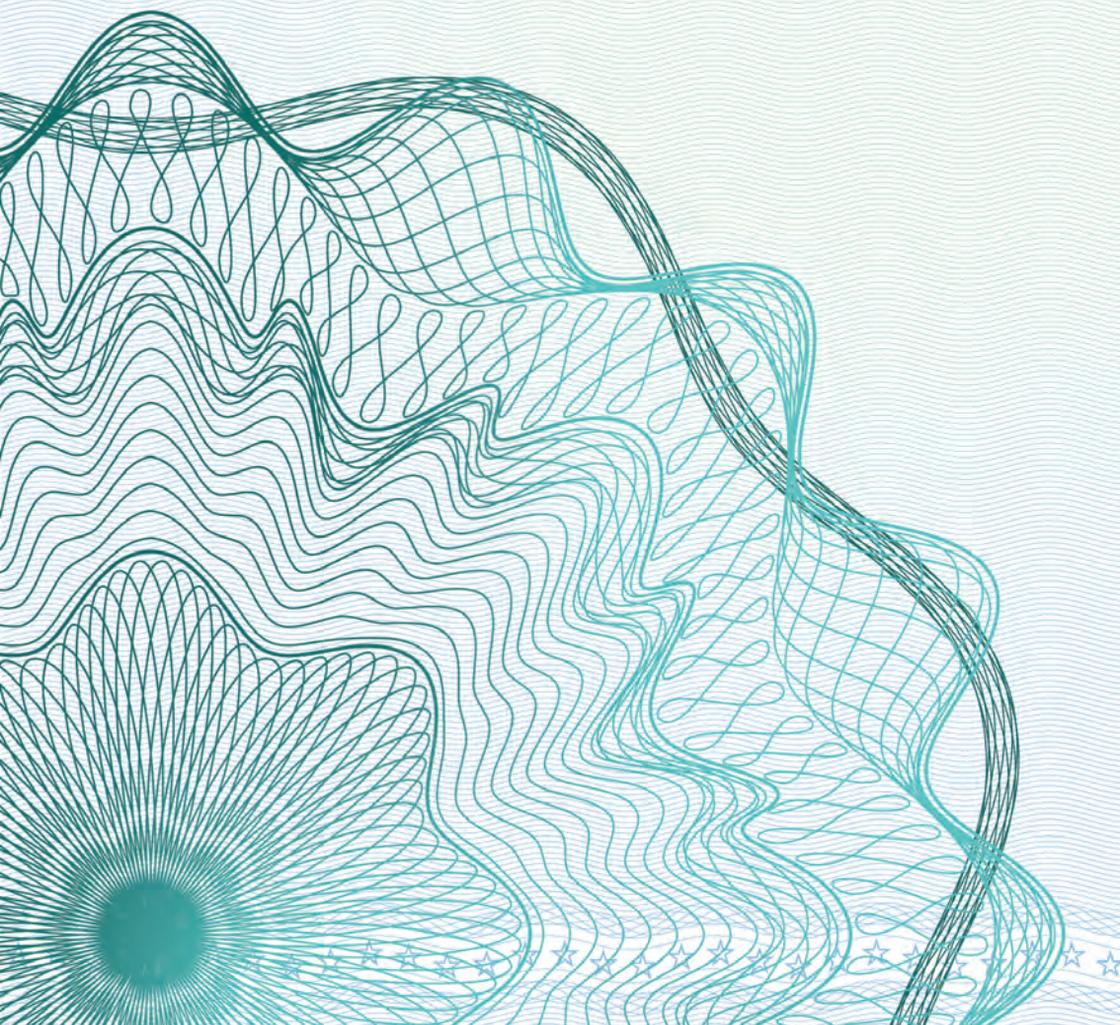
SOUZA, G. S. de.; KANNEBLEY JR., S.; DINIZ, E. M. Política fiscal e crescimento de longo prazo no Brasil: evidências para dados do orçamento função. *Pesquisa e Planejamento Econômico*, v. 40, n. 1, p. 41-84, abr., 2010.

SUTHERLAND, D.; PRICE, R.; JOUMARD, L.; NICQ, C. Performance indicators for public spending efficiency in primary and secondary education. *OECD Economics Department Working Paper*, n. 546, 2007.



PARTE II

A EXPERIÊNCIA INTERNACIONAL NA
AVALIAÇÃO DO GASTO PÚBLICO



The background features a complex, abstract design of teal lines. It includes a series of concentric, wavy lines that resemble a topographical map or a stylized sunburst pattern in the lower-left corner. The lines are dense and create a sense of depth and movement across the page.



Capítulo 4

Revisões de Despesas na OCDE

Marc Robinson

4.1 Introdução

Este capítulo examina as práticas de revisão de despesas¹ em seis países da Organização para Cooperação e Desenvolvimento Econômico (OCDE), para identificar as principais características do processo e tirar conclusões a respeito da melhor maneira de estruturar a análise de gastos. Cinco dos seis países abordados neste capítulo – Austrália, Canadá, Dinamarca, Holanda e Reino Unido – foram selecionados porque têm processos de revisão das despesas mais bem desenvolvidos do que a maioria de seus pares da OCDE. O sexto país – a França – foi escolhido porque passou recentemente por um grande episódio de revisão de despesas, com o qual foram aprendidas lições importantes.

A primeira seção deste capítulo descreve a natureza e os objetivos da revisão de despesas, enfatizando que se trata de uma ferramenta usada não somente para reduzir os gastos agregados, mas também capaz de aumentar o espaço disponível para priorizar novas despesas consistentes com a manutenção de uma política fiscal agregada sólida. A seção apresenta uma tipologia de práticas de revisão de despesas que é baseada em seu foco e cobertura. Um ponto importante apresentado neste contexto é que a revisão de despesas está interessada não somente em identificar ganhos de eficiência (reduzindo o desperdício), mas também em aumentar a priorização – e, portanto, a eficácia – dos gastos públicos. A seção discute as funções institucionais e distingue revisão de despesas da análise de gastos, incluindo avaliações que podem ser usadas para fornecer uma base de informações mais sólidas para a revisão de despesas.

1 A expressão *revisão de despesas* é a tradução literal da expressão inglesa *spending review*.

Da seção 4.2 à 4.7 são apresentados estudos de casos nacionais. Cada um deles começa com uma descrição das características e objetivos básicos da revisão de despesas no país em questão. A descrição é seguida de (i) uma discussão sobre processos e funções institucionais; (ii) uma análise da base de informações da revisão de despesas; e (iii) uma avaliação geral do sistema, que identifica pontos fortes e pontos fracos.

A seção 4.8 apresenta uma síntese, o principal objetivo para tirar conclusões a respeito de como a revisão de despesas pode ser organizada para se tornar mais eficaz. Neste contexto, a seção analisa se a revisão das despesas deve visar a ser global ou seletiva, quais funções a liderança política e o serviço público devem ter, e o que pode ser feito para fortalecer a base de informações. A seção também aponta áreas importantes onde há grandes diferenças de abordagem para a revisão de despesas entre os países, mas onde não é possível identificar melhores práticas. Essa conclusão não causa surpresa: o projeto de uma boa revisão de despesas também precisa levar em conta características nacionais e não pode adotar o mesmo modelo para todos os lugares.

4.1.1 O que é uma revisão de despesas?

Uma revisão de despesas é um processo institucionalizado para a revisão do *cenário base de gastos* com o principal objetivo de identificar opções de economia a serem analisadas no processo orçamentário. *Cenário base de gastos* nesse contexto significa gasto em projetos ou programas existentes (serviços ou pagamentos de transferências existentes). Portanto, a revisão de despesas não está relacionada à avaliação orçamentária de *novas* propostas de despesas.

A identificação de opções de economia é uma parte essencial do processo de revisão de despesas. No entanto, esse processo pode ser designado para identificar opções de aumento ou de redução de financiamentos destinados aos programas existentes. No que se refere a opções de redução de financiamento, a revisão de despesas visa a identificar opções de economia orçamentária que podem ser realizadas melhorando a eficiência ou reduzindo gastos ineficazes ou de baixa prioridade. Nenhuma opção de economia identificada durante as revisões de despesas é, por definição, opção de economia

específica, em oposição a cortes orçamentários indiscriminados ou gerais (ver Box 4.1 e a discussão logo em seguida).

A revisão de despesas é um processo explicitamente destinado a contribuir para as decisões orçamentárias. Por esse motivo, a revisão de despesas é necessariamente um processo comandado e gerenciado – no nível burocrático – pelo Ministério das Finanças (MF), possivelmente em parceria com alguns órgãos centrais que talvez tenham, em determinados países, uma função fundamental na preparação do orçamento, como o Gabinete do primeiro-ministro ou o do presidente.

Box 4.1 – Definição de revisão da despesa: a abordagem da OCDE

[Revisão de despesas é] a análise retrospectiva de um programa, política ou organização... que é encomendado *ex-ante* com o objetivo específico de identificar uma economia orçamentária em todo o governo... A revisão de despesas pode ser usada para reduzir o déficit e/ou criar espaço fiscal para programas de prioridade mais alta por meio de reestruturação ou corte de atividades. Embora as revisões de despesas também possam examinar a eficácia dos programas (como melhorar o desempenho e o projeto de um programa), seu diferencial é o foco na identificação e extração de economia de recursos por meio do processo orçamentário... [Há] um vínculo explícito com a tomada de decisões orçamentárias, tanto em termos de análise e recomendação em níveis de financiamento alternativo (p. ex., metas da revisão de despesas) como em termos do processo orçamentário.

Fonte: OECD (2011, p. 3).

Embora a revisão de despesas tenha como principal objetivo informar decisões orçamentárias sobre o nível de financiamento para programas e projetos existentes, ela também pode ter objetivos mais amplos de melhoria de desempenho. Seria o caso, por exemplo, se o processo de economia exigisse explicitamente que o gerenciamento do órgão tivesse de lidar com problemas de desempenho identificados durante a revisão de despesas e responder, quando apropriado, com melhorias no desenho do programa ou no gerenciamento do processo, ou em ambos. Contudo, um processo de revisão que visa

unicamente à melhoria do desempenho, sem o objetivo explícito de mudar os níveis de financiamento dos programas existentes, não constitui uma revisão de despesas.

4.1.2 Revisão de despesas e priorização dos gastos

A revisão de despesas é uma ferramenta para melhorar a priorização dos gastos. *Priorizar gastos* significa alocar recursos limitados do governo para os programas que oferecem maiores benefícios à sociedade (eficiência alocativa).

Em relação a novas propostas de despesas, uma boa priorização de gastos significa garantir que somente essas propostas que oferecem à sociedade o maior valor de dinheiro sejam financiadas no orçamento.

Com relação ao *cenário de referência* de gastos, uma boa priorização significa assegurar que recursos escassos não continuem a ser desperdiçados em processos de produção ineficientes ou em programas existentes – que sejam inerentemente ineficazes ou de prioridade baixa. Também significa fornecer financiamento adicional necessário quando a expansão de um programa existente é socialmente importante.

Um problema amplamente reconhecido é que a preparação do orçamento governamental tende a focar desproporcionalmente sobre a consideração de novas propostas de despesas, com pouca revisão de *cenário de referência* de gastos. A revisão de despesas – com seu foco específico neste cenário – visa a lidar com esse problema.

Na maioria dos países, a melhoria da priorização dos gastos requer não somente uma revisão de despesas, mas também a melhoria dos processos existentes para o escrutínio de novas propostas de despesas. Contudo, esses processos não estão incluídos no escopo deste capítulo, que foca exclusivamente na revisão de despesas.

4.1.3 Revisão de despesas e gastos governamentais agregados

A revisão de despesas aumenta o escopo para novas despesas de alta prioridade. É geralmente difícil encontrar espaço fiscal para novos projetos ou programas prioritários – pelo menos sem gerar uma taxa de crescimento insustentável dos gastos públicos agregados. A revisão de despesas fornece um meio sistemático de ajustar o cenário de referência de gastos ineficientes e de baixa prioridade

para aumentar o espaço para as novas prioridades. Isso também pode ser um meio importante de aumentar o nível de despesa com infraestrutura, ajustando seletivamente o gasto atual.

O benefício da revisão de despesas é verdadeiro em períodos de forte crescimento na economia e nas receitas públicas. Mesmo nessas circunstâncias, é útil aumentar o espaço para uma nova despesa prioritária. Nesse contexto, a revisão de despesas reduz o risco de, durante *períodos favoráveis*, o governo usar aumentos cíclicos na receita unicamente para financiar novas prioridades, levando à deterioração no saldo orçamentário estrutural.

Além de aumentar o escopo para uma nova despesa de alta prioridade, a revisão de despesas às vezes serve para reduzir os gastos públicos agregados. Nesse caso, qualquer economia realizada por meio da revisão de despesas será alocada nos gastos priorizados pelo governo em vez de ser alocada para financiar novas iniciativas de despesa. A revisão de despesas tende a ser usada dessa forma quando o objetivo da política fiscal do governo é reduzir déficits e dívidas (durante períodos de consolidação fiscal).

Em resumo, os governos usam a revisão de despesas (i) para aumentar o espaço fiscal a novos gastos prioritários ou (ii) para reduzir o gasto público agregado.

4.1.4 Quais são os tipos de economia de recursos?

A revisão de despesas oferece dois tipos de ganhos orçamentários: ganhos de eficiência e economias dos produtos². *Os ganhos de eficiência* são obtidos mudando-se a forma como os serviços são prestados, de modo que esses serviços (produtos) forneçam a mesma quantidade e qualidade a um custo mais baixo – em outras palavras, economia resultante da melhoria da eficiência técnica. A despesa reduzida obtida pelo corte de serviços não é qualificada como economia da eficiência.

As economias dos produtos – um novo termo introduzido neste capítulo – são aquelas obtidas pela redução ou pela eliminação de serviços ou pagamentos de transferências que são ineficazes ou de baixa prioridade. Aqui, esse tipo de economia é designado *economias*

2 A expressão *economias dos produtos* é tradução literal da expressão inglesa *output savings*.

dos produtos porque produtos são serviços fornecidos à comunidade e porque as economias dos produtos envolvem reduções deliberadas e direcionadas na quantidade ou na qualidade desses serviços.

Como citado anteriormente, as opções de economia identificadas pela revisão de despesas são, em princípio, específicas. Isso significa que um governo sabe como a redução dos gastos básicos será alcançada, ou seja, quais serviços serão reduzidos (no caso das economias dos produtos) ou quais mudanças de redução de custos serão feitas nos processos de negócios (no caso de ganhos de eficiência). As reduções dos gastos obtidas por meio da revisão da despesa são, portanto, diferentes dos cortes não específicos, definidos como cortes que o governo impõe aos ministérios sem que os ministros saibam de antemão como essas reduções serão implementadas (Box 4.2).

Box 4.2 – Exemplos de técnicas de redução de gastos não específicos

As seguintes técnicas orçamentárias são amplamente usadas para impor reduções de gastos não específicos aos ministérios. Elas contrastam com o objetivo da revisão de despesas, de realizar reduções de gastos por meio de economia específica:

- **Cortes gerais:** Esta técnica envolve a aplicação de uma redução de porcentagem-padrão para todos os orçamentos de despesa ministerial com base única (p. ex., cortar todos os orçamentos ministeriais em cerca de 7% no ano seguinte).
- **Dividendo de produtividade:** Também conhecido em alguns países como um dividendo em eficiência, envolve a aplicação contínua (a cada ano) de uma pequena redução percentual (p. ex., 1% para cada orçamento ministerial, baseada na justificativa de que os ministérios deveriam ser, de modo geral, capazes de aumentar a eficiência em pelo menos 1% anualmente). Normalmente, determinadas categorias dos gastos ministeriais, que não são potencialmente abertas aos ganhos de eficiência, são excluídas da cobertura de dividendos de produtividade (como pagamentos de transferências para terceiros).

A revisão de despesas busca deliberadamente reduzir a dependência do governo – quando as reduções de gastos são inevitáveis – para cortes não específicos. O principal foco deste livro está nas técnicas e métodos para a melhoria da eficiência técnica nos gastos públicos (p. ex., pelo uso de metodologias formais, como a Análise Envoltória de Dados (DEA³)). A revisão de despesas é vinculada a este trabalho porque ele fornece uma metodologia para usar, no processo orçamentário, as informações fornecidas pela análise da eficiência. Contudo, a revisão de despesas também é um processo mais amplo, precisamente porque pode focar não somente nos ganhos de eficiência, mas também nas economias dos produtos. Em outras palavras, a revisão de despesas é uma ferramenta de eficiência técnica e alocativa.

4.1.5 Foco e escopo da revisão de despesas

Qualquer processo de revisão de despesas compreende um conjunto de revisões de tópicos. Existem três tipos principais de revisão de tópicos:

- *Revisões de programas*: esse tipo examina os programas específicos (ou seja, categorias de serviços ou pagamentos de transferências específicos) e pode fornecer ganhos de eficiência (reduzindo os custos da prestação dos serviços do programa) ou economias dos produtos (reduzindo os serviços prestados pelo programa), ou ambos.
- *As revisões do processo* examinarão minuciosamente processos de negócio específicos usados na produção de serviços públicos (p. ex., processos de aquisição, sistemas de tecnologia da informação [TI] e práticas de gestão de recursos humanos). Essas revisões ajudam a obter ganhos de eficiência ao invés de economias dos produtos.
- *As revisões dos órgãos* examinam um órgão público inteiro (ministério ou outro órgão) e podem, em princípio, abranger todos os programas e processos do órgão.

Revisões de programas e processos também são denominadas revisões *estratégicas* e *funcionais*, respectivamente (OCDE, 2011, p. 3-4). Elas podem ser específicas de um órgão ou podem ser

3 Análise Envoltória de Dados é a tradução literal do termo em inglês *Data Envelopment Analysis* (DEA).

horizontais. Uma *revisão de programa horizontal* examina um grupo de programas relacionados, realizado por dois ou mais órgãos, enquanto uma *revisão de processo horizontal* analisa um domínio particular dos processos de negócios em vários (ou em todos os) órgãos governamentais (p. ex., uma revisão das práticas de aquisição de grande parte do governo).

Os processos de revisão de despesas nunca revisam todos os programas governamentais e todos os processos de negócios, mesmo quando os governos, às vezes, descrevem suas revisões das despesas como *globais*. Seria impraticável revisar tudo. Portanto, o termo *global* não deveria ser levado ao pé da letra. No entanto, há uma real diferença entre revisões globais das despesas e revisões seletivas das despesas.

Uma *revisão seletiva das despesas* é limitada a uma lista específica de tópicos – programas, processos ou órgãos – e é estimulada no início de cada rodada de revisão de despesas.

Em uma *revisão global das despesas*, por outro lado, o escopo não é limitado a uma lista *ex ante* de tópicos. As equipes de revisão são solicitadas a examinar todos os ministérios com a expectativa de que identificarão, na medida do possível, as opções de economia mais importantes. Uma revisão global das despesas deve ter um escopo maior – e gerar mais economia – do que uma revisão seletiva.

Os processos de revisão de despesas diferem em sua abrangência. Algumas revisões de despesas focam somente nos *gastos orçamentários* (no gasto que é legalmente autorizado na lei orçamentária anual), enquanto outras revisões também cobrem os *gastos obrigatórios* (portanto, gastos como benefícios de seguro social, autorizados pela legislação em vigor, também são conhecidos como *gastos estatutários*). Outra questão importante da abrangência é se a revisão de despesas no nível governamental nacional inclui transferências governamentais nacionais para governos subnacionais, como estados ou províncias.

4.1.6 Processo e funções institucionais

O processo da revisão de despesas varia enormemente entre países. Em alguns, é um processo contínuo (anual ou periódico); em outros, é um processo *ad hoc*. As principais diferenças nas funções institucionais podem ser vistas nos níveis políticos e burocráticos.

No nível burocrático, as funções desempenhadas pelo MF (e outros órgãos centrais relevantes), pelos ministérios que efetuam as despesas e pelos peritos externos diferem em estágios cruciais do processo de revisão de despesas – como decidir o que será revisado e como conduzir as revisões. Essas diferenças dizem respeito, em parte, ao poder do MF sobre os ministérios que efetuam as despesas.

No nível político, os países diferem com relação a quais instituições ou líderes políticos devem determinar as opções de revisão de despesas a serem adotadas no orçamento. Essas diferenças dependem parcialmente de onde reside o *real* poder orçamentário do país (em oposição ao poder legal formal para aprovar o orçamento), seja com o legislativo, executivo, instituições ou líderes específicos do governo (o presidente, o primeiro-ministro ou o ministro das Finanças).

4.1.7 Revisão das despesas em oposição à análise dos gastos

Os estudos de caso neste capítulo também examinam a base de informações usada para a revisão de despesas. Distinta da revisão de despesas, a *análise dos gastos* refere-se a análises da eficiência, eficácia ou relevância do gasto (p. ex., prioridade e duplicação). Toda revisão de despesas envolve algum tipo de análise dos gastos. Contudo, a revisão de despesas é mais do que uma análise dos gastos, porque identifica explicitamente as opções para níveis alternativos de financiamento (incluindo opções de economia).

Somente a análise de um volume limitado de gastos pode ser realizada dentro do cronograma da maioria dos ciclos de revisão das despesas (normalmente, de vários meses a um ano). Uma boa revisão de despesas deve, portanto, ser capaz de aproveitar uma análise de gastos de alta qualidade que é realizada separadamente – e, antes, do processo de revisão de despesas. Isso aponta para uma questão crucial da revisão de despesas: a qualidade da análise de gastos que apoia o processo.

Uma questão chave neste contexto diz respeito a (i) até que ponto a revisão de despesas é capaz de aproveitar a análise formal de gastos ou (ii) se é obrigada a contar basicamente com a análise informal de gastos. Análise *formal* de gastos, neste contexto, significa uma análise de gastos que envolve a aplicação de técnicas analíticas formais, como uma avaliação de resultados (ver Box 4.3).

Box 4.3 – Exemplos de técnicas de análise formal de gastos

A seguir estão algumas das técnicas de análise de gastos que podem ser usadas para apoiar a revisão de despesas:

- **Análise de processo de negócio:** uma forma de análise de eficiência baseada na revisão de processos de negócios existentes e em sua modificação, de forma a reduzir o custo do fornecimento de produtos.
- **Padrões de referência de custos:** processo de análise de eficiência baseado na comparação dos custos de um produto ou processo com os custos dos mesmos produtos ou de similares produzidos por outras empresas.
- **Avaliações de resultados:** avaliações para garantir até que ponto os resultados pretendidos de um programa ou projeto tenham sido alcançados.
- **Avaliações de processos:** tipo de avaliação que revisa a implementação de um programa ou projeto, geralmente com o objetivo de identificar mudanças na gestão ou formulação de políticas para melhorar a eficácia.
- **Análise lógica do programa:** processo que examina a forma como um determinado programa ou projeto é apoiado para obter seu resultado pretendido e que pergunta se, considerando teoria e experiência relevantes, tem chances de ser eficaz; também conhecida como *avaliação baseada em teorias* (BANCO MUNDIAL, 2004, p. 10).

A pergunta sobre a base de informações da revisão de despesas também levanta duas outras questões intimamente relacionadas: Como os indicadores de desempenho disponíveis provaram ser úteis no processo de revisão de despesas? Até que ponto a estrutura do desempenho orçamentário – onde ela existe – impactou a eficácia do processo de revisão de despesas?

Como mencionado anteriormente, os sistemas de avaliação e classificação da eficácia e eficiência dos gastos governamentais não

constituem, na maioria dos casos, uma revisão de despesas porque a identificação das opções de economia não é um dos recursos desses tipos de sistemas. Somente os processos de revisão criados para desenvolver opções de economia explícitas para a tomada de decisões do governo podem ser considerados como revisão de despesas. Portanto, sistemas de classificação e avaliação são tidos mais como uma parte da *base de informações* que está disponível para a revisão de despesas (ver Box 4.4).

A lógica dessa distinção é que, por exemplo, o fato de um programa específico ter sido classificado ou avaliado como ineficaz é informação útil para a revisão de despesas, mas não significa automaticamente que o programa em questão represente uma opção de economia. Reduzir o financiamento do programa pode ser muito inapropriado: a resposta apropriada às dificuldades do programa pode ser a reformulação desse programa ou até mesmo um financiamento *adicional*. Somente uma revisão das despesas pode fazer esse discernimento (ver Box 4.4).

Box 4.4 – Avaliação de desempenho e sistemas de classificação não constituem uma revisão de despesas

- A *Ferramenta de Classificação para Avaliação do Programa*, utilizada nos Estados Unidos entre 2002 e 2009, não foi um processo de revisão de despesas porque sua função era atribuir classificações como *eficaz* e *ineficaz* aos programas do governo local e não para identificar opções de economia (MARK; PFEIFFER 2011; GILMOUR, 2006; WHITE, 2011).
- Um *sistema de avaliação em todo o governo*, como existiu no Canadá por muitos anos e na Austrália nas décadas de 1980 e 1990, requer ministérios para avaliar seus programas e sistemas de acordo com um ciclo de revisão. Requer também que todos os programas sejam revisados em um período de cinco anos (MACKAY, 1998; 2004; SECRETARIA DO CONSELHO DO TESOURO, 2004; 2009).

Os estudos de casos dos países a seguir são apresentados em uma ordem que reflete, em linhas gerais, o amadurecimento do processo de revisão de despesas, ou seja quão bem desenvolvidos e bem estabelecidos estão esses processos nos países em questão. Portanto, os estudos começam com a Holanda que, sem dúvida, tem o processo de revisão de despesas mais duradouro e eficaz. Para uma visão geral comparativa dos seis países da OCDE, consulte a Tabela 4.1 no final deste capítulo.

4.2 Holanda

Características básicas, objetivos e foco

Na Holanda, a revisão de despesas formal existe desde o início da década de 1980 – inicialmente como *Procedimento de Reconsideração*; a partir de 1995, como *Revisão de Políticas Interdepartamentais* (IPR); e finalmente, em 2010, como *Revisão Global de Despesas* (CER). A Holanda tem um sistema de tetos de gastos plurianuais que cobre, a cada período parlamentar de quatro anos, as prioridades dos gastos estabelecidas em acordos de coalizão. A CER de 2010 teve um papel fundamental no debate da eleição daquele ano sobre medidas de economia orçamentária e no acordo de coalização subsequente sobre tetos de gastos. Há hoje um consenso entre todos os partidos para adotar um ciclo de revisão de despesas regular de quatro anos, tendo um processo como a CER de 2010 conduzido durante a preparação para cada eleição.

Refletindo a forte pressão da União Europeia para reduzir os déficits orçamentais, que surgiram durante a crise financeira global, a CER de 2010 focou exclusivamente nos objetivos orçamentários e mais especificamente na redução de gastos agregados. Em oposição, o foco da IPR não foi somente na identificação de opções orçamentárias, mas também na melhoria do desempenho de modo geral.

O processo de revisão de despesas incluiu uma análise não somente das opções de economia de gastos, mas também das opções de cortes de gastos tributários e de aumentos em cobranças dos usuários. As opções de gastos adicionais geralmente não são consideradas (com poucas exceções) como parte do processo de revisão de despesas.

Essa revisão de despesas tem sido seletiva e não global. A CER de 2010 teve 20 tópicos de revisão e focou principalmente nas economias dos produtos. A maior parte das revisões de tópicos da CER foi de revisões de programas (tanto específicas de órgãos quanto horizontais) e não de revisões de processos. Consistente com uma norma antiga, a CER de 2010 exigiu que cada força-tarefa de revisão apresentasse opções capazes de gerar pelo menos 20% de redução nos gastos – ao longo de quatro anos – no programa sob revisão. Ainda não se sabe se essa norma será mantida durante as CER futuras.

Nenhum gasto do governo central está, em princípio, fora dos limites para revisão de despesas na Holanda. Os gastos obrigatórios foram o foco principal da CER de 2010, assim como o foram outras questões politicamente delicadas, como a dedução de juros de hipoteca.

Processo e funções institucionais

Os tópicos de revisão são formalmente decididos pelo Gabinete, seguindo as recomendações feitas ao ministro das Finanças pela Inspeção de Orçamento no Ministério (MF), em consulta com os ministérios que efetuam as despesas. A CER de 2010 foi supervisionada pelo primeiro-ministro e dois primeiros-ministros adjuntos. No nível burocrático, a supervisão do processo foi de responsabilidade da Comissão de Revisão de Despesas Interdepartamentais, liderada pela Inspeção de Orçamento.

As revisões dos tópicos são realizadas pelas forças-tarefa de revisão, com termos de referência e processos uniformes definidos pelo MF e acordados pelo Gabinete. Durante a CER de 2010, as 20 forças-tarefa foram compostas essencialmente por funcionários públicos (em serviço ou aposentados), incluindo representantes do Ministério das Finanças. As forças-tarefa conseguiram contratar especialistas externos (dentre eles, especialistas de universidades e instituições de pesquisas públicas). Seguindo a tradição e para evitar conflitos de interesse, a presidência da força-tarefa era independente, ou seja, não era do MF nem dos ministérios que efetuam as despesas abordadas pela revisão. As forças-tarefa da revisão de despesas sempre têm dado total liberdade na identificação de opções, sem direito a vetar aquelas apresentadas por uma minoria de seus membros.

A intensa participação dos funcionários do Ministério das Finanças com sólidas habilidades políticas e conhecimento detalhado de portfólio – especialmente da Inspeção de Orçamento – tem sido essencial para o sucesso da revisão de despesas. Para maximizar sua influência sobre o processo, o MF insistiu em nomear funcionários de alto nível – incluindo o diretor de Orçamento – como funcionários das forças-tarefa de revisão. Os representantes do Ministério tendem a desempenhar funções muito ativas em forças-tarefa, inclusive propondo opções de economia específicas para análise.

Base de informações da revisão da despesa

De modo geral, a análise informal tem sido mais importante no processo holandês do que a análise formal dos gastos, e a revisão de despesas usa principalmente a análise de gastos que já foi realizada, em oposição a uma análise especialmente encomendada para apoiar a despesa. (Ver Box 4.5).

Box 4.5 – Estudo de caso: reforma da Assistência Social

Um dos muitos resultados bem-sucedidos de revisão de despesas na Holanda, sob o sistema da IPR no final da década de 1990, foi uma grande reforma do esquema de assistência social. Esse sistema proporcionou às administrações municipais, que administram benefícios sociais, grandes incentivos para controlar os custos das reivindicações. Ele conseguiu isso dividindo o orçamento central para benefícios entre os municípios com base em uma fórmula (um sistema baseado na população municipal, nível de renda e fatoraçoão do sucesso das políticas de reemprego) e permitindo que os municípios retivessem qualquer excedente depois de pagar os direitos aos benefícios sociais. Após essas mudanças, o número de requerentes de benefícios baixou de 336.000 em dezembro de 2003 para 311.000 (excluindo pessoas com mais de 65 anos de idade) em setembro de 2006.

Geralmente, as forças-tarefa de revisão encontram indicadores de desempenho analisados isoladamente (em oposição à análise de gastos formal), de valor limitado em seu trabalho. As avaliações

realizadas pelos ministérios que efetuam as despesas também comprovaram ser de valor muito limitado para forças-tarefa de revisão de despesas. Foram dadas duas razões para esse limite: muitas avaliações são consideradas de baixa qualidade e, como regra geral, elas não avaliam a eficácia de uma política. As informações limitadas disponíveis sobre a eficácia do programa são percebidas como uma grande limitação do processo de revisão de despesas, embora esta percepção varie entre áreas de políticas.

Avaliação

O processo de revisão de despesas da Holanda tem sido um dos mais bem-sucedidos em termos de capacidade de identificar opções concretas de economia – incluindo economia de programas – frequentemente adotadas pela liderança política. A cultura política de consenso e a ampla aceitação da necessidade de restringir as despesas têm sido fatores decisivos para esse sucesso. Outro fator importante é que as instituições de serviços públicos relevantes, intimamente envolvidas no processo de revisão de despesas, são consideradas imparciais e tecnicamente sólidas, o que dá credibilidade às opções de economia identificadas pelas forças-tarefa.

Um apoio político sólido e de alto nível para a revisão de despesas, incluindo o do primeiro-ministro e o dos líderes dos partidos políticos na coalização em vigor, também tem sido considerado crucial para forçar os ministérios que efetuam despesas a *passar a bola* e cooperar. Finalmente, o fortalecimento da base de informações, especialmente por meio do desenvolvimento de uma avaliação mais relevante, é visto como prioridade na melhoria da eficácia do processo de revisão de despesas.

4.3 Canadá

Características básicas, objetivos e foco

O Canadá tem realizado, desde 2007, um processo anual de revisão de despesas implementado inicialmente na forma de *Revisão Estratégica* (SR), que se tornou *Revisão Estratégica e Operacional* (SOR) em 2011. Um episódio anterior de revisão de despesas, o processo

Revisão do Programa de 1994 a 1996, também foi importante (ver Box 4.6). Desde o início, o principal foco do processo de revisão de despesas tem sido orçamentário. No entanto, o governo também identificou a melhoria do desempenho como uma meta do processo. Não houve, contudo, mecanismos explícitos para exigir que os órgãos solucionassem os problemas de desempenho identificados durante o processo de revisão de despesas.

Na Revisão Estratégica, o principal objetivo era criar um espaço fiscal adicional para novas prioridades da despesa, e o governo alegou, durante os quatro anos do processo, ter realocado toda a economia de recursos para novas iniciativas de despesa. Consistente com esta alegação, durante quatro anos os ministérios que efetuam as despesas tiveram autorização para apresentar opções de novas despesas financiadas pela economia gerada: as chamadas *propostas de reinvestimento*. Em 2011, na SOR, o foco mudou para consolidação fiscal e um controle gradual dos gastos agregados. A economia era alocada principalmente à linha de base e os órgãos não tinham mais permissão para apresentar propostas de reinvestimento.

Na SR e na SOR, os órgãos tinham sido proibidos de apresentar opções de *gastar para economizar*, ou seja, opções que exigiam uma despesa significativa (como novos sistemas de TI) para obter ganhos de eficiência. Na SR, o foco era maior nos ganhos de eficiência e nas economias dos produtos. A SOR envolveu certa mudança de ênfase nos ganhos de eficiência.

Durante todo esse tempo, o processo de revisão de despesas canadense tem sido do tipo global. No caso da SR, contudo, a cobertura global foi obtida em um ciclo de quatro anos, com cerca de um quarto dos órgãos federais revisados todos os anos. Em um processo de SOR, mais intensivo, todos os órgãos foram incluídos em um único ano, em preparação para uma consolidação fiscal substancial no orçamento de 2012. O processo tem sido limitado à chamada despesa de programa direto, que exclui os principais benefícios de seguro social e as transferências para províncias e territórios.

Os processos de SR e SOR têm sido quase inteiramente revisões de órgãos. Contudo, foram incluídas determinadas revisões de processos horizontais, como a revisão (de 2008 a 2009) da política de gestão dos recursos humanos centrais.

O governo definiu uma meta de economia agregada para o processo de SOR de 2011 a 2012, mas não o fez durante os quatro anos da SR. As metas no âmbito dos órgãos foram aplicadas, contudo, na SR e na SOR. Na SR, essa meta tomou a forma de um requisito para que todo órgão identificasse opções de economia que totalizassem pelo menos 5% de sua despesa com programas. Na SOR, tinham que ser apresentados dois conjuntos de opções, um para corte de 5% e outro para corte de 10%.

Box 4.6 – Revisão do programa canadense em meados da década de 1990

O Canadá é reconhecido pelo grande sucesso do processo de revisão da despesa da Revisão do Programa, que ocorreu durante as rodadas dos orçamentos de 1995 e 1996. A revisão do programa objetivava explicitamente a consolidação fiscal para refrear os altos déficits e reduzir a dívida. Foram estabelecidas duras metas de economia específicas dos órgãos – até 50% no caso do Ministério dos Transportes e de 15% a 25% para a maioria dos outros ministérios. O processo de revisão do programa baseou-se em revisões dos órgãos e foi orientado por seis testes (critérios de avaliação do programa). O processo foi supervisionado por uma subcomissão do Gabinete. O primeiro-ministro “apoiou visível e firmemente seu ministro das Finanças”, contra os ministros que efetuam despesas (GOOD, 2007). Os gastos públicos foram reduzidos em cerca de 10% em dois anos (BOURGON, 2009).

Processos e funções institucionais

Os processos de SOR e SR foram supervisionados no nível político por uma subcomissão do Gabinete. No nível burocrático, o processo é coordenado pela Secretaria do Conselho do Tesouro (TBS), instituição canadense responsável pela alocação dos gastos do processo de preparação do orçamento que não se reporta a um único ministro, mas a uma comissão de ministros (o Conselho do Tesouro). No Canadá, o Departamento das Finanças é responsável pela política macroeconômica e fiscal agregada e, portanto, não está fortemente envolvido no processo de revisão da despesa.

Como observado antes, a revisão da despesa no Canadá é principalmente um processo de revisões dos órgãos, essencialmente descentralizadas: cada órgão faz sua própria revisão e cria suas próprias opções de economia, sem a participação direta da equipe da TBS. Depois disso, os órgãos apresentam ao governo as propostas das revisões. Embora não seja feita uma consulta entre a equipe da TBS e a equipe do ministério que efetua as despesas durante a preparação das propostas da revisão do órgão, a principal contribuição da equipe da TBS é feita depois que os ministérios que efetuam despesas enviam suas revisões. Nesse estágio, a equipe da TBS analisa as opções de economia propostas e assessora os ministros quanto ao efeito geral e à adequabilidade das opções. Normalmente, a equipe da TBS não apresenta aos ministérios opções alternativas de economia. O processo descentralizado é visto como vantajoso pelos participantes.

O conhecimento dos programas e processos dos órgãos, por parte dos oficiais da TBS, constituiu parte importante do sucesso do processo de revisão da despesa no país, processo que exige uma grande parcela do tempo da equipe da Secretaria do Conselho do Tesouro. Peritos externos também têm papel de consultoria.

O perigo potencial de um processo de revisão da despesa descentralizado como o do Canadá é o de os ministérios que efetuam as despesas deixarem de apresentar opções de economia ou apresentarem deliberadamente opções que não são realistas – na esperança de serem rejeitadas. Esse risco é atenuado de três formas principais: (i) o processo impõe metas mínimas de economia no nível do órgão; (ii) as apresentações departamentais insatisfatórias são rejeitadas e os órgãos em questão são solicitados a repetir a tarefa; e (iii) os ministros dos órgãos devem participar ativamente das revisões, aprovar as propostas e apresentá-las aos colegas do Gabinete.

As revisões dos órgãos são primeiramente apresentadas e analisadas pela subcomissão do Gabinete que supervisiona o processo de revisão da despesa. Essa subcomissão tende a desempenhar o papel mais importante na decisão das opções de economia que serão adotadas. Após essa etapa, o presidente do Conselho do Tesouro apresenta as opções recomendadas à Comissão de Planejamento e Prioridades do Gabinete (presidida pelo primeiro-ministro) para revisão. As decisões finais sobre todas as propostas cabem ao primeiro-ministro.

Base de informações da revisão da despesa

Os participantes acreditam que há necessidade de fortalecer a base de análise de gastos da revisão da despesa. Embora o Canadá tenha exigido por muitos anos que os ministérios que efetuam despesas avaliassem seus programas, as avaliações não demonstraram ser muito úteis para as decisões orçamentárias em geral e para a revisão da despesa em particular. Como resultado, está sendo feito um esforço ativo de melhoria na importância das avaliações para fins orçamentários. Os indicadores de desempenho do programa – relativamente bem desenvolvidos no Canadá – são considerados como fonte útil de informações no processo de revisão da despesa, embora seja difícil julgar o quanto eles foram realmente úteis.

Avaliação

Os 5,2 bilhões de dólares canadenses (CAD) em economias contínuas reivindicados como resultado do processo de SOR são equivalentes a pouco menos de 2% da despesa direta do programa federal, aos quais devem ser adicionados os CAD 2,8 bilhões de economia reivindicados como resultado do processo de SR. Entretanto, essa estimativa ainda não foi verificada de forma independente.

A abordagem para a revisão da despesa foi influenciada por duas características nacionais. A primeira é a ampla aceitação da importância da sustentabilidade e da disciplina nos gastos públicos. A segunda é a tradição de descentralização na administração pública. De acordo com os participantes do processo, a abordagem descentralizada da revisão da despesa – com a elaboração das opções de economia deixada para os ministérios que efetuam as despesas – tem duas grandes vantagens: mobilizou o conhecimento dos ministérios que efetuam as despesas quanto às melhores opções de economia; e, ao garantir a propriedade do órgão, aumentou as possibilidades de implementação real das opções de economia.

A dominação do processo de revisão da despesa segundo as revisões dos órgãos, em vez das revisões seletivas dos programas, limita a capacidade dos responsáveis pelas decisões do orçamento central de concentrar-se, durante os processos de revisão da despesa, nos programas considerados particularmente questionáveis.

4.4 Austrália

Características básicas, objetivos e foco

Os processos sistemáticos de revisão da despesa, que são essencialmente gerenciados pela Comissão de Revisão de Despesas (ERC, *Expenditure Review Committee*) do Gabinete, existem na Austrália desde 1976. A ERC é a comissão de orçamento do Gabinete e não uma comissão exclusivamente dedicada à revisão dos gastos básicos. A última rodada de atividades de revisão da despesa, conhecida como *Revisão Global das Despesas* (CER, *Comprehensive Expenditure Review*), funcionou durante três ciclos do orçamento a partir de 2008-2009.

Embora certo nível de revisão da despesa seja contínuo, a atividade intensa de revisão da despesa concentrou-se em três períodos-chave: final da década de 1970, com a consolidação fiscal sob um governo conservador; meados da década de 1980, com a consolidação fiscal e uma mudança de prioridades sob um governo social-democrata; e o período de 2008-2010, com uma mudança de prioridades e a preparação para a consolidação fiscal, sob um governo social-democrata. O processo australiano de revisão da despesa é totalmente voltado para a identificação de economias orçamentárias e não é formalmente vinculado a processos mais amplos de melhoria do desempenho.

O poder orçamentário na Austrália está fortemente concentrado no governo executivo. Embora o primeiro-ministro seja poderoso, o Gabinete – particularmente a ERC – toma muitas das principais decisões orçamentárias. O ministro das Finanças tem um papel de liderança na ERC e na preparação do orçamento de forma mais geral. Além do Departamento de Finanças (DOF), que tem responsabilidade direta pela preparação do orçamento detalhado, o Departamento do Primeiro-Ministro e Gabinete (PM&C, *Department of the Prime Minister and Cabinet*) tem grande influência nas prioridades dos gastos.

A revisão da despesa tem focado de forma consistente as economias dos produtos e não exclusivamente, ou principalmente, os ganhos de eficiência. No entanto, às vezes foi dedicado um esforço considerável à busca por opções de ganhos de eficiência, tais como

a importante revisão de tecnologia da informação (TI) realizada na CER de 2008-2010. A revisão da despesa também englobou, de forma consistente, tanto os gastos obrigatórios quanto os gastos orçamentários. Em relação às transferências para o governo subnacional, ela não engloba as *transferências desvinculadas* (transferências de fundos que não estão vinculados a fins específicos), mas às vezes analisa as *transferências "vinculadas"*. Aparentemente não foram definidas metas de economias durante a CER de 2008-2010. Entretanto, as metas de economias específicas da carteira foram uma característica-chave do processo em meados da década de 1980.

O processo australiano utilizou as revisões dos órgãos e também as revisões de processos e programas. A ERC é que toma a decisão *ex ante* sobre quais programas ou processos serão revisados. Entretanto, o foco da revisão da despesa não se limitou necessariamente a esses tópicos: durante a última CER, além dos tópicos de revisão identificados, os oficiais do DOF realizaram uma busca mais abrangente por opções de economia.

Processos e funções institucionais

O processo de revisão da despesa é supervisionado no nível político pela ERC. No nível burocrático, o DOF e o PM&C – que tendem a cooperar estreitamente – supervisionam o processo. As revisões da despesa são realizadas principalmente pela burocracia. Embora às vezes tenham sido lideradas por pessoas proeminentes de fora (como uma revisão de TI de 2008), ainda assim as equipes de revisão são invariavelmente compostas por servidores públicos. Contudo, especialistas externos são frequentemente usados como consultores. Apenas ocasionalmente (2008) o público foi expressamente convidado a contribuir com propostas de economias.

Cada ministério que efetua despesas conduz sua própria revisão interna e identifica suas opções de economia para cada rodada do orçamento. Ao mesmo tempo, o DOF elabora opções de economia – durante a CER de 2008-2010, criou, para este fim, uma Força de Trabalho para Revisão das Despesas. O Departamento desenvolveu fortes competências em análise da política de gastos e um conhecimento detalhado dos programas do órgão, essenciais

para sua função no processo de revisão da despesa. As opções de economia criadas tanto pelos ministérios que efetuam as despesas quanto pelo DOF são apresentadas à ERC, que normalmente decide quais opções serão adotadas no orçamento.

Base de informações da revisão da despesa

Um sistema de revisões estratégicas foi introduzido em 2007 como forma de análise de gastos de determinados programas ou processos, com o objetivo, em parte, de apoiar a revisão da despesa. As revisões estratégicas não devem ser consideradas como revisões da despesa, uma vez que não têm a finalidade de identificar opções de economia do orçamento. As revisões estratégicas são realizadas conjuntamente por equipes do ministério que efetuou as despesas e do DOF, e baseiam-se principalmente na análise informal dos gastos.

Apesar das informações úteis geradas, o DOF considera que o processo de revisão estratégica tem um sucesso limitado e acredita ser necessário mais para fortalecer a base de informações da revisão da despesa. Essa necessidade é reforçada pela visão do DOF de que as avaliações feitas pelos próprios ministérios que efetuam as despesas têm um valor limitado para a revisão da despesa, em parte porque sua qualidade média é classificada como baixa. De um modo geral, deve-se dizer que não é feito um amplo uso de técnicas *formais* de análise de gastos para apoiar a revisão da despesa na Austrália.

Avaliação

O processo foi considerado bastante eficaz na criação de opções de economia para ajudar a conseguir a consolidação fiscal e para levar a despesa a novas prioridades. O sucesso do sistema reflete parcialmente o consenso político e o dos eleitores sobre a importância de manter a sustentabilidade fiscal e, desse modo, controlar os gastos agregados. No nível burocrático, o sistema australiano de revisão da despesa, que coloca uma responsabilidade considerável nos próprios ministérios que efetuam as despesas para identificar opções de economia, reflete as tradições descentralizadas da administração pública. Do mesmo modo, o sucesso do sistema reflete o maior poder exercido pelos órgãos centrais sobre assuntos de importância para todo o governo, incluindo o nível de gastos agregados.

4.5 Dinamarca

Características básicas, objetivos e foco

Desde meados dos anos 1980 a Dinamarca tem um sistema de revisão da despesa conhecido como *estudos especiais*. Esses estudos fazem parte do processo normal de preparação do orçamento anual, embora em alguns anos não tenham sido realizados. Em princípio, os estudos especiais podem recomendar aumentos no financiamento dos programas existentes, mas na prática esse resultado não é comum e o foco é voltado para as opções de economia. O sistema não tem vínculo formal com processos mais amplos de melhoria do desempenho do governo.

Os governos dinamarqueses geralmente são de coalizão e os acordos sobre quais programas devem ser cortados ou devem receber maior financiamento normalmente são definidos em acordos pré-eleitorais. A Comissão Econômica do Gabinete, presidida pelo ministro e apoiada pelo Ministério das Finanças (MF), é o principal órgão de tomada de decisões executivas e orçamentárias. O MF, responsável no nível burocrático pela preparação do orçamento, é uma instituição poderosa em relação aos ministérios que efetuam as despesas.

Normalmente, são feitos de 10 a 15 estudos especiais por ano, mas às vezes o número é significativamente maior, incluindo o ano atual. Historicamente, o foco principal do processo de estudos especiais tem sido o aumento do espaço fiscal para novas prioridades dos gastos. Entretanto, o foco atual mudou para uma redução dos gastos agregados, com vistas à consolidação fiscal. A mudança levou a um aumento tanto no número de estudos especiais quanto no valor dos gastos que eles abrangem (p. ex., um estudo dos gastos em defesa está em andamento e um sobre a força policial foi recentemente concluído).

A maioria dos estudos especiais refere-se a revisões dos órgãos ou revisões de programas, mas foram realizadas algumas revisões de processos, tais como a dos processos de aquisições do governo. O foco principal dos estudos especiais tem sido, ao longo dos anos, mais em ganhos de eficiência do que em economias dos produtos. As revisões que visam reconsiderar os programas com base na sua eficácia ou na prioridade tendem a ocorrer em uma base *ad hoc*, separada do

processo de estudos especiais, que enfoca principalmente os gastos orçamentários e, mais raramente, a previdência social e outros gastos obrigatórios.

No passado, normalmente não eram definidas metas quantitativas para o valor das opções de economia identificadas pelos estudos especiais. Essa abordagem, porém, mudou recentemente e as metas de economia tornaram-se o foco dos últimos estudos especiais principais.

Processos e funções institucionais

Por recomendação do ministro das Finanças, os tópicos dos estudos especiais são selecionados pelos membros da Comissão Econômica do Gabinete que, por sua vez, recebe recomendações do MF. São reservados aproximadamente dois meses para a realização dos estudos especiais, supervisionados e gerenciados pelo Ministério das Finanças. As forças-tarefa conjuntas do MF e do ministério que efetua as despesas geralmente são estabelecidas para revisões em larga escala. As de menor escala, ao contrário, normalmente são realizadas somente por oficiais do MF (embora com a assessoria dos oficiais do ministério que efetua as despesas). Não existe unidade organizacional no Ministério das Finanças, responsável pelos estudos especiais. Em comparação com outros países analisados neste capítulo, a Dinamarca é diferente por confiar muito mais fortemente em consultores externos para efetuar uma parte significativa do trabalho detalhado de revisão.

Uma característica particular do sistema dinamarquês de estudos especiais é a falta de processos ou de critérios de revisão padronizados que se apliquem a todas as revisões. O Gabinete aprova termos de referência formais para estudos especiais em larga escala (ou seja, aqueles executados por forças-tarefa conjuntas), mas não para os estudos de menor escala conduzidos internamente no MF.

Para as revisões de maior escala gerenciadas por forças-tarefa específicas, é a força-tarefa que apresenta recomendações para o ministro das Finanças e a Comissão Econômica do Gabinete. Em princípio, essas recomendações devem ser baseadas no consenso entre o MF e o ministério que efetua as despesas em questão, mas se não for alcançado um consenso, podem ser apresentadas

recomendações separadas. A comissão geralmente toma a decisão final sobre as opções de economia que serão adotadas no orçamento.

Como em todos os lugares, a resistência dos ministérios que efetuam as despesas ao processo de revisão às vezes constitui um problema. O Ministério das Finanças acredita que o uso de consultores externos contribui um pouco para isso, devido à condição independente dos consultores. Contudo, a pressão das lideranças políticas é, em última análise, o fator mais importante para compelir os ministérios que efetuam as despesas a concordarem com o processo de estudos especiais.

Base de informações da revisão da despesa

A análise formal dos gastos não é muito usada no processo de estudos especiais e nenhuma dessas análises (p. ex., avaliações de resultados, estudos de padrões de referência) é comissionada especificamente como contribuição para os estudos especiais. Os indicadores de desempenho em geral não foram considerados muito úteis para os estudos especiais. As avaliações feitas pelos ministérios que efetuam as despesas nem sempre são verificadas pelo MF ou não estão disponíveis e, quando ocorre a verificação, demonstram ter valor limitado para a revisão da despesa.

Avaliação

Uma vantagem do sistema dinamarquês é a sua flexibilidade. Se o governo e o Ministério das Finanças veem um problema orçamentário em um determinado ministério ou programa, eles podem estabelecer rapidamente um estudo especial para analisar a questão. Ao mesmo tempo, entretanto, existe um sentimento generalizado de que a adoção dos estudos especiais precisa se tornar mais sistemática, com processos mais padronizados estabelecidos para orientá-los.

O grau de confiança nos consultores externos constitui um problema, por dois motivos: contribui para a falta de um organismo da equipe do MF com conhecimentos detalhados dos órgãos e dos programas; e os consultores externos de qualidade são caros. É provável que sejam feitas grandes reformas no processo de estudos especiais nos próximos anos.

4.6 Reino Unido

Características básicas, objetivos e foco

Ao analisar a revisão da despesa no Reino Unido, pode-se facilmente ser confundido pela terminologia. As Revisões das Despesas (SRs) periódicas têm sido uma característica das finanças públicas por mais de 20 anos – as revisões ocorrem a cada dois anos, entre 1998 e 2004, e depois a cada três anos, entre 2007 e 2010. Essas SRs não devem ser confundidas com as revisões da despesa no sentido definido em outra parte deste capítulo.

O processo local é, essencialmente, o de preparação do orçamento plurianual e, como tal, inclui um intenso foco nas decisões sobre novas despesas.⁴ Além disso, como o Tesouro do Reino Unido (o Ministério das Finanças) observa, antes de 2007 as SRs eram quase totalmente “voltadas para a alocação de aumentos incrementais na despesa” (TESOURO DO REINO UNIDO, 2006, p. 24), com pouco exame dos gastos básicos. Como foi enfatizado na introdução deste capítulo, a revisão da despesa no sentido mais usado enfoca explicitamente os gastos básicos, não as novas despesas.

Para evitar confusão, todas as referências em letras minúsculas de *revisão da despesa* nesta seção devem ser interpretadas como referentes a revisões dos gastos básicos. Por outro lado, quando usado com iniciais maiúsculas, *Revisão da Despesa* (ou SR) refere-se ao termo conforme usado no Reino Unido (o processo de preparação do orçamento plurianual).

Importantes revisões da despesa de todo o governo – revisões dos gastos básicos – foram realizadas em 2004, 2007 e 2010. A de 2004 tomou a forma de uma importante revisão isolada e com foco na eficiência (a Revisão de Gershon, ver Box 4.7). Ela não somente identificou ganhos de eficiência em potencial como também negociou reduções do orçamento específicas, ministério por ministério, com base nesses ganhos. As revisões dos gastos básicos de 2007 e 2010 fizeram parte das SRs oficiais realizadas nesses anos, mas foram

4 O Reino Unido tem um sistema de orçamento pelo qual uma parte dos gastos públicos fica sujeita a compromissos plurianuais por parte do governo para os ministérios que efetuam as despesas.

descritas pelo governo como Revisões Globais das Despesas (CSRs), para ressaltar o fato de que pretendiam revisar os gastos básicos.

Box 4.7 – Estudo de caso: a Revisão de Gershon

A Revisão de Gershon da Eficiência do Setor Público em 2004 identificou os departamentos que efetuaram despesas – e negociou acordos que, segundo estimou, gerariam 21,5 bilhões de libras em ganhos de eficiência até 2008. Estimativas independentes, mas altamente conservadoras do Escritório Nacional de Auditoria (NAO, *National Audit Office*), apontam que em 2007 o programa havia gerado economias comprováveis de pelo menos 10 bilhões de libras, equivalentes a cerca de 2,4% do total dos gastos do governo central em 2007-2008. As principais áreas de ganhos de eficiência foram as seguintes:

- Reformas de aquisições, tais como a agregação de pedidos para aumentar o poder de negociação com os fornecedores.
- Tratamento mais eficiente (principalmente por meio do uso de TI) das interações com o público em áreas de serviços tais como impostos, alfândega e serviços do governo local.
- Redução do tempo gasto pelos profissionais da linha de frente (p. ex., médicos e enfermeiros) em tarefas administrativas.
- Melhoria dos processos de *suporte* (serviços de suporte corporativo).

Durante a CSR de 2010, o foco era em reduções dos gastos agregados para conseguir a consolidação fiscal. Já o objetivo da Revisão de Gershon de 2004 e da CSR de 2007 era criar uma margem de manobra fiscal adicional para financiar as novas prioridades das despesas.

Anteriormente a 2010, o processo de SR mais amplo – ou seja, o processo de preparação do orçamento plurianual – sempre combinava objetivos orçamentários com objetivos mais amplos de melhoria do desempenho. Em particular, as três primeiras SRs (1998, 2002 e 2004) foram combinadas com um sistema de Acordos de Serviços Públicos de metas de desempenho de alto nível e

focadas em resultados. A CSR de 2010, por sua vez, tinha um foco essencialmente centrado em objetivos orçamentários.

A CSR de 2010 era fortemente voltada para a obtenção de economias dos produtos, bem como ganhos de eficiência, enquanto a ênfase da Revisão de Gershon de 2004 e da CSR de 2007 era principalmente em ganhos de eficiência. No entanto, não foram definidas metas de economias específicas dos departamentos antes da CSR de 2010. Na CSR de 2007, todos os departamentos deveriam apresentar opções de ganhos de eficiência de pelo menos 3% ao ano.

Tanto em 2010 como em 2007, as CSRs foram abrangentes, ou seja, não estavam limitadas por lista *ex ante* de tópicos de revisão da despesa. A CSR de 2010 englobou quase todos os gastos governamentais – orçamentários, obrigatórios e transferências para o governo subnacional – e também os gastos fiscais. A CSR de 2007 e a Revisão de Gershon de 2004 excluíram a previdência social e a maioria dos outros gastos obrigatórios.

Processos e funções institucionais

A CSR de 2010 foi presidida por uma recém-criada Comissão de Despesa Pública (PEX) do Gabinete. O papel central dessa comissão no processo de revisão da despesa era um reflexo do governo de coalizão (historicamente pouco usual). Já a CSR de 2007 foi essencialmente controlada pelo chanceler do Exchequer (o ministro das finanças britânico), com uma participação bem limitada do Gabinete ou mesmo do primeiro-ministro. No nível burocrático, a revisão da despesa é essencialmente gerenciada pelo Tesouro, uma instituição muito poderosa em relação aos ministérios que efetuam as despesas.

Durante as CSRs de 2007 e de 2010, a principal fonte de opções de economia eram os próprios ministérios que efetuam as despesas (*departamentos* na terminologia britânica), que deviam conduzir revisões de despesas internas e, em seguida, apresentar propostas formais detalhando as opções de economia para o Tesouro. Os departamentos eram instruídos a aplicar um conjunto de critérios-padrão claramente definidos, tais como: se a atividade é essencial para atender às prioridades do governo, se o governo precisa financiar essa atividade e se a atividade oferece valor econômico substancial. Depois de receber as propostas dos departamentos,

os oficiais do Tesouro também injetavam suas próprias opções de economia. A CSR de 2010 ainda incluiu um convite para sugestões do público sobre opções de economia – o chamado Desafio da Despesa – que até então não tinha feito parte do processo.

A revisão de 2010 era apoiada por um grupo de especialistas independentes (*Independent Challenge Group*), mas a grande maioria dos membros era de funcionários públicos. O papel do Tesouro é substancial e exigente, e a CSR de 2010 consumiu, nas palavras de um importante ministro, um *esforço enorme* do Tesouro. Com o passar dos anos, o Tesouro construiu um grande conhecimento das áreas de serviços dos ministérios que efetuam as despesas, o que foi essencial para a gestão do processo de revisão. Em resumo, portanto, tanto as revisões da despesa de 2007 quanto as de 2010 foram principalmente realizadas por funcionários públicos. Contudo, isso não ocorreu na Revisão de Gershon de 2004, quando tanto o chefe quanto a maioria da equipe de revisão eram externos ao governo.

Base de informações da revisão da despesa

As CSRs de 2007 e de 2010 foram apoiadas por importantes estudos de programas e processos de negócios. Entretanto, os estudos foram todos realizados de forma isolada e se basearam em análises informais dos gastos. O governo não tem um programa contínuo de análise formal de gastos elaborado para apoiar o processo de despesas, tal como avaliação de programas, e aparentemente as avaliações realizadas pelos ministérios que efetuam as despesas por sua própria iniciativa não são muito usadas na tomada de decisões orçamentárias.

O Reino Unido é um dos países mais avançados do mundo no desenvolvimento de indicadores de desempenho do setor público – inclusive na difícil área de indicadores de eficácia (resultados) – e aparentemente esses indicadores são utilizados como variáveis no processo *de revisões de despesa na Inglaterra*.

Avaliação

Embora o processo de revisão da despesa incorporado na CSR de 2010 certamente tenha alcançado seu objetivo de facilitar importantes cortes na despesa pública agregada, o Reino Unido tem

um longo caminho a percorrer em relação ao estabelecimento de um sistema eficaz de revisão *contínua* dos gastos básicos. Grande parte do processo usado em 2010 teve de ser criada pela primeira vez, já que precisavam ser feitos cortes importantes no programa e os processos anteriores de revisão da despesa eram voltados quase exclusivamente para os ganhos de eficiência. Entretanto, o Reino Unido tem pelo menos duas vantagens importantes que apoiarão uma revisão da despesa eficaz no futuro: o poder do Tesouro e a ampla aceitação pública e política da necessidade de restrição dos gastos (apesar do debate sobre o fato de a política fiscal atual ser muito pró-cíclica).

Finalmente, embora muitas das economias conseguidas por meio de revisões da despesa tenham sido feitas explicitamente – incluindo, em 2010, os programas que seriam abolidos –, em cada caso importantes áreas tinham sido imprecisas sobre a origem exata de uma parte das economias do orçamento reivindicadas pelo governo. Particularmente, os críticos atacaram tanto a Revisão de Gershon de 2004 quanto a CSR de 2007 por sua falha em detalhar todos os ganhos de eficiência que elas alegaram ter identificado.

4.7 França

Características básicas, objetivo e foco

A França oferece um estudo de caso de um recente processo ambicioso de revisão da despesa, cujos resultados reais são um pouco desapontadores. O processo em questão foi a Revisão Geral de Políticas Públicas (*Révision générale des politiques publiques* [RGPP]), constituída de duas rodadas: RGPP 1 (2007–08) e RGPP 2 (2010–11).

A RGPP, formalmente encerrada em maio de 2012, foi explicitamente proposta como um processo de revisão da despesa, mas também tinha dois objetivos mais amplos de melhoria de desempenho: a melhoria do atendimento aos clientes e a modernização da gestão de recursos humanos do serviço público. Na revisão dos gastos básicos, a RGPP excluiu essencialmente a consideração de novas propostas de gastos, salvo as opções de *gastar para economizar*.

O poder orçamentário na França está fortemente concentrado no poder executivo: o poder orçamentário recai principalmente no

presidente e no primeiro-ministro. A função do ministro das Finanças é dividida entre um ministro sênior – o ministro da Economia e das Finanças – e um ministro júnior – o ministro do Orçamento. Refletindo essa disposição, no nível burocrático existe o Ministério do Orçamento, que faz parte do *superministério* de Economia e Finanças.

O objetivo orçamentário da RGPP durante as duas rodadas, mais do que a criação de espaço fiscal para novos gastos, era a redução do crescimento projetado nos gastos agregados. Esse objetivo refletia o desejo do governo de corrigir a posição estruturalmente pouco sólida das finanças públicas. Para esse fim, o governo declarou em 2007 uma política de redução do serviço público por desgaste natural, que se baseou na não substituição de um em cada dois servidores públicos aposentados. Essa política proporcionou a estrutura do processo da RGPP (ver a citação a seguir).

A intenção inicial era que a RGPP 1 deveria ter um foco bem amplo tanto em ganhos de eficiência quanto em economias dos produtos. Na prática, no entanto, os planejadores decidiram quase imediatamente focar em ganhos de eficiência e, até de forma mais restrita, em ganhos de eficiência resultantes de reorganização administrativa. O escopo restrito da RGPP foi um foco importante de crítica, que gerou este comentário por parte do Escritório da Administração do auditor:

O processo da RGPP não constitui uma “revisão dos programas” do tipo empreendido em outros países. A RGPP deveria inicialmente ter focado a revisão dos fundamentos dos programas públicos a fim de evitar um foco unicamente em insumos e processos. Na prática, no entanto, o foco foi [...] em “fazer mais com menos”, sem reexaminar as missões dos programas. (COURS DES COMPTES 2010, p. 19-20).

A RGPP consistiu principalmente em revisões dos órgãos – não em revisões de programas ou processos –, embora um pequeno número de revisões horizontais de processos tenha sido realizado (p. ex., gestão de recursos humanos). A cobertura das revisões dos órgãos pretendia ser abrangente, começando na RGPP 1 com uma revisão dos próprios ministérios e, depois, estendendo-se na RGPP2 aos órgãos associados.

Box 4.8 – Estudo de caso: serviços centralizados

Uma das principais medidas de economia da RGPP era a criação de centros de serviços unificados (serviços centralizados) em alguns setores de prestação de serviços públicos. Por exemplo, esses centros fornecem serviços para os quais os clientes anteriormente eram obrigados a tratar com duas organizações ou mais.

- Centros de serviços unificados para questões fiscais, com 390 unidades estabelecidas em todo o país no final de 2010.
- Centros de serviços unificados para quem busca emprego, com a criação de uma nova organização (*Pôle emploi*) como resultado da fusão de duas empresas que antes já forneciam esses serviços.

A RGPP orientou-se pelo que eram, na verdade, metas quantitativas de ganhos de eficiência. Especificamente, a equipe de revisão foi solicitada a identificar ganhos de eficiência suficientes para garantir que a política de não substituição de um em cada dois servidores públicos aposentados não afetaria os serviços. Além dessa meta, a RGPP 2 definiu a de redução de 10% nos custos não salariais da administração (a ser atingida em 2013).

Na prática, a cobertura dos gastos da RGPP era pequena: visava principalmente aos custos operacionais e de pessoal. Os pagamentos de previdência social foram excluídos da revisão.

Processos e funções institucionais

O processo da RGPP foi supervisionado por uma Comissão de Monitoramento da RGPP (*Comité de suivi de la RGPP*) presidida conjuntamente pelos chefes dos escritórios do presidente e do primeiro-ministro, incluindo também o ministro do Orçamento. As revisões foram feitas por equipes de auditoria (*Equipes d'audit*) formadas pelo pessoal do Ministério do Orçamento e por consultores privados. Na primeira rodada da RGPP, as equipes de auditoria tiveram 11 meses para o trabalho e com mandatos muito bem definidos, que incluíam uma lista detalhada de perguntas-padrão (critérios de

revisão). Particularmente durante a RGPP 1, houve (i) pouca consulta por parte das equipes de auditoria aos ministérios que efetuam as despesas em revisão e (ii) pouca oportunidade para que os ministérios que efetuam as despesas apresentassem opções de reforma.

As equipes de auditoria apresentaram suas opções de economia ao Ministério do Orçamento para verificação antes de apresentá-las à liderança política. Elas foram, então, primeiro consideradas pela Comissão de Monitoramento da RGPP, que decidia quais opções seriam apresentadas ao organismo responsável pela decisão final, o *Conseil de la Modernisation des Politiques Publiques* (Conselho de Modernização), chefiado pelo presidente.

Base de informações da revisão da despesa

Em geral, a análise formal de gastos foi pouco usada no processo de RGPP. Durante a condução desse processo, houve uma percepção crescente de que a fragilidade da avaliação e de outra análise formal de gastos estava reduzindo a eficácia da revisão da despesa. Refletindo essa percepção, a RGPP tomou medidas para reformar o processo de avaliação, incluindo a criação de uma comissão interministerial de avaliação e de uma unidade responsável pelo desenvolvimento de avaliações. Infelizmente, pouco parece ter sido feito desde então para implementar tal decisão.

A França desenvolveu um amplo conjunto de indicadores de desempenho dos programas como um componente de reformas orçamentárias que entraram em operação em 2006. Esses indicadores incluem muito mais os de eficácia dos programas do que o geralmente observado em sistemas de orçamentos de programas. Conforme a afirmação do governo, os indicadores devem, em princípio, ter sido úteis para a RGPP. Na prática, porém, o estreitamento do foco da RGPP, bem como a não realização de revisões efetivas dos programas e da abordagem de economias dos produtos e ganhos de eficiência, significou que o valor potencial da estrutura de indicadores não foi percebido.

Em princípio, a classificação dos programas orientados a resultados do orçamento, que foi apresentada nas reformas orçamentárias de desempenho de 2006, deveria ter sido também uma ferramenta crucial em termos de revisão da despesa. A criação de programas

orçamentários deveria ter facilitado as revisões dos programas e o governo afirmou que este seria o caso – mais genericamente, afirmou que a RGPP foi criada diretamente com base na nova estrutura orçamentária de desempenho. Na verdade, isso parece contestável, ante a não realização de revisões efetivas de programas por parte da RGPP. Na prática, porém, essa falta de revisão de programas levou a RGPP a adotar uma abordagem tradicional centrada nos dados de entrada em relação a economias orçamentárias.

Avaliação

O governo alegou uma economia bruta total acumulada das duas rodadas do processo de RGPP, equivalente a 3,4% dos gastos públicos de 2011. Contudo, a natureza precisa da maior parte das economias não foi explicitada e a precisão da estimativa do governo em termos de economias foi amplamente contestada. O *Cours des Comptes* (Tribunal de Contas) considerou que a RGPP tinha *impacto orçamentário limitado*.

O processo de RGPP não teve um desempenho de acordo com seus objetivos ambiciosos (principalmente) por causa do apoio insuficiente para a contenção de gastos. Isso tem gerado pressão nos proponentes da RGPP para uma reformulação exclusivamente relacionada à eficiência administrativa e à qualidade de serviço.

No nível do processo, um grande problema foi a natureza excessivamente centralizada da RGPP e a participação limitada dos ministérios que efetuam despesas. Esse problema não só limitou o fluxo ascendente de possíveis opções de economia, mas também tornou a implementação das decisões de RGPP mais difícil, porque não havia um título de propriedade dos ministérios que efetuam despesas quanto às opções de economia identificadas.

4.8 Considerações finais

Observações gerais

Os estudos de caso analisados neste capítulo indicam que a revisão da despesa não é uma invenção recente. Alguns países têm realizado tais revisões de forma contínua ou periódica, há décadas.

A revisão da despesa é geralmente considerada como abordagem a uma grande deficiência na capacidade do processo orçamentário de atingir uma boa priorização de gastos; isto é, uma revisão insuficiente no cenário de referência de gastos.

Em cada um dos seis países analisados existem rotinas e requisitos bem desenvolvidos para o escrutínio de novas propostas de despesas no processo orçamentário. Apesar de esses processos poderem, sem dúvida, ser ainda mais aperfeiçoados, os países em questão geralmente consideraram a melhoria de processos para a revisão de cenários de referência de gastos ainda mais importante para o aumento da eficiência na alocação de recursos.

Os estudos de caso indicam que nem tudo o que se denominou revisão da despesa é, na verdade, uma revisão da despesa definida como a revisão de cenário de referência de gastos para fins orçamentários. Por exemplo, a maior parte das *Revisões da Despesa* britânicas focou principalmente a alocação de novas despesas e não incluiu uma revisão sistemática no cenário de referência de gastos.

Embora, em princípio, a revisão da despesa possa identificar opções para *aumentar* e *reduzir* os recursos financeiros para serviços existentes, na prática seu enfoque tem sido principalmente na identificação de opções de economia (o número limitado de exceções sobre esse enfoque é discutido mais adiante, neste documento). Isso é indiscutível, uma vez que não existe necessidade de uso do processo de revisão da despesa para garantir que as opções de aumento das despesas em serviços existentes sejam identificadas, já que os ministérios que efetuam despesas podem ser usados para a condução dessa etapa de modo sistemático.

Na maioria dos casos, as economias geradas das rodadas de revisão da despesa não ultrapassaram 2% a 3% das despesas do governo. Apesar de não serem tão grandes, as economias são, todavia, bastante úteis para fins de realocação ou controle de gastos agregados.

Ocasionalmente, porém, a revisão da despesa tem sido usada para a obtenção de cortes mais profundos. O exemplo mais recente é a Revisão Global das Despesas do Reino Unido em 2010, quando o governo estimou que reduziria os orçamentos departamentais e não a ajuda às áreas de saúde e além-mar em até uma média de

19% por mais de quatro anos. Voltando mais no tempo, a Revisão do Programa do Canadá, na metade da década de 1990, reduziu as despesas em cerca de 10% por mais de dois anos.

As economias geradas com a revisão da despesa não são instantâneas. Na maioria dos casos, são necessários vários anos para a implementação das opções de economia, mesmo com um processo dinâmico. No caso de ganhos de eficiência, novos processos precisam ser apresentados em conjunto com projetos de TI às vezes significativos. Tanto os ganhos de eficiência quanto as economias dos produtos geralmente requerem reduções de pessoal – e tais alterações demandam tempo para a devida ocorrência, principalmente se a segurança do emprego no serviço público significar que as reduções de pessoal precisam ser obtidas principalmente via contração natural.

A experiência destaca a natureza exigente do processo de revisão da despesa, mesmo se esta buscar ser seletiva em vez de global. A revisão bem-sucedida de despesas requer um compromisso que exige muito tempo por parte do pessoal do Ministério das Finanças (e quaisquer outros órgãos centrais relevantes). A revisão da despesa não pode simplesmente ser deixada para os ministérios que efetuam despesas, mesmo quando bastante envolvidos (ver outra discussão a seguir).

Na prática, a revisão da despesa não tem sido buscada exclusivamente como um instrumento para reduzir os gastos públicos agregados. A necessidade e a urgência de consolidação fiscal causaram, sem dúvida, maior uso da revisão da despesa para tal finalidade nos últimos dois ou três anos. Contudo, nas décadas anteriores, existem muitos exemplos de governos em busca da revisão da despesa, principalmente como um mecanismo que possibilitasse a criação de um espaço fiscal adicional para as novas despesas contempladas. Um exemplo notável é o processo de Revisão Estratégica do Canadá, em operação entre 2007 e 2010, quando o governo consistentemente enfatizou que todas as economias do processo de revisão estavam sendo reinvestidas.

Além disso, a revisão da despesa não está associada a determinado grupo no espectro político, mas tem sido usada tanto por governos conservadores quanto pelos de centro-esquerda.

A revisão da despesa não tem sido usada o suficiente durante os períodos de receitas favoráveis e expansão fiscal. Apesar de alguns governos terem continuado a operar tais sistemas durante os chamados bons tempos, o processo era geralmente menos intenso na ausência de uma grande pressão fiscal. Assim, por exemplo, o número de revisões realizadas a cada ano sob os sistemas holandês e dinamarquês diminuiu durante a década de 1990.

Foco e cobertura de revisão da despesa

Revisões seletivas de despesas são muito mais comuns do que revisões globais. Essa afirmação é relevante, apesar da recente série de revisões globais decorrentes de pressões de consolidação fiscal pós-crise. Conforme descrito na introdução deste capítulo, uma revisão *seletiva* de despesas está limitada a uma lista específica de tópicos de revisão, definidos no início de cada rodada. Uma revisão *global* das despesas, por outro lado, não está limitada por nenhuma lista desse tipo e visa ir mais longe na identificação das opções de economia.

As revisões globais das despesas são interessantes quando um governo deseja obter grandes reduções de gastos agregados em um curto período de tempo. Elas também funcionam quando um novo governo assume com prioridades de gastos radicalmente diferentes daquelas de seu predecessor e quando precisa criar um espaço fiscal considerável para suas próprias prioridades. Até mesmo uma revisão seletiva de despesas exige bastante dos funcionários do Ministério das Finanças. Uma revisão global das despesas, como as realizadas no Reino Unido em 2010 e no Canadá em 2011, é um processo verdadeiramente exaustivo durante o qual é difícil para o MF encontrar o tempo necessário para suas outras responsabilidades. As revisões globais das despesas são às vezes inevitáveis. Por exemplo, um governo que se sinta obrigado a implementar grandes reduções em um gasto agregado, geralmente desejará fazer tais reduções no início de seu período no cargo, em vez de arrastá-las ao longo dos anos. Contudo, a menos que tais motivos especiais existam, é melhor evitar as revisões globais.

Sob condições mais normais, a revisão seletiva de despesas é uma abordagem melhor. O interesse em evitar a pressão extrema de uma revisão global das despesas tem a vantagem agregada de que

as revisões de tópicos realizadas durante cada rodada do processo podem ser mais aprofundadas e de melhor qualidade.

A revisão da despesa é mais bem executada como um processo contínuo e não como um exercício pontual. Funciona melhor executar algumas revisões da despesa a cada ano ou a cada dois ou três anos. Essa é uma lição aprendida durante a década de 1980, quando revisões *ad hoc* de despesas em alguns países, incluindo o Reino Unido e os Estados Unidos, geraram resultados relativamente decepcionantes. Quando é um processo contínuo, ela pode apresentar um grupo de pessoal de revisão qualificado, em vez de depender de pessoas novas. Os processos de revisão também podem ser aprimorados e aperfeiçoados com o tempo. Além disso, um processo contínuo possibilita que cada rodada seja seletiva, obtendo, porém, uma cobertura global da revisão por alguns anos, conforme exemplificado na abordagem da Avaliação Estratégica do Canadá, que abrange todos os órgãos em um ciclo de quatro anos. Finalmente, com um processo contínuo de revisão da despesa, os países acham viável autorizar uma análise formal de gastos – que demanda algum tempo para sua realização, tal como as avaliações de resultados – para uso em futuras rodadas.

Se o governo desejar fazer grandes economias, a revisão da despesa não deverá ficar restrita à pesquisa para fins de ganhos de eficiência: deverá também buscar identificar economias dos produtos. A experiência ensina que, apesar de os governos gostarem muitas vezes de invocar as imensas economias que podem ser geradas com a redução do desperdício, em geral as revisões de eficiência não fornecem grandes reduções nos gastos, a curto prazo. Assim, se um governo desejar criar um espaço fiscal considerável para novas despesas ou para obter importantes reduções em gastos agregados, ele deverá também procurar identificar programas ineficazes ou de baixa prioridade e que possam ser abolidos ou reduzidos. Mesmo nos países mais bem geridos, números significativos de programas desse tipo são sempre encontrados, incluindo aqueles que podem ter sido importantes anos atrás, quando foram introduzidos, mas que agora deixaram de ter sua relevância.

Os sistemas de revisão da despesa mais bem-sucedidos têm sido os que revisam gastos obrigatórios e gastos do orçamento. A restrição da revisão da despesa para gastos do orçamento limita bastante as

economias que podem ser potencialmente identificadas, uma vez que exclui a segurança social e muitos outros pagamentos de transferência, responsáveis por uma grande parte dos gastos do governo na maioria dos países. No contexto brasileiro, logicamente, alguns gastos obrigatórios importantes surgem dos requisitos constitucionais muito difíceis de mudar, reduzindo assim de algum modo o possível escopo de revisão da despesa para fins de nova priorização dos gastos. Contudo, no Brasil, como em outros países, a maior parte dos gastos obrigatórios é exigida pela legislação ordinária – que pode ser alterada pelo parlamento. Mesmo onde as disposições constitucionais tornem obrigatório o uso de uma determinada quantia em dinheiro em um setor específico (p. ex., educação), a revisão da despesa permanece eficaz como uma ferramenta para garantir que os recursos sejam alocados aos programas mais úteis dentro do setor em questão.

A inclusão de transferências *não vinculadas* dentro do processo de revisão da despesa é provavelmente inapropriada (isto é, o uso de transferências não condicionadas criadas apenas para garantir que o governo subnacional tenha recursos suficientes para realizar suas responsabilidades; em outras palavras, para contrapor o denominado desequilíbrio fiscal vertical). Na medida em que o governo nacional usa subsídios *vinculados* (condicionados) ao governo subnacional com o objetivo de alcançar seus próprios objetivos políticos, a inclusão de tais transferências dentro do escopo de revisão da despesa parece convincente.

No contexto de um processo seletivo de revisão da despesa, a liderança política deve manter o direito de identificar programas e processos específicos que deseja ver examinados durante uma futura rodada. Geralmente, a liderança política e o MF consideram os programas ou processos específicos como carentes de uma revisão da despesa, por exemplo, como se sua eficácia e relevância tivessem se tornado uma questão de debate público. Por esse motivo, deverá haver sempre um mecanismo por meio do qual o Gabinete (ou outra instituição relevante) possa indicar os programas ou processos específicos a serem revisados.

Revisão da despesa e opções para novas despesas

Na maioria dos países, o processo de revisão da despesa tem focado exclusivamente a identificação das opções de economia e

tem sido mantido separadamente dos processos que consideram propostas de novas despesas.

Em alguns casos, porém, os processos têm permitido aos ministérios que efetuam despesas apresentar opções de novos gastos, financiados por economias geradas por meio da revisão. Um ótimo exemplo ocorreu no Canadá, onde os ministérios que efetuam despesas puderam apresentar, como parte de suas Revisões Estratégicas durante os quatro anos (2007 a 2010), as denominadas *opções de reinvestimento* que o governo poderia aceitar ou rejeitar. Esse mecanismo foi criado para fornecer um incentivo adicional aos ministérios que efetuam despesas para uma abordagem da tarefa de identificar as opções de economia de modo relevante e, ao mesmo tempo, convencê-los de que ao apresentarem opções atraentes de realocação, seriam capazes de manter alguns dos recursos financeiros ou todos os recursos que perderiam de outro modo, no processo de Revisão Estratégica.

As opções de *gastar para economizar* constituem uma exceção generalizada à proibição de propostas de novas despesas apresentadas como parte do processo de revisão. Alguns ganhos de eficiência podem ser obtidos apenas se houver gastos antecipados, tais como investimentos em tecnologias que economizam custos (p. ex., sistemas de TI que reduzem custos com mão de obra). Pelo fato de tal investimento ser tão óbvio, a maioria dos processos de revisão da despesa tem permitido a apresentação das opções de economia que exigem gastos antecipados significativos.

Funcionários públicos *versus* peritos externos no processo de revisão da despesa

Na maioria dos casos, a revisão da despesa nos países abordados neste capítulo é realizada principalmente por funcionários públicos. Apesar de o uso de peritos externos ser bem generalizado, na maioria dos casos eles atuam como assessores (ou, menos comumente, membros) das equipes de revisão, que são por si só controladas e predominantemente compostas por funcionários públicos. Nesse sentido, a prática contemporânea de revisão da despesa difere da abordagem usada em alguns países na década de 1980 – exemplificada pela Comissão Grace nos Estados Unidos, sob a qual as revisões da despesa eram realizadas em uma base *ad hoc* por comissões lideradas

por ilustres empresários e constituídas, principalmente, por pessoas de fora. A experiência naquele momento foi considerada bem decepcionante, em grande medida, porque as pessoas de fora não tinham o conhecimento detalhado suficiente do governo para fazer o trabalho do modo apropriado.

Por isso, hoje existe uma aceitação geral de que o serviço público deve desempenhar um papel central na condução das revisões da despesa. Essa abordagem faz sentido quando se trata de um processo contínuo, em vez de um exercício puramente pontual. Esse ponto é salientado pelos problemas que a Dinamarca – único país analisado onde consultores externos realizam grande parte do trabalho de revisão de tópicos – tem com a falta de um conhecimento contínuo dentro do MF.

Todavia, os peritos do setor privado podem potencialmente desempenhar um importante papel na busca de ganhos de eficiência. É principalmente na área de revisão de programas e, mais especificamente, na busca de economias dos produtos, que o papel do serviço público é essencial – de modo geral, é onde se pode encontrar o grau de conhecimento detalhado dos programas de governo necessários para apoiar esse tipo de revisão. Por sua vez, a revisão do processo, tal como a revisão nos processos de aquisição ou sistemas de TI, é uma área onde o trabalho de revisão da despesa é geralmente mais genérico e onde a capacidade técnica do setor privado é mais diretamente aplicável ao governo. Este é particularmente o caso em revisões direcionadas à identificação de opções para melhorar a eficiência de órgãos governamentais que fornecem serviços padrão de massa, tais como o fisco, a alfândega e os serviços de visto (aos quais o cientista político James Q. Wilson se referiu como órgãos de produção). Em órgãos desse tipo, as soluções de *reengenharia de negócios* desenvolvidas no setor privado são em geral igualmente aplicáveis ao governo, o que explica por que um importante exemplo recente de revisão da despesa executada principalmente por pessoas de fora – a revisão de eficiência de Gershon, no Reino Unido – é geralmente considerado como tendo sido um sucesso.

Para a revisão da despesa realizada principalmente pelos funcionários públicos ser bem-sucedida, é essencial que o MF (e quaisquer outros órgãos centrais envolvidos na gestão do

processo) tenha uma política sólida, além de qualificações e conhecimento técnico e financeiro. É possível melhorar a eficácia do processo de preparação do orçamento na realocação dos gastos somente se o Ministério das Finanças conseguir fornecer à liderança política sênior um assessoramento de alto nível em relação às prioridades e opções de economia. Se o MF for constituído quase exclusivamente por funcionários peritos em contabilidade e gestão financeira, mas que sabem pouco sobre política, o ministério não poderá desempenhar esse papel. Tem sido crucial para o sucesso dos processos contemporâneos de revisão da despesa a transformação dos MF em organizações com aptidões políticas mais sólidas, onde os funcionários responsáveis pelos orçamentos de ministérios específicos que efetuam despesas devem adquirir um entendimento detalhado das políticas e dos próprios programas.

MF e funções dos ministérios que efetuam despesas na condução da revisão da despesa

Na grande maioria dos casos, a gestão e a condução do processo de revisão da despesa são, em nível burocrático, de responsabilidade do MF ou órgão equivalente. Todavia, a atribuição de papéis na liderança burocrática dessa revisão deve refletir as estruturas institucionais do país em questão. É precisamente por essa razão que na França os gabinetes do presidente e do primeiro-ministro desempenharam papéis tão importantes na supervisão da RGPP. Do mesmo modo, na Austrália, o Departamento do Primeiro-Ministro e o Gabinete têm geralmente colaborado bastante com o Departamento de Finanças. Em geral, o mais importante para os órgãos centrais envolvidos na preparação do orçamento é realizar uma abordagem coordenada e cooperativa sobre a gestão do processo de revisão da despesa.

Nesse contexto, nenhum dos países destacados neste capítulo tem Ministério de Planejamento como o existente no Brasil e em muitos outros países. Se tal tipo de ministério existisse em algum dos seis países, eles obviamente precisariam estar envolvidos na gestão do processo de revisão da despesa.

Existem outras diferenças entre os países quanto aos respectivos papéis do MF (ou de outros órgãos centrais relevantes) e dos

ministérios que efetuam despesas na *condução real* de revisões de tópicos, isto é: no trabalho detalhado de análise de programas e processos para identificar as opções de economia. Em uma extremidade do espectro, as equipes de revisão da despesa da RGPP da França (*equipes de auditoria*) são compostas exclusivamente pelo pessoal do órgão central (MF) e peritos externos, sendo que as equipes não incluem representante dos ministérios que efetuam despesas. Na outra extremidade, o Canadá, cada ministério tem sido responsável pela realização de sua própria revisão, sem participação direta da Secretaria do Conselho do Tesouro na equipe de revisão. Entre esses dois extremos está a Holanda (e, até certo ponto, a Dinamarca), onde o trabalho de revisão da despesa é realizado por forças-tarefa conjuntas do MF e dos ministérios que efetuam despesas.

É um erro do Ministério das Finanças tentar realizar revisões de tópicos sem a participação direta dos ministérios que efetuam despesas. Estes não só têm um conhecimento detalhado incomparável de seus próprios programas e processos, como também são os que precisam implementar todas as opções de economia que o governo decidir adotar. Se os ministérios que efetuam despesas pelo menos não entenderem a lógica das opções de economia que devem implementar, a implementação poderá se revelar bastante difícil. Esse é o motivo principal pelo qual a abordagem centralizada do processo de RGPP na França foi subsequentemente reconhecida como um engano.

Alguns analistas poderiam sugerir que deixar a condução das revisões para os ministérios que efetuam despesas torna todo o processo dependente do empenho desses ministérios em participar com a apresentação de sólidas opções de economia. Afinal, os ministérios que efetuam despesas tendem a proteger seus orçamentos e programas. Assim, por que eles deveriam oferecer unilateralmente seus próprios programas para serem cortados? O problema da falta de cooperação dos ministérios que efetuam despesas surge mesmo no contexto da abordagem das forças-tarefa conjuntas (conforme experiência demonstrada na Holanda). A questão geral levantada diz respeito a quais pressões ou incentivos pôr em prática para garantir a cooperação dos ministérios que efetuam despesas (ver outra discussão a seguir).

Se os ministérios que efetuam despesas conduzirem suas próprias revisões da despesa sem a participação direta do MF, este deverá desempenhar a função de contestação das propostas desses ministérios e, assim, ser obrigado a apresentar opções alternativas de economia. Essas opções de economia precisam ser testadas de modo rigoroso e independente, para fins de verificação de sua viabilidade e impacto orçamentário, e esse papel é mais bem desempenhado pelo MF, juntamente com quaisquer outros órgãos centrais relevantes.

A criação de uma unidade específica de revisão da despesa dentro do Ministério das Finanças pode facilitar bastante o processo. Essas unidades existem em alguns dos países destacados, mas não em todos. Logicamente, nenhuma unidade de revisão pode efetuar sozinha todo o trabalho do MF na revisão da despesa. Os analistas de orçamento do setor devem necessariamente desempenhar uma função importante.

Garantia de cooperação dos ministérios que efetuam despesas

Mecanismos para incentivar a cooperação dos ministérios que efetuam despesas na identificação das melhores opções de economia são essenciais. A experiência demonstra que a cooperação pode ser incentivada por uma mistura das seguintes medidas:

- Definir objetivos para o valor mínimo das opções de economia a serem encontradas por cada revisão de tópico ou ministério que efetua despesas. A maioria dos países que usaram essa abordagem definiu objetivos mínimos uniformes que se aplicam a todas as revisões de tópicos ou ministérios (p. ex.: 5% para todos os ministérios). Contudo, é possível, principalmente durante uma rígida revisão global das despesas, definir objetivos diferenciados, em que se exige dos ministérios com menor prioridade a identificação de mais opções de economia do que daqueles com maior prioridade.
- Ter pressão dos principais líderes políticos (p.ex.: presidente, primeiro-ministro ou Gabinete) nos ministérios que efetuam despesas e que não estiverem cooperando. Mesmo quando objetivos mínimos de economia são definidos, os ministérios que efetuam despesas podem *jogar*, fornecendo opções óbvias e politicamente impossíveis. Assim, é essencial que qualquer

ministério que efetue despesas e se comporte desse modo seja imediatamente disciplinado pela liderança política.

- Permitir que os ministérios que efetuam despesas apresentem opções de realocação, juntamente com suas opções de economia. Essa medida fornece a esses ministérios certa esperança de poderem recuperar uma parte significativa de quaisquer reduções em seus orçamentos causadas pelo processo de revisão da despesa. Além disso, pelo fato de o valor total das opções de realocação apresentadas por um ministério específico que efetua despesas não poder exceder o valor das opções de economia apresentadas por tal ministério, essa abordagem não significa abrir a porta para inúmeras propostas de novas despesas.
- Ter o envolvimento do MF na identificação e análise de opções de economia. Conforme discutido anteriormente, a existência de assessoria por parte de um perito independente para contestar os ministérios que efetuam despesas criará uma pressão adicional considerável nesses ministérios por melhoria na qualidade das opções de economia por eles apresentadas.

Revisão da despesa e a liderança política

As diferenças institucionais entre os países tornam impossível generalizar sobre o papel apropriado de instituições específicas em nível político, tais como o Gabinete ou parlamento, no processo de revisão da despesa. Pode-se dizer que a tomada de decisões políticas e orçamentárias deve desempenhar um papel importante no processo.

A revisão da despesa não pode ser bem-sucedida sem um sólido apoio dos principais tomadores de decisões políticas e orçamentárias. Conforme mencionado na seção anterior, a pressão da liderança política é essencial para garantir a cooperação de ministérios que efetuam despesas. A liderança política deve também desempenhar um papel central na promoção do serviço público e na compreensão pública dos objetivos gerais que o governo está procurando obter via revisão da despesa. Se, por exemplo, o objetivo principal for redefinir as prioridades dos gastos, a liderança política deverá explicar ao público com cuidado que o exercício não visa à redução de serviços, mas é essencial para o financiamento de iniciativas das novas despesas contempladas.

Cabe à liderança política tomar a decisão final sobre quais opções de economia devem ser implementadas. Essa abordagem é particularmente relevante em relação às decisões de eliminação ou redução de programas, mas também se aplica às opções mais importantes de ganhos de eficiência. Contudo, o modo como as opções de economia são apresentadas aos líderes políticos para fins de decisão e a forma como as decisões são na verdade tomadas variam entre os países.

A base de informações para revisão da despesa

A disponibilidade de boas informações é uma questão importante na revisão da despesa. Do ponto de vista do MF e de outros órgãos centrais envolvidos, os problemas na obtenção de informações suficientes sobre a eficácia, a eficiência e a relevância de programas e processos em revisão limitam sua capacidade de tomar decisões fundamentadas sobre opções de economia.

Na maioria dos países analisados, os ministérios das Finanças enfatizam a necessidade de melhoria da quantidade e qualidade na análise formal de gastos, inclusive por meio de mais avaliação e análise de eficiência, a fim de dar suporte ao processo de revisão da despesa. Em relação à avaliação, o problema ainda é o fato de muito pouco estar sendo realizado, ou de grande parte da avaliação realizada pelos ministérios que efetuam despesas não ser muito útil para fins de tomada de decisões orçamentárias.⁵ Os limites de curto prazo aplicados às revisões da despesa geralmente impossibilitam a realização de uma análise formal de gastos, o que, em muitos casos, demanda um tempo considerável⁶ como parte do próprio processo de revisão da despesa. Por isso, é importante desenvolver uma análise mais formal de gastos que possa ser realizada separadamente, e antes da revisão da despesa – mas que seja criada para funcionar como uma entrada ao processo.

5 Um motivo para isso é que as avaliações autorizadas pelos ministérios que efetuam despesas tendem a ser mais focadas na orientação dos ministérios sobre como melhorar a formulação e gestão dos programas do que na orientação do Ministério das Finanças sobre o fato de os programas precisarem ser cortados ou não.

6 Por exemplo, um espaço de 18 meses é considerado tempo muito curto para uma avaliação dos resultados.

A disponibilidade de bons indicadores de desempenho é crucial para o desenvolvimento de uma boa análise formal de gastos, a fim de dar suporte ao processo de revisão da despesa. Contudo, é necessário ser claro sobre o potencial papel dos indicadores e sobre suas limitações. Os indicadores de desempenho, por si só, raramente fornecem informações claras e conclusivas em termos de eficácia e eficiência. Eles precisam passar principalmente por uma análise formal de gastos, antes de poderem fornecer a base de informações necessárias a uma boa revisão da despesa.

A classificação dos programas do orçamento – parte integrante do sistema de orçamento – facilita a revisão da despesa, que inclui a revisão dos programas e pesquisas de economias dos produtos, apesar de não ser tão essencial para a revisão focada exclusivamente em ganhos de eficiência. Isso ocorre porque uma boa classificação dos programas de orçamento, baseada em resultados, identifica de imediato os programas e subprogramas em que a revisão da despesa precisará de enfoque. Além disso, indica quanto dinheiro está sendo gasto nos programas. O MF não pode efetuar uma boa revisão da despesa, exceto se souber exatamente os serviços (e transferências) prestados por cada ministério que efetua despesas e quanto está sendo gasto em cada um desses serviços. Em um sistema orçamentário tradicional, o MF geralmente não sabe com precisão o que cada ministério faz, além daqueles evidentes (educação escolar no caso do Ministério da Educação, tratamentos médicos no caso do Ministério da Saúde). Um orçamento dos programas torna esse conhecimento bem mais claro.

Sob as condições econômicas atuais, a revisão da despesa está se tornando cada vez mais importante. Além dos países analisados neste capítulo, muitos outros estão nos estágios iniciais de desenvolvimento de processos de revisão da despesa. Desse modo, as técnicas usadas nas revisões da despesa serão provavelmente aperfeiçoadas nos próximos anos.

Tabela 4.1 – Tabela comparativa das práticas mais recentes de revisão da despesa em seis países da OCDE

	Holanda	Canadá	Austrália	Dinamarca	Reino Unido	França
Nome do processo de revisão da despesa	Revisão Global de Despesas (CER)	Revisão Estratégica e Operacional (SOR)	Revisão Global de Despesas (CER)	Estudos Especiais	Revisão Global das Despesas (CSR)	Revisão Geral de Políticas Públicas (RGPP)
Última rodada	2010	2011	2008-2010	2011-2012	2010	2010-2011
Contexto da política fiscal	Consolidação fiscal	Consolidação fiscal	Consolidação fiscal	Consolidação fiscal	Consolidação fiscal	Incerta
Objetivos + foco						
Cobertura	Seletiva	Global	Global	Seletiva	Global	Global
Melhoria do desempenho?	Não	Sim	Não	Não	Sim	Sim
Objetivo principal	Reduzir despesas agregadas	Reduzir despesas agregadas	Reduzir despesas agregadas	Reduzir despesas agregadas	Reduzir despesas agregadas	Reduzir despesas agregadas
Foco em ganhos de eficiência?	Não	Sim	Sim	Sim	Sim	Sim
Foco em economias dos produtos?	Sim	Sim	Sim	Não	Sim	Não
Principal tipo de revisão	Programa	Órgão	Programa	Programa	Órgão	Órgão
Objetivos de economia	Sim	Sim	Não	Sim	Não	Sim
Funções institucionais						
<i>Escolha ex ante de tópicos de revisão</i>	Gabinete	n.a.	Comissão de Gabinete	Comissão de Gabinete	n.a.	n.a.
Forças-tarefa conjuntas de revisão?	Sim	Não	Não	Sim	Não	Não
Identificação de opções de economia	Forças-tarefa conjuntas	Ministérios que efetuam despesas	Ministérios que efetuam despesas MF	Forças-tarefa conjuntas	Ministérios que efetuam despesas MF	MF
Base de informações						
Análise de gastos	Principalmente informais	Principalmente informais	Principalmente informais	Principalmente informais	Principalmente informais	Principalmente informais
Uso de indicadores de desempenho	Sim	Sim	Sim	Sim	Sim	Sim
Amplo sistema de avaliação do governo	Sim	Sim	Não	Sim	Não	Não

Fonte: Elaboração própria.
Nota: n.a. = não se aplica.

REFERÊNCIAS

Gerais

AUSTRALIA GOVERNMENT. *Report of the review of the measures of agency efficiency*. Canberra: Australian Government, 2011.

DIAMOND, J. *Budget system reform in emerging economies: the challenges and the reform agenda*. Washington, DC: International Monetary Fund, 2006. (IMF Occasional Paper n. 245).

DOGGETT, J. *Beyond the blunt instrument: the efficiency dividend and its alternatives*. Sydney: Centre for Policy Development, 2010. (Occasional Paper n. 11).

GILMOUR, J. *Implementing OMB's program assessment rating tool (PART): meeting the challenges of integrating budgeting and performance*. Washington, DC: IBM Centre for the Business of Government, 2006.

LUINAUD, M.; WILHELMSSON, T. *Literature review: automatic cuts of productivity dividends*. Paris: Organisation for Economic Co-operation and Development, 2012. (Working Paper).

MARK, K.; PFEIFFER, J. R. *Monitoring and evaluation in the United States government: an overview*. Washington, DC: World Bank, 2011. (Evaluation Capacity Development Working Paper n. 26).

ORGANIZATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD). *Value for money in government: public administration after "new public management"*. Paris: OECD, 2010.

_____. *Typology and implementation of spending reviews: discussion paper*. Paris: Organisation for Economic Co-operation and Development, 2011. (Working Paper).

WHITE, J. *Playing the wrong PART: the program assessment rating tool and the functions of the President's Budget*. *Public administration review*, v. 72, n. 1, p. 112-121, 2011.

WORLD BANK. *Monitoring and evaluation: some tools, methods, and approaches*. Washington, DC: World Bank, Operations Evaluation Department, 2004.

Austrália

BARRETT, C. *Australian expenditure and strategic reviews*. Paris: Organisation for Economic Co-operation and Development Network on Performance and Results, nov. 2011.

COMMONWEALTH OF AUSTRALIA. *Strategic review of indigenous expenditure*. Canberra: Commonwealth of Australia, 2010.

_____. *Strategic review of small and medium agencies in the attorney-general's portfolio*. Canberra: Commonwealth of Australia, 2012.

DEPARTMENT OF FINANCE AND DEREGULATION (DOFD). *Evaluation and strategic reviews*. Current Issue Brief, 2010.

_____. *Expenditure review principles*. 2012. Disponível em: <<http://www.finance.gov.au/budget/budget-process/expenditure-review-principles.html>>.

GERSHON, P. *Review of the Australian government's use of Information and communication technology*. Canberra: Commonwealth of Australia, 2008.

GRANT, P. *Strategic review of the administration of Australian government grant programs*. Canberra: Commonwealth of Australia, 2008.

MACKAY, K. *The development of Australia's evaluation system*. Washington DC: World Bank, 1998. (Evaluation Capacity Development Working Paper n. 4).

_____. *Two generations of performance evaluation and management system in Australia*. Washington DC: World Bank, 2004. (Evaluation Capacity Development Working Paper n. 11).

WANNA, J.; KELLY, J.; FORSTER, J. *Managing public expenditure in Australia*. St. Leonards, Australia: Allen & Unwin, 2001.

Canadá

BOURGON, J. *Program review: the government of Canada's experience eliminating the deficit, 1994-99 – a Canadian case study*. London: Institute for Government, 2009.

ENNS, D. *Strategic reviews: lessons learned and future directions*. Ottawa: Financial Management Institute of Canada, nov. 2010.

INTERVIEW with David Enns, Deputy Assistant Secretary, Expenditure Management, Treasury Board Secretariat of Canada, set. 2012.

GOOD, D. A. *Politics of public money: spenders, guardians, priority setters, and financial watchdogs inside the canadian government*. Toronto: University of Toronto Press, 2007.

GOVERNMENT OF CANADA. *The budget plan 2008: responsible leadership*. Ottawa: Public Works and Government Services Canada, 2008.

_____. *Canada's Economic Action Plan: Budget 2009*. Ottawa: Public Works and Government Services Canada, 2009.

_____. *Canada's Economic Action Plan Year 2: Budget 2010*. Ottawa: Public Works and Government Services Canada, 2010.

_____. *Jobs, growth, and long-term prosperity: Economic Action Plan 2012*. Ottawa: Public Works and Government Services Canada, 2011a.

_____. *The next phase of Canada's Economic Action Plan: a low-tax plan for jobs and growth*. Ottawa: Public Works and Government Services Canada, 2011b.

TREASURY BOARD SECRETARIAT (TBS). *Evaluation function in the government of Canada*. 2004. Disponível em: <http://www.tbs-sct.gc.ca/cee/pubs/func-fonc-eng.asp#s2_3>.

_____. *Strategic reviews website*. 2008. Disponível em: <<http://www.tbs-sct.gc.ca/sr-es>>.

_____. *Policy on evaluation*. 2009. Disponível em: <<http://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=15024§ion=text>>.

Dinamarca

BLÖNDAL, J. R.; RUFFNER, M. Budgeting in Denmark. *OECD Journal on Budgeting*, v. 4, n. 1, p. 49-79, 2004.

GINNERUP, R.; JØRGENSEN, T. B.; JACOBSEN, A. M.; REFSLUND, N. Performance Budgeting in Denmark. *OECD Journal on Budgeting*, v. 7, n. 4, p. 1-24, 2007.

INTERVIEW with David Fjord Nielsen, Chief Consultant, Danish Ministry of Finance, set. 2012.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD). *Value for money: country assessment of Denmark*. Handout

at the 32nd Annual Meeting of OECD Senior Budget Officials. Luxembourg City, jun. 2011.

França

COUR DES COMPTES. *Rapport sur la situation et les perspectives des finances publiques*. Paris, 2009.

_____. *Rapport sur la situation et les perspectives des finances publiques*. Paris, 2010.

_____. Finances publiques. In: _____. *Rapport annuel 2010*. Paris, 2011a.

_____. *Rapport sur la situation et les perspectives des finances publiques*. Paris, 2011b.

_____. *La mise en oeuvre de la loi organique relative aux lois de finances (LOLF): un bilan pour de nouvelles perspectives*. Paris, 2011c.

_____. *La situation et les perspectives des finances publiques*. Paris, 2012.

DIRECTION GÉNÉRALE DE LA MODERNISATION DE L'ÉTAT (DGME). RGPP: Sixième Conseil de Modernisation des Politiques Publiques. *Les Cahiers de la DGME 5*, DGME, Paris, 2011.

DUBERTRET, J. *Révision générale des politiques publiques (RGPP) et politique budgétaire: Mise en oeuvre concrète d'une 'Spending review' en France*. Paris: Organisation for Economic Co-operation and Development Network on Performance and Results, nov. 2011.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD). *Examens de l'OCDE sur la gouvernance publique: France, une perspective internationale sur la révision générale des politiques publiques*. Paris: OECD, 2012.

PRIMAT, H. *Cinq ans de revue générale des politiques publiques (RGPP): un échec politique et administrative*. Paris : Terra Nova, 2011.

RÉPUBLIQUE FRANÇAISE. *Révision générale des politiques publiques: guide méthodologique*. Paris : République Française, 2007.

_____. *RGPP: 2ème rapport d'Etape au président de la république*. 2009. Disponible em: <<http://www.modernisation.gouv.fr/>>.

_____. *RGPP: 2ème rapport d'Etape au président de la république*. 2010a. Disponible em: <<http://www.modernisation.gouv.fr/>>.

_____. *RGPP: 4ème Conseil de Modernisation des Politiques Publiques*, juin 2010. 2010b. Disponible em: <<http://www.modernisation.gouv.fr/>>.

_____. *RGPP: 5ème Conseil de Modernisation des Politiques Publique*, mars 2011. Disponible em: <<http://www.modernisation.gouv.fr/>>.

Holanda

BERGER, B. *Spending cuts without the cheese slicer*. Public Financial Management Blog, dez. 2011. Disponible em: <<http://blog-pfm.imf.org/>>.

DEBETS, R. Performance budgeting in the Netherlands. *OECD Journal on Budgeting*, v. 7, n. 4, p. 1-20, 2009.

DEBETS, R.; VOSSERS, H. *Program budgeting in the Netherlands*. Disponible em: <http://siteresources.worldbank.org/INTMEXICO/Resources/H.Volseer_Paper.pdf>.

INTERVIEW with Peter Van den Berg, Deputy Budget Director, Netherlands Ministry of Finance, ago. 2012.

NETHERLANDS MINISTRY OF FINANCE. *Budget practices in the Netherlands*. Inspectorate of Budget, Office of Policy Research, Amsterdam, 2010. Disponible em: <<http://www.government.nl/documents-and-publications/leaflets/2010/07/27/publication-budget-practices-in-the-netherlands.html>>.

VAN DEN BERG, P. *Spending reviews*. Paris: Organisation for Economic Co-operation and Development Network on Performance and Results, nov. 2011.

Reino Unido

CABINET OFFICE. *Commentary to back office benchmark information 2009/10*. London: Cabinet Office Efficiency and Reform Group, 2010.

GERSHON, P. *Releasing resources to the front line: independent review of public sector efficiency*. London: H.M. Stationery Office, 2004.

H. M. TREASURY. *Releasing the resources to meet the challenges ahead: value for money in the 2007 comprehensive spending review*. Cm 6889. London: The Stationery Office, 2006.

_____. *2007 comprehensive spending review: guidance for departmental submissions*. London: H.M. Treasury, 2007a.

_____. *Meeting the aspirations of the british people: 2007 pre-budget report and comprehensive spending review*. Cm 7227. London: H.M. Stationery Office, 2007b.

_____. *The spending review framework*. Cm 7872. London: H.M. Stationery Office, 2010a.

_____. *Spending review 2010*. Cm 7942. London: H.M. Stationery Office, 2010b.

NATIONAL AUDIT OFFICE. *The efficiency programme: a second review of progress*. London: H.M. Stationery Office, 2007.

TREASURY COMMITTEE, HOUSE OF COMMONS. *The 2007 comprehensive spending review: prospects and processes sixth report of session 2006-07*. London: H.M. Stationery Office, 2007.

_____. *Spending review 2010: sixth report of session 2010-11*. London: H.M. Stationery Office, 2010a. (v. I Report, Together with Formal Minutes).

_____. *Spending review 2010: sixth report of session 2010-11*. London: H.M. Stationery Office, 2010b. (v. II, Oral and Written Evidence).



Capítulo 5

Revisão das Despesas Públicas

Ian Lienert

5.1 Introdução

A Revisão das Despesas Públicas (PER¹) é um estudo diagnóstico que ajuda os países a compreenderem desafios econômicos e da despesa, proporcionando perspectivas para orientar o diálogo sobre políticas e recomendar mecanismos mais eficazes para alocar recursos públicos disponíveis. As PERs têm sido parte do trabalho econômico e setorial do Banco Mundial há mais de quatro décadas e são feitas em colaboração com os países. Oferecem um insumo importante para o planejamento público, processos orçamentários e tomada de decisões estratégicas, e têm contribuído para reformas fiscais e de políticas adotadas em países de renda baixa e média.²

Os objetivos deste capítulo são:

- Discutir mais plenamente o que é uma PER e descrever sua estrutura inerente típica;
- Examinar brevemente a evolução das PERs nos últimos 40 anos;
- Fazer comentários sobre o conteúdo variado das PERs e a diversidade das questões sobre despesa analisadas;
- Comparar as PERs com revisões da despesa;³

1 A expressão *Revisão de Despesas Públicas* é a tradução literal da expressão inglesa *Public Expenditure Review (PER)*.

2 O Banco Mundial normalmente não proporciona assessoramento em política nem realiza PERs em países de alta renda. Entretanto, já foram feitas PERs em alguns países de renda mais alta, incluindo a Polônia e a República Eslovaca.

3 Ver Capítulo 4 deste livro, o qual examina experiências em revisão da despesa em seis países da Organização de Cooperação e Desenvolvimento Econômico (OCDE).

- Ilustrar, em quatro estudos de casos, como foram abordadas questões de despesa agregada e de setores específicos; e
- Examinar lições aprendidas na preparação das PERs e melhorar sua eficácia.

A Seção 5.2 examina a questão *O que é uma PER?* e descreve o escopo e a estrutura típicos da análise dos gastos públicos. A seguir, discute como as PERs evoluíram. A Seção 5.3 analisa mais plenamente questões relacionadas com o orçamento e a despesa, comumente examinadas nas PERs. A Seção 5.4 compara as PERs com revisões da despesa realizadas em alguns países da Organização de Cooperação e Desenvolvimento Econômico (OCDE). A Seção 5.5 examina certas PERs recentes em vários países de renda média: México, Peru, Federação Russa e Turquia. Esses casos de países ilustram como as PERs têm abordado tanto questões macrofiscais (incluindo criação de espaço fiscal para novas prioridades da despesa) quanto à eficiência da despesa em certos setores (a saber, educação, saúde e infraestrutura). Cada subseção resume as principais constatações e as recomendações dessas quatro PERs. A Seção 5.6 utiliza as principais lições aprendidas da realização das PERs levando em consideração os elementos e uma PER de alta qualidade, bem como os fatores que aumentam a eficácia das PERs.

Este capítulo também abrange as principais questões e métodos para realizar uma PER e refere-se a documentos prontamente disponíveis para orientar a preparação de uma PER. As conclusões da Seção 5.7 ressaltam a diversidade de questões examinadas nas PERs realizadas em muitos países, que vão desde grandes economias – como o Brasil, cujas instituições orçamentárias são bem desenvolvidas – até países de baixa renda, com capacidade técnica e administrativa limitada para implementar reformas de longo alcance de políticas e instituições do gasto público propostas nas PERs.

5.2 O que é uma PER, seu escopo e estrutura de análise?

Ferramenta para analisar a alocação e a eficiência dos gastos

Normalmente uma PER analisa o nível e a alocação dos gastos públicos, avalia a sua eficácia e a equidade e identifica gargalos no intuito de melhorar a eficácia da despesa. As PERs fazem

recomendações para aprimorar a gestão orçamentária, as alocações de gastos entre setores amplos e, especialmente, a eficiência da despesa dentro dos setores ou subsetores. As opções e sugestões oferecidas nas PERs lançam luz sobre o desenvolvimento estratégico de um país e as escolhas de políticas da despesa, bem como a estratégia própria do Banco Mundial de assistência aos países e o seu trabalho contínuo no setor econômico.

Estrutura de uma análise de PER

Uma PER é formada sobre uma estrutura analítica, cujo conteúdo substantivo incluiria uma combinação dos seguintes componentes:⁴

- **Análise de médio prazo da sustentabilidade fiscal e da renda:** discussão sobre a evolução de receitas agregadas, despesa e equilíbrios fiscais do setor público consolidado. A análise pode incluir uma avaliação que considera se o país está gerando renda adequada sem distorções, equitativa e sustentável;
- **Combinação público-privada de fornecimento de bens e serviços:** análise da justificação da intervenção pública após identificação de questões relacionadas com a falha do mercado e redistribuição;
- **Alocação estratégica de recursos:** avaliação de prioridades do gasto público – entre as funções e dentro de cada uma – dados as restrições dos recursos e objetivos da distribuição;
- **Eficiência e eficácia da despesa:** análise para saber se os resultados da despesa pública estão sendo alcançados de forma eficaz. Essa análise pode incluir um exame da combinação de insumos (p.ex., remunerações *versus* operações e manutenção) ou as alocações de capital *versus* despesas correntes dentro dos setores ou subsetores;
- **Incidência e equidade:** o impacto de programas da despesa sobre grupos alvo, especialmente pessoas de baixa renda, e se a despesa é equitativa; e

4 Os seus seis marcadores baseiam-se em Pradhan (1996) e na obra *What is a good PER: the substance* (O que é uma boa PER: a substância), disponível no *website* do Banco Mundial: <<http://web.worldbank.org/EXTERNAL/TOPICS/EXTPUBLICSECTOR>>.

- **Instituições:** o grau em que as instituições e processos orçamentários promovem a disciplina fiscal, eficiência técnica e das alocações, e equidade na composição da despesa.

Apesar do uso generalizado das PERs no Banco Mundial nas últimas três décadas, não há diretrizes formais ou obrigatórias para sua implementação, em parte devido à ampla heterogeneidade de questões e enfoques das PERs. No entanto, o Banco Mundial tem procurado documentar essas práticas. Por exemplo, algumas diretrizes preliminares para trabalhos de gasto público (BANCO MUNDIAL, 2001) recomendaram uma estratégia bem articulada para a Análise e Suporte da Despesa Pública (PEAS⁵). Cada PEAS selecionaria algumas ou todas as seis questões anteriores destacadas ao formular a estratégia de análise da despesa. O conteúdo de cada PER seria adaptado ao seguinte: (i) disposição do governo de realizar uma PER; (ii) capacidade do governo de implementar as recomendações da PER; (iii) solidez das instituições de supervisão interna (legislativo, auditoria externa, mídia, grupos de cidadãos); e (iv) grau de dependência da ajuda.

Ampliação do escopo e aumento da variedade de questões abordadas nas PERs

As PERs foram originalmente elaboradas como ferramenta de avaliação interna do Banco Mundial. Em 1979 foram realizadas somente três PERs, em comparação com cerca de 40 por ano, hoje. No correr do tempo, a PER evoluiu para incluir questões além da análise da despesa nos níveis agregados e setoriais. Embora muitas PERs ainda enfoquem tais questões, algumas PERs recentes – especialmente as preparadas nos países de renda média – examinam questões fiscais mais amplas, inclusive sustentabilidade fiscal ou espaço fiscal. As PERs atualmente atribuem mais ênfase ao papel das instituições orçamentárias, tais como a estrutura jurídica, dispositivos organizacionais ou as regras dos procedimentos orçamentários. As questões de economia política também são levadas em consideração, porque as PERs podem discutir questões politicamente sensíveis, tais como reformas de subsídios.

5 A expressão *Análise e Suporte da Despesa Pública* é a tradução literal da expressão inglesa *Public Expenditure Analysis and Support (PEAS)*.

Em algumas PERs utiliza-se uma visão abrangente do governo: são analisadas não somente questões orçamentárias e da despesa do governo central, mas também alocações extraorçamentárias ou subnacionais, porque frequentemente são importantes para os resultados das políticas. Nessas PERs o escopo da análise pode também incluir instituições do setor público mais amplo. Em certos países, isso reflete a importância de empresas estatais.⁶ Finalmente, os acordos-limite público-privados, tais como parcerias público-privadas, são discutidos nas PERs que examinam o alcance do financiamento privado das políticas públicas.

Algumas PERs enfatizam questões de *microdespesas*, especialmente ao examinarem a eficiência da despesa. Quando o objetivo é identificar melhorias no efeito da despesa sobre a prestação de serviços públicos, o enfoque pode estar na eficácia da despesa e não em sua eficiência. Por exemplo, os estrangulamentos do fluxo de recursos para os beneficiários finais podem ser examinados por meio de levantamentos de rastreamento do gasto público (PETS) para identificar as limitações da oferta e da procura nas unidades de prestação de serviços.

Em suma, é amplo o alcance das questões examinadas em diferentes PERs, indo da análise microfiscal a uma análise detalhada de um ou mais setores da despesa pública. Na prática, o conteúdo de cada PER adapta-se às necessidades de cada país.

5.3 Questões tipicamente abordadas nas PERs e ferramentas analíticas específicas

Questões macrofiscais: Espaço fiscal e estratégia orçamentária de médio prazo

Em alguns países – especialmente nas economias emergentes –, as questões de sustentabilidade e espaço fiscais revestem-se de importância especial, dado que a despesa adicional frequentemente resulta em um aumento dos déficits fiscais e da dívida pública. Nos países de renda média, a PER pode examinar as opções de políticas disponíveis para a criação de espaço fiscal voltado a prioridades urgentes da despesa pública, assegurando ao mesmo tempo que

6 As empresas estatais são especialmente importantes nas PERs que enfocam a infraestrutura. Por exemplo, o setor elétrico do México e o setor ferroviário da Rússia são dominados por empresas estatais (ver respectivas PERs, resumidas adiante, neste capítulo).

a posição fiscal de médio prazo do país seja sustentável. Na PER os cenários fiscais de médio prazo, preparados sob várias hipóteses, podem proporcionar aos formuladores de políticas escolhas para nova priorização de políticas do gasto e medidas tributárias.

Neste contexto, algumas PERs dedicam um capítulo à análise do sistema tributário. Essa análise examina frequentemente o escopo da melhoria das receitas. As despesas tributárias também podem ser analisadas. Um tema recorrente nesse tipo de PER é a necessidade de simplificar o sistema tributário e ampliar sua base, especialmente mediante a redução de isenções e outros privilégios fiscais. Algumas PERs enfocam determinados impostos, inclusive o escopo do aumento ou redução de certas taxas tributárias, ao passo que outras examinam medidas para aumentar a receita por meio de melhor administração de impostos e de não impostos.

Questões institucionais no nível público central

As responsabilidades institucionais da gestão orçamentária central variam entre os países. No ramo executivo do governo, alguns países têm ministérios distintos para o planejamento, orçamento, e gestão e contabilidade financeiras públicas. No entanto, não há consenso se um país deve ter ministérios distintos para o planejamento, orçamento e finanças ou ter apenas um *superministério*. Há argumentos em favor de fundir as funções de planejamento e orçamentação sob um único ministério, especialmente se o país tiver uma Estrutura dos Gastos de Médio Prazo (MTEF), na qual os orçamentos correntes e de capital estejam integrados. Entretanto, a fusão de dois ou mais ministérios centrais pode ser impossível devido à existência de limitações legislativas ou políticas que impedem essa combinação.⁷ Nesses países, a PER pode aceitar o sistema existente de responsabilidades centrais pela gestão orçamentária. No entanto, se houver possibilidade de escolhas políticas, a PER pode propor opções para acordos sobre mudança organizacional.

7 Na Índia, por exemplo, a constituição determina uma distinção entre despesas correntes e de capital. Em países com governos de coalizão, os presidentes podem manter ministérios distintos para o planejamento, orçamento ou finanças, a fim de equilibrar o poder entre os partidos políticos da coalizão. Um exemplo extremo é o da República Democrática do Congo, que dispõe de quatro ministérios centrais distintos para planejamento, orçamento, finanças e economia.

A estrutura jurídica da gestão orçamentária e do gasto pode ser examinada em uma PER. Múltiplas leis e regulamentações são normalmente inerentes aos sistemas de gestão orçamentária e financeira pública. Devido a isso, a PER poderá abordar apenas as questões importantes nas quais são necessárias mudanças na lei ou nas regulamentações. Uma lista não exaustiva de questões que poderiam ser abordadas em uma revisão das leis orçamentárias existentes – ou em uma lei ou regulamentações novas – inclui o seguinte:

- Mudança de normas processuais para a preparação, execução, apresentação de relatórios e auditoria orçamentais, especialmente se forem incompletas, ambíguas, incoerentes ou demasiadamente complexas para serem implementadas na prática; leis referentes ao sistema orçamentário também determinam as responsabilidades institucionais em matéria de orçamento;⁸
- Introdução de novos requisitos de relatórios fiscais para melhorar a transparência;
- Modificação ou introdução de normas fiscais que limitam agregados orçamentários; e
- Melhoria da contabilização do executivo (governo) por parte do legislativo (parlamento). Por exemplo, uma lei pode obrigar o governo a atualizar periodicamente uma estratégia orçamentária de médio prazo.

Processos orçamentários anuais do governo central

As PERs ressaltam frequentemente a necessidade de reforçar a relação entre as metas socioeconômicas estratégicas do país e seu orçamento nacional. A preparação orçamentária é um exercício elaborado que envolve vários ministérios e órgãos públicos. Em alguns países, as PERs identificam se as estimativas orçamentárias são motivadas mais pela inércia do que pelo pensamento estratégico. Isto se aplica especialmente ao caso dos países que dependem de orçamentos com rubricas detalhadas nas quais se dá ênfase à gestão de insumos, tais como pessoal, viagens e materiais,

8 Para obter mais detalhes, ver Lienert e Moo-Kyung (2004), que examina disposições legislativas sobre orçamento de 13 países da OCDE.

e não à forma como a alocação orçamentária contribui para a realização das metas sociais e econômicas do país. Nesses países a PER pode recomendar a introdução do sistema orçamentário baseado no desempenho (reconhecendo ao mesmo tempo que a implementação de tal sistema é um processo lento). Ao exigir que a despesa setorial se enquadre no teto da despesa total, um sistema orçamentário baseado no desempenho permite que as estratégias setoriais dos ministérios da despesa sejam mais bem integradas no orçamento anual. O sistema geralmente requer relatórios que indiquem se os recursos gastos em cada programa do orçamento alcançaram seus objetivos.

As PERs podem ressaltar a necessidade de o processo de formulação orçamentária incorporar não somente as diferentes estratégias setoriais específicas, mas também os objetivos de políticas públicas que afetam o conjunto da economia. Entre estes figuram: (i) metas macrofiscais de alto nível, tais como consolidação fiscal ou redução da dívida; (ii) metas políticas, tais como descentralização de responsabilidades orçamentárias para diminuir os níveis do governo; ou (iii) metas específicas para a melhoria do desenvolvimento humano ou realização dos Objetivos de Desenvolvimento do Milênio.

As implicações dos gastos de médio prazo das diferentes estratégias setoriais e objetivos de políticas públicas que afetam o conjunto da economia precisam ser claramente indicadas e incorporadas no orçamento anual. Nesse contexto pode ser útil se diferentes grupos – não somente os diversos ministérios públicos, mas também os governos municipais e a sociedade civil – participarem da formulação das estratégias setoriais e do orçamento global.

Quanto à formulação do orçamento anual, um problema comumente abordado pelas PERs é o início tardio do ciclo orçamentário, resultando em tempo limitado para o legislativo examinar as propostas orçamentárias do governo. Os ciclos orçamentários curtos podem perpetuar o processo orçamentário pela inércia e crença nos ministérios da despesa de que o conteúdo do orçamento não é importante. Para enfrentar essa crença, as PERs podem sinalizar a necessidade de o governo indicar claramente suas prioridades, de forma que possam ser traduzidas em alocações

orçamentárias anuais. Para isso é necessário examinar como o Conselho de Ministros do país chega a acordos sobre propostas relacionadas com a despesa.

Orçamento duplo – a prática de tomar separadamente decisões sobre despesas correntes e despesas de capital – tem sido abordada nas PERs. Dois ministérios distintos podem estar envolvidos: o Ministério das Finanças, que prepara as estimativas sobre as despesas correntes, e o Ministério do Planejamento (ou equivalente), que prepara o plano de investimento público ou o orçamento. Poderão surgir incoerências, salvo se houver uma estreita coordenação entre os dois órgãos e um processo integrado. Por exemplo, os investimentos de capital podem ser planejados sem assegurar plenamente que a despesa recorrente futura, proveniente do projeto de investimento, seja atendida. As PERs podem ressaltar como o capital físico (estradas, escolas, hospitais etc.) se deteriora prematuramente em consequência de manutenção precária ou falta de materiais (p. ex., livros escolares, medicamentos). As PERs podem também destacar a importância de especificar claramente as responsabilidades dos ministérios das Finanças, do Planejamento e da Despesa, especialmente no tocante à avaliação, planejamento e financiamento do projeto de investimento. Uma Estrutura de Gastos de Médio Prazo (MTEF) bem planejada, na qual os tetos das despesas correntes e de capital sejam coerentes com os recursos disponíveis, é sumamente útil para promover a integração orçamentária, a coordenação entre setores e a previsibilidade do financiamento.

De modo geral, a execução orçamentária recebe atenção considerável nas PERs. Essa é a fase em que os recursos dos órgãos públicos e dos governos locais são realmente transferidos para a prestação de bens e serviços públicos. É também o palco onde se tornam evidentes as consequências de uma formulação orçamentária deficiente. Por exemplo, quando os programas de despesas orçamentárias são subfinanciados devido a estimativas orçamentárias irreais, os formuladores de políticas que executam o orçamento tomam decisões arbitrárias a respeito de quais programas devem ser financiados.

Em alguns países, especialmente os que têm órgãos de supervisão ineficientes, a alocação de fundos não é transparente. Uma demonstração importante de uma execução orçamentária deficiente

é o fato de os fundos públicos não alcançarem os beneficiários pretendidos. Algumas PERs discutem e resumem as principais questões de gestão das finanças públicas, incluindo procedimentos de execução orçamentária, aquisições públicas, controle interno e auditoria, contabilidade pública, auditoria externa e papel do parlamento e da sociedade civil. Exemplos recentes incluem as PERs para a Indonésia (BANCO MUNDIAL, 2007) e Turquia (BANCO MUNDIAL, 2006b).

O setor público ampliado

Um problema comum em algumas PERs é a falta de abrangência de orçamentos públicos centrais. Para fins de transparência, o orçamento em cada nível governamental precisa incorporar informações sobre todas as receitas e gastos dos órgãos públicos. Não é incomum constatar despesas não incluídas no orçamento de órgãos públicos autônomos ou fundos extraorçamentários legalmente estabelecidos. Em alguns países as despesas extraorçamentárias podem ser produto de acordos informais e não transparentes. Subsídios implícitos para empresas estatais deficitárias (p. ex., mediante regimes tributários especiais ou garantias governamentais) ou para despesas militares ocultas podem constituir um problema em alguns países. Uma questão estreitamente relacionada é a destinação de recursos para despesas específicas. Quando há destinação generalizada é difícil fazer realocações orçamentárias devido ao fato de uma parcela considerável da despesa orçamentária não estar disponibilizada para outros usos.

As PERs podem recomendar a ampliação do alcance do orçamento para incluir todas as taxas de uso e despesas extraorçamentárias, bem como comunicação transparente de dívidas garantidas e outros passivos públicos. A identificação de riscos fiscais, inclusive aqueles associados a parcerias público-privadas, precisa constar de documentos orçamentários anuais.

Orçamento público de entes subnacionais e responsabilidades da despesa

O grau em que esferas públicas mais baixas aumentam as próprias receitas, gastam recursos orçamentários, incorrem em déficits ou acumulam dívidas pode ser especificado na constituição ou nas

leis. O nível apropriado de participação do governo estadual ou provincial e do governo local ou municipal em responsabilidades orçamentárias e gestão financeira é uma questão altamente política. Embora algumas PERs examinem questões centrais e locais específicas (p. ex., as respectivas responsabilidades e gestão de recursos para a educação), é menos frequente examinarem opções de descentralização da gestão financeira, salvo se houver solicitação específica de autoridades políticas para que a PER considere opções amplas para aumentar a delegação regional ou municipal de responsabilidades fiscais. Geralmente disposições constitucionais e jurídicas para a governança descentralizada são aplicadas como fixas. Por essa razão, em alguns países federais as PERs têm sido realizadas em um ou mais estados ou províncias selecionados (p. ex., China, Índia, Indonésia, México e Paquistão).

Um aspecto específico da descentralização é o sistema de transferências intergovernamentais – incondicionais e condicionais. Essas transferências podem ser analisadas em uma PER, especialmente nas PERs de setores específicos, em países nos quais os governos central e subnacional têm responsabilidade conjunta da despesa, como por exemplo, na educação, saúde ou infraestrutura. O grau em que essa análise é realizada na PER depende do interesse das autoridades em desenvolver um diálogo nacional sobre essas questões.

Papel do governo *versus* setor privado na prestação de serviços

Outra questão importante discutida nas PERs é se os governos deveriam gastar dinheiro em atividades que podem ser mais bem desempenhadas pelo setor privado. Se o governo estiver envolvido em atividades não necessárias, esse envolvimento prejudica o enfoque estratégico da despesa pública. No setor da saúde, por exemplo, se o papel do governo for expandido para prestar uma ampla gama de serviços de saúde, incluindo cuidados especializados para toda a população, essa responsabilidade adicional limita a capacidade de o governo focar funções de saúde pública mais prementes, tais como financiamento de serviços de saúde para pessoas de baixa renda ou regulamentação de serviços de saúde privada. No trabalho da PER sobre economias outrora planejadas centralmente, foi utilizada a abordagem de árvore decisória

para determinar se certas funções governamentais deveriam ser descartadas, reduzidas ou mantidas (ver MANNING; PARISON 2001). Embora as PERs não proponham uma visão de privatização extrema, algumas PERs demonstram que o setor privado pode produzir certos resultados de forma mais eficiente.

Dados e ferramentas da análise da despesa

Muitas PERs começam focando a alocação da despesa pública total por agregados amplos. As parcelas da despesa são frequentemente comparadas com as de outros países e são identificadas as anomalias principais. Com frequência são utilizados dados das Estatísticas de Finanças Públicas (GFS)⁹. Os dados das GFS são classificados por função (defesa, educação, saúde etc.) e por categoria econômica (salários, despesa corrente não relacionada com salários etc.). A maioria dos países latino-americanos publica dados gerais sobre a despesa pública no *GFS Yearbook* do Fundo Monetário Internacional (FMI) – ou seja, os dados são consolidados por unidades federais e centrais, estaduais, municipais e inclusive unidades extraorçamentárias.

Ao fazer comparações internacionais, Pradhan (1996) argumenta que é preferível começar com a classificação funcional, porque é difícil analisar a composição econômica da despesa em um nível agregado (p. ex., a conta do total de salários ou a despesa total de investimentos) sem conhecimento prévio da composição funcional. No âmbito de uma determinada área (ou setor) de função, a PER pode identificar padrões de gastar de menos ou gastar demais de categorias econômicas da despesa. A PER pode ressaltar desequilíbrios na mescla de insumos, tais como financiamento inadequado de custos correntes de novos investimentos em hospitais e instalações de saúde. Segundo observam várias PERs, a despesa com remunerações tem aumentado rapidamente em comparação com materiais essenciais (p. ex., livros escolares ou medicamentos), para os quais a despesa pode ser totalmente inadequada.

Em alocações orçamentárias anuais, muito poucos países utilizam as classificações de despesa das GFS. A maioria dos países adota a lei

9 A expressão *Estatísticas de Finanças Públicas* é a tradução literal da expressão inglesa *Government Finance Statistics (GFS)*.

do orçamento anual, sendo a despesa classificada no primeiro nível por unidade administrativa (ministérios responsáveis pela despesa, órgãos públicos, parlamento, judiciário, escritório de auditoria externa etc.). O segundo nível da despesa pode ser classificado por programa ou produto ou alternativamente por insumo (versão local das categorias econômicas das GFS). Quando classificada por programa, a despesa não pode ser comparada entre os países.

Uma PER *macro* requer dados sobre agregados orçamentários para o governo em geral (a saber, dados confiáveis para todos os níveis de governo, incluindo suas unidades extraorçamentárias). Embora uma análise de dados sobre despesa disponíveis nas GFS seja um ponto de partida adequado, dados desagregados são necessários nos níveis de setor, subsetor ou programa para identificar: (i) ineficiências; (ii) número excessivo de funcionários públicos; (iii) subfinanciamento da despesa não relacionada à remuneração; e (iv) investimentos públicos ineficazes etc. Estudos sobre a eficiência da despesa requerem dados sobre insumos e produtos ou resultados da despesa, custos unitários e indicadores de desempenho. Em alguns setores há também necessidade de dados sobre prestação de serviços do setor privado e despesas ocultas. Levantamentos de domicílios ou outros dados detalhados da despesa são necessários para realizar análises da incidência de benefícios e para identificar o acesso aos serviços públicos e privados de diversos grupos de renda (p. ex., quintis).

Em muitos países, os dados da despesa desagregada são frequentemente incompletos, indisponíveis ou disponíveis somente após intervalo considerável. Essa falta de dados é um desafio de grandes proporções para a PER. Mesmo disponíveis, os dados talvez não sejam comparáveis entre os anos, áreas geográficas, setores ou classificações econômicas. Em resposta a tais desafios, o Banco Mundial desenvolveu a ferramenta de dados BOOST,¹⁰ que apoia a criação, uso e publicação de bancos de dados sobre a despesa desagregada (Box 5.1). Em 2011, 17 países construíram bancos de dados BOOST e dois países (Quênia e Moldávia) os disponibilizaram publicamente, oferecendo assim meios para seus cidadãos monitorarem e analisarem a despesa pública no nível de base.

¹⁰ A sigla BOOST refere-se à: *Business, Organized Global Education, Opportunity, Science and Technology*.

Box 5.1 – A Iniciativa BOOST

O que é o BOOST?

O BOOST foi desenvolvido pelo Banco Mundial para coletar e compilar dados sobre despesa pública e apresentá-los em um formato de fácil utilização. O banco de dados é usado para examinar tendências em alocações de recursos públicos e eficiência da despesa.

Como funciona o BOOST?

Coleta dados detalhados sobre despesa pública do tesouro nacional e de outros sistemas de registro de órgãos públicos. Os dados brutos são coligidos no nível mais desagregado para tirar proveito da profundidade total da classificação orçamentária do país e dos sistemas contábeis. O banco de dados visa a abranger todos os setores, unidades da despesa e tipos de gasto registrado em diversos sistemas. O banco de dados de cada país é especial e normalmente contém informações sobre o orçamento aprovado e a despesa real, discriminada nas seguintes dimensões: (i) nível governamental; (ii) unidade administrativa, incluindo unidades subnacionais da despesa; (iii) classificação econômica; (iv) classificação funcional; (v) classificação programática e (vi) fonte de financiamento. O banco de dados pode ser complementado com indicadores socioeconômicos, se esses dados estiverem disponibilizados no nível nacional ou descentralizado.

Benefícios do BOOST

Um banco de dados coerente e detalhado proporciona alcance para melhorar a qualidade e a profundidade da análise da despesa. Uma vez estabelecido o banco de dados, os usuários podem examinar o desvio entre as despesas orçamentárias planejadas e reais, bem como identificar ineficiências da despesa. Os governos podem aumentar a transparência quando decidem disponibilizar dados de fácil utilização publicamente. Esse acesso possibilita a organizações da sociedade civil utilizar os dados, melhorando assim a qualidade do assessoramento em políticas.

Os bancos de dados têm sido utilizados por PERs que analisam a qualidade da despesa pública em setores específicos, tais como educação e saúde, como a PER da Guatemala (BANCO MUNDIAL, 2012a). Na Moldávia, a despesa por aluno e outros dados sobre educação fornecidos pelo BOOST foram usados para gerar fronteiras de eficiência por meio da Análise de Envolvimento de Dados (discutida posteriormente neste capítulo). Os pontos relativos da eficiência foram comparados com indicadores da necessidade (índice nacional de privação) e do que foi recebido (montante da despesa por estudante). Os resultados foram mapeados para revelar diferenças entre cada distrito do país.¹¹ Em 2012 os pilotos do BOOST foram lançados em dois estados do Brasil: Minas Gerais e Rio Grande do Sul. Prevê-se que os dados, uma vez disponibilizados, sejam integrados aos processos de planejamento plurianual na esfera estadual e apoiem processos de participação existentes em cada estado, incluindo monitoramento de decisões sobre execução orçamentária e eficiência da despesa.

Eficiência, eficácia e equidade da despesa

No âmago de muitas PERs está uma análise da eficiência, eficácia e equidade da despesa pública. A Parte III deste livro trata de conceitos e mensuração da eficiência. Nas PERs dois conceitos comumente usados são *eficiência da alocação* e *eficiência operacional* ou *eficiência técnica*. Eficiência da alocação refere-se à priorização estratégica de gastos nas políticas, programas e projetos, ao passo que a eficiência operacional ou técnica diz respeito à consecução de produtos ao custo mais baixo possível. A eficiência operacional é avaliada nas PERs medindo-se o seguinte: (i) *eficiência do custo* – volume ideal de insumos usados para obter um determinado nível de resultado; e (ii) *eficiência interna* – combinação ideal de insumos alocados para gerar produtos ou resultados dentro de um setor, dados os preços dos insumos e sua produtividade marginal.

11 Em: <[http://web.worldbank.org/WBSITE/EXTERNAL/COUNTRIES/ECAEXT/MOLDOVA BOOST](http://web.worldbank.org/WBSITE/EXTERNAL/COUNTRIES/ECAEXT/MOLDOVABOOST)>, a figura tem áreas verdes que mostram distritos de alto desempenho e as áreas vermelhas que indicam distritos que requerem ação corretiva.

Eficiência da despesa dentro de setores sociais

A comparação da eficiência *entre* os setores não é direta, porque muitos setores produzem bens e serviços não prontamente comercializáveis. Na defesa nacional e relações exteriores, por exemplo, é difícil medir produtos. Além disso, no caso de certas despesas, as externalidades – custos e benefícios externos ao setor – são difíceis de serem mensuradas. De modo geral, as estimativas da eficiência da despesa por meio dos setores são demasiadamente tênues para a formulação de políticas e alocação orçamentária. Em contraste, a comparação da eficiência (ou taxas de retorno) da despesa *dentro de* um único setor é mais fácil do que entre setores. Essa é a razão pela qual muitas PERs analisam a despesa dentro de um setor ou subsetor (p. ex., a PER da Rússia, examinada aqui nos estudos de caso, discute rodovias e ferrovias).

Como a educação, saúde e outras despesas sociais são geralmente um grande componente da despesa pública total e porque é possível medir a eficiência nesses setores, as PERs frequentemente focam a eficiência da despesa dos setores sociais. Psacharopoulos e Patrinos (2002) estabelecem que os retornos da despesa no ensino fundamental são maiores do que os do ensino médio e superior. Igualmente, constatou-se que a despesa pública em cuidados preventivos da saúde, tais como prevenção de doenças transmissíveis (p. ex., imunização infantil), é mais eficiente do que a despesa com cuidados curativos da saúde. Várias PERs concluíram que a despesa pública se inclina em favor de níveis mais altos de escolaridade e cuidados curativos da saúde (p. ex., a PER de 2007 da Indonésia).

A Análise Envoltória de Dados (DEA¹²) tem sido usada em várias PERs para calcular eficiências técnicas. Neste método são calculadas a fronteira de eficiência e a distância da fronteira a cada unidade (escolas, hospitais, postos médicos etc.). O método DEA transforma os insumos e produtos (resultados alcançados) de uma unidade em uma única medida de eficiência (ver detalhes no Capítulo 8). As unidades que utilizam os recursos da melhor forma possível são classificadas com eficiência de 100%; são unidades referenciais sobre a fronteira da eficiência. O método DEA identifica a eficiência relativa de todas as unidades, o que permite estabelecer metas de melhoria da eficiência para as unidades classificadas com menos de 100%.

12 Análise Envoltória de Dados é a tradução literal do termo em inglês *Data Envelopment Analysis* (DEA).

Esse tipo de análise foi usado no setor da educação na PER da Guatemala (BANCO MUNDIAL, 2012a). A fronteira da eficiência baseou-se em vários indicadores de produção, incluindo taxas de conclusão e notas de testes, bem como em produtos tais como custo unitário do ensino fundamental e coeficientes professor-aluno. Indicadores socioeconômicos vinculados ao aproveitamento escolar, tais como taxas de alfabetização de adultos e taxas de subnutrição, foram usados como variáveis de controle. A análise constatou apenas uma fraca relação entre o nível da despesa e o aproveitamento escolar – resultado encontrado em estudos realizados em outras partes. Na PER das Filipinas (BANCO MUNDIAL, 2011a) as fronteiras da eficiência foram estabelecidas no ensino fundamental e médio. O método DEA foi também aplicado ao setor de saúde das Filipinas, utilizando a mortalidade materno-infantil como indicador de produtos e as despesas públicas nacional e local *per capita* como insumos, ao mesmo tempo em que como controle para o nível da despesa em saúde privada. Em ambos os setores foram feitas comparações de eficiência com outros países asiáticos. No Brasil, este método foi usado em uma PER do ensino municipal (Box 5.2).

Box 5.2 – Brasil: Uso do DEA no ensino municipal

Metodologia

A taxa de aprovação das escolas municipais foi usada como principal indicador de produto. Embora três insumos – despesas administrativas, despesas com professores e despesa de capital por aluno – tenham sido examinados, quando os resultados foram apresentados o enfoque concentrou-se na pontuação de eficiência do DEA para despesas administrativas por aluno. Foram analisadas cinco categorias demográficas, desde municípios *muito pequenos* a cidades *muito grandes* (com mais de 500.000 habitantes).

Principais conclusões

- As despesas administrativas e as taxas de aprovação de estudantes estão fortemente correlacionadas entre si nos municípios, ao passo que nos estados as despesas administrativas maiores não estão associadas a melhores resultados.

(continua)

(continuação)

- Em contextos quase idênticos, os municípios mostram grandes diferenças na eficiência do uso da despesa pública na educação. O Nordeste estava bem representado nos municípios que eram 100% eficientes.
- Os recursos do Fundo de Desenvolvimento da Educação Básica (Fundef) foram bem utilizados: municípios eficientes não somente utilizaram recursos do Fundef para aumentar salários, mas também os aplicaram no apoio administrativo e de pessoal.
- A eficiência não é necessariamente maior entre os municípios mais ricos.
- Alguns municípios grandes têm obrigações importantes em matéria de pensões de professores (inativos).

Recomendações sobre políticas

- Formular intervenções de políticas que se baseiem no sucesso de municípios com melhor desempenho: aqueles que melhoraram a qualidade dos professores proporcionaram aos professores maior apoio administrativo e pedagógico, investiram adequadamente na atualização e racionalização da infraestrutura escolar e instituíram serviços de transporte para as crianças.
- Estabelecer padrões operacionais mínimos para escolas e secretarias municipais.
- Permitir aos governos federal e estadual assumir a liderança na criação de oportunidades para transmitir o sucesso de municípios de alto desempenho a municípios em atraso.
- Alavancar programas federais para induzir mudanças no comportamento municipal proporcionando incentivos e assistência técnica.

Eficiência da despesa na infraestrutura – exemplo das rodovias

As PERs de infraestrutura utilizam as constatações de que as taxas de retorno de gastos rodoviários são geralmente mais elevadas para a manutenção de rodovias existentes do que para a construção de novas. A manutenção inadequada impõe grandes custos correntes e de capital e a negligência na manutenção rotineira relativamente barata pode aumentar os problemas do pavimento das estradas. Essa negligência resulta na substituição de toda a pavimentação da rodovia e a um custo mais elevado do que a manutenção regular. Na América Latina calcula-se que para cada dólar não gasto em manutenção, são necessários de US\$ 3 a US\$ 4 para reconstrução prematura (citado em PRADHAN, 1996, p. 12). Embora a análise da taxa de retorno possa detectar essas ineficiências da despesa, no caso de projetos de infraestrutura podem existir grandes benefícios políticos na construção de novas estruturas em comparação com a manutenção das existentes. Por essa razão, as PERs levam em consideração questões institucionais que envolvem o planejamento da infraestrutura e a tomada de decisões com vistas a melhorar os incentivos, a fim de evitar a alocação indevida de recursos.

Eficácia da despesa: o dinheiro atinge os beneficiários pretendidos?

Eficácia refere-se ao fato de um programa de despesa alcançar ou não os objetivos declarados. Alguns países têm grandes discrepâncias entre despesa orçada e despesa real devido à *perda* em vários estágios da execução orçamentária. Essas discrepâncias tornam difícil medir a eficácia, pelo menos utilizando dados orçamentários, que constituem um indicador precário da despesa real.

As Pesquisas de Rastreamento do Gasto Público (*Public Expenditure Tracking Surveys* – PETS¹³) foram desenvolvidos na década de 1990 pelo Banco Mundial para rastrear o fluxo de recursos públicos através da hierarquia administrativa: do orçamento público central aos prestadores de serviços na linha de frente (instalações

13 A expressão *Pesquisas de Rastreamento do Gasto Público* é a tradução literal da expressão inglesa *Public Expenditure Tracking Surveys* (PETS).

ou pessoal de saúde, escolas, professores e outros).¹⁴ As PETS têm sido usadas nas PERs ou para complementá-las em vários países africanos, asiáticos e latino-americanos, inclusive no Brasil (Box 5.3), Colômbia, Equador e Peru.

Box 5.3 – As PETS no Sistema Único de Saúde do Brasil

Estrutura analítica

O estudo sobre levantamentos de rastreamento do gasto público examinou a prestação de serviços e a qualidade do gasto em saúde mediante o seguinte: (i) análise do planejamento e sistema orçamentário; (ii) rastreamento dos fluxos financeiros entre os diferentes níveis do governo, incluindo as unidades de prestação de serviços de saúde; (iii) avaliação do modo como os recursos foram usados nas unidades de saúde; e (iv) comparação dos recursos usados com os resultados alcançados em termos da quantidade e qualidade dos serviços de saúde produzidos.

Coleta e comparação de dados

A coleta de dados foi feita em quatro esferas: Ministério da Saúde, secretarias estaduais da Saúde, secretarias municipais da Saúde e unidades de saúde (hospitais e clínicas ambulatoriais). Seis estados, 18 municípios e 76 unidades de saúde foram escolhidos para obter uma amostragem diversificada. Os levantamentos utilizaram três fontes complementares de dados: (i) questionário estruturado para gestores; (ii) entrevistas com pessoal técnico; e (iii) relatórios oficiais e outros documentos. Os questionários foram testados no campo antes da coleta de dados. Seis equipes de entrevistadores receberam treinamento em métodos de coleta de dados. Os questionários tinham sete componentes dos quais os dados foram retirados, classificados e tabulados em um banco de dados para facilitar a análise quantitativa.

(continua)

14 Para obter mais detalhes, ver *Public Expenditure Tracking and Facility Surveys: A General Note on Methodology* (Rastreamento do gasto público e levantamento de instalações: uma nota geral sobre metodologia) e outros documentos disponibilizados no website <<http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/PETS>>.

(continuação)

Recomendações principais

- Sincronizar o planejamento, orçamentação e gestão de informação.
- Orientar o sistema orçamentário para o desempenho.
- Desenvolver dispositivos organizacionais que confirmem às unidades de saúde maior autonomia de gestão.
- Reforçar a responsabilização, por exemplo, mediante a introdução de contratos de gestão que incentivem os administradores a focar metas e resultados específicos.
- Estabelecer sistemas robustos de monitoramento e avaliação.
- Fortalecer e profissionalizar a capacidade administrativa.

Fonte: Banco Mundial (2006a).

As PETS identificam não somente o grau de perda de recursos – a diferença entre o montante de recursos desembolsados no nível central e o montante de recursos realmente recebidos pelas instalações de serviços – mas também o montante da perda em cada estágio da cadeia de prestação de serviços. São também identificadas as fontes potenciais de perda. Essas fontes incluem deficiências institucionais, tais como inadimplência de procedimentos de aquisições, dispositivos ineficientes de prestação de serviços, corrupção e fraude. Os levantamentos também podem destacar o desperdício de recursos, tais como absenteísmo de professores, e podem identificar estrangulamentos importantes na execução de gastos (tanto no tesouro nacional quanto nos ministérios públicos). Embora as PETS ofereçam uma fonte rica de informação sobre a qualidade da prestação de serviços, elas requerem um bom conhecimento de procedimentos orçamentários, formais e informais, e de questionários bem formulados, além de recursos consideráveis.

Equidade da despesa pública

Equidade refere-se à distribuição de benefícios da despesa pública. Se uma parcela desproporcional dos benefícios for absorvida por

pessoas em melhores condições financeiras, a despesa pública é considerada injusta ou regressiva. Em contrapartida, quando os pobres se beneficiam de forma mais do que proporcional, a despesa é descrita como pró-pobres ou progressiva.

As PERs, especialmente as PERs setoriais, podem identificar grupos de renda que se beneficiam de políticas da despesa pública. Com base nos levantamentos de domicílios, os beneficiários são estratificados por nível de renda a fim de analisar a equidade da despesa social. Por exemplo, a PER da educação na Polônia, em 2009, proporcionou evidência de que subsídios ao ensino superior (ensino gratuito para os estudantes do turno do dia) beneficiavam desproporcionalmente os mais abastados. O relatório recomendou teste de renda verificada para os beneficiários de subsídios do ensino superior (RODRIGUEZ; HERBST 2009).

A Análise de Incidência de Benefícios (BIA)¹⁵ é uma ferramenta usada em algumas PERs para examinar questões relacionadas com a equidade da despesa (ver Box 5.4). As BIAs têm demonstrado que a despesa pública no ensino fundamental é pró-pobres por três razões: primeiro, os pobres tendem a ter mais filhos em idade do ensino fundamental do que os não pobres; segundo, as crianças pobres têm maior probabilidade de frequentar escolas públicas de ensino fundamental do que crianças em melhores condições, as quais tendem a frequentar escolas de ensino fundamental privadas; e terceiro, crianças de domicílios pobres têm menos probabilidade de se matricularem no ensino médio do que crianças de domicílios de alta renda. Igualmente, a despesa pública em cuidados básicos da saúde é normalmente mais pró-pobres do que a despesa em cuidados secundários. Portanto, a despesa com ensino fundamental e cuidados preventivos de saúde pode favorecer as metas, tanto de eficiência quanto de equidade.

15 A expressão *Análise de Incidência de Benefícios* é a tradução literal da expressão inglesa *Benefit Incidence Analysis (BIA)*.

Box 5.4 – Análise da Incidência de Benefícios (BIA)

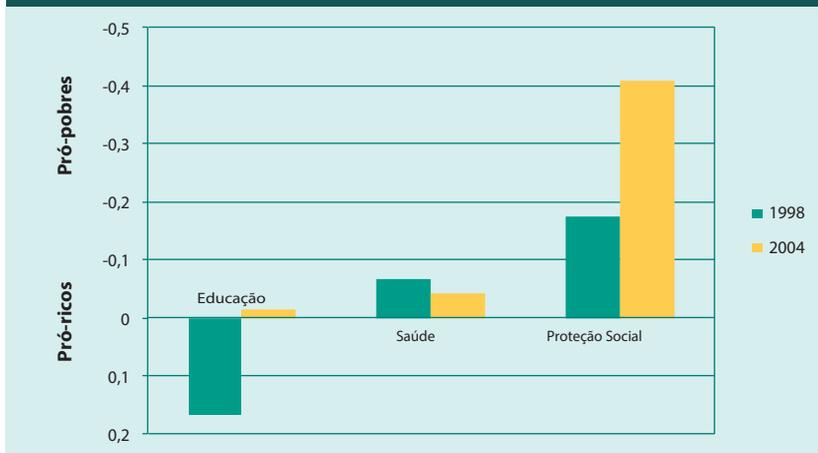
Uma BIA examina a parcela de benefícios de um programa de despesas que é recebida por diferentes grupos. Os únicos dados necessários para uma BIA são variáveis que definem o grupo e uma estimativa dos benefícios que cada grupo recebe. Uma BIA tem três etapas básicas:

- Identificar aqueles que utilizam o serviço. Uma pesquisa de domicílios é uma boa fonte de informação.
- Calcular os custos unitários dividindo o total da despesa pública para a prestação do serviço (p. ex.: educação de alunos do ensino fundamental) pelo número total de beneficiários ou usuários (ou seja, o número total de estudantes que completaram o ensino fundamental ou que se matricularam no ensino fundamental).
- Agregar usuários em grupos relevantes por renda ou despesa, região ou gênero para a análise comparativa. Essa etapa envolve primeiramente ordenar os usuários usando uma medida bem-estar, tais como a renda para definir quantis. A análise mais profunda examina essas quantis e investiga se os beneficiários residem em área urbana ou rural, por grupos de gênero etc.

Outra medida para entender qual grupo populacional se beneficia da despesa em determinados serviços é o coeficiente de concentração ou quasi-Gini. Essa medida mostra se os grupos mais pobres recebem proporcionalmente mais ou menos benefícios.¹⁶ A PER da Costa Rica demonstrou que a despesa com Educação e Proteção Social (excluídas as pensões contributivas) beneficiou mais os pobres no período entre 1998 e 2004 (Gráfico 5.1).

¹⁶ O quasi-Gini baseia-se nos gastos públicos recebidos que cada segmento da população recebe (em geral são usados quintis), dos mais pobres para os mais ricos. Quando os domicílios mais pobres recebem uma parcela desproporcionalmente elevada de despesas, o quasi-Gini é negativo. O quasi-Gini varia de -1 a 1. Os valores negativos denotam despesas pró-pobres (*progressivas*) e os valores positivos indicam que os relativamente ricos são os mais beneficiados pela despesa (*despesa regressiva*).

Gráfico 5.1 – Costa Rica: Despesa social (quasi-Gini), 1998 e 2004



Fonte: Banco Mundial (2008). Tabela 2.3.

5.4 De que modo uma PER difere de uma Revisão de Despesa?

Uma revisão da despesa é definida como “um processo institucionalizado para a revisão de um cenário base de gastos com o objetivo principal de identificar opções para níveis alternativos de financiamento para consideração no processo orçamentário” (ver Capítulo 4). Revisões de despesas e PERs diferem nos seguintes aspectos:

- **Objetivos e escopo:** As PERs são adaptadas às necessidades da análise de despesas dos formuladores de políticas. Algumas PERs concentram-se em questões macrofiscais, ao passo que outras são extremamente focadas em um ou alguns setores de despesas. PERs que examinam questões *micro* de eficiência geralmente incluem uma análise de despesa setorial nas esferas nacional e subnacional. Em contrapartida, o foco das revisões de despesas é restrito: elas reexaminam as políticas de despesa *existentes*, especialmente no âmbito do governo central. Mais especificamente, as revisões de despesas têm o objetivo explícito de auxiliar os formuladores de decisão no processo de preparação do orçamento acerca do futuro nível de financiamento para os programas existentes;

- **Foco da análise:** As PERs podem começar com uma visão geral das despesas por função ou por tipo econômico, ou ambos, e são frequentes as comparações internacionais entre amplas categorias de despesas. As PERs geralmente enfocam os setores sociais (educação e saúde) ou a infraestrutura, áreas nas quais o Banco Mundial tem ampla experiência. Por sua vez, as revisões de despesas geralmente examinam todas as despesas (ou sua maior parte) dos órgãos governamentais, com foco na identificação da eficiência da despesa ou na economia dos produtos (Nota do tradutor: a expressão *economias dos produtos* é tradução literal da expressão inglesa *output savings*);
- **Vínculo com o orçamento anual:** Uma PER é um documento analítico com recomendações para melhorias na elaboração e implementação da política; ela não está necessariamente vinculada ao ciclo do orçamento anual. Já o objetivo central de uma revisão de despesa é reexaminar a fundamentação para as políticas de gastos iniciais, ou seja, em relação ao cenário de referência de gastos com o objetivo de modificá-las no próximo orçamento anual. Em alguns casos, a revisão está explicitamente ligada à garantia de que as despesas futuras, em conjunto, não excedam os subtotais da Estratégia de Despesas a Médio Prazo (MTEF);
- **Uso de ferramentas analíticas:** As PERs utilizam uma ampla gama de ferramentas analíticas (p. ex.: análise de custo-benefício, BIA, DEA). Em alguns casos, são preparados bancos de dados detalhados, tais como o BOOST. É possível usar um número menor de ferramentas nas revisões de despesas, nas quais o foco está frequentemente voltado para a compreensão das ligações entre os resultados da despesa (para os quais foram desenvolvidos indicadores de desempenho) e o tamanho das alocações de orçamento para a despesa do programa de linha de base;
- **Responsabilidade dos ministérios que efetuam as despesas:** As revisões de despesa são proveitosas quando os ministérios responsáveis pela despesa e não apenas o do orçamento central ou o Ministério das Finanças, apresentam propostas para alterar as políticas de despesa. Embora as

PERs setoriais envolvam totalmente ministérios (tais como Agricultura, Educação, Saúde ou Obras Públicas), pode haver menos interação dos ministérios que efetuam as despesas setoriais com o orçamento central ou Ministério das Finanças (em comparação com a interação das revisões de despesa);

- **Recomendações sobre políticas:** Tanto as PERs quanto as revisões de despesa geralmente incluem recomendações sobre políticas para a obtenção de ganhos de eficiência. Ambas podem também recomendar a revisão das políticas de definição de preços e de recuperação de custos para a prestação de serviços governamentais e para a melhoria dos dispositivos institucionais para as despesas. Embora as recomendações das PERs possam focar as melhorias na política de despesas, as revisões de despesa têm mais probabilidade de sugerir opções para cortes de financiamento; e
- **Regularidade das atualizações:** Uma revisão de despesa é geralmente preparada em um ciclo anual regular ou plurianual (p. ex.: três anos); ou seja, ela geralmente não é uma revisão isolada. Em contrapartida, as PERs são em geral revisões únicas e não exercícios contínuos. Embora alguns países tenham preparado várias PERs durante um período de 5 a 10 anos, em muitos países o intervalo entre duas PERs sucessivas é em geral longo.¹⁷

5.5 Estudos de Caso de PERs: México, Peru, Rússia e Turquia

México – PER de infraestrutura de 2005

Experiência anterior com PERs: No período 2000-2005, o Banco Mundial apoiou ativamente a agenda de reforma de finanças públicas, inclusive a reforma tributária e fiscal, a sustentabilidade fiscal e a gestão fiscal em regime de descentralização. Como os estados e municípios administram quase a metade da despesa pública, considerou-se importante examinar as responsabilidades federais e estaduais ou municipais nos setores-chave. Por esse motivo, foi

17 Na lista de 89 PERs regulares realizadas no período 2001-2007 (ver DEOLALIKAR, 2008, p. 3), somente oito países tiveram PERs repetidas nesses seis anos.

realizado em 2003 um PER subnacional para o estado de Veracruz. Esse PER foi seguido de outro, em 2004, que analisou a incidência de benefícios nos diferentes níveis de renda e a distribuição de despesas federais nos estados. Em 2009 foi realizado outro PER setorial para agricultura e desenvolvimento rural.

Objetivos: Embora os PERs de 2003-2004 tenham dado alguma atenção ao investimento em infraestrutura pública, houve forte necessidade de realizar uma análise mais completa. O objetivo principal do PER de Infraestrutura (IPER) de 2005 foi apresentar opções para tratar das necessidades de serviços de infraestrutura, inclusive (i) melhorias na concepção do programa e alocações orçamentárias; (ii) ganhos de eficiência na prestação de serviços; e (iii) maior participação e financiamento do setor privado (BANCO MUNDIAL, 2005).

Escopo: O IPER forneceu uma visão geral da despesa dos setores de infraestrutura, órgãos e da despesa agregada. Essa visão geral preparou o cenário para a avaliação do desempenho do setor de infraestrutura. Foi realizada uma análise aprofundada dos processos de planejamento, elaboração do orçamento e coordenação para os subsetores de transporte, eletricidade e abastecimento de água. Foram examinados os mecanismos para recuperação dos custos, a qualidade e a eficiência dos serviços de infraestrutura e a eficácia dos programas para incentivar a participação e o financiamento privados, assim como os papéis dos governos federal e estadual.

Análise utilizada: O desempenho da infraestrutura foi analisado mediante o uso de indicadores de resultados dos setores. Com relação à eletricidade, por exemplo, foram desenvolvidos indicadores para examinar o acesso (em comparação com outros países) e a qualidade do serviço (interrupções do serviço, número de reclamações por mil clientes e horários de conexão). Para calcular as necessidades de despesa com infraestrutura, foram utilizadas técnicas simples e avançadas, desde padrões de referência – por exemplo, exercícios de custos para definir metas, tais como quanto custaria para elevar a infraestrutura do México *per capita*, por unidade do Produto Interno Bruto (PIB) e por quilômetro quadrado ao mesmo nível dos líderes da América Latina – mediante modelos econômicos econométricos ou de engenharia. Embora o relatório não desenvolva modelos sofisticados, ele apresenta alguns resultados de pesquisa. Na *abordagem do crescimento*, por exemplo, a análise constatou que o México poderia

obter lucros significativos se aumentasse a cobertura de infraestrutura, mas essa alteração exigiria aumentos substanciais de investimento.

Principais constatações: Houve progresso constante no aumento, em todo o país, da cobertura de estradas, eletricidade, água e saneamento; em 2005 os níveis estiveram entre os mais elevados da América Latina. Embora ainda tenham sido observadas algumas lacunas no acesso aos serviços de infraestrutura, notadamente nas comunidades de baixa renda, rurais e indígenas, os principais desafios foram a melhoria da qualidade dos serviços e a eficiência operacional, tendo em vista, especialmente, que a qualidade média da infraestrutura do México era ligeiramente inferior à de outras importantes economias da América Latina e de algumas do Leste Asiático (Gráfico 5.2) Em 2005, o maior hiato de qualidade ocorreu no fornecimento de eletricidade e o menor, em portos e estradas de ferro. A necessidade de aprimorar a qualidade ficou mais evidente para as estradas, onde havia necessidade urgente de reparos devido ao crescente volume de tráfego e ao envelhecimento da malha rodoviária, particularmente das rodovias administradas por estados e municípios. Em infraestrutura, a ausência de políticas de subsídios abrangentes – e a multiplicidade de atores nas esferas federal, estadual e municipal – produziu uma grande variação no grau de recuperação de custos e de subsídios entre os setores e regiões. Foram estabelecidas tarifas muito abaixo dos custos para algumas áreas e usuários, até mesmo para aqueles que podiam pagar.

Constatou-se que a forma mais comum de subsídio para usuários eram tarifas reduzidas para algumas categorias deles nos setores de eletricidade, abastecimento de água e saneamento. As tarifas das estradas com pedágio eram elevadas em comparação com os padrões internacionais. Dentre os setores de infraestrutura, a despesa pública recebeu peso muito elevado para novas construções e reformas, ao passo que as atividades regulares de manutenção receberam recursos financeiros insuficientes. No caso das estradas, os gastos federais com manutenção foram insuficientes para manter a malha em boas condições. Para alcançar melhor qualidade, acesso e eficiência em serviços de infraestrutura são necessários dispositivos institucionais mais robustos, com maior coordenação e planejamento intra e intersetorial e maior responsabilização.

Gráfico 5.2 – Qualidade da infraestrutura. Países selecionados, 2003

Fonte: Banco Mundial (2005).

Nota: O Gráfico baseia-se na Tabela 1 da IPER, que é derivada do Fórum Econômico Mundial (2004), que investiga grandes usuários industriais. Os indicadores de qualidade para todos os setores de infraestrutura estão incluídos (o transporte aéreo e tecnologias da informação e da comunicação não fazem parte da IPER do México). A avaliação da pesquisa é feita em uma escala de 1 (subdesenvolvida e ineficiente) a 7 (desenvolvida e eficiente como a melhor do mundo).

Principais recomendações de política: Uma mensagem central da PER foi que os recursos gastos em infraestrutura devem ser mais bem utilizados. Foram feitas sete sugestões específicas:

- Fazer uso mais eficaz dos recursos dos contribuintes no investimento público e concentrar o foco em áreas que o setor privado não possa financiar. Reduzir o papel do governo no financiamento do setor elétrico e de estradas com pedágio, os quais fizeram demandas fiscais consideráveis (muitas delas fora do orçamento ou eventuais);
- Aumentar a eficiência das despesas, especialmente em três áreas: manutenção e recuperação, gargalos estratégicos em segmentos da infraestrutura e ampliação dos serviços básicos à população de baixa renda;
- Adotar medidas para melhorar a concepção dos programas de investimento e a escolha de projetos de infraestrutura;
- Mobilizar mais financiamento privado para a infraestrutura, tanto de fontes nacionais quanto internacionais;

- Revisar a concepção e o funcionamento dos aumentos de crédito federal, necessários para atrair vultosas somas de financiamento privado;
- Criar órgãos independentes para a regulamentação de tarifas e qualidade dos serviços, especialmente nos setores em que esses órgãos estão em grande parte ausentes, notadamente em eletricidade e recursos hídricos; e
- Aumentar a responsabilização e a informação sobre desempenho, especialmente para obter informações sistemáticas acerca da qualidade dos resultados dos projetos.

Peru – PER 2012

Experiência anterior com PERs: Uma PER regular foi concluída em 2002. Em 2010 o Banco Mundial auxiliou as autoridades na preparação de um documento de política (*O Processo de Descentralização e suas Ligações com a Eficiência das Despesas Públicas*). Em 2011, o pessoal do Banco Mundial preparou uma série de notas sobre políticas acerca de questões específicas dos setores e de gestão das finanças públicas para a próxima administração. Em 2012, a PER foi preparada em estreita colaboração com as autoridades peruanas.

Objetivos: O principal objetivo da PER 2012 (BANCO MUNDIAL, 2012b) foi contribuir para as discussões políticas sobre questões de gastos públicos. Como já haviam transcorrido 10 anos desde a última PER, era necessário fazer uma avaliação. A revisão procurou abordar três questões importantes: (i) evoluções dos gastos públicos e receitas durante a década anterior; (ii) o impacto da descentralização na mudança para a elaboração do orçamento com base em resultados; e (iii) até que ponto as despesas públicas são eficientes em termos técnicos e de alocação.

Escopo: A PER tem um foco macrofiscal. Ela identifica os principais desafios enfrentados pelos gastos públicos, particularmente as desigualdades de acesso aos serviços públicos e a sua baixa qualidade. Ela examinou de que modo a estabilização macroeconômica e a política fiscal estão relacionadas. A adequação da receita e o sistema tributário foram revisados, seguidos de uma análise da alocação, da eficiência e da equidade dos gastos. As notas sobre políticas

preparadas em 2011 podem ser consideradas o equivalente aos *capítulos sobre setores* de uma PER regular.

Análise utilizada: A PER analisou as tendências do PIB, receita total e despesa total e seus principais componentes. Embora o foco fosse a exploração completa dos dados macrofiscais do Peru, foram feitas algumas comparações internacionais, inclusive para a despesa com educação, saúde e despesa social. Foram utilizadas técnicas econométricas para determinar a relação entre crescimento econômico e receitas tributárias e entre crescimento e gastos públicos, além de identificar os países pares do Peru (em termos de PIB *per capita* e população). Foram criados dois índices¹⁸ sobre a eficiência da despesa com educação, saúde e transporte e foram feitas as correlações entre os subíndices.

Principais constatações:

- A sólida gestão macroeconômica havia trazido estabilidade econômica e rápido crescimento do produto durante a década anterior. Políticas fiscais cautelosas contribuíram para esse resultado;
- O sistema tributário arrecada um volume limitado de receita, em parte porque a base tributária é prejudicada por inúmeras isenções. Além disso, existe forte dependência dos tributos indiretos, o que torna o sistema tributário injusto;
- A taxa de pobreza caiu significativamente entre 2005 e 2010. O resultado não se deveu totalmente ao rápido crescimento da despesa pública total. As baixas taxas de execução orçamentária, limitações de gestão e as restrições da capacidade administrativa e de infraestrutura também foram fatores que contribuíram;
- O programa de transferência condicionada de renda (*Juntos*), voltado para os pobres das áreas rurais, parece ter tido efeito positivo na redução da pobreza rural, aumentando o uso dos serviços de saúde e assegurando que as crianças ingressem e concluam o ensino fundamental;

18 Os dois índices – o Índice de Desempenho do Setor Público e o Índice de Eficiência do Setor Público – são baseados em Afonso, Schuknecht e Tanzi (2005).

- Houve grande progresso na melhoria da cobertura dos serviços públicos básicos em saúde e educação. Entretanto, o acesso aos serviços públicos foi altamente desigual entre os distintos grupos de pessoas e a análise demonstrou diferenças significativas no desempenho e na eficiência do governo entre as regiões;
- A introdução de elaboração de orçamento baseada em resultados em 2008 produziu um foco bem-vindo em resultados, monitoramento e avaliação;
- O investimento público aumentou substancialmente (de 3% para 6% do PIB), mas havia necessidade de melhorar o planejamento, gestão e coordenação do investimento entre as esferas de governo e o setor privado;
- O planejamento, execução e contabilidade do investimento não eram vinculados ao sistema de gestão financeira do governo; e
- O mecanismo para a alocação das receitas com recursos naturais exacerbou as desigualdades regionais.

Principais recomendações de política: A PER reiterou recomendações específicas (comunicadas pela primeira vez nas notas sobre política de 2011) para a melhoria da responsabilização nos setores de educação e saúde para o aprimoramento da gestão das finanças públicas. Em questões macrofiscais, a PER comunicou as seguintes mensagens:

- Manter uma postura de política fiscal cautelosa;
- Ampliar a base tributária reduzindo as isenções de impostos e não aumentando ainda mais as taxas de impostos;
- Simplificar ainda mais o sistema tributário;
- Dependendo menos dos trabalhadores assalariados para obter receitas do imposto de renda de pessoas físicas;
- Se forem contempladas as mudanças nas taxas de impostos sobre recursos naturais, examinar cuidadosamente as consequências de médio prazo e o possível impacto adverso sobre o investimento;
- Reorientar a despesa pública para que haja maior redução da pobreza, diminuição das grandes desigualdades de renda e melhoria da qualidade do serviço público;

- Basear o sistema orçamentário em resultados mais pragmáticos e evitar que ele se torne mais um requisito de relatório para os órgãos de primeira linha;
- Aprimorar o planejamento do investimento, inclusive com o aumento da capacidade de governos locais, para que todos os projetos de investimento sejam avaliados adequadamente e a fragmentação do investimento seja evitada;
- Assegurar que as despesas recorrentes sejam adequadas para manter o capital existente;
- Criar uma interface entre o sistema de investimento público e o sistema de gestão financeira do governo; e
- Revisar as relações intergovernamentais com o objetivo de reduzir as desigualdades da despesa pública no âmbito local, inclusive modificando a distribuição desigual de receitas dos recursos naturais.

Rússia – PER 2012

Este estudo de caso enfoca o capítulo da PER intitulado *Fortalecimento do Orçamento de Capital nos Setores Rodoviário e Ferroviário*.

Experiência anterior com PER: Há muitos anos, o Banco Mundial fornece assistência analítica e técnica ao Departamento de Política Orçamentária do Ministério das Finanças (MF). As várias questões incluem (i) avaliação dos riscos fiscais; (ii) preparação de um Estudo de Sustentabilidade Fiscal de Longo Prazo; (iii) realização de análises macroeconômicas e fiscais regulares; e (iv) fornecimento de consultoria para fortalecer o orçamento e a gestão das finanças públicas. Com o início da crise global em 2008 e o aumento das restrições orçamentárias, o diálogo enfocou medidas para aumentar a eficiência das despesas públicas. Em 2011, o Banco Mundial finalizou uma Revisão de Despesa Social acerca da eficiência da despesa pública subnacional em saúde, educação e assistência social (BANCO MUNDIAL, 2011b). Essa revisão foi seguida em 2012 de uma PER que enfocou questões macrofiscais e dois setores (BANCO MUNDIAL, 2012c).

Objetivos: A PER de 2012 complementou a revisão do setor social de 2011. O objetivo principal da PER foi auxiliar o MF na identificação de

oportunidades de economia e ganhos de eficiência em duas grandes categorias de despesa: infraestrutura de transportes e emprego no setor público. Relatórios analíticos separados de antecedentes sobre aqueles dois setores já haviam sido preparados para as autoridades russas; os dois foram integrados à PER 2012.

Escopo: A PER analisou primeiro a disciplina fiscal agregada e a gestão das finanças públicas em geral. Essa análise forneceu o contexto para a revisão do escopo do espaço fiscal adicional para a despesa em dois subsetores de infraestrutura, particularmente rodovias (onde havia escassez de recursos) e ferrovias. No setor de transportes, embora a PER da Rússia tenha examinado somente rodovias e ferrovias, a IPER do México examinou também os portos.

Análise utilizada: A análise da PER fez comparações internacionais relevantes, fornecendo assim padrões de referência. Para rodovias, por exemplo, a PER ressaltou que, na década de 1990, a China gastou por ano 3,5% do PIB em rodovias, em comparação com 2,8% em 2000 e apenas 1,5% da Rússia em 2009. Com relação aos custos unitários, foram feitas comparações com a Finlândia, um país com condições climáticas semelhantes. Houve também análise de planejamento de manutenção de rodovias e ferrovias, intensidade do tráfego rodoviário e o escopo para o aumento das tarifas dos usuários e financiamento da infraestrutura rodoviária ou ferroviária. O relatório identificou o hiato de financiamento para a manutenção de rodovias. A Estratégia de Transportes do governo também foi analisada.

Principais constatações:

Rodovias – Com escassez crônica de recursos, a infraestrutura rodoviária deteriorou-se significativamente durante o período 2000-2011. Dos quase 50 mil quilômetros de rodovias federais, somente cerca de um terço foi considerado estar em condições boas ou razoáveis. Embora os custos reais de manutenção por quilômetro das rodovias na Rússia e na Finlândia fossem semelhantes, os dois países têm resultados em termos de condições rodoviárias muito diferentes. Além disso, os custos unitários reais de despesa por quilômetro de rodovia eram muito inferiores aos custos unitários oficiais. A maioria dos padrões de projeto de rodovias públicas era baseada em cargas máximas por eixo de menos de 10 toneladas, o que é altamente inadequado para os caminhões europeus, resultando em uma vida útil relativamente curta da superfície

das rodovias. O relatório constatou um vultoso hiato de financiamento para a manutenção e reconstrução de rodovias embora houvesse espaço para um programa de manutenção totalmente financiado. Para os projetos de infraestrutura rodoviária, eram necessários recursos financeiros adicionais; esses recursos poderiam vir de pedágios, títulos de infraestrutura, entidades de aumento de crédito local e novos projetos financiados por instituições financeiras internacionais.

Ferrovias – A Ferrovia Russa (RDZ) tem *status* de monopólio como proprietária da infraestrutura ferroviária nacional e fornecedora de locomotivas. A RDZ domina o mercado de frete ferroviário. O financiamento da RDZ continuará a ser determinado principalmente pela estrutura de tarifas, indexação anual de tarifas e apoio financeiro do governo. Os subsídios para a operação de frete foram introduzidos como uma medida temporária em 2009 e continuaram no orçamento federal em 2010. A estratégia de transportes do governo prevê, primeiramente, a modernização do transporte ferroviário (até 2015) e, em segundo lugar, a expansão da malha ferroviária (2016-2030). As principais questões que os formuladores de política precisam abordar são (i) as compensações entre o crescimento periférico e nacional para o investimento em infraestrutura ferroviária; (ii) as distâncias eficientes para as diversas modalidades de transporte e uma análise da velocidade com que os novos trilhos devem ser construídos; e (iii) a questão da escassez de recursos financeiros das extremamente necessárias manutenção e recuperação, tendo em vista especialmente a grande expansão planejada nas malhas ferroviárias.

Principais recomendações de política:

Rodovias

Garantir o fornecimento adequado de financiamento.

- Aumentar as cobranças dos usuários das rodovias (impostos de consumo sobre os combustíveis e taxas de licenciamento de veículos);
- Aumentar gradualmente a despesa com manutenção, recuperação e reconstrução de rodovias;
- Introduzir um fundo de manutenção de estradas de segunda geração; e

- Estabelecer uma estrutura de financiamento em várias camadas e levar em conta a assistência financeira de *International Financial Institutions* (IFIs).

Fortalecer as práticas de gestão e planejamento:

- Revisar os objetivos e o cronograma da Estratégia de Transportes;
- Realizar um exercício de classificação funcional de rodovias e melhorar a gestão dos ativos rodoviários;
- Introduzir contratos de manutenção de estradas baseado em desempenho; e
- Revisar os padrões de projetos de veículos e de estradas e fortalecer os controles da sobrecarga de caminhões.

Aumentar a estrutura institucional:

- Revisar a Lei de Aquisições e aumentar a competição para a aquisição de construção de estradas; e
- Fortalecer a estrutura jurídica e institucional para parcerias público-privadas.

Ferrovias

Para aumentar a dotação orçamentária:

- Dar prioridade à manutenção e recuperação;
- Ajustar as taxas de frete para assegurar a recuperação dos custos; e
- Revisar o processo para a indexação das tarifas de frete.

Fortalecer as práticas de gestão e planejamento:

- Revisar os objetivos e o cronograma da primeira fase da Estratégia de Transportes;
- Assegurar que o investimento em infraestrutura ferroviária nas regiões periféricas se justifica do ponto de vista financeiro e econômico; e
- Revisar a necessidade de 10 mil quilômetros de trilhos de alta velocidade.

Para melhorar a estrutura institucional:

- Introduzir o financiamento obrigatório do serviço público; e
- Revisar o requisito legal de que uma empresa de transporte deva fornecer serviços em toda a rede.

Turquia – PER 2006

Experiência anterior com PERs: A PER macrofiscal de 2006 para a Turquia (BANCO MUNDIAL, 2006b) seguiu uma PER do setor de educação que foi concluída em 2005.

Objetivos: Os objetivos da PER eram rever os desenvolvimentos e perspectivas macrofiscais. Os objetivos específicos eram examinar (i) as fontes de consolidação fiscal e a qualidade do ajuste fiscal; (ii) as opções de ganhos de eficiência na despesa pública, questões horizontais e específicas de setores; e (iii) as instituições orçamentárias e gestão das finanças públicas.

Escopo: A PER teve um escopo amplo: a cobertura institucional estendeu-se para além do governo geral sempre que a disponibilidade de dados permitiu (empresas estatais e instituições de seguridade social têm um efeito importante sobre o desenvolvimento do setor público na Turquia). O lado da receita do orçamento também foi examinado, assim como os gastos tributários. O escopo para ganhos de eficiência foi examinado quanto à despesa com o serviço público e a gestão dos investimentos públicos (questões horizontais) e em quatro setores específicos: despesas com educação, saúde, previdência social e desenvolvimento rural. Finalmente, foram avaliadas as reformas da gestão das finanças públicas.

Análise utilizada: A PER foi primordialmente uma análise estática, na qual os dados macrofiscais e setoriais foram analisados em tabelas e gráficos. Foram feitas comparações internacionais, principalmente com países da OCDE (da qual a Turquia é membro), mas foram feitas também algumas comparações com outros países emergentes, inclusive o Brasil. Prepararam-se cenários de referência e alternativos de política fiscal, cenários que ajudaram a identificar o espaço fiscal. A PER não utilizou ferramentas específicas (p. ex.: BIA, DEA) para examinar a eficiência da despesa em setores específicos.

Principais constatações:

- Durante o período 2001-2005 um ajuste fiscal ambicioso serviu de base para um declínio substancial na taxa de dívida pública e para o forte desempenho econômico;
- Em 2006, ocorreram dois importantes desafios fiscais: manutenção da disciplina fiscal e criação do espaço fiscal necessário para atender aos prementes desafios de desenvolvimento;
- Foram necessárias reformas fiscais estruturais para manter a disciplina fiscal e, ao mesmo tempo, criar espaço fiscal para a despesa com o aumento do crescimento e redução de impostos no futuro;
- A composição das receitas mudou substancialmente no sentido da tributação indireta. A base para os impostos diretos era pequena devido às numerosas isenções e os gastos tributários para 15 provisões para impostos foram calculados em 1,6% do PIB por ano;
- Poderiam ser obtidos ganhos de eficiência horizontais se as reformas do serviço público fossem direcionadas para o alcance de melhor desempenho e, ao mesmo tempo, contenção da pressão sobre a folha salarial;
- Quanto ao investimento público, regras de processamento pouco claras levaram ao excesso de investimento e prazos muito longos para a execução de projetos;
- Ainda não havia uma estrutura normativa adequada para incentivar as parcerias público-privadas. No setor energético, embora o investimento privado respondesse por 50% da geração de eletricidade, esse percentual só ocorreu porque o governo havia assumido novos passivos contingentes importantes na forma de garantias e contratos de exploração;
- Várias reformas da gestão das finanças públicas haviam sido adotadas, especialmente a promulgação em 2003 da Lei de Gestão e Controle das Finanças Públicas (PFMC) e melhor cobertura, formulação, execução, contabilidade, auditoria e aquisição orçamentária;

- Foram feitos avanços notáveis nas taxas de matrícula no ensino fundamental. Entretanto, ainda existiam diversos hiatos educacionais e houve uma necessidade de aumentar a despesa com educação para ampliar a parcela de gastos não relacionados a pessoal, a fim de melhorar a eficiência, obter melhor equilíbrio regional em serviços educacionais e aumentar a autonomia da gestão nos estabelecimentos de ensino;
- A introdução do seguro saúde universal havia melhorado o acesso à saúde, mas os resultados na saúde eram ruins, apesar da despesa relativamente alta; e
- A reforma da previdência de 2006, se integralmente implementada, eliminaria o déficit das aposentadorias em 2040. Esse resultado exigiria a manutenção dos impostos sobre a folha de pagamento, tanto para empregados quanto para empregadores, em nível elevado em comparação com os padrões internacionais.

Principais recomendações de política:

- Racionalizar os impostos, ampliar a base tributária e estabelecer uma estrutura de gastos com impostos que promova a responsabilização e a transparência;
- Fortalecer os incentivos para que os servidores públicos alcancem resultados e, ao mesmo tempo, controlar a massa salarial, harmonizando as escalas de salários e reformando os auxílios;
- Adotar métodos de avaliação eficientes para projetos plurianuais de investimento público e assegurar a provisão adequada para a despesa com manutenção;
- Incentivar o setor privado a desempenhar um papel mais relevante no financiamento, desenvolvimento e operação do investimento em infraestrutura, desde que os riscos de passivos contingentes adicionais do governo nas parcerias público-privadas sejam reduzidos;
- Enfrentar a grande agenda inacabada de gestão das finanças públicas, focando primeiramente a criação de estruturas de auditoria interna, promulgando uma lei do Tribunal de

Contas coerente com a Lei PFMC¹⁹ e trazendo os recursos extraorçamentários e rotativos remanescentes para o controle orçamentário e financeiro;

- Estimular as escolas a limitarem a capacidade das salas de aula dentro de níveis aceitáveis; fornecer autonomia de recursos financeiros para as escolas e, ao mesmo tempo, introduzir a responsabilização pelos resultados educacionais; conceder às instituições de educação superior mais autonomia sobre os recursos financeiros e modernizar o sistema de exame para ingresso nas universidades, de modo a eliminar a vantagem dos cursos privados preparatórios para o exame;
- Implementar economia de custos e medidas para aumento da eficiência em saúde (a PER continha várias recomendações específicas para o setor de saúde); e
- Alterar as políticas de previdência social caso fique decidido reduzir as taxas de contribuição previdenciária para alcançar superávit de previdência no futuro.

5.6 Lições aprendidas de experiências da PER

A experiência de preparar PERs nas últimas décadas forneceu os principais ingredientes para a identificação de uma PER de alta qualidade. Entretanto, mesmo que a PER tenha alta qualidade, sua eficácia depende primordialmente da vontade das autoridades do país e da capacidade para implementar as principais mensagens da PER.

Ingredientes de uma PER de alta qualidade

Os tópicos a seguir descrevem algumas características-chave de uma PER de alta qualidade:

- **Dados de alta qualidade e oportunos são importantes.** Muitas PERs exploram dados das Estatísticas de Finanças Públicas (GFS) para despesas por funções e principais categorias econômicas, o que permite o estudo das tendências na alocação de despesas. Para o trabalho setorial, são necessários dados detalhados confiáveis, inclusive para despesa por unidade orçamentária

19 A sigla PFMC significa *Public Financial Management and Control*.

(na esfera central e nas diversas esferas subnacionais), indicadores de resultados, custos unitários, indicadores de qualidade etc. Talvez seja necessário um esforço considerável para estabelecer um banco de dados multidimensional adequado, como nos moldes dos dados do BOOST. Os dados da pesquisa também são de grande ajuda. Se houver suspeita de perdas de receita, as PERs são úteis para revelar as discrepâncias entre o orçamento e os dados do resultado;

- **Técnicas específicas são necessárias para a análise de gastos.** As PERs podem ser enriquecidas com uma série de técnicas, que precisam ser utilizadas em análises de eficiência, eficácia e equidade da despesa, inclusive BIAs, DEAs e outras, para avaliar a eficiência e o impacto da despesa sobre os pobres;
- **PERs requerem sólidas habilidades analíticas e conhecimento setorial especializado.** As PERs requerem pessoal competente com habilidades sólidas para analisar as questões da política de despesas. O conhecimento especializado de peritos de setores (agricultura, educação, saúde, infraestrutura etc.) é de grande ajuda. O Banco Mundial possui vantagem comparativa no fornecimento desse conhecimento;
- **Comparações internacionais em PERs são muito úteis.** Pela comparação dos principais resultados da despesa ou custos unitários com os resultados dos países em estágio semelhante de desenvolvimento, é possível estabelecer metas para examinar os resultados da despesa de um país do mesmo nível, levando em conta as limitações de tal comparação, dadas as diferenças entre os países em termos de instituições e capacidade;
- **A análise de PER precisa ser oportuna.** A análise de PER precisa estar atualizada e ser fornecida às autoridades do país de maneira oportuna.²⁰ Em PERs com projeções de desenvolvimentos de despesas futuras nos termos de opções diferentes (por exemplo, a PER da Turquia revisada na seção anterior), as propostas de políticas relacionadas a elas precisam

20 O Banco Mundial (2009) alerta os gerentes de PERs para definirem recursos de tempo e orçamento realistas para (i) missões para coletar dados; (ii) missões para discutir e disseminar resultados; e (iii) acompanhar as discussões da política. Os atrasos na realização de cada etapa estão incluídos nos Termos de Referência de uma PER.

ser apresentadas às autoridades do país antes que as projeções fiquem obsoletas e, de preferência, em um ponto crítico do ciclo de preparação do orçamento anual;

- **Colaboração com organizações multilaterais evita incoerência na consultoria sobre políticas.** As recomendações sobre políticas das PERs lideradas pelo Banco Mundial precisam ser coerentes com as de outras organizações. Em 2003, o Banco Mundial e o FMI fizeram um acordo sobre as diretrizes para a colaboração em um trabalho sobre o gasto público, que foi bom em muitos países, inclusive Bulgária, Peru e Turquia.²¹ Da mesma forma, não é raro órgãos de desenvolvimento regional participarem de PERs. Por exemplo, o Banco de Desenvolvimento Interamericano (BID) contribuiu para a PER do Peru de 2002 e a PER da Guatemala de 2012;
- **Outras instituições de desenvolvimento podem ter recomendações úteis.** Tanto os organismos financiados com recursos públicos quanto privados realizam análises e avaliações da qualidade da despesa pública. Por exemplo, o Instituto de Desenvolvimento Além-mar do Reino Unido (mediante seu programa Cape²²) prepara análises do efeito geral da ajuda e do gasto público em uma série de países. As PERs devem levar em consideração as recomendações sobre políticas das organizações voltadas para pesquisa; e
- **É necessária uma estratégia para a preparação, finalização, disseminação e acompanhamento.** Para tanto, o Banco Mundial preparou notas de orientação para o trabalho de PER em desenvolvimento humano (cobrindo educação, saúde e proteção social), agricultura e silvicultura.²³ As notas fornecem

21 Para obter detalhes, veja Banco Mundial e FMI (2003), que apresenta as seguintes diretrizes: (i) o FMI deve ser o organismo principal para aspectos de política macroeconômica agregada, inclusive despesa e receita do setor público; (ii) o Banco Mundial deve ser o organismo principal em questões relacionadas à composição e eficiência dos gastos públicos, inclusive a reforma das empresas estatais, reforma do serviço público, reforma da previdência e proteção social; e (iii) a gestão do gasto público é uma área de responsabilidade compartilhada.

22 Cape é o Centro de Ajuda e Despesa Pública. Ver <<http://www.odi.org.uk/programmes/aid-public-expenditure>>.

23 Ver Banco Mundial (2009) para PERs de desenvolvimento humano, Banco Mundial (2011c) para PERs sobre agricultura, e Fowler et al. (2011) para orientação sobre PERs de silvicultura.

uma lista de verificação de tópicos que devem ser considerados em PERs setoriais, referências e recursos úteis e exemplos de boas PERs. Em 2001 foi elaborada uma nota de orientação para PERs *macrofiscais* (BANCO MUNDIAL, 2001) que enfatizou a necessidade de uma abordagem estratégica para PERs.

Eficácia das PERs

As PERs contribuem para o entendimento de uma ampla variedade de questões de políticas de gastos. Elas oferecem aos MF, ministérios responsáveis pela despesa, e aos governos uma base sólida para a alteração das políticas fiscais em geral ou políticas de despesa específicas de setores.

Embora nenhuma análise independente tenha examinado a eficácia das PERs, as constatações das PERs, em alguns países, produziram alterações nas políticas de gastos.²⁴ A PER das Filipinas (BANCO MUNDIAL, 2011a) é um exemplo. A PER propôs alterações na política sobre redistribuição e transparência dos gastos; essas propostas foram adotadas e refletidas no orçamento de 2011.

Contudo, a eficácia de uma PER – seu impacto na melhoria da eficiência e equidade da despesa – depende da vontade política de um país e da capacidade administrativa para implementar as principais mensagens das PERs. As PERs têm probabilidade de serem mais eficazes quando são reconhecidos os seguintes elementos:

- **As restrições políticas e institucionais podem ser substanciais.** As restrições institucionais são atualmente analisadas de forma mais explícita nas PEIRs,²⁵ particularmente para as estruturas jurídicas que são inerentes às políticas de despesa (ver, por exemplo, as recomendações da PER rodoviária e ferroviária da Rússia, na seção anterior). Entretanto, é difícil para as PEIRs avaliar todas as restrições

24 Em 1998, o Departamento de Avaliação das Operações (OED) do Banco Mundial conduziu uma avaliação aprofundada da eficácia das PERs. Desde então, não tem havido análise aprofundada da eficácia da PER comparável, embora Deolaliker (2008) tenha destacado a importância das comparações entre custos unitários em nível micro e tenha defendido o maior uso de PERs em esferas subnacionais.

25 As Revisões das Despesas Públicas e Institucionais (PEIRs) enfocam particularmente as restrições institucionais.

políticas e administrativas que impedem a implementação das recomendações de uma PER;

- **A propriedade da PER por parte do país é importante.** Nos últimos anos, o Banco Mundial vem procurando ampliar o nível de participação nas PERs. Nos países de renda média, as PERs tendem a ser conduzidas pelo país ou em conjunto.²⁶ Porém, a propriedade política de alto nível talvez não seja adequada quando a visão para a reforma não é totalmente compartilhada dentro do executivo político,²⁷ pela administração pública ou com outros atores, tais como o parlamento ou os sindicatos de servidores públicos;
- **As opções de políticas devem ser viáveis e específicas.** As PERs revisadas neste capítulo ilustram a diversidade de propostas de políticas. Em PERs setoriais as recomendações de mudanças nas políticas são geralmente bastante específicas e direcionadas para (i) aumentar a eficiência ou a equidade (ou ambas); ou (ii) alinhar mais as políticas setoriais à boa prática internacional. Em contrapartida, para as PERs macrofiscais, tais como as que analisam o espaço fiscal para novas políticas de despesa, a tendência é apresentar opções de políticas realistas em lugar de adotar uma linha rígida para medidas específicas de aumento de receita ou reorientação de despesa, para as quais as autoridades políticas nacionais exercerão suas prerrogativas.

5.7 Considerações finais

As PERs continuam em demanda pelas autoridades dos países para a análise dos gastos públicos e desenvolvimentos fiscais, estabelecendo mecanismos eficazes e transparentes para a alocação de recursos públicos e melhoria da eficiência da despesa. Não há

26 O Banco Mundial (2001) identifica três protótipos de PERs: (i) *conduzida em conjunto ou pelo país*, na qual o Banco Mundial fornece suporte substancial para a liderança cliente-país; (ii) *participativa liderada pelo Banco Mundial*, na qual a gestão geral da PER é efetuada pelo Banco Mundial, mas há substancial participação do país; e (iii) *totalmente interna*, na qual a coleta e análise de dados são realizadas pelo pessoal do Banco Mundial ou consultores externos, com participação limitada do país.

27 Isso está ilustrado no Capítulo 4 na discussão da revisão da despesa da França, 2007-2011, na qual a iniciativa presidencial não foi totalmente aceita em outros setores do governo.

uma estrutura única para todas as PERs, que são adaptadas às necessidades específicas de cada país. Não obstante, as PERs enfocam (i) questões macrofiscais, tais como a sustentabilidade de médio prazo de agregados do orçamento ou a alocação estratégica dos recursos orçamentários ou (ii) os setores de despesa, onde há forte ênfase na eficiência das políticas de despesas setoriais, a equidade e a incidência de despesa, além da escolha entre o fornecimento público ou privado de bens e serviços. Algumas PERs combinam questões macrofiscais e específicas de setores, reunindo diversas questões de despesa em um pacote coerente para opções de reforma da despesa.

O conteúdo das PERs é bastante variável e reflete a diversidade das questões de gastos entre os países e as diferenças de prioridades de despesas fiscais e públicas das autoridades nacionais. Cada vez mais, as PERs incluem uma análise das instituições orçamentárias, especialmente as estruturas jurídica e normativa para despesa. A gestão e as práticas de planejamento das despesas agora também recebem mais foco do que em PERs anteriores. As questões de economia política são incluídas de alguma forma.

É impossível precisar o efeito exato das PERs nas prioridades e políticas de despesa dos países porque essa análise exigiria saber o que teria acontecido na ausência da PER. Entretanto, com o benefício de mais de 30 anos na preparação de PERs, os principais ingredientes de uma PER de alta qualidade já são conhecidos. Ademais, a eficácia de uma PER para reorientar os resultados da despesa aumenta quando as opções de alteração de políticas são claras, viáveis e cientes das restrições políticas e administrativas nacionais. O êxito de uma PER na alteração de políticas de gastos depende da aceitação das propostas por parte das autoridades políticas de um país – sua disposição, não apenas para aprovar a alteração da política, como também para assegurar sua total implementação.

REFERÊNCIAS

- AFONSO, A.; SCHUKNECHT, L.; TANZI, V. Public sector efficiency: an international comparison. *Public Choice*, v. 123, n. 3, p. 321-347, 2005.
- DEOLALIKAR, A. B. *Lessons from the World Bank's public expenditure reviews 2000-2007 for improving the effectiveness of public spending*. Washington, DC: Brookings Institution, 2008.
- FOWLER, M. ABBOTT, P.; AKROYD, S.; CHANNON, J.; DODD, S. *Forest sector public expenditure reviews: review and guidance note*. Program on Forests. Washington, DC, 2011. Disponível em: <<http://www.profor.info/knowledge/forest-sector-public-expenditure-reviews-toolkit>>.
- KHEYFETS, I.; MASTRUZZI, M.; MEROTT, D.; SONDERGAARD, L. *A new data tool to boost public spending efficiency*. Europe and Central Asia Knowledge Brief 43. Washington, D.C: Banco Mundial, 2011.
- LIENERT, I.; MOO-KYUNG, J. The legal framework for budget systems: an international comparison. *OECD Journal on Budgeting*, Special Issue, v. 4, n. 3, 2004.
- MANNING, N.; PARISON, N. *Modernization of the federal civil service: methodology for functional reviews*. Paper produced for the Russian Federation Government. Washington, DC; Moscow: World Bank, 2001.
- OPERATIONS EVALUATION DEPARTMENT (OED). *The impact of public expenditure reviews: an evaluation*. Washington, D.C: Banco Mundial, 1998. (Report n. 18573, nov. 13).
- PSACHAROPOULOS, G.; PATRINOS, H. *Returns to investment in education: a further update*. Washington, DC: World Bank, 2002. (Policy Research Working Paper 2881).
- PRADHAN, S. *Evaluating public spending: a framework for public expenditure reviews*. Washington, DC: World Bank, 1996. (Discussion Paper n. 323).
- RODRIGUEZ, A.; HERBST, M. *Public sector expenditure review for the education sector in Poland*. Washington, DC: World Bank, 2009.
- WORLD BANK. Draft guidelines for the world bank's work on public expenditure analysis and support (including PERs). In: _____. *Public sector group, poverty reduction, and management network*. Washington, DC, 2001.

_____. Policy Report. v. I. In: _____. *Brazil municipal education: resources, incentives, and results*. Washington, DC, 2002a.

_____. Research Report. v. II. In: _____. *Brazil municipal education: resources, incentives, and results*. Washington, DC, 2002b.

_____. *Mexico: infrastructure public expenditure review*. World Bank, Finance, Private Sector and Infrastructure Unit, Latin America and the Caribbean Region. Washington, DC, 2005. (Report 33483-MX).

_____. *Improving the quality of health spending resource management in Brazil's Unified Health System (SUS)*. World Bank, PREM Unit, Latin America and the Caribbean Region. Washington, DC, 2006a. (Report 36601-BR).

_____. *Turkey public expenditure review*. World Bank, PREM Unit, Europe and Central Asia Region. Washington, DC, 2006b. (Report 36764-TR).

_____. *Indonesia public expenditure review – spending for development: making the most of Indonesia's new opportunities*. Jakarta, 2007.

_____. *Costa Rica public expenditure review: enhancing the efficiency of expenditures*. World Bank, Central America Department, PREM Unit, Latin America and Caribbean Region. 2008. (Report 40774-CR).

_____. *Preparing PERs for human development*. Core Guidance Note, Human Development Network. 2009. Disponível em: <<http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTEDUCATION/Guidance>>.

_____. *Philippines public expenditure review: strengthening public finance for more inclusive growth*. Washington, DC, 2011a. (Report 55695).

_____. *Russian Federation: social expenditure and fiscal federalism*. World Bank, Human Development Sector Unit, Europe and Central Asia Region. Washington, DC, 2011b. (Report 54392-RU).

_____. *Practitioners' toolkit for agriculture public spending analysis*. Washington, DC, 2011c. (Report 60015-GLB).

_____. *Guatemala public expenditure review*. World Bank, PREM Unit, Latin America and Caribbean Region, Washington, DC, 2012a. (Report 67363-GT).

_____. *Public expenditure review for Peru*. World Bank, PREM Unit, Latin America and Caribbean Region. Washington, DC, 2012b. (Report 62586-PE).

_____. *Russia public expenditure review*. World Bank, PREM Unit, Europe and Central Asia Region, Washington, DC, 2012c. (Report 58836-RU).

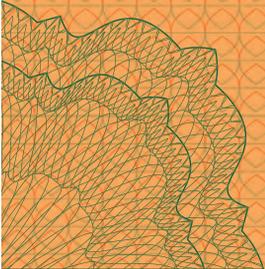
WORLD BANK. International Monetary Fund. *Bank/Fund Collaboration on Public Expenditure Issues*. IMF Fiscal Affairs Department and World Bank Poverty Reduction and Economic Management (PREM) Network. Washington, DC, 2003.

WORLD ECONOMIC FORUM. *Global Competitiveness Report, 2004-2005*. Edited by Michael E. Porter, Klaus Schwab, and Xavier Sala-i-Martin. New York: Oxford University Press, 2004.

PARTE III

O QUE É EFICIÊNCIA?





Capítulo 6

Conceitos sobre eficiência

Enlinson Mattos
Rafael Terra

6.1 Introdução

O problema central da teoria econômica é sabidamente a escassez. Por isso, em um mundo com restrições, os economistas se devotam a desenvolver uma teoria que forneça sugestões de políticas para tirar maior proveito dos recursos disponíveis. De fato, sob a ótica da teoria econômica, eficiência tem a ver exatamente com isso: obter o maior benefício possível de uma quantidade fixa de recursos.

Estudos sobre as condições suficientes para resultados eficientes deram origem ao Primeiro Teorema do Bem-Estar Social. De acordo com o teorema, se todos os bens forem transacionados em um mercado com preços definidos publicamente, e consumidores e produtores agirem competitivamente, então a alocação e a produção de bens serão Eficientes de no sentido de Pareto. Economistas neoclássicos procuram elaborar políticas no sentido de garantir as condições necessárias para atingir a eficiência, como mercados domésticos competitivos, livre comércio, mercado de trabalho não sindicalizado, impostos não distorcivos etc.

Apesar de o Primeiro Teorema do Bem-Estar assegurar eficiência sob determinadas condições, ele não garante que a alocação será justa, i.e., é possível que a distribuição de recursos seja assimétrica entre os agentes. Nesse caso, a teoria econômica prevê que é possível atingir qualquer alocação Pareto ótima por meio de transferências *lump-sum* de riqueza, contanto que as preferências e os conjuntos de produção sejam convexos, os mercados sejam completos, os preços sejam conhecidos publicamente e os agentes ajam competitivamente. A essa conclusão dá-se o nome de Segundo Teorema do Bem-Estar Social: se a sociedade preferir uma distribuição de riqueza mais

igualitária, o governo poderia atuar no sentido de equalizar as dotações dos agentes por meio de impostos e transferências *lump-sum*, e ainda assim lograr um resultado Eficiente de Pareto.

Note que impostos *lump-sum* não distorcem os comportamentos dos agentes, i.e., não provocam um *efeito substituição*, pois não alteram os incentivos para consumirem um ou outro bem, ou ofertarem mais ou menos trabalho. Na prática, no entanto, a aplicação de impostos *lump-sum* é extremamente complicada. Um imposto do tipo *head-tax*, por exemplo, prevê a cobrança de um montante fixo a todos os agentes. É fácil perceber que tal imposto é regressivo e dificilmente obteria apoio dos agentes para sua implementação. Uma alternativa seria um imposto de montante fixo incidente de acordo com características dos contribuintes que determinam maiores rendimentos. Ocorre que tal tipo de imposto dificilmente seria factível. Um imposto em função (crescente) do QI, por exemplo, seria razoável, mas de difícil implementação, pois os agentes não teriam incentivos para revelarem seus verdadeiros quocientes de inteligência.

Portanto, para realizar redistribuições ou outras atividades o governo é obrigado utilizar outros tipos de impostos, dentre os quais se destacam aqueles sobre bens e sobre a renda. Tais impostos geram distorções e resultam em ineficiências. Ainda que estudos sobre tributação ótima se dediquem a descrever as condições sob as quais tais distorções são minimizadas, alguma distorção é inevitável para que o governo cumpra suas funções básicas.

Esses tipos de distorções ou ineficiências estão no cerne da Teoria Microeconômica. No entanto, existem outros conceitos de eficiência relacionados especificamente com a atividade produtiva. Nesse caso, a preocupação é com o uso eficiente dos recursos escassos. Esses conceitos se baseiam no pressuposto de que a relação entre insumos e produtos não é determinística como preconizado pela Teoria Microeconômica tradicional, e.g., a quantidade de produto produzida com um montante fixo de insumos varia de acordo com fatores motivacionais. Esse tipo de ineficiência pode existir mesmo se as condições básicas do Primeiro Teorema forem válidas e a despeito do objetivo de maximização de lucros das firmas.

Este capítulo procura descrever os conceitos básicos de eficiência na produção, fazendo uma diferenciação entre eficiência, eficácia

e produtividade, e depois definindo os vários tipos de eficiência relacionados à atividade produtora. Em seguida, são apresentados conceitos de eficiência na produção de bens e serviços pelo setor público. Nesse sentido, o problema de definição dos objetivos do setor público é apresentado, assim como o indicador de desempenho do setor público de Afonso, Schuknecht e Tanzi (2005), que procura refletir os objetivos corretos associados à maximização do bem-estar da população.

6.2 Conceitos básicos de eficiência na produção de bens

Produção é o processo pelo qual insumos são combinados de uma maneira específica para serem transformados em produtos, o qual é levado a cabo por unidades produtoras denominadas pela teoria econômica como *firmas*, ou como unidades de decisão (DMUs) pelas ciências da gestão. Esse processo não é determinístico e pode variar devido a diversos fatores. Nesse sentido se inserem conceitos como eficiência e produtividade, que refletem uma preocupação em obter o maior retorno econômico possível com uma quantidade escassa de recursos. Comumente se confundem – erroneamente – as definições de produtividade e eficiência. Por exemplo, um produtor é dito mais ou menos eficiente, quando na verdade o intuito era classificá-lo como produtivo.

Produtividade

A Produtividade Média pode ser definida como a razão entre produtos e insumos – ou quantas unidades de produto podem ser produzidas por unidade de insumo. Essa medida é frequentemente utilizada para aferir a performance de uma firma, indústria ou economia. Tendo em vista que uma firma pode usar vários insumos para produzir diversos produtos, o cálculo da produtividade se torna mais complexo, pois exige que se usem medidas agregadas de insumos e produtos, e nesse caso teremos uma medida de Produtividade Total dos Fatores.

Modelos macroeconômicos de crescimento lidam constantemente com a questão dos diferenciais de produtividade entre países. Por ser difícil modelar o progresso técnico que leva a aumentos de produtividade, esta é tratada como resíduo de uma função de

produção que pode ser atribuído a diversos fatores. Fried, Lovell e Schmidt (2008) destacam o estado da tecnologia de produção, o ambiente em que ela se dá, a eficiência e a escala de operação das firmas como os principais fatores. Determinar as causas de aumentos ou quedas na produtividade é essencial para saber quais práticas incentivar ou desincentivar por meio de políticas públicas ou atos privados. Somente os fatores ligados ao ambiente de produção – incluindo fatores institucionais, culturais e religiosos – não podem ser ajustados livremente para a melhoria da produtividade. Os demais fatores, no entanto, estão todos sob o controle de agentes privados ou públicos.

Tipos de eficiência

Há diversos conceitos de eficiência e é importante diferenciá-los. Um deles é o de eficiência técnica, que pode ser definida pela ótica do produto como a diferença entre o montante efetivamente produzido com certa quantidade de insumos e o montante factível de ser produzido, dada a tecnologia disponível. Esse montante factível de produção sob a ótica do produto é descrito pela Fronteira de Possibilidades de Produção (FPP). Sob outra ótica, a do insumo, a eficiência técnica pode ser entendida como a diferença entre a quantidade de insumos efetivamente utilizada para produzir determinado nível de produto e o montante mínimo factível de insumos necessários para produzir esse mesmo nível de produto com a tecnologia de produção disponível. Nesse caso, a combinação de insumos factíveis para a produção de um dado nível de produto almejado é descrita pela Fronteira de Custos.

O conceito de eficiência de escala está associado à eficiência técnica. A ineficiência de escala surge quando a firma opera em uma escala desfavorável – aquela em que a produtividade média dos insumos não é máxima. Sob a ótica do produto, um exemplo refere-se à exploração de recursos comuns, como no caso da indústria pesqueira. Considere que a área onde opera uma firma permite a produção de até 20 toneladas/dia de peixe e que os insumos da firma sejam somente dois barcos pesqueiros com capacidade de produzir 4 toneladas/dia de peixe, cada. Em um primeiro momento, somente essa firma opera na área, mas no período seguinte mais quatro firmas idênticas passam

a explorar a área com dois barcos idênticos, cada uma. A produção de cada firma nesse segundo momento será igual a 4 toneladas/dia (2 toneladas/dia de pescado por barco). Nessa situação, os retornos à escala para cada firma serão decrescentes. Note que todas as firmas operam eficientemente, i.e., sobre suas fronteiras de produção, produzindo o máximo possível com os seus insumos (dois barcos).

Mas se cada firma reduzisse o número de barcos para um, todas as firmas poderiam manter a mesma produção total de 4 toneladas/dia, cada uma. Portanto, a escala de cada firma após o aumento da competição é desfavorável. E haveria um ganho em reduzir a escala de operação.

No longo prazo, as funções de produção das firmas pesqueiras devem apresentar retornos constantes, isto é, as firmas podem sempre ajustar suas escalas de operação, de modo que aumentos ou diminuições na quantidade de insumos utilizados levem a variações equiporcionais no produto. No curto prazo, no entanto, esse ajuste pode não ocorrer, por uma série de razões, resultando em uma produção ineficiente em relação à função de produção de longo prazo. A eficiência técnica associada à tecnologia com retornos constantes é denominada *eficiência técnica global* e pode ser decomposta em um componente de eficiência associada à escala atual de operação, por vezes denominada *eficiência técnica pura local* ou *eficiência sob retornos variáveis a escala*, e outro componente que mede a distância entre a escala atual e a ideal (com retornos constantes). Tal distância pode ser medida a partir da razão entre a produtividade máxima (obtida na escala ótima) e a produtividade efetivamente obtida pela firma se ela fosse tecnicamente eficiente em sua escala atual de operação.

Outro conceito é o de eficiência alocativa, o qual está relacionado ao objetivo e ao comportamento da firma. Se busca maximizar os lucros, a firma irá escolher a quantidade a ser produzida, de modo que a diferença entre as receitas e os custos seja máxima. Sob outra perspectiva, se a firma busca minimizar os custos, ela irá escolher o montante e a proporção de insumos que produzam o total desejado ao mínimo custo. É possível haver eficiência do ponto de vista técnico mesmo que o lucro não seja o máximo ou o custo seja o mínimo possível. Nesse sentido, qualquer desvio em relação ao ponto de lucro ou custo ótimos,

ainda que haja eficiência técnica, representará ineficiência alocativa, no sentido de que uma melhoria paretiana seria possível.

Um último conceito é o de eficiência social, que está relacionado ao montante e à proporção ótimas de bens do ponto de vista social, i.e., aquela cesta de bens produzida que maximiza a utilidade da sociedade. Sob a ótica do produto, a utilidade percebida pela sociedade será máxima em um ponto em que não se poderá produzir mais de um bem sem reduzir a produção de outro, ou seja, um ponto sobre a fronteira de possibilidades de produção. Mas a combinação ótima de produtos do ponto de vista social pode ou não ser igual à produção ótima do ponto de vista alocativo. A presença de externalidades, bens públicos ou poder de mercado pode fazer com que a combinação de produtos não seja ótima do ponto de vista social, ainda que o seja do ponto de vista alocativo (em que a firma só considera a própria função custo ou receita). Na ausência de tais imperfeições de mercado, e na presença de um sistema de preços e mercados completos, a firma eficiente do ponto de vista alocativo também será eficiente do ponto de vista social.

Eficiência x eficácia

Outro conceito relacionado à Teoria da Produção – e que deve ser diferenciado de eficiência – é o de eficácia, que se refere à capacidade de atingir um objetivo proposto. Em termos técnicos, é a razão entre o produto planejado (ou desejado) e o produto efetivamente observado. Por exemplo, se o objetivo de uma firma for produzir uma quantidade Y em cada período e ela produz $\hat{Y} = 0,9Y$ no primeiro período e $\hat{Y} = Y$ no segundo, pode-se dizer que a mesma atuou de forma mais eficaz no segundo período em comparação ao primeiro. No entanto, isso não quer dizer que a firma produziu da melhor maneira possível ou com a menor quantidade de insumos. É possível que ela tenha sido eficiente ao produzir $0,9Y$ no período 1, e não o tenha sido no período 2, quando produziu Y .

Aspectos conceituais da mensuração da eficiência

Quando falamos em *mensurar a eficiência*, referimo-nos normalmente à eficiência técnica. Nesse sentido, Fried, Lovell e Schimidt (2008) destacam três problemas centrais para a mensuração. O primeiro se refere a quais insumos e produtos

devem ser considerados. O segundo, à ponderação que deve ser dada a cada insumo e produto sob análise quando vários produtos são produzidos com vários insumos. O terceiro se refere a como determinar a tecnologia de produção ótima.

O primeiro problema está intrinsicamente ligado à fonte das ineficiências. Há divergências teóricas entre as visões de Leibenstein (1966) e Stigler (1976) em relação ao que se pode considerar como fonte de ineficiência. Leibenstein (1966) argumenta que a eficiência alocativa não é tão importante quanto preconizado pela teoria econômica e que a maior fonte de ineficiência se deve ao que se convencionou chamar de X-efficiency (doravante denominada eficiência-X). Essa ineficiência se deve basicamente a três fatores: (i) eficiência motivacional interna à firma; (ii) eficiência motivacional externa; e (iii) eficiência de insumos sem mercados. Segundo o autor, a relação entre o emprego de insumos e a produção não é determinística, e a motivação desempenha um importante papel para a determinação da produção eficiente. Os trabalhadores e mesmo as firmas podem não agir com o afincamento necessário para produzir o máximo possível com os recursos disponíveis.

Do ponto de vista dos trabalhadores (agentes), essa questão motivacional intrafirma pode ser explicada pelo problema inerente a sua relação com a firma (principal), pela existência de contratos incompletos e pelas dificuldades de monitoramento dos agentes. A informação acerca do esforço empregado na produção é assimétrica entre os agentes e os principais. Se os custos de monitoramento dos agentes são proibitivos e os agentes têm responsabilidade limitada no resultado da firma, sem um mecanismo de incentivos adequado, os agentes maximizarão sua utilidade empregando um esforço mínimo. Esse mecanismo de incentivo adequado está associado à realização de contratos completos, i.e., que definam legalmente as consequências para cada possível estado da natureza. Mas a elaboração de tais contratos pode ter custos proibitivos, tornando-os ineficazes. Portanto, variações na eficiência-X estão associadas à habilidade dos principais em desenhar mecanismos (ou contratos) que incentivem o esforço dos agentes.

Os problemas motivacionais externos se referem predominantemente às características do ambiente institucional que podem desincentivar a operação eficiente da firma. Por exemplo,

em uma economia centralmente planejada (e.g. socialista) não há motivação para a firma ser eficiente, tendo em vista que sua sobrevivência não depende disso, diferentemente do que ocorre nos mercados competitivos. Mesmo em uma economia de mercado, a falta de competição pode determinar menor motivação para adotar as melhores práticas, pois estas não são tão essenciais para a sobrevivência das firmas. A motivação de firmas em mercados concentrados para produzir na fronteira de eficiência técnica pode ser substituída por outros objetivos, dando lugar à ineficiência-X. Outro exemplo é o de uma economia de mercado que tributa todo o lucro que exceder determinado patamar. Nesse caso, mudam os incentivos à maximização de lucros, e não necessariamente haverá adoção das práticas produtivas eficientes.

O último fator destacado como determinante da ineficiência se refere à existência de insumos sem mercado, ou ao acesso desigual das firmas a insumos com mercados. Tais fatores podem resultar em maior ou menor produção para um nível de insumos fixos. Por exemplo, em determinadas localidades as informações sobre as melhores práticas gerenciais ou melhores tecnologias podem estar mais facilmente disponíveis, conferindo uma vantagem produtiva em relação às firmas de outras localidades.

Stigler (1976) critica essa noção de eficiência-X e defende que o que esse conceito considera como ineficiência é, de fato, erro de especificação da função de produção. Para o autor, todos esses fatores deveriam entrar como insumos da função de produção. Especificando completamente a função de produção, o conceito de ineficiência-X não mais existiria: só desperdício. Para o autor, somente a ineficiência alocativa seria relevante para a teoria econômica.

O segundo problema para mensurar a eficiência destacado por Fried, Lovell e Schmidt (2008) surge quando múltiplos insumos são utilizados ou múltiplos produtos são produzidos. Nesse caso, uma ponderação natural para cada insumo ou produto é dada por seu preço. No entanto, variações nos preços podem ocorrer ao longo do tempo e entre firmas. A existência de monopólios ou monopsônios pode se refletir nos preços de mercado e distorcer os pesos. Ademais, como já mencionado, existem fatores externos não precificados pelo mercado que interagem com os insumos empregados na atividade

produtiva e aumentam a produção em relação àquelas firmas onde tais fatores externos não estão disponíveis.

O terceiro e último dos problemas de mensuração da eficiência destacados por Fried, Lovell e Schmidt (2008) se refere à determinação da fronteira de possibilidades de produção. Empiricamente, o exercício de estimar essa fronteira se dá por meio de comparações entre o nível de produção e a utilização de insumos de diversas firmas. No entanto, não há como saber se a unidade produtora que se utiliza das melhores práticas verificadas entre seus pares é, de fato, eficiente, no sentido de que nenhuma melhoria poderia ser atingida, dado o estado tecnológico. Dessa forma, a fronteira tecnológica empírica só pode ser considerada uma fronteira de melhores práticas, e não de eficiência tecnológica. Entretanto, à medida que o número de firmas utilizadas no exercício de avaliação empírica aumenta, a fronteira de melhores práticas tende a se aproximar da fronteira eficiência verdadeira.

Isso não quer dizer que a estimação dessas fronteiras seja desprovida de sentido. Mesmo diferindo do ideal de produção, as fronteiras empíricas permitem identificar a fonte das ineficiências, separando os determinantes *controláveis* dos *incontroláveis*. Com isso, políticas públicas e privadas podem ser implementadas para a diminuição da ineficiência. Ao mensurar sua eficiência periodicamente e identificar os fatores que podem ser alterados para diminuir as ineficiências, a firma procurará melhorar sua performance produtiva, pois com isso melhorará seus indicadores financeiros. Em termos agregados, esse comportamento terá reflexos no crescimento macroeconômico.

6.3 Conceitos básicos de eficiência na produção pública de bens

Afonso, Schuknecht e Tanzi (2006) destacam que a eficiência na produção de bens de um modo geral se baseia na relação entre quantidade de insumos e produtos ou custos e benefícios, e requer geralmente: (i) uma estimativa de custos; (ii) uma estimativa de produtos; e (iii) a comparação entre os dois. Com isso, para medir a eficiência na produção de bens, pode-se comparar essa relação produto/custo de uma determinada firma, país ou unidade de decisão (DMU) com um padrão ideal em que a produção é ótima do ponto de vista teórico; ou, na falta desse padrão, comparar com outras firmas, países ou unidades de decisão.

No caso da produção de bens públicos, esse procedimento também é válido. Mas há algumas dificuldades adicionais envolvidas. Primeiramente, como será visto adiante, a mensuração dos custos da atividade pública é complicada. Além disso, os objetivos do governo são distintos daqueles das firmas. Para que o benefício à sociedade seja máximo, além do *quanto produzir* de um bem, o setor público deve decidir qual bem produzir e de que forma.

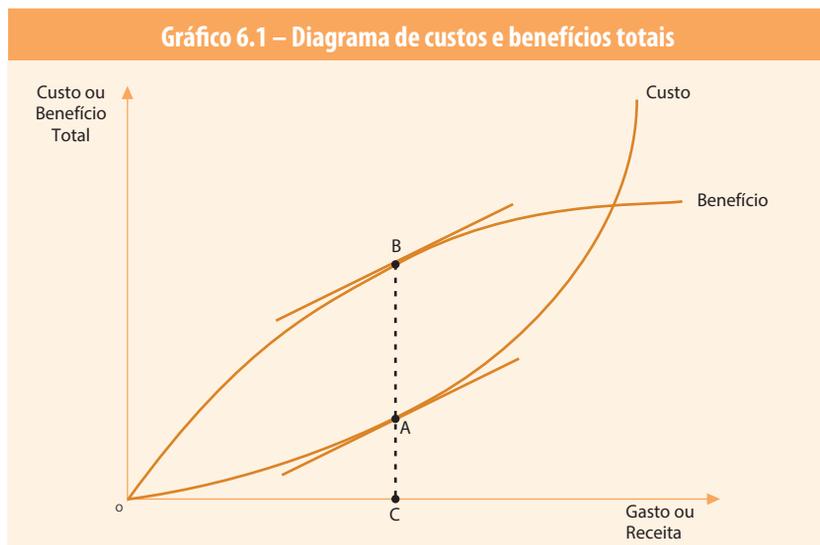
6.3.1 O problema da mensuração dos custos de produção de bens pelo setor público

Do ponto de vista de um agente privado, como uma firma, pode-se assumir uma curva de oferta de insumo perfeitamente elástica. Isto é, para produzir determinado montante de produto, a firma pode obter quantidades adicionais de determinado insumo aos mesmos preços das quantidades anteriores. O mesmo não se verifica no caso do gasto público, que é financiado pelas receitas de impostos distorcivos. Para obter receitas adicionais, o governo deve aumentar os impostos, que terão custos marginais crescentes e normalmente maiores do que o custo médio. De acordo com Afonso, Schuknecht e Tanzi (2006), tais impostos distorcivos normalmente apresentarão: (i) peso morto, associado ao efeito substituição que reflete a mudança nos incentivos decorrente dos impostos; (ii) custos adicionais de administração do sistema tributário; e (iii) custos adicionais aos contribuintes. Dessa forma, o custo marginal de arrecadação de uma unidade monetária adicional (que serve de insumo) é crescente e pode ser consideravelmente superior ao de uma unidade monetária. Diferentemente da produção privada, em que a oferta de insumos é perfeitamente elástica, para bens públicos o mesmo não se verifica.

Do ponto de vista social, o gasto público gera benefícios que devem ser levados em conta na mensuração da eficiência, assim como todos os custos associados à tributação. Obviamente, a incorporação de tais custos é problemática, o que leva a maioria dos estudos empíricos a focar na eficiência do uso da receita, i.e., na eficiência técnica do gasto público ou receita arrecadada.

No Gráfico 6.1 podem-se observar no eixo das abscissas os gastos e receitas, e no eixo das ordenadas, os benefícios e os custos decorrentes dos impostos que incluem o peso morto, os custos administrativos e

os custos adicionais infligidos aos contribuintes, inclusive subornos, caso o sistema tributário seja corrupto. Nesse caso, assume-se que a unidade de medida para custos e benefícios é a mesma.



Fonte: baseado em Afonso, Schuknecht e Tanzi (2006).

A curva de benefícios mostra que quanto maior o gasto público, maior é o benefício social; mas o benefício social marginal é decrescente (a segunda derivada é negativa), o que é bastante razoável, dado que o gasto público é utilizado para produzir bens e serviços que proporcionam utilidades marginais decrescentes aos indivíduos. A curva de custos mostra uma relação positiva entre custos e receitas. Quanto mais o governo arrecadar, maiores serão os custos impostos ao contribuinte, e esses custos crescem a taxas crescentes (a segunda derivada é positiva).

Supondo que o orçamento seja equilibrado, do ponto de vista da sociedade, o nível ótimo de arrecadação dar-se-á quando o benefício marginal de uma unidade de gasto público for igual ao custo marginal de arrecadação de uma unidade monetária. O nível ótimo é dado pelo ponto C, em que as inclinações das curvas de custos e benefícios totais são iguais e, portanto, a distância AB é máxima. Em pontos à esquerda de C, um aumento de uma unidade monetária

na arrecadação (e por consequência no gasto) gera um benefício adicional maior do que o custo marginal. Portanto, do ponto de vista social há espaço para aumentar a arrecadação. Já em pontos à direita de C há incentivos para o governo diminuir a arrecadação, pois o custo adicional que ela impõe à sociedade é maior do que o benefício marginal.

Dahlby (2008) caracteriza algebricamente a condição de produção ótima de bens públicos, em que o governo escolhe a alíquota tributária e os gastos com provisão de bens para maximizar o bem-estar social. Para verificar essa caracterização, considere uma função arrecadação total $R(\tau, G)$, em que τ é um vetor de n alíquotas incidentes sobre o mesmo número de bases tributárias, e G é o vetor de m bens e serviços providos pelo setor público. A restrição do governo é dada por $R(\tau, G) = C(G)$, em que $C(G)$ representa o custo total de provisão dos m bens. Note que esse custo difere daquele representado no Gráfico 6.1 e se refere somente ao custo de produção. Note também que a provisão de bens pelo setor público deve afetar as receitas tributárias. Com alíquotas tributárias fixas, os níveis de arrecadação afetarão a demanda ou a oferta de bases tributárias, mas o sinal é ambíguo. Ademais, tais bases são interdependentes, de modo que um aumento em τ_i afetará a receita obtida da base tributária j .

O benefício gerado pela provisão desses bens é capturado pela função de bem-estar social $W(U_1, \dots, U_p)$. Por simplicidade, assume-se aqui que as utilidades individuais são uma função das alíquotas τ e bens públicos G . Portanto, a função de bem-estar social pode ser representada por $W(\tau, G)$. É razoável supor que, em geral, $\partial W(\tau, G) / \partial \tau_i < 0$, pois um aumento na alíquota eleva os preços para os consumidores e reduz os preços percebidos pelos produtores. Adicionalmente, supõe-se que $\partial W(\tau, G) / \partial G_j > 0$, ou seja, quantidades adicionais de bens públicos, em média, aumentarão as utilidades individuais e o bem-estar social. O problema do governo pode ser representado por:

$$\max_{\tau, G} W(\tau, G)$$

$$s.a. R(\tau, G) = C(G) \quad (6.1)$$

O Lagrangeano é representado por:

$$\mathcal{L} = W(\tau, G) + \lambda(R(\tau, G) - C(G)) \quad (6.2)$$

As condições de primeira ordem nesse caso são dadas por:

$$\frac{\partial W(\tau, G)}{\partial \tau_i} + \lambda \frac{\partial R(\tau, G)}{\partial \tau_i} = 0 \quad i = 1, \dots, n \quad (6.3)$$

$$\frac{\partial W(\tau, G)}{\partial G_j} + \lambda \left(\frac{\partial R(\tau, G)}{\partial G_j} - \frac{\partial C(G)}{\partial G_j} \right) = 0 \quad j = 1, \dots, m \quad (6.4)$$

Rearranjando (6.3) temos:

$$CMFPS_{\tau_i} = - \frac{\partial W(\tau, G) / \partial \tau_i}{\partial R(\tau, G) / \partial \tau_i} = \lambda \quad i = 1, \dots, n \quad (6.5)$$

Em que $CMFPS_{\tau_i}$ é o custo marginal social de financiamento do setor público e representa o custo de arrecadar uma unidade monetária adicional da base tributária i . A condição (6.5) nos diz que, no ótimo, o $CMFPS_{\tau_i}$ deve ser igual entre as bases tributárias.

Substituindo (6.5) em (6.4) e rearranjando:

$$\frac{\partial W(\tau, G)}{\partial G_j} + CMFPS_{\tau_i} \left(\frac{\partial R(\tau, G)}{\partial G_j} - \frac{\partial C(G)}{\partial G_j} \right) = 0 \quad (6.6)$$

Denotando o benefício marginal social $\frac{\partial W(\tau, G)}{\partial G}$ por BSM_{G_j} , a receita marginal $\frac{\partial R(\tau, G)}{\partial G_j}$ por R_{G_j} , o custo marginal $\frac{\partial C(G)}{\partial G_j}$ por CM_{G_j} e rearranjando resulta em:

$$BSM_{G_j} = CMFPS \left(CM_{G_j} - R_{G_j} \right) \quad (6.7)$$

Essa condição para provisão ótima de bens públicos via impostos distorcivos é também conhecida como regra de Atkinson-Stern (1974). De acordo com essa regra, o Custo Social Marginal (CMS_{G_j}) é tal que $CMS_{G_j} = CMFPS \left(CM_{G_j} - R_{G_j} \right)$, ou seja, é dado pela diferença entre o custo de se produzir uma unidade adicional de bem público e o aumento na arrecadação proporcionado por essa unidade adicional, multiplicado pelo custo de financiamento do setor público. Note que

no Gráfico 6.1, CMS_{G_j} representa o custo total e a condição de ótimo (6.7) está associada ao ponto C do gráfico.

Outra dificuldade na mensuração do custo da provisão dos bens providos pelo setor público se refere à dificuldade em se considerarem os possíveis usos alternativos dos recursos. Por exemplo, escolas, hospitais e prédios públicos em geral poderiam ser alugados para o setor privado, e para obter os verdadeiros custos de cada serviço provido, tais fatores deveriam ser levados em consideração. Outro fator comumente ignorado e que leva a uma significativa subestimação dos verdadeiros custos de provisão de determinado bem é a necessidade de um governo instituído para que o bem seja produzido. Por exemplo, normalmente considera-se que o custo de provisão de defesa, saúde etc. é dado pelo montante destinado pelo governo a essas áreas, e o custo de ter um governo instituído é ignorado.

6.3.2 Eficiência no gasto público e objetivos do governo

Ao analisar a eficiência da despesa pública, é possível observar uma produção do bem público eficiente do ponto de vista técnico, e.g. o produto é máximo com um montante de despesas (ou insumos) fixo, mas não eficiente do ponto de vista social, i.e., não produz o maior bem-estar social factível. Se o governo produzir uma quantidade desproporcionalmente grande de um bem que não gera muita utilidade aos indivíduos, mesmo de forma tecnicamente eficiente, ele estará sendo ineficiente do ponto de vista social. Isso pode ocorrer, por exemplo, quando um governo gasta muito, e de forma tecnicamente eficiente, com defesa, mas pouco com saúde, que em geral é um bem mais valorizado pelas pessoas.

Poder-se-ia argumentar que se o governo gastar mais em defesa, em uma sociedade democrática, deve ser porque os indivíduos preferem mais gastos com esse bem, pois os políticos eleitos tendem a reproduzir as preferências de seus eleitores. No entanto, uma série de fatores pode contribuir para que os políticos ajam de forma distinta da exigida por seus eleitores. A captura do Estado por grupos privados e a corrupção podem fazer com que sejam perseguidos outros interesses que não aqueles que maximizam o bem-estar social, e esses interesses podem ser compatíveis com a eficiência técnica de determinados

tipos de bens. Portanto, a eficiência técnica de determinado bem analisado isoladamente não fornece muita informação relevante para uma análise da qualidade do governo como um todo. A análise deve ser mais abrangente e definir precisamente quais objetivos a sociedade mais valoriza e deseja ver atendidos.

Afonso, Schuknecht e Tanzi (2006) destacam que também é possível que os objetivos certos sejam perseguidos, mas de forma ineficiente. Mesmo que o governo defina os objetivos corretos, estabelecendo prioridades de gasto com categorias específicas de bens públicos, é possível que os recursos sejam capturados dentro dessas categorias para a produção de bens e serviços menos valorizados pela sociedade. Por exemplo, dado um orçamento para a saúde considerado ótimo pela sociedade, grupos de prestadores de serviço com influência sobre a divisão do orçamento da saúde podem pressionar para que os recursos sejam direcionados para hospitais de alta tecnologia que atendam regiões ricas, em vez de hospitais voltados aos cuidados básicos atendendo regiões mais pobres. No âmbito da educação, o orçamento pode ser, em sua maioria, direcionado para o ensino superior, para atender alunos de maior poder aquisitivo, enquanto a educação pública básica, que atende crianças e jovens mais pobres, é relegada a segundo plano. Considerando que serviços de saúde e educacionais proporcionem utilidade marginal decrescente aos indivíduos, assumindo-se que o bem-estar é dado pela soma de utilidades, este seria maior se os bens fossem direcionados aos mais pobres e com menor capacidade de obtê-los no setor privado.

Outra forma de captura de recursos destinados a determinada área pode ocorrer se grupos de prestadores de serviços públicos forem organizados o suficiente para ficarem com a maior parte dos recursos. Serviços como educação, por serem trabalho-intensivos, apresentam normalmente grande parcela de gastos com pessoal. Essa parcela pode ficar ainda maior se os prestadores públicos de serviços educacionais pressionarem por salários por meio de ameaças de greve. O aumento na importância da folha de pagamentos pode diminuir os recursos disponíveis para gastar em outros insumos, como instalações escolares e material pedagógico. Como esses insumos são complementares à mão de obra, os salários pagos podem facilmente exceder a produtividade marginal do trabalho. Note que nesse caso pode haver eficiência técnica no sentido de a combinação de insumos

utilizada gerar o maior nível de produto factível. Mas com a mesma despesa pública seria possível aumentar a quantidade de insumos disponível e produzir mais, ou seja, a despesa pública é ineficiente.

Dessa forma, se os objetivos são corretos, mas recursos de determinada área são destinados para usos que gerarão menor bem-estar do que outras alternativas, ou, ainda, se há captura de recursos para pagamento de salários por grupos influentes dentro dessas áreas, então o governo deve ser mais específico na definição de seus objetivos, atribuindo metas para cada componente a fim de maximizar o bem-estar social. Ocorre, no entanto, que pode ser inviável tamanho grau de detalhamento nos objetivos do governo.

Essa discussão demonstra que a despesa na produção de determinado tipo de bem não deve ser confundida com o benefício ou resultado gerado. A despesa, direcionada para os insumos errados ou menos adequados, irá produzir menos bens públicos ou de pior qualidade, resultando em menores benefícios para a sociedade.

Um fator crucial para a eficiência do gasto público se refere à existência de corrupção. Em um ambiente permeado por casos de corrupção, parte do dinheiro público é utilizada para fins indevidos, como a produção de bens e serviços de qualidade inferior, ou, ainda, o atendimento de interesses de terceiros. Por exemplo, fraudes em licitações constituem um grande problema para as administrações públicas brasileiras. Considerando que contratos celebrados entre uma parte pública e outra privada para a aquisição de bens ou serviços normalmente envolvem montantes vultosos, os agentes privados têm incentivos para burlar os procedimentos legais e corromper os agentes públicos encarregados da escolha dos vencedores, especialmente na ausência de mecanismos de inibição de práticas ilícitas, como a aplicação de penas severas. Portanto, a corrupção na administração pública pode ser considerada o exemplo mais óbvio de ineficiência da despesa pública, pois determina menor produção de bens públicos, ou bens de qualidade inferior.

A ineficiência pode se dever também a fatores motivacionais, culturais e ambientais, que refletirão na produção de bens e serviços públicos. A eficiência decorrente desses fatores é a já conhecida eficiência-X (ou *X-efficiency*) e está relacionada com a eficiência técnica. No setor público, tais fatores podem ser ainda mais importantes na

determinação da eficiência. A motivação para o esforço, por exemplo, pode decrescer com a estabilidade no emprego característica do serviço público. Com isso, a produção de bens e serviços, dados os mesmos insumos, tende a ser menor. De fato, a estabilidade no emprego pode ser um dos principais determinantes da ineficiência técnica. Ademais, o monitoramento do esforço do servidor público é mais complexo do que o do funcionário do setor privado, pois há maior dificuldade em se mensurar o produto da atividade.

Iniciativas têm sido tomadas em todo o mundo no sentido de tornar mais eficiente a administração pública. Para isso, os governos têm elaborado mecanismos de incentivo à produção eficiente de bens e serviços. Um princípio importante nesse sentido é o de *accountability*, pelo qual as informações devem ser tornadas públicas para que os eleitores possam acompanhar o desempenho dos governos eleitos e dos servidores públicos. De preferência, o desempenho das administrações públicas deve ser mensurável, permitindo a realização de comparações entre unidades de governo e até dentro de um mesmo governo. Tal iniciativa incentiva a responsabilização dos representantes da população, que buscarão melhorar seus indicadores ao exigir mais empenho por parte dos servidores, a fim de se reelegerem.

Para incentivar o esforço dos servidores, pode-se recorrer a um sistema de pagamento por mérito baseado em indicadores de produtividade. Em 2008, a rede estadual de Educação de São Paulo adotou esse mecanismo para estimular o empenho dos professores em melhorar a qualidade da educação local. Outro expediente, utilizado na administração pública federal, é a concessão de cargos comissionados (conhecidos como DAS) aos servidores com desempenho destacado, o que na prática equivale a uma promoção. Em teoria, os servidores se esforçariam para receber essa *promoção* e o aumento salarial associado. No entanto, no Brasil tais cargos podem ser distribuídos livremente entre servidores e não servidores, e muitas vezes são usados para favorecer determinadas pessoas por razões políticas ou não ligadas ao melhor interesse público.

Por fim, do ponto de vista social, o objetivo do governo deve ser maximizar o bem-estar da sociedade, o que envolve produzir os bens

e serviços certos de forma eficiente. Para isso, os mecanismos de incentivo adequados devem ser implementados e os objetivos do governo devem refletir o bem-estar da população. Nesse sentido, a definição e a mensuração de tais objetivos, são cruciais e são analisados na próxima seção.

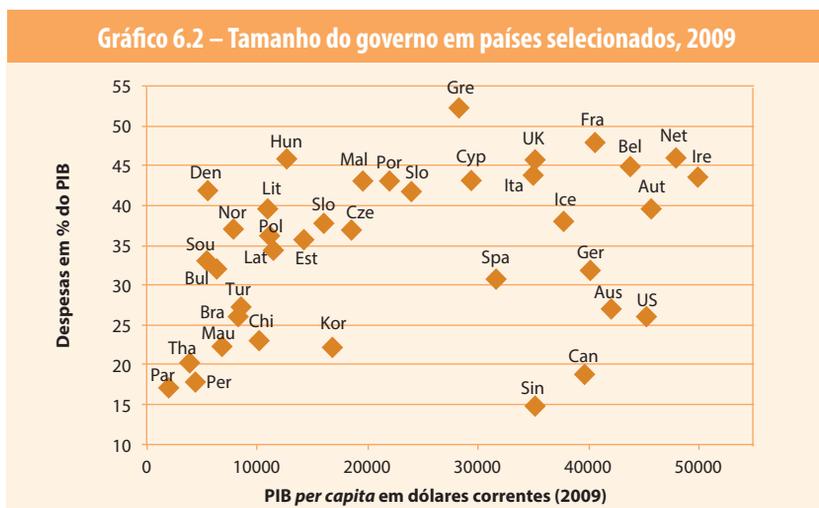
6.3.3 Um indicador de desempenho do setor público

A elaboração de indicadores de desempenho do setor público é crucial para o monitoramento das atividades, para a mensuração do bem-estar social e para a determinação da eficiência do setor público, tanto do ponto de vista técnico quanto do social. Em especial, os indicadores de desempenho do setor público permitem calcular a eficiência da despesa pública, em que o indicador representa o produto, e a despesa, o insumo da função de produção.

O indicador de desempenho do setor público é normalmente multidimensional, pois deve refletir os objetivos corretos, isto é, aqueles associados a um maior bem-estar da população. Para cada dimensão deve existir um subindicador que meça os resultados da atividade correspondente e que consiste em uma *proxy* para determinado objetivo do setor público. Por exemplo, se o objetivo é uma população altamente qualificada, o indicador multidimensional pode ser composto por uma média ponderada entre subindicadores, como o desempenho de alunos do ensino básico em matemática e leitura, a média de anos de escolaridade, o atraso escolar, o percentual da população com ensino superior, a produção científica e tecnológica, entre outros.

As medidas de desempenho do setor público podem ser *macro* ou *micro*. Medidas *macro* buscam fornecer um panorama geral do desempenho do setor público e do bem-estar proporcionado à população. Com isso, pode-se verificar a eficiência da despesa pública total. Já as medidas de desempenho *micro* se destinam a medir o desempenho do setor público em áreas específicas e a eficiência do gasto público nessas áreas. Tais medidas são limitadas, pois não levam em consideração o peso dessas áreas específicas no bem-estar total da população. Todavia são importantes, pois permitem identificar detalhadamente onde se encontram os problemas na provisão de bens e serviços públicos.

As medidas *macro* de desempenho permitem verificar, por exemplo, o bem-estar geral da população, se o gasto público total tem algum impacto sobre esse bem-estar e se o gasto público é eficiente. O Gráfico 6.2 apresenta a relação entre o Produto Interno Bruto (PIB) *per capita* e o percentual dos gastos públicos totais no PIB. O PIB *per capita* pode ser entendido como uma medida (ainda que imperfeita) do bem-estar da população. Como se pode notar, aparentemente há uma relação positiva. A correlação é igual a 0,33. Há exceções como Estados Unidos, Canadá, Austrália, Alemanha e Singapura, onde a participação do setor público é relativamente pequena e a renda *per capita* é alta. No entanto, Tanzi e Schuknecht (1997; 2000) investigam a relação entre a participação do setor público e o nível de bem-estar medido por indicadores multidimensionais, e não encontram evidências de que o gasto público aumente o bem-estar. Esse resultado sugere que os gastos públicos não sejam eficientes.



Fonte: versão atualizada de Afonso, Schuknecht e Tanzi (2006) com base nas informações de World Development Indicators (2009).

Obs.: Austrália-Aus; Áustria-Aut; Bélgica-Bel; Brasil-Bra; Bulgária-Bul; Canadá-Can; Chile-Chi; Chipre-Cyp; República Checa-Cze; Dinamarca-Den; Estônia-Est; França-Fra; Alemanha-Ger; Grécia-Gre; Hungria-Hun; Islândia-Ice; Irlanda-Ire; Itália-Ita; Coreia do Sul-Kor; Letônia-Lat; Lituânia-Lit; Malta-Mal; Ilhas Maurício-Mau; Holanda-Net; Noruega-Nor; Paraguai-Par; Peru-Per; Polônia-Pol; Portugal-Por; Singapura-Sin; Eslováquia-Slo; Eslovênia-Slo; África do Sul-Sou; Espanha-Spa; Tailândia-Tha; Turquia-Tur; Reino Unido-UK; Estados Unidos-US. As siglas estão em inglês.

Uma medida de desempenho multidimensional é proposta por Afonso, Schuknecht e Tanzi (2005).

Considerando um indicador composto (PSP) que depende de indicadores sociais e econômicos (I) de j áreas do governo, tem-se:

$$PSP_i = \sum_{j=1}^n PSP_{ij} \quad (6.8)$$

Em que i denota a unidade de análise (país, estado ou município), e $PSP_{ij} = f(I)$.

Uma melhoria no desempenho do setor público depende da melhoria nos k indicadores econômicos e sociais:

$$\Delta PSP_{ij} = \sum_{i=k}^n \frac{\partial f}{\partial I_k} \Delta I_k \quad (6.9)$$

De acordo com Afonso, Schuknecht e Tanzi (2005), os indicadores de desempenho podem ser de dois tipos: indicadores oportunidade, no sentido de que buscam medir o grau de igualdade de oportunidades, e indicadores musgravianos, que abrangem as tarefas do governo de alocação, distribuição e estabilização. O indicador agregado de desempenho do setor público é composto por sete subindicadores. Os quatro primeiros são indicadores de *oportunidades* e abrangem resultados administrativos, educacionais, de saúde e de infraestrutura.

Cada subindicador pode ser composto por outros elementos. O que mede os resultados administrativos, por exemplo, inclui indicadores de corrupção, burocracia, qualidade do sistema judiciário e de economia subterrânea. O indicador de saúde abrange taxa de mortalidade, esperança de vida, entre outros. O de educação proposto pelos autores inclui desempenho de estudantes em testes padronizados e taxa de matrícula no ensino médio, mas poderia ainda abranger a taxa de analfabetismo, a média de anos de escolaridade, entre outros. O indicador de infraestrutura, como definido pelos autores, reflete a qualidade das comunicações e da infraestrutura de transporte. Na visão de Afonso, Schuknecht e Tanzi (2005), uma população qualificada (educada), com acesso adequado a serviços de saúde e infraestrutura e com uma administração bem organizada,

que não dificulte a livre iniciativa e o acesso a diversos serviços, tem mais oportunidades e, portanto, o bem-estar social é maior.

Os três subindicadores restantes refletem as tarefas musgravianas do governo. Esses indicadores buscam refletir a interação governo/mercado. A tarefa distributiva pode ser mensurada pelo indicador de distribuição de renda e a tarefa de estabilização pode ser captada pelo coeficiente de variação do crescimento do produto e pela inflação média do período analisado. A tarefa alocativa está relacionada à eficiência alocativa e é medida pelo desempenho econômico em termos de crescimento do PIB e taxa de desemprego.

Esses subindicadores são utilizados para calcular um indicador multidimensional de desempenho do setor público. Os pesos conferidos a cada indicador são iguais, mas poderiam diferir de acordo com algum critério. Uma possibilidade é definir os pesos de acordo com o montante gasto. A esses indicadores de desempenho do setor público Afonso, Schuknecht e Tanzi (2005) dão o nome de *eficiência do Setor Público (PSE)*:

$$PSE_i = \frac{PSP_i}{PEX_i} \quad (6.10)$$

em que PEX_i são normalizados entre as i unidades, e

$$\frac{PSP_i}{PEX_i} = \sum_{j=1}^n \frac{PSP_{ij}}{PEX_{ij}} \quad (6.11)$$

Os gastos, nesses casos, podem ser entendidos como os insumos e variam para cada um dos indicadores de desempenho (PSP). Para os indicadores de oportunidade, os insumos podem ser: (i) o gasto público com consumo (que serve de *proxy* para os insumos do indicador de desempenho administrativo); (ii) despesa com saúde (*proxy* de insumos para o desempenho em saúde); (iii) despesas de educação (para desempenho em educação).

Para os indicadores musgravianos, os insumos podem ser representados por: (i) transferências ou subsídios (que atuam como *proxies* para os insumos do indicador de distribuição de renda); (ii) despesa total (*proxy* para o insumo do indicador de estabilidade econômica e para o indicador de eficiência econômica alocativa).

Finalmente, para o cálculo da eficiência do gasto público é possível também usar como referência funções de produção (sob a ótica do produto) ou funções custo (sob a ótica do insumo). Os princípios são os mesmos utilizados para o cálculo de eficiência da firma – o que difere são os insumos e os produtos, e os objetivos dos setores privado e público. Empiricamente, no entanto, não há diferenças importantes quanto ao método de estimar.

6.4 Considerações finais

Este capítulo apresentou os conceitos básicos de eficiência na produção de bens. Primeiramente, foram definidos os conceitos de eficiência econômica, no sentido de Pareto. Depois foram definidos os conceitos de produtividade e eficácia. Estes foram diferenciados do conceito de eficiência produtiva, a qual pode ser classificada em: (i) eficiência técnica; (ii) alocativa; (iii) total; (iv) de escala; e (v) social.

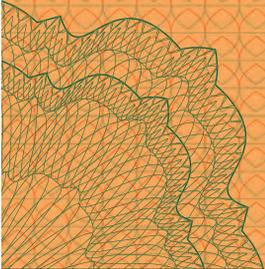
Em seguida, foi introduzido o conceito de eficiência sob a ótica do setor público. Foi mostrado, por exemplo, que a aplicação de impostos distorcivos gera custos adicionais. Como os governos, em geral, têm como objetivo maximizar o bem-estar da população, na decisão acerca do quanto produzir bens e serviços públicos, devem considerar os custos e benefícios associados.

Se a maximização do bem-estar é a meta final, os governos devem definir os objetivos corretos relacionados a um maior nível de satisfação social. Com isso, é possível obter indicadores mensuráveis do desempenho do setor público que representem uma medida de bem-estar. Por último, com esses indicadores é possível que o governo racionalize o uso dos recursos, calculando indicadores de eficiência do gasto do setor público.

No próximo capítulo serão apresentados os fundamentos microeconômicos do cálculo da eficiência na produção e que complementarão os conceitos aqui apresentados.

REFERÊNCIAS

- AFONSO, A.; SCHUKNECHT, L.; TANZI, V. Public sector efficiency: an international comparison. *Public Choice*, v. 123, n. 3-4, p. 321-347, 2005.
- _____. Public sector efficiency: evidence for new EU member states and emerging markets. *European Central Bank working paper series*, n. 581, 2006.
- ATKINSON, A.; STERN, N. Pigou, taxation and public goods. *Review of Economic Studies*, v. 41, p. 119-128, 1974.
- DAHLBY, B. *The marginal cost of public funds: theory and applications*. Cambridge: The MIT Press, 2008.
- FRIED, H.; LOVELL, C.; SCHIMIDT, S. *The measurement of productive efficiency and productivity growth*. New York: Oxford University Press, 2008.
- LEIBENSTEIN, H. Allocative efficiency vs. "X-efficiency". *American Economic Review*, v. 56, n. 3, p. 392-415, 1996.
- STIGLER, G. J. The Xistence of X-efficiency. *American Economic Review*, v. 66, n. 1, p. 213-216, maio 1976.
- TANZI, V.; SCHUKNECHT, L. Reconsidering the fiscal role of government: the international perspective. *American Economic Review*, v. 87, n. 2, p. 164-168, 1997.
- _____. *Public spending in the 20th century: a global perspective*. Cambridge: Cambridge University Press, 2000.



Capítulo 7

Fundamentos microeconômicos da mensuração de eficiência

Enlison Mattos
Rafael Terra

Este capítulo apresenta os fundamentos microeconômicos da mensuração da eficiência. Para isso, primeiramente, é necessário revisar os conceitos básicos da Teoria da Produção, fundamentais para demonstrar microeconomicamente a existência de ineficiência por parte das unidades produtoras. Portanto, na seção 7.1 são apresentados os elementos básicos da Teoria da Produção, como a função de produção, as produtividades marginais e médias, isoquantas e taxa marginal de substituição técnica e retornos à escala. Também são abordados os problemas duais de maximização de lucros e minimização de custos, sendo o primeiro analisado em um arcabouço orientado para o produto e o segundo, em um arcabouço voltado aos insumos.

Em seguida, na seção 7.2, são apontados os fundamentos microeconômicos da mensuração da eficiência, os quais utilizam os conceitos previamente descritos na seção 7.1. São apresentados, primeiramente, indicadores de eficiência estáticos, descrevendo os diferentes tipos de eficiência sob as óticas do insumo e do produto, e, finalmente, os indicadores de mudança na eficiência técnica ao longo do tempo, também sob as óticas do insumo e do produto.

7.1 Elementos básicos da teoria da produção

7.1.1 Produção

A atividade principal de uma firma é sabidamente transformar insumos em produtos. No âmbito da teoria econômica, essa função é formalizada em um modelo abstrato que mapeia a transformação

de insumos em produtos por meio de uma tecnologia de produção. Essa função de transformação tem a seguinte forma geral:

$$Q = F(X) \quad (7.1)$$

Em que $Q = \{q_1, \dots, q_p\}$ representa o vetor de p bens produzidos por uma firma, $X = \{x_1, \dots, x_m\}$ é o vetor que denota os m insumos utilizados, e $F(X)$ é a tecnologia de produção.

Para facilitar a apresentação dos conceitos, consideraremos um caso particular com apenas um bem a ser produzido por meio da transformação de diversos insumos. Tal relação de transformação é comumente denominada *função de produção*.

$$q = f(x_1, \dots, x_m) \quad (7.2)$$

Essa função mostra a quantidade máxima de produto que a firma pode obter com a utilização dos insumos $\{x_1, \dots, x_m\}$.

Produtividade marginal

A produtividade marginal de um insumo é dada pelo acréscimo no produto q resultante de uma variação de uma unidade de insumo, com a quantidade dos demais insumos mantida constante.

O produto marginal do insumo i (PMg_i) é tal que:

$$PMg_i = \frac{\partial q}{\partial x_i} = f_i \quad (7.3)$$

A produtividade marginal de um dado insumo pode ser crescente ou decrescente, mas é por hipótese sempre positiva ou nula se supusermos que pode haver livre descarte dos insumos. Matematicamente, a produtividade marginal decrescente é caracterizada pelas seguintes condições:

$$PMg_i = \frac{\partial q}{\partial x_i} = f_i > 0 \quad (7.4)$$

$$\frac{\partial PMg_i}{\partial x_i} = \frac{\partial^2 f(x_1, \dots, x_m)}{\partial x_i^2} = f_{ii} < 0 \quad (7.5)$$

A produtividade marginal será crescente se

$$\frac{\partial PMg_i}{\partial x_i} = \frac{\partial^2 f(x_1, \dots, x_m)}{\partial x_i^2} = f_{ii} > 0 \quad (7.6)$$

A produtividade marginal de um insumo pode mudar com a alteração na utilização de outros insumos. Por exemplo:

$$\frac{\partial f(x_1, \dots, x_m)}{\partial x_i \partial x_j} \leq 0 \quad (7.7)$$

Isto é, a produtividade marginal do insumo i pode mudar de acordo com o nível do insumo j empregado na produção. Normalmente, o termo $\partial f(x_1, \dots, x_m) / \partial x_i \partial x_j$ é positivo. Para ver isso, basta pensar em x_i como trabalho e em x_j como capital. Nesse caso, a produtividade marginal do trabalho deve ser maior quanto maior o estoque de capital empregado na produção.

Produtividade média

A produtividade média (PMe) é a medida mais comum e simples usada para mensurar eficiência. No entanto é uma medida imprecisa, pois não leva em consideração o formato da função de produção ou se os retornos à escala são constantes ou variáveis. A produtividade média é dada pela razão entre o produto e o total de insumos utilizados na produção:

$$PMe_i = \frac{q}{x_i} = \frac{f(x_1, \dots, x_m)}{x_i} \quad (7.8)$$

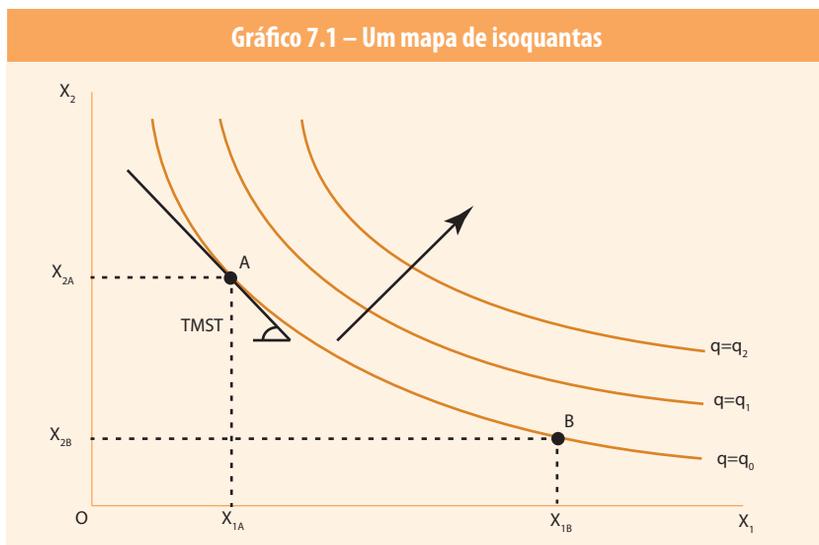
Note que a produtividade média é também função da quantidade de outros insumos usados na atividade de produção, pois a produtividade marginal de cada insumo pode variar de acordo com o nível de utilização dos demais insumos. Ademais, pode-se medir a produtividade média em relação ao total de insumos de todos os tipos. No entanto, para agregar os insumos, é necessário trazê-los antes a uma mesma base.

Isoquantas e a taxa marginal de substituição técnica

As isoquantas consistem em um modo conveniente de representar as combinações de insumo que geram determinado nível de produto. Para uma função de produção com insumos x_1 e x_2 , as isoquantas podem ser representadas graficamente, o que também é possível se houver mais insumos, contanto que estes sejam considerados fixos ou neutros em relação a x_1 e x_2 .

Matematicamente, uma isoquanta com insumos x_1 e x_2 pode ser representada por:

$$f(x_1, x_2) = q_0 \quad (7.9)$$



Fonte: Elaboração própria.

Como se pode notar pelo Gráfico 7.1, no ponto A é empregada uma combinação de insumos (x_{1A}, x_{2A}) que gera o nível de produto q_0 . No ponto B a combinação de insumos utilizada na produção é (x_{1B}, x_{2B}) e o nível de produto é o mesmo, igual a q_0 . O gráfico mostra ainda que a produção da firma é maior quanto mais a nordeste se encontra a isoquanta.

Taxa Marginal de Substituição Técnica

A inclinação das isoquantas no Gráfico 7.1 mostra a qual taxa um insumo pode ser trocado por outro, de modo a manter o nível de produto inalterado. A essa inclinação damos o nome de Taxa Marginal Substituição Técnica (TMST). Ao longo da isoquanta $q = q_0$:

$$TMST_{12} = - \left. \frac{dx_2}{dx_1} \right|_{q=q_0} \quad (7.10)$$

O valor dessa taxa de troca entre insumos depende da quantidade de insumos utilizada na produção de q . Além disso, a $TMST_{12}$ é dada pela razão entre as produtividades marginais dos insumos x_1 e x_2 . Note que a diferenciação total de (7.9) resulta em

$$dq = \frac{df}{dx_1} dx_1 + \frac{df}{dx_2} dx_2 = PMg_1 dx_1 + PMg_2 dx_2 \quad (7.11)$$

Sobre uma mesma isoquanta $dq = 0$, o ganho de produto auferido em virtude do aumento do insumo x_1 é anulado pela perda de produto em virtude da redução em x_2 . Portanto, (7.11) pode ser rearranjado de forma a obter a referida relação entre as produtividades marginais e a TMST:

$$- \left. \frac{dx_2}{dx_1} \right|_{q=q_0} = TMST_{12} = \frac{PMg_1}{PMg_2} \quad (7.12)$$

Retornos à escala

Na seção anterior, vimos o conceito de produtividade marginal dos insumos, o qual informa o acréscimo de produto induzido por uma variação em determinado insumo, com os outros insumos mantidos constantes. Outra pergunta que se deveria fazer é como o nível de produto seria afetado por um aumento simultâneo em todos os insumos e na mesma proporção. Essa pergunta nos remete ao conceito de retornos à escala.

Uma firma pode apresentar uma tecnologia de produção com retornos crescentes, decrescentes ou constantes à escala. Se um aumento de $k\%$ na utilização de todos os insumos aumenta o

produto q em mais do que $k\%$, dizemos que a função de produção apresenta retornos crescentes à escala. Se esse mesmo aumento produzir uma variação no produto menor do que $k\%$, dizemos que a tecnologia de produção apresenta retornos decrescentes à escala. Se o aumento for exatamente igual a $k\%$, a firma apresenta uma função de produção com retornos constantes à escala.

Formalmente, se a função de produção é dada por $q = f(x_1, \dots, x_m)$, o aumento proporcional na utilização do insumo igual a k produz a seguinte função de produção:

$$q = f(kx_1, \dots, kx_m) = k^\alpha f(x_1, \dots, x_m) \quad (7.13)$$

Se $\alpha = 1$, a função de produção apresenta retornos constantes à escala; se $\alpha < 1$ os retornos são decrescentes; e se $\alpha > 1$, os retornos são crescentes.

Teoricamente, é possível que uma função de produção apresente retornos crescentes, decrescentes e constantes, dependendo da escala em que esteja operando. Isto é, dependendo da quantidade de insumos que esteja sendo empregada para fins de produção.

Retornos constantes

A hipótese de retornos constantes à escala é considerada empiricamente plausível e, portanto, é amplamente utilizada em modelos microeconômicos. Tal hipótese faz sentido se pensarmos que uma firma produzindo q em uma planta pode produzir $2q$ se dobrar todos os insumos empregados na produção, ao abrir uma nova planta idêntica.

Como visto anteriormente, tecnologias com retornos constantes à escala têm coeficientes $\alpha = 1$ em (7.13), ou seja, tais funções de produção são homogêneas de grau 1. É comum a utilização de funções de produção homogêneas de grau 1 em livros-texto e modelos microeconômicos devido às suas propriedades características. As principais funções de produção homogêneas podem ser do tipo linear, de Coeficientes Fixos, Cobb-Douglas e CES.

Uma propriedade interessante das funções homogêneas de grau α é a de que suas derivadas são homogêneas de grau $\alpha - 1$, e,

portanto, funções homogêneas de grau 1 terão derivadas de grau 0. Logo, para uma tecnologia homogênea de grau 1, temos para um insumo i qualquer

$$PMg_i = \frac{\partial f(x_1, \dots, x_m)}{\partial x_i} = \frac{\partial f(kx_1, \dots, kx_m)}{\partial x_i} \quad (7.14)$$

Voltando ao exemplo com dois insumos x_1 e x_2 , se $k = 1/x_1$, podemos reescrever os produtos marginais da seguinte forma:

$$PMg_1 = \frac{\partial f\left(1, \frac{x_2}{x_1}\right)}{\partial x_1} \text{ e } PMg_2 = \frac{\partial f\left(1, \frac{x_2}{x_1}\right)}{\partial x_2} \quad (7.15)$$

Nesse caso, o produto marginal de qualquer insumo depende somente da razão entre x_2/x_1 e não dos valores absolutos desses insumos. Isso quer dizer que a função de produção é homotética, e suas isoquantas são expansões radiais umas das outras.

A elasticidade de substituição é dada pela seguinte expressão:

$$\begin{aligned} \sigma &= \frac{\Delta(x_2/x_1)}{(x_2/x_1)} \bigg/ \frac{\Delta(TMST)}{TMST} = \frac{\partial(x_2/x_1)}{\partial(TMST)} \frac{TMST}{(x_2/x_1)} \\ &= \frac{\partial \ln(x_2/x_1)}{\partial \ln(TMST)} \end{aligned} \quad (7.16)$$

Note que σ tem valor positivo, pois a TMST e a razão (x_2/x_1) se movem na mesma direção. Se σ tem um valor alto, a TMST não é muito sensível a variações em x_2/x_1 e a isoquanta será mais achatada. Já no caso de σ apresentar um valor baixo, a isoquanta será bastante inclinada e a TMST será sensível a mudanças em (x_2/x_1) . Uma propriedade interessante da classe de funções homotéticas é a de que a elasticidade-substituição dos fatores é constante. No caso de uma função Cobb-Douglas, $\sigma = 1$, no caso linear, $\sigma = \infty$, e $\sigma = 0$ no caso de uma função Leontief.

7.1.2 O problema da firma

Uma vez apresentados os elementos básicos da teoria da produção, cabe perguntar qual é o objetivo principal da firma. Naturalmente,

a firma procurará auferir o maior ganho possível com sua atividade produtiva. Esse objetivo de maximização dos ganhos pode ser representado por um problema de maximização dos lucros, i.e., a firma procurará produzir um montante que maximize a diferença entre receita e custos, ou de minimização de custos, em que cabe a ela escolher o montante de insumos que reduza os custos ante um nível de produção desejado. Tais problemas são duais, no sentido de que a abordagem é distinta, mas a escolha ótima deve ser a mesma.

Minimização de custos

Os custos para uma empresa, do ponto de vista econômico, referem-se à melhor remuneração que os insumos deveriam receber no mercado, o que está relacionado à questão de custos de oportunidade. Outra forma de se pensar em custos econômicos é considerar o preço do aluguel desses insumos para utilização no processo produtivo.

O custo de produção C é dado pela função:

$$C = w_1 x_1 + \dots + w_m x_m \quad (7.17)$$

em que $\{x_1, \dots, x_m\}$ são os insumos utilizados na produção, e $\{w_1, \dots, w_m\}$ são os preços dos insumos.

O problema da firma é o de minimização restrita dos custos. A firma escolherá a quantidade de insumos a ser utilizada para produzir q_0 unidades do bem, de forma que os custos sejam minimizados. Matematicamente, o problema pode ser representado por:

$$\begin{aligned} \min_{x_1, \dots, x_m} C(x_1, \dots, x_m, w_1, \dots, w_m) \\ \text{s.t. } f(x_1, \dots, x_m) = q_0 \end{aligned} \quad (7.18)$$

O lagrangeano será, portanto,

$$\mathcal{L} = w_1 x_1 + \dots + w_m x_m + \lambda [q_0 - f(x_1, \dots, x_m)] \quad (7.19)$$

As condições de primeira ordem para um insumo x_i serão

$$\frac{\partial \mathcal{L}}{\partial x_i} = w_i + \lambda \frac{\partial f(x_1, \dots, x_m)}{\partial x_i} = 0 \quad (7.20)$$

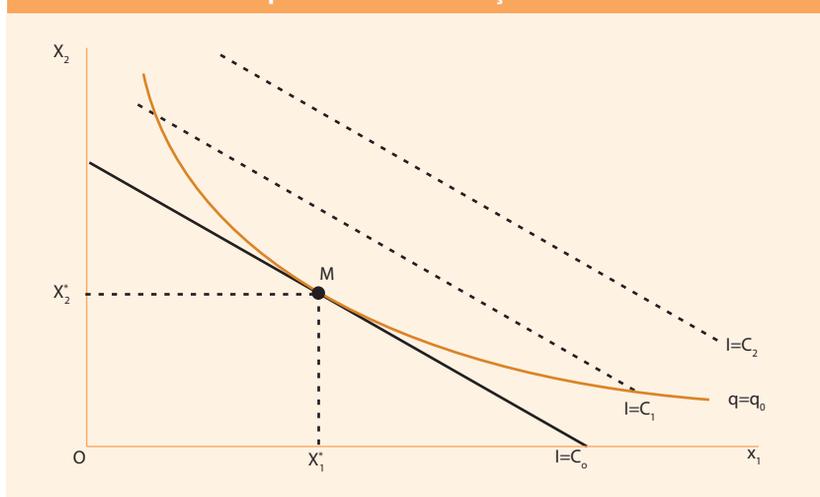
$$\frac{\partial \mathcal{L}}{\partial \lambda} = q_0 - f(x_1, \dots, x_m) = 0 \quad (7.21)$$

Tomando dois insumos x_i e x_j quaisquer, é possível ver que teremos, no ótimo:

$$\frac{w_i}{w_j} = \frac{\partial f(x_1, \dots, x_m) / \partial x_i}{\partial f(x_1, \dots, x_m) / \partial x_j} = TMST_{ij} \quad (7.22)$$

Isto é, a firma minimizadora de custos deve escolher insumos tais que a TMST entre dois insumos quaisquer seja igual à razão entre seus preços. Recorrendo novamente ao caso em que a firma usa somente os insumos x_1 e x_2 , é possível ilustrar esse problema graficamente. No Gráfico 7.2 podem-se verificar retas paralelas de isocusto cuja equação é definida como $x_2 = (C/w_2) - (w_1/w_2)x_1$, em que C designa o nível do custo total. No gráfico, nota-se que para uma dada quantidade de produto q_0 desejada pela firma, a escolha de insumos que irá minimizar os custos será (x_1^*, x_2^*) , que permite produzir a um custo total igual a C_0 , o mais baixo dentre as alternativas factíveis.

Gráfico 7.2 – O problema de minimização de custos da firma



Fonte: Elaboração própria.

A função custo total é expressa como uma função da quantidade total de produto produzida q . Para obtê-la em função de q basta determinar a escolha ótima dos insumos e substituí-la na função de produção. Com isso, pode-se obter a utilização ótima de insumos em termos de unidades de produto q . Em seguida, substituindo as expressões dos insumos em termos de q na função custo, obtemos a função custo total $C = C(W, q)$, na qual $W = \{w_1, w_2, \dots, w_n\}$.

No caso da função de produção com retornos constantes à escala, a função custo total é proporcional ao produto. Dessa maneira, o custo de se produzir k unidades de produto é igual ao custo de se produzir uma unidade de produto vezes k .

$$C(W, q) = qC(W, 1) \quad (7.23)$$

A ocorrência de um progresso técnico no período t permite que os insumos necessários para produzir uma unidade de produto no período $t-1$ viabilizem a produção de uma quantidade $A(t)$ de unidades de produto, em que $A(t) > 1$.

$$\begin{aligned} C_i(W, A(t)) &= A(t)C_i(W, 1) \\ &= C_{i-1}(W, 1) \end{aligned} \quad (7.24)$$

A função custo total pode ser reescrita como:

$$C_i(W, q) = qC_i(W, 1) = \frac{C_0(W, q)}{A(t)} \quad (7.25)$$

Dessa forma, os custos caem ao longo do tempo a uma taxa igual à taxa de mudança tecnológica.

Maximização de lucros

Como mencionado anteriormente, existe uma dualidade entre os problemas de maximização de lucros e minimização de custos. A escolha ótima do nível de produção será a mesma para os dois problemas. No de maximização de lucros, a firma irá escolher q tal que o lucro seja máximo. O lucro é definido como:

$$\pi = R(q) - C(q) \quad (7.26)$$

em que $R(y)$ é a função receita, tal que $R(q) = p(q)q$, e $C(q)$ é a função custo total. O problema da firma será maximizar o lucro:

$$\max_q \pi = R(q) - C(q)$$

$$\frac{\partial \pi}{\partial q} = R'(q) - C'(q) = 0 \quad (7.27)$$

Portanto, o nível de produto ótimo deve ser aquele que iguala a receita marginal ao custo marginal. A condição suficiente para a existência de um ponto máximo da função lucro é a de que:

$$\left. \frac{\partial \pi(q)}{\partial q} \right|_{q=q^*} = \left. \frac{\partial \pi'(q)}{\partial q} \right|_{q=q^*} < 0 \quad (7.28)$$

Ou seja, o lucro marginal deve ser crescente para $q < q^*$ e decrescente para $q > q^*$.

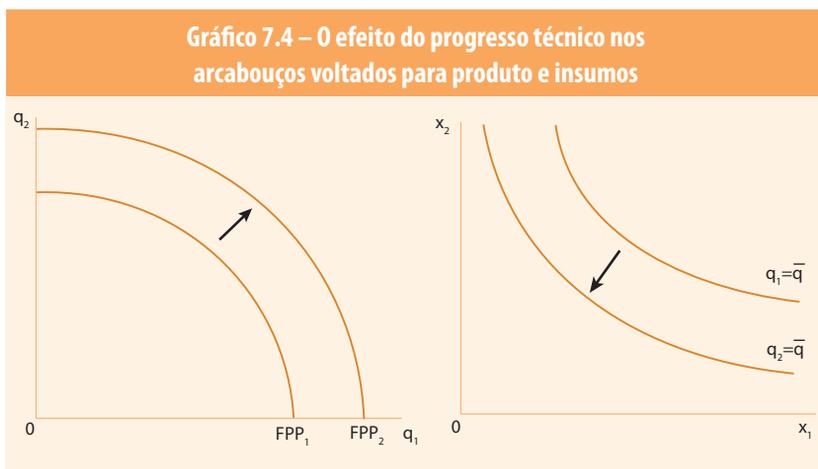
Fronteira de Possibilidades de Produção e maximização de lucros

Normalmente as firmas produzem mais de um bem e enfrentam *trade-offs* na produção desses bens. Pode-se pensar também em uma economia que produz dois bens e que deve decidir quanto produzir de cada um deles. O diagrama da Fronteira de Possibilidades de Produção (FPP), no Gráfico 7.3, mostra as combinações factíveis de produção dos dois bens com um montante fixo de insumos empregado no processo produtivo.

por $q_2 = (\bar{R} / p_2) - (p_1 / p_2)q_1$. As retas de isoreceita podem ser observadas no Gráfico 7.3. É possível produzir eficientemente, i.e., sobre a FPP, sem a receita ser máxima, como no ponto B do Gráfico 7.3, em que a reta de isoreceita cruza a fronteira, mas a receita é igual a R_1 . No entanto, a receita é maximizada no ponto A, em que (q_1^o, q_2^o) e a receita é igual a \bar{R} . Nesse ponto, a TMT é igual aos preços relativos dos produtos $(-p_1 / p_2)$.

Progresso técnico

A posição da FPP em um dado período é estabelecida pelo estado da tecnologia de produção e a quantidade de fatores (insumos) empregados na produção. Se houver uma melhoria tecnológica que aumente a produtividade dos fatores, com os insumos mantidos constantes, a fronteira deve se deslocar para a direita. Já em um arcabouço orientado para o insumo, com o nível de produção mantido fixo, as isoquantas devem se deslocar para a esquerda (para dentro). O Gráfico 7.4 mostra os efeitos do progresso técnico para os dois arcabouços (para produtos e insumos).



Fonte: Elaboração própria.

7.2 Definições e medidas de eficiência econômica

Esta seção apresenta as definições e medidas de eficiência econômica.¹ Para isso, consideram-se primeiramente cenários estáticos, isto é, em que não há progressos técnicos que desloquem as fronteiras de possibilidades de produção ou as isoquantas associadas à função de produção.

Em seguida, são apresentadas as medidas de eficiência de Malmquist (1953) que medem o crescimento na produtividade ao longo do tempo. Essas medidas podem ter outros usos: comparar as produtividades entre países com diferentes tecnologias de produção ou quaisquer outros casos que exigem a comparação de diferentes funções de produção.

7.2.1 Medidas de eficiência econômica estática

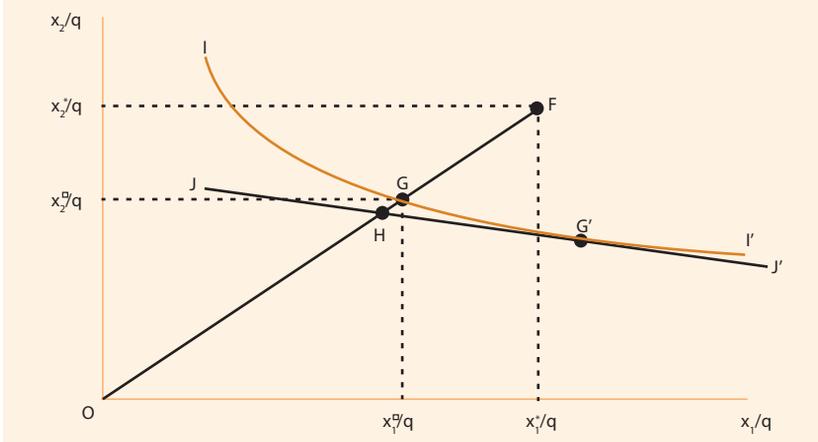
As medidas de eficiência econômica estática podem ser avaliadas sob a ótica dos insumos ou dos produtos. A rigor não há uma ótica melhor que a outra. De fato, as duas produzirão resultados iguais, dependendo da hipótese que se faça sobre os retornos à escala da função de produção.

Sob a ótica dos insumos

Primeiramente, analisemos as medidas de eficiência econômica sob a ótica dos insumos. Como ponto de partida, pode-se considerar uma firma com uma função de produção com retornos constantes à escala, usando insumos para produzir q no ponto F apresentado no Gráfico 7.5. Para produzir essa quantidade, qualquer combinação de x_1 e x_2 sobre a isoquanta II' do gráfico é suficiente. A firma em questão está produzindo q no ponto F, usando insumos x_1^* e x_2^* . Mas é possível para essa firma reduzir o emprego de cada um dos insumos proporcionalmente (mantendo a mesma razão de insumos x_1^*/x_2^*) até o ponto G, em que os insumos são x_1^o e x_2^o , e ainda assim manter a produção em q .

1 Descrições dos fundamentos teóricos e das técnicas econométricas de estimação de eficiência técnica podem ser encontrados também em Johnes (2004).

Gráfico 7.5 – Eficiência técnica sob a ótica do insumo



Fonte: Elaboração própria.

Com esse cenário, é possível definir as três medidas de eficiência de Farrell (1957) sob a ótica dos insumos. A primeira medida se refere à *eficiência técnica*, e é dada pela seguinte razão dos segmentos de reta do Gráfico 7.5:

$$ET_i = \frac{\overline{OG}}{\overline{OF}} \quad (7.29)$$

À medida que ET_i tende a 1, a firma tende à eficiência técnica (ou eficiência técnica máxima); e quanto mais próximo de 0, mais ineficiência técnica a firma apresenta. A ineficiência técnica (IT_i), portanto, pode ser expressa por:

$$IT_i = 1 - \frac{\overline{OG}}{\overline{OF}} \quad (7.30)$$

Essa é a medida da proporção à qual os insumos poderiam ser reduzidos mantendo a mesma razão entre si, sem reduzir a produção q .

A segunda medida de eficiência sob a ótica do insumo destacada por Farrell (1953) é a *eficiência alocativa*. Esta se refere à eficiência associada ao custo mínimo de produção de q . Mesmo que a firma reduza seus insumos proporcionalmente e passe a produzir q no ponto G, no qual há eficiência técnica, este ponto não é

necessariamente aquele em que o custo de produção é minimizado. Se realmente esses pontos não coincidirem, haverá espaço para uma mudança na combinação (ou razão) de insumos que manterá a produção eficiente de q do ponto de vista técnico e ainda reduzirá os custos de produção, aumentando o bem-estar econômico.

No Gráfico 7.5 pode-se verificar que a razão de insumos utilizada na produção em G leva a um custo de produção maior do que aquele representado pela reta de isocusto JJ' . No ponto G' verifica-se tanto a eficiência técnica quanto a eficiência alocativa (em termos de minimização de custos). Nesse ponto, a reta isocusto JJ' tem inclinação igual à dos preços relativos dos insumos (w_1 / w_2) a preços competitivos (a firma não é monopsonista) e, portanto, temos a condição de ótimo $TMST = w_1 / w_2$.

Os custos de produção em G' são dados pelo segmento \overline{OH} , isto é, uma fração $\overline{OH} / \overline{OG}$ dos custos de produzir em G. A medida de eficiência alocativa será dada por:

$$EA_i = \frac{\overline{OH}}{\overline{OG}} \quad (7.31)$$

A medida de ineficiência alocativa será dada por:

$$IA_i = 1 - \frac{\overline{OH}}{\overline{OG}} \quad (7.32)$$

Finalmente, Farrel propõe uma medida de *eficiência geral* sob a ótica do insumo, medida por:

$$EG_i = \frac{\overline{OH}}{\overline{OF}} \quad (7.33)$$

A medida de ineficiência geral é dada por:

$$IG_i = 1 - \frac{\overline{OH}}{\overline{OF}} \quad (7.34)$$

O segmento \overline{HF} é uma medida de redução de custos que poderia ser atingida se a firma mudasse seu ponto de produção de F para o ponto de minimização de custos G' .

Sob a ótica do produto

As medidas de eficiência sob a ótica do produto são similares às aquelas definidas sob a ótica dos insumos. A princípio o cenário assumido é o mesmo, de uma firma com retornos constantes à escala, mas para facilitar a análise, consideramos que um insumo x é utilizado em uma quantidade fixa para produzir dois produtos (q_1, q_2) . Nesse caso, podemos analisar os diferentes tipos de eficiência em um cenário voltado para o produto com base em uma fronteira de possibilidades de produção.

No Gráfico 7.6 pode-se verificar a fronteira de possibilidades produção dos bens q_1 e q_2 . A firma está produzindo no ponto F um nível de produto $Q'' = (q_1'', q_2'')$, com insumo $x = x^*$ fixo. Mas com o mesmo nível de insumo, a firma poderia produzir $Q' = (q_1', q_2')$ no ponto K, no qual haveria eficiência técnica.

Portanto, a medida de eficiência técnica sob a ótica do produto é dada por:

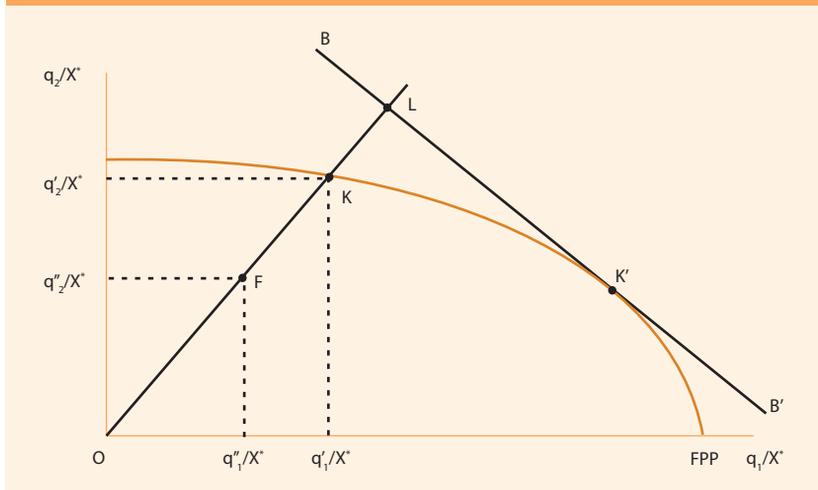
$$ET_o = \frac{\overline{OF}}{\overline{OK}} \quad (7.35)$$

Já a medida de ineficiência sob a ótica do produto é dada por:

$$IT_o = 1 - \frac{\overline{OF}}{\overline{OK}} \quad (7.36)$$

Essa medida informa a qual proporção se poderia aumentar o produto (mantendo a mesma razão – mix – de produtos) sem aumentar o nível de insumos utilizados para produzir $Q'' = (q_1'', q_2'')$ no ponto F.

Gráfico 7.6 – Eficiência técnica sob a ótica do produto



Fonte: Elaboração própria.

A reta de isorreceita BB' está também representada no Gráfico 7.6. Sua inclinação é dada pela razão entre os preços dos dois produtos p_1 / p_2 e a receita obtida com os produtos ao longo dessa reta é igual a \bar{R} . A inserção da reta de isorreceita permite averiguar a eficiência alocativa sob a ótica do produto. Note-se que essa reta tangencia a FPP no ponto K' , indicando que não há nenhuma receita superior a \bar{R} factível de ser obtida com a tecnologia disponível. Como a inclinação da curva isorreceita é igual à Taxa Marginal de Transformação no ponto ótimo K' , então vale a condição $TMT = p_1 / p_2$. Nesse ponto, tanto a eficiência técnica quanto a alocativa são atingidas pela firma.

Mesmo que a firma opere em condições de eficiência técnica no ponto K , a receita não será máxima,² e ainda haverá espaço para que ela altere seu *mix* de produtos sobre a FPP e aumente o nível de receitas. No ponto L temos a mesma receita \bar{R} , igual à do ponto K' , e a razão de produtos igual àquela em K e F . Com isso pode-se comparar o nível de receitas que se está abrindo mão de receber ao produzir em K ao invés de em K' .

2 Note-se que a receita cresce na direção nordeste do Gráfico 7.6.

A medida de eficiência alocativa é dada por:

$$EA_o = \frac{\overline{OK}}{\overline{OL}} \quad (7.37)$$

Portanto, os pontos K e K' são tecnicamente eficientes, mas o ganho de receita que a firma poderia obter produzindo a proporção correta de produtos (maximizadora de receitas) é igual à distância \overline{KL} .

A eficiência total sob a ótica do produto combina as duas medidas de eficiência e é dada por:

$$EG_o = \frac{\overline{OF}}{\overline{OL}} \quad (7.38)$$

A medida de ineficiência geral será:

$$IG_o = 1 - \frac{\overline{OF}}{\overline{OL}} \quad (7.39)$$

A distância \overline{FL} é o montante que a firma poderia obter a mais de receita ao se mover do ponto F para o ponto de maximização de receita K'.

Finalmente, é importante destacar que, sob retornos constantes à escala, as medidas de eficiência são idênticas, sob as óticas do insumo e do produto, mas essas medidas podem diferir se os retornos à escala forem variáveis.

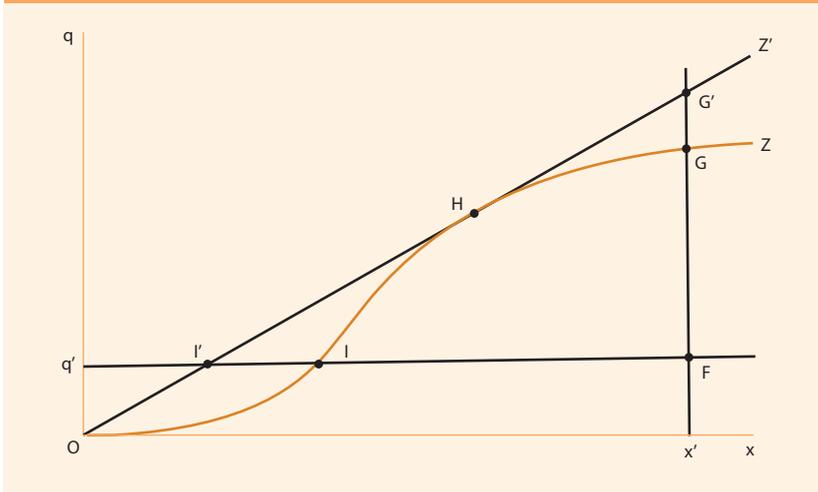
Introduzindo retornos variáveis à escala

Até aqui se considerou que a função de produção tinha retornos constantes à escala. Mas, muitas vezes, a firma pode estar sendo ineficiente por ter o *tamanho errado*. A ineficiência originada do *tamanho errado* compõe a *ineficiência técnica*.

Para ver isso, basta construir um diagrama simples de uma função de produção com retornos variáveis à escala (crescentes em determinadas escalas e decrescentes em outras), em que a firma usa apenas um insumo x para produzir um produto q . No Gráfico 7.7 pode-se verificar esse cenário, em que a fronteira de possibilidades de produção é dada por \overline{OZ} . A firma opera inicialmente no ponto F, ou

seja, produz q' usando x' insumos. No mesmo gráfico, consideramos para efeito de comparação uma função de produção com retornos constantes $\overline{OZ'}$ que seria a função de produção ideal associada ao equilíbrio competitivo de longo prazo.

Gráfico 7.7 – Eficiência técnica e eficiência de escala em um arcabouço com retornos variáveis à escala



Fonte: Elaboração própria.

A medida de eficiência sob a ótica do insumo, i.e., mantendo o produto fixo e variando o insumo, é dada pela distância horizontal do ponto de produção até a fronteira. Tomando como referência primeiramente a função de produção com retornos constantes $\overline{OZ'}$, a medida de eficiência técnica é dada por:

$$ET_i^{RC} = \frac{q' I'}{q' F} \quad (7.40)$$

Sob a ótica do produto, mantendo como referência a tecnologia com retornos constantes, a eficiência técnica é medida pela distância vertical até a fronteira. Portanto, temos

$$ET_o^{RC} = \frac{x' F}{x' G'} \quad (7.41)$$

Se tomarmos como referência a tecnologia com retornos variáveis, sob a ótica do insumo, a eficiência técnica será medida pela distância horizontal até a fronteira:

$$ET_i^{RV} = \frac{\overline{q' I}}{\overline{q' F}} \quad (7.42)$$

Sob a ótica do produto, a eficiência técnica da firma com função de produção com retornos variáveis à escala é dada pela distância vertical até a fronteira:

$$ET_o^{RV} = \frac{\overline{x' F}}{\overline{x' G}} \quad (7.43)$$

Se considerarmos uma função de produção com retornos variáveis à escala, as eficiências técnicas sob as óticas do insumo e do produto podem diferir, enquanto essas duas medidas serão iguais para uma firma com tecnologia com retornos constantes. A eficiência técnica será maior se a fronteira de produção tiver retornos variáveis comparativamente a retornos constantes, independentemente da ótica que se utilizar – insumo versus produto. A única exceção se dá no ponto H, onde as eficiências técnicas são idênticas para as duas tecnologias e a produtividade média é máxima.

Portanto, caso a firma apresente retornos variáveis à escala, a utilização da fronteira com retornos constantes como referência tecnológica irá atribuir à ineficiência técnica o que de fato é uma ineficiência de escala, i.e., causada por uma *escala* desfavorável. Os segmentos $\overline{II'}$ e $\overline{GG'}$ no Gráfico 7.7 refletem as diferenças de utilização de insumos e produção de produtos resultantes de uma *escala* desfavorável.

A eficiência técnica baseada em uma tecnologia com retornos constantes pode ser decomposta em dois termos: um referente à eficiência técnica baseada em uma tecnologia com retornos variáveis e outro referente à eficiência de escala. Sob a ótica do insumo temos, então:

$$ET_i^{RC} = \frac{\overline{q' I'}}{\overline{q' F}} = \frac{\overline{q' I}}{\overline{q' F}} \frac{\overline{q' I'}}{\overline{q' I}} = ET_i^{RV} \times \frac{\overline{q' I'}}{\overline{q' I}} \quad (7.44)$$

em que o primeiro termo se refere à eficiência técnica sob a ótica do insumo e retornos variáveis. O segundo termo pode ser considerado uma medida de *eficiência de escala*, pois mede a ineficiência devido à divergência entre a escala real de operação em I e a escala mais produtiva. Logo, (7.44) pode ser reescrito como:

$$ET_i^{RC} = ET_i^{RV} \times ET_i^{Escala} \quad (7.45)$$

Sob a ótica do produto, o cálculo é similar. Para decompor a medida de eficiência técnica, basta notar que:

$$ET_o^{RC} = \frac{\overline{x'F}}{\overline{x'G'}} = \frac{\overline{x'F}}{\overline{x'G}} \frac{\overline{x'G}}{\overline{x'G'}} \quad (7.46)$$

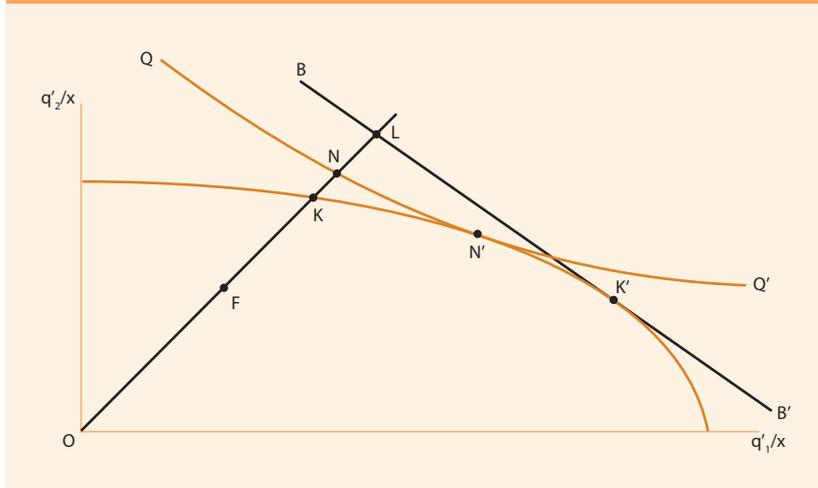
Que pode ser representado por:

$$ET_o^{RC} = ET_o^{RV} \times ET_o^{Escala} \quad (7.47)$$

Eficiência social

Outra medida de eficiência menos conhecida e utilizada pela literatura empírica se refere à *eficiência social*. Para obter uma medida desse tipo de eficiência, consideramos um arcabouço orientado para produto em que dois produtos q_1 e q_2 são produzidos com apenas um insumo fixo x , tal como no Gráfico 7.6. No Gráfico 7.8, adicionamos preferências sociais representadas por QQ' , o que determina as diferentes combinações de produto que geram o mesmo nível de utilidade.

Gráfico 7.8 – Eficiência social sob a ótica do produto



Fonte: Elaboração própria.

Se a firma opera inicialmente no ponto F, ela poderia aumentar proporcionalmente sua produção até o ponto K sobre a fronteira. Nesse caso, a firma seria eficiente do ponto de vista técnico, mas outros pontos tecnicamente eficientes poderiam ser atingidos e o bem-estar seria maior. Como se pode verificar pelo gráfico, a curva de *iso-bem-estar* mais alta que se pode atingir dadas as possibilidades de produção se dá no ponto N', que é, portanto, o ótimo do ponto de vista social. O nível de utilidade em N e N' é o mesmo do ponto de vista da sociedade, i.e., $W(N) = W(N')$. Desse modo, uma medida de eficiência do ponto de vista social é dada por:

$$ES_o = \frac{\overline{OF}}{\overline{ON}} \quad (7.48)$$

Note-se que é possível ser eficiente do ponto de vista social sem ser eficiente do ponto de vista alocativo, como no ponto N', por exemplo. Da mesma forma, é possível ser eficiente em termos alocativos sem o ser do ponto de vista social, como no ponto K'.

A ineficiência social será dada então por:

$$IS_o = 1 - \frac{\overline{OF}}{\overline{ON}} \quad (7.49)$$

7.2.2 Introduzindo melhorias tecnológicas – as medidas de eficiência de Malmquist

Até agora nenhuma consideração foi feita em relação ao papel da tecnologia de produção no desempenho das firmas. No entanto, é possível que entre dois períodos ocorra um progresso tecnológico que permita à firma produzir o mesmo montante com menos insumos ou, com os mesmo insumos, produzir mais produtos. Esta seção apresenta os indicadores de Malmquist sob as óticas do produto e dos insumos, os quais medem o crescimento na produtividade resultante de um aumento na eficiência ou de um progresso tecnológico.

Sob a ótica do produto

Para obter uma medida de eficiência de Malmquist sob a ótica do produto, consideramos novamente um arcabouço orientado para produto com retornos constantes, em que a firma produz dois produtos q_1 e q_2 usando um único insumo x ao longo de dois períodos s (base) e t (final). Esse arcabouço está representado no Gráfico 7.9, em que se podem observar as fronteiras de possibilidades de produção nos dois períodos.

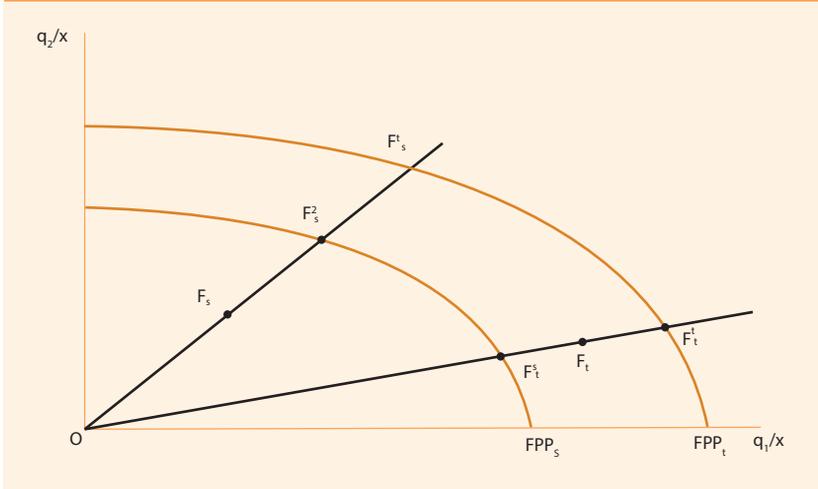
A firma produz inicialmente (período s) no ponto F_s , claramente abaixo da fronteira de eficiência FPP_s . A eficiência técnica sob a ótica do produto no período s segundo a metodologia vista até aqui é:

$$ET_s = \frac{\overline{OF_s}}{\overline{OF_s^s}} \quad (7.50)$$

Já no período t , a firma produz no ponto F_t , também interior à fronteira FPP_t . Portanto, a firma está operando em t de forma ineficiente, e a eficiência técnica pode ser representada por:

$$ET_t = \frac{\overline{OF_t}}{\overline{OF_t^t}} \quad (7.51)$$

Gráfico 7.9 – Mudança na eficiência sob a ótica do produto ao longo do tempo



Fonte: Elaboração própria.

Seja $D_o^s(x, q_{1s}, q_{2s})$ o inverso do montante máximo que o produto pode ser aumentado no período s usando um insumo x constante. O termo $D_o^s(x, q_{1s}, q_{2s})$ representa a eficiência técnica ET_s da produção no período s e $D_o^t(x, q_{1t}, q_{2t})$ a eficiência técnica ET_t no período t .

Para examinar o modo como a produtividade mudou ao longo do tempo basta usar uma das tecnologias nos períodos s ou t como referência. Usando a tecnologia em s como referência e considerando a produção no período t no ponto F_t , o máximo que se conseguiria produzir em t usando a tecnologia disponível em s seria F_t^s .

Usando a tecnologia do período s como referência, a eficiência técnica no ponto F_t é denotada por:

$$D_o^s(x, q_{1t}, q_{2t}) = \frac{\overline{OF}_t}{\overline{OF}_t^s} \quad (7.52)$$

O índice de crescimento de produtividade, denominado *Índice de produtividade Malmquist sob a ótica do produto* com referencial tecnológico em s , pode ser representado por:

$$M_o^s(x, q_{1t}, q_{2t}) = \frac{D_o^s(x, q_{1t}, q_{2t})}{D_o^s(x, q_{1s}, q_{2s})} = \frac{\overline{OF}_t / \overline{OF}_t^s}{\overline{OF}_s / \overline{OF}_s^s} \quad (7.53)$$

Se, no entanto, usarmos a tecnologia do período t como referência, temos o seguinte índice de Malmquist sob a ótica do produto:

$$M_o^t(x, q_{1t}, q_{2t}) = \frac{D_o^t(x, q_{1t}, q_{2t})}{D_o^t(x, q_{1s}, q_{2s})} = \frac{\overline{OF}_t / \overline{OF}_t^t}{\overline{OF}_s / \overline{OF}_s^t} \quad (7.54)$$

Não há um melhor referencial tecnológico a ser utilizado, em t ou s . Ao invés, recomenda-se a utilização de uma média ponderada dos dois indicadores.

$$\begin{aligned} M_o &= [M_o^s M_o^t]^{1/2} = \left[\frac{D_o^s(x, q_{1t}, q_{2t})}{D_o^s(x, q_{1s}, q_{2s})} \frac{D_o^t(x, q_{1t}, q_{2t})}{D_o^t(x, q_{1s}, q_{2s})} \right]^{1/2} \\ &= \left[\frac{\overline{OF}_t / \overline{OF}_t^s}{\overline{OF}_s / \overline{OF}_s^s} \frac{\overline{OF}_t / \overline{OF}_t^t}{\overline{OF}_s / \overline{OF}_s^t} \right]^{1/2} \end{aligned} \quad (7.55)$$

Multiplicando (7.55) por

$$[D_o^s(x, q_{1s}, q_{2s}) / D_o^t(x, q_{1t}, q_{2t})] \cdot [D_o^t(x, q_{1t}, q_{2t}) / D_o^s(x, q_{1s}, q_{2s})]$$

e rearranjando resulta em:

$$\begin{aligned} M_o &= \frac{D_o^t(x, q_{1t}, q_{2t})}{D_o^s(x, q_{1s}, q_{2s})} \left[\frac{D_o^s(x, q_{1t}, q_{2t})}{D_o^t(x, q_{1t}, q_{2t})} \frac{D_o^s(x, q_{1s}, q_{2s})}{D_o^t(x, q_{1s}, q_{2s})} \right]^{1/2} \\ &= \frac{\overline{OF}_t / \overline{OF}_t^t}{\overline{OF}_s / \overline{OF}_s^s} \left[\frac{\overline{OF}_t / \overline{OF}_t^s}{\overline{OF}_t / \overline{OF}_t^t} \frac{\overline{OF}_s / \overline{OF}_s^s}{\overline{OF}_s / \overline{OF}_s^t} \right]^{1/2} \end{aligned} \quad (7.56)$$

O primeiro termo fora dos colchetes consiste na razão entre as eficiências técnicas em t e s , portanto, mostra a mudança na eficiência técnica entre os dois períodos. Se esse termo é igual a 1, isso significa que não houve aumento de eficiência. Se for maior (menor) que 1, implica uma melhora (piora) na eficiência técnica entre os dois períodos.

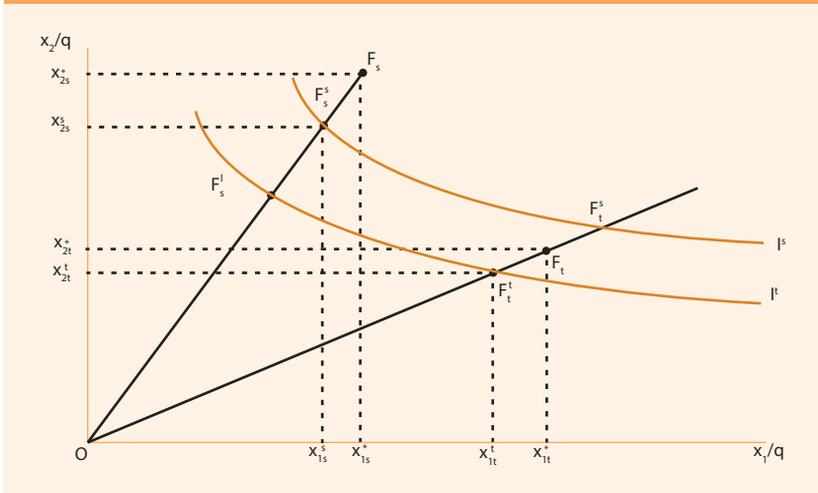
O segundo termo (entre colchetes) mede a mudança na tecnologia de produção entre s e t . Consiste na média geométrica da mudança tecnológica entre s e t avaliadas nos níveis de insumo x . Esse termo será igual a 1 quando não houver mudança tecnológica, e será maior (menor) que 1 se a mudança tecnológica tiver um efeito positivo (negativo).

Sob a ótica do insumo

Para verificar a mudança na eficiência ao longo do tempo em um arcabouço orientado para o insumo, consideramos uma firma produzindo q no ponto F_s usando insumos (x_{1s}^*, x_{2s}^*) . A firma poderia produzir q no período s usando menos insumos (x_{1s}^s, x_{2s}^s) . Portanto, a firma está sendo tecnicamente ineficiente no período s .

No período t há um progresso tecnológico. Note-se no Gráfico 7.10 que a isoquanta I^t está abaixo de I^s , indicando que em t é possível produzir q com uma quantidade de insumos inferior. Mesmo assim, no período t , a firma opera de forma tecnicamente ineficiente no ponto F_t , ainda que esteja em melhor situação do que a verificada no período s .

Gráfico 7.10 – Mudança na eficiência sob a ótica do insumo ao longo do tempo



Fonte: Elaboração própria.

No período s a *função distância* de insumo pode ser representada por:

$$D_i^s(x_{1s}, x_{2s}, q) = \frac{\overline{OF_s}}{OF_s^s} \quad (7.57)$$

Essa medida é a recíproca da medida de eficiência de Farrel para eficiência técnica sob a ótica do insumo.

Pode-se também, nesse caso, tomar como referência as tecnologias em s ou t . Tomando como referência a tecnologia em s e avaliando o ponto de produção F_t , o insumo observado em t relativo ao mínimo requerido pela tecnologia em s será:

$$D_i^s(x_{1t}, x_{2t}, q) = \frac{\overline{OF_t}}{OF_s^s} \quad (7.58)$$

Comparando com a medida de eficiência da produção em $F_{s,t}$ temos o índice de Malmquist sob a ótica do insumo com referencial tecnológico em s :

$$M_i^s = \frac{D_i^s(x_{1t}, x_{2t}, q)}{D_i^s(x_{1s}, x_{2s}, q)} = \frac{\overline{OF_t}/\overline{OF_t^s}}{\overline{OF_s}/\overline{OF_s^s}} \quad (7.59)$$

Se, no entanto, adotarmos o referencial tecnológico do período t , o aumento na eficiência técnica para um dado nível de produto q é medido a partir de F_t por:

$$D_i^t(x_{1t}, x_{2t}, q) = \frac{\overline{OF_t}}{OF_t^t} \quad (7.60)$$

Partindo do ponto $F_{s,t}$ e usando a tecnologia em t como referência, obtemos uma medida da proporção de redução de insumos possível com a nova tecnologia disponível em t :

$$D_i^t(x_{1s}, x_{2s}, q) = \frac{\overline{OF_s}}{OF_s^t} \quad (7.61)$$

O índice de Malmquist sob a ótica do insumo, usando a tecnologia em t como referência, pode então ser expresso por:

$$M_i^t = \frac{D_i^t(x_{1t}, x_{2t}, q)}{D_i^t(x_{1s}, x_{2s}, q)} = \frac{\overline{OF_t}/\overline{OF_t^t}}{\overline{OF_s}/\overline{OF_s^t}} \quad (7.62)$$

De forma similar ao verificado para o índice de Malmquist sob a ótica do produto, sob a ótica do insumo esse índice é dado pela média geométrica dos índices com referencial tecnológico em s e t :

$$M_i = \left[M_i^s M_i^t \right]^{1/2} = \left[\frac{D_i^s(x_{1t}, x_{2t}, q)}{D_i^s(x_{1s}, x_{2s}, q)} \frac{D_i^t(x_{1t}, x_{2t}, q)}{D_i^t(x_{1s}, x_{2s}, q)} \right]^{1/2}$$

$$= \left[\frac{\overline{OF}_t / \overline{OF}_t^s}{\overline{OF}_s / \overline{OF}_s^s} \frac{\overline{OF}_t / \overline{OF}_t^t}{\overline{OF}_s / \overline{OF}_s^t} \right]^{1/2} \quad (7.63)$$

Multiplicando (7.63) por

$$[D_i^t(x_{1t}, x_{2t}, q) / D_i^s(x_{1s}, x_{2s}, q)] \cdot [D_i^s(x_{1s}, x_{2s}, q) / D_i^t(x_{1t}, x_{2t}, q)]$$

e rearranjando resulta em:

$$M_i = \frac{\overline{OF}_t / \overline{OF}_t^t}{\overline{OF}_s / \overline{OF}_s^s} \left[\frac{\overline{OF}_t / \overline{OF}_t^s}{\overline{OF}_t / \overline{OF}_t^t} \frac{\overline{OF}_s / \overline{OF}_s^q}{\overline{OF}_s / \overline{OF}_s^t} \right]^{1/2} \quad (7.64)$$

O primeiro termo fora dos colchetes representa a mudança na eficiência técnica entre s e t . Esse termo será igual a 1, se não houver melhora na eficiência técnica. E será menor (maior) que 1 se houver melhora (piora) na eficiência técnica. Note-se que a interpretação desse termo sob a ótica do insumo é contrária à interpretação sob a ótica do produto.

O segundo termo (entre colchetes) mede a mudança na tecnologia de produção entre s e t . Esse termo será menor (maior) que 1 se a mudança tecnológica tiver um efeito positivo (negativo).

7.3 Considerações finais

As medidas de eficiência técnica apresentadas neste capítulo proveem uma fundamentação microeconômica para a mensuração da eficiência. As análises sob as óticas do produto ou dos insumos são apenas duas maneiras distintas de abordar o mesmo problema. Não há uma recomendação acerca da melhor maneira de abordar o problema. Empiricamente, a decisão da abordagem teórica dependerá do problema em mãos.

Uma limitação importante em relação aos fundamentos

microeconômicos da mensuração da eficiência se refere ao fato de que estes se baseiam em pressupostos de que a unidade produtora seja maximizadora de lucros ou minimizadora de custos. Tais suposições se adéquam bem aos problemas de mensuração da eficiência no setor privado, mas não necessariamente aos problemas de mensuração da eficiência no setor público. É possível, por exemplo, que o objetivo do burocrata administrador de empresas ou órgãos públicos seja aumentar o bem-estar da população. Nesse caso, ele poderá sacrificar eficiência em troca de aumento de bem-estar. É possível também que o objetivo do burocrata seja maximizar seu orçamento para aumentar seu poder e prestígio (NISKANEN, 1971).

Ainda assim, é possível calcular a eficiência técnica empiricamente. No entanto, do ponto de vista teórico, não se pode mensurar a eficiência total ou alocativa, pois não se sabe exatamente o objetivo do burocrata. Mesmo a eficiência social seria de difícil mensuração, pois também há divergências na teoria sobre o objetivo de maximização do bem-estar social do burocrata.

REFERÊNCIAS

FARREL, M. The measurement of productive efficiency. *Journal of the Royal Statistical Society, Series A*, v. 120, p. 253-281, 1957.

JOHNES, J. Efficiency measurement. In: JOHNES, G.; JOHNES, J. (Ed.). *International handbook on the economics of education*. Cheltenham: Edward Elgar Publishing Ltd., 2004. p. 613-742.

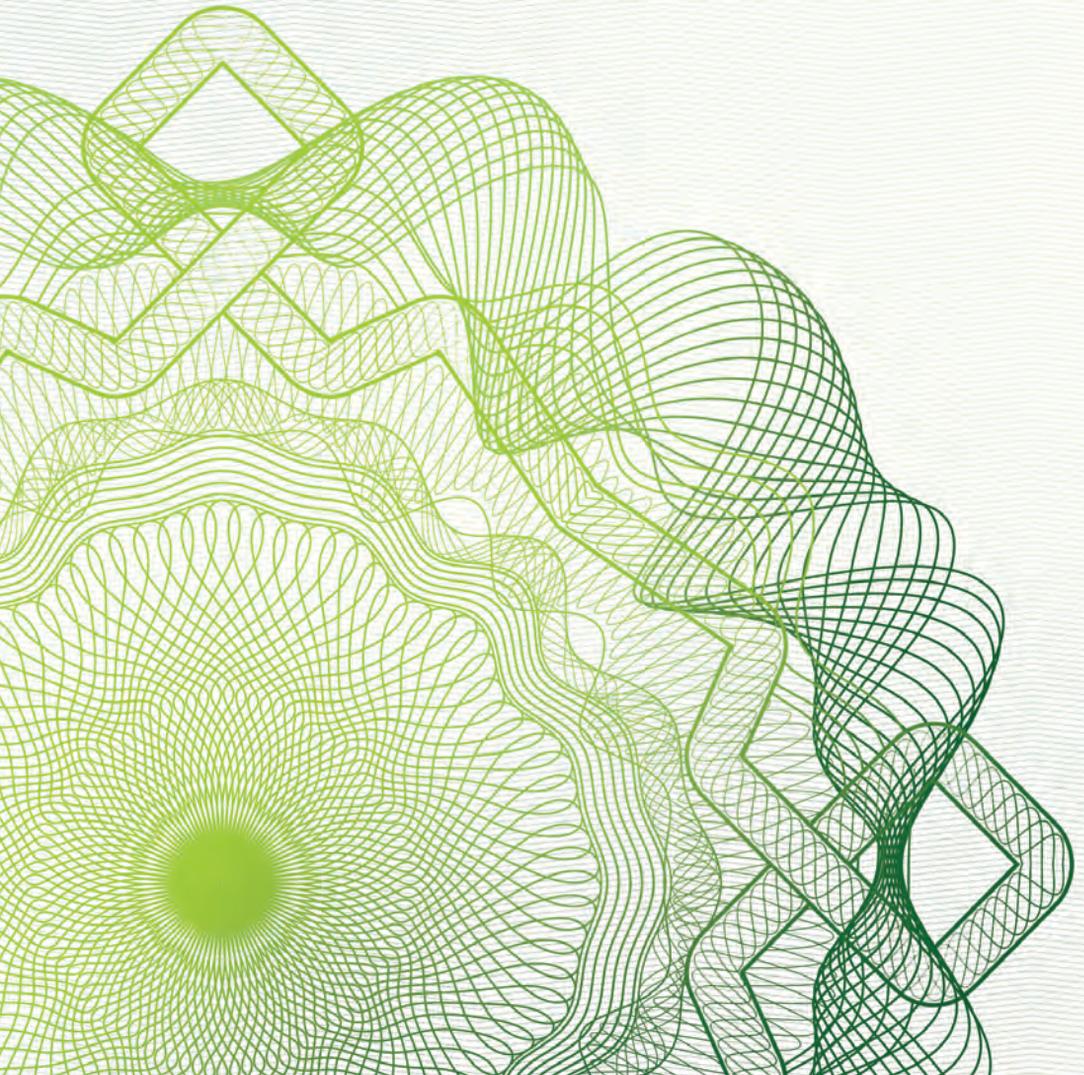
MALMQUIST, S. Index numbers and indifference surfaces. *Trabajos de Estatística*, v. 4, p. 209-242, 1953.

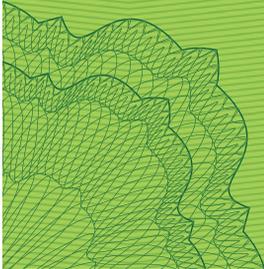
NISKANEN, W. A. *Bureaucracy and representative government*. Chicago: Aldine-Atherton, 1971.

SNYDER, C.; NICHOLSON, W. *Microeconomic theory*. Mason: Thompson South-Western, 2008.

PARTE IV

COMO MEDIR A EFICIÊNCIA?





Capítulo 8

Modelos não paramétricos: Análise Envoltória de Dados (DEA)

Rogério Boueri

8.1 Introdução

Neste capítulo será introduzida de forma intuitiva a metodologia de Análise Envoltória de Dados (DEA¹) para análise da fronteira eficiente. Após a explicação sobre a metodologia DEA, haverá um aprofundamento dos detalhes técnicos de seu funcionamento e uma discussão a respeito de suas várias versões, incluindo modelos com ótica nos produtos e nos insumos, modelos com e sem retornos de escala e análise de dados em painel.

A análise parte do pressuposto de que existe certo número de unidades decisórias (DMUs)² que convertem insumos (*inputs*) em produtos (*outputs*) ou resultados (*outcomes*). Assim, por exemplo, um hospital utiliza seus médicos, enfermeiros, leitos, medicamentos etc. como insumos para gerar produtos, tais como atendimentos ambulatoriais, internações e cirurgias. Dessa forma, se houver dados disponíveis para essas variáveis em vários hospitais, será possível avaliar a eficiência relativa das diversas unidades hospitalares a partir de uma comparação entre os seus gastos com insumos e a geração de produtos.³

Outro aspecto a ser ressaltado é a diferença entre *produtos* e *resultados*. Muito embora o objetivo das unidades decisórias seja a obtenção de resultados e não a mera geração de produtos, os primeiros são de mais difícil avaliação, uma vez que a quantidade de fatores externos que contribuem para a sua produção pode ser muito grande.

1 Análise Envoltória de Dados é a tradução literal do termo em inglês *Data Envelopment Analysis* (DEA).

2 Do inglês, *decision making units*.

3 Outros exemplos, em diversos setores de atividade, poderiam ser citados: a eficiência dos municípios na prestação de serviços públicos poderia ser avaliada utilizando-se o gasto municipal (ou diversas de suas modalidades) como insumo e a evolução do IDH como resultado (BOUERI, 2007) ou uma rede de agências bancárias poderia ser avaliada contrastando-se a utilização de capital e de mão de obra de cada uma delas (insumos) com suas respectivas captações de depósitos e clientes (TECLES; TABAK, 2010).

Um hospital especializado em cirurgias cardíacas, por exemplo, existe não com o objetivo de realizar tais cirurgias, que são o produto gerado pelo hospital, mas sim para reduzir o número de óbitos decorrentes desse tipo de doença (resultado). Acontece, no entanto, que o número de óbitos depende de uma infinidade de fatores sobre os quais o *staff* do hospital não tem nenhum controle, tais como a dieta habitual da população atendida, seu estilo de vida, sua renda e aderência aos tratamentos prescritos. A capacidade de o hospital realizar as internações e cirurgias necessárias está muito mais sob o controle de seus profissionais.

8.2 Índices de produtividade

Suponha um processo produtivo simples, no qual apenas um insumo seja utilizado para a geração de um único produto. Nesse caso, a avaliação da produtividade seria uma tarefa razoavelmente simples, para a qual bastaria que se registrassem corretamente os montantes utilizados do insumo e o total produzido.

Um exemplo ilustrativo dessa aferição ocorre quando se avalia a *eficiência* de determinado veículo em relação ao seu consumo de combustível. Nesse caso, eles são comparados por quantos quilômetros rodam com um litro de combustível consumido.⁴ Assim, o índice de produtividade tem o total produzido (quilômetros rodados) no numerador e a quantidade de insumo (litros de combustível consumidos) no denominador. A partir dessa computação relativamente simples, denomina-se de eficiente aquele veículo que possui a maior produtividade, ou seja, aquele que roda maior quilometragem com uma unidade de combustível.

Observe que, como destacado nesse exemplo, os conceitos de produtividade e de eficiência são relacionados, porém distintos. A produtividade se apresenta como um conceito absoluto, enquanto a eficiência é uma grandeza relativa, ou comparativa. Ainda seguindo o exemplo, um carro que percorresse 10 quilômetros com um litro de combustível talvez fosse considerado eficiente na década de 1970. Hoje, um veículo com essa performance não poderia mais ostentar esse adjetivo.

4 Esse exemplo, apesar de ilustrativo, deixa de fora vários aspectos importantes da eficiência energética de um veículo. Não é considerada, por exemplo, a carga total transportada pelo veículo.

A situação fica mais complicada quando existem mais de um insumo e produto. Nesse caso, a criação de um índice de produtividade requereria o emprego de pesos que ponderassem as utilizações dos diferentes insumos e a produção dos diferentes produtos. Para tornar a discussão mais clara, suponha uma situação na qual seja necessário avaliar n DMUs, que estão engajadas em um processo produtivo no qual utilizem m insumos diferentes para gerar s tipos de produtos distintos. Poder-se-ia, então, criar um índice virtual de produção para cada unidade, de modo a se obter uma agregação dos produtos gerados:

$$IVP_i = \mu_1 q_{1,i} + \mu_2 q_{2,i} + \dots + \mu_s q_{s,i} \quad (8.1)$$

Onde $q_{j,i}$ é a quantidade produzida pela i -ésima DMU do produto j e μ_j é o peso atribuído ao produto de número j na construção do índice. Analogamente, um índice virtual de insumos também poderia ser computado:

$$IVI_i = v_1 x_{1,i} + v_2 x_{2,i} + \dots + v_m x_{m,i} \quad (8.2)$$

Nesse caso, $x_{j,i}$ é a quantidade utilizada do insumo j pela i -ésima DMU e v_j é o peso atribuído ao insumo de número j na construção do índice. A obtenção do índice de produtividade ocorreria pela simples divisão do índice virtual de produtos pelo índice virtual de insumos:

$$\theta_i = \frac{IVP_i}{IVI_i} = \frac{\mu_1 q_{1,i} + \mu_2 q_{2,i} + \dots + \mu_s q_{s,i}}{v_1 x_{1,i} + v_2 x_{2,i} + \dots + v_m x_{m,i}} \quad (8.3)$$

Como já citado, o principal problema desse índice está na seleção dos vetores de pesos μ e v . Como os escolher de forma não arbitrária, ou seja, como selecionar tais pesos sem favorecer nenhuma DMU específica *a priori* e de modo a poder compará-las?

8.3 Modelo CCR com ótica nos insumos

Para resolver esse problema, Charnes, Cooper e Rhodes (1978) propuseram um método não arbitrário de escolha dos vetores de pesos. Desde então, esse modelo tem sido denominado DEA-CCR. A ideia é deixar a própria amostra escolher os pesos para cada

unidade. Matematicamente, esse processo pode ser descrito pelo seguinte problema de maximização:

$$\left\{ \begin{array}{l} \text{Max} : \theta_i = \frac{\mu_{1i}q_{1i} + \mu_{2i}q_{2i} + \dots + \mu_{si}q_{si}}{v_{1i}x_{1i} + v_{2i}x_{2i} + \dots + v_{mi}x_{mi}} \\ \text{s.t.} : \frac{\mu_{1i}q_{1j} + \mu_{2i}q_{2j} + \dots + \mu_{si}q_{sj}}{v_{1i}x_{1j} + v_{2i}x_{2j} + \dots + v_{mi}x_{mj}} \leq 1, \forall j = 1, 2, \dots, n \\ \mu_{1i}, \mu_{2i}, \dots, \mu_{si} \geq 0 \text{ e } u_i \neq \bar{0} \\ v_{1i}, v_{2i}, \dots, v_{mi} \geq 0 \text{ e } v_i \neq \bar{0} \end{array} \right. \quad (8.4)$$

A função objetivo do problema é o nível de eficiência da unidade i . Então o problema sugere que sejam escolhidos os melhores pesos (μ e v) possíveis para a unidade, mas também estabelece restrições. Note que a solução diferirá para cada DMU,⁵ sempre buscando o conjunto de pesos mais favoráveis a cada DMU em particular.

O primeiro conjunto de restrições determina que os pesos escolhidos, quando utilizados sobre os registros de qualquer outra unidade, não podem gerar valores superiores à unidade. Observe que essa restrição se aplica inclusive à própria DMU i . Esse conjunto de restrições é denominado *restrições de consistência*, e é utilizado para que o problema tenha significado comparativo. As demais restrições são as chamadas *condições de positividade*, e asseguram que os pesos sejam todos não negativos e que pelo menos um insumo e um produto sejam positivamente ponderados na construção do índice.

É possível provar que o problema fractal (não linear) expresso em (8.4) é equivalente ao seguinte problema linear:

$$\left\{ \begin{array}{l} \text{Max} : \theta_i = \omega_{1i}q_{1i} + \omega_{2i}q_{2i} + \dots + \omega_{si}q_{si} \\ \text{s.t.} : \vartheta_{1i}x_{1i} + \vartheta_{2i}x_{2i} + \dots + \vartheta_{mi}x_{mi} = 1 \\ \omega_{1i}q_{1j} + \omega_{2i}q_{2j} + \dots + \omega_{si}q_{sj} \leq \vartheta_{1i}x_{1j} + \vartheta_{2i}x_{2j} + \dots + \vartheta_{mi}x_{mj}, \forall j = 1, 2, \dots, n \\ \omega_{1i}, \omega_{2i}, \dots, \omega_{si} \geq 0 \text{ e } \omega_i \neq \bar{0}_s \\ \vartheta_{1i}, \vartheta_{2i}, \dots, \vartheta_{mi} \geq 0 \text{ e } \vartheta_i \neq \bar{0}_m \end{array} \right. \quad (8.5)$$

5 Daí o subscrito i nos pesos, denotando pesos individuais para cada unidade.

Nessa transformação do problema fractal para o problema linear, acrescenta-se a *restrição normalizadora*, que assegura que a soma ponderada dos insumos seja igual a 1. A grande vantagem da formulação linear é que ela é de mais fácil resolução que a versão fractal. Na época em que o método foi desenvolvido, a facilidade computacional era uma propriedade muito importante. Embora hoje, com o avanço dos processos computacionais, seja possível a resolução direta do problema fractal, a literatura da DEA se desenvolveu com ênfase no problema linear e, por isso, a maioria das interpretações sobre o método é proveniente dessa formulação.

A solução do problema (8.5) é obtida por meio de programação linear. Note que como existem no problema n DMUs, será necessária a resolução de n problemas, um para cada unidade. A cada uma dessas unidades será assinalado um conjunto próprio de pesos que serão, em geral, diferentes dos pesos das outras unidades.

Observe o seguinte exemplo, no qual seis DMUs que produzem a mesma quantidade de um único produto (q) utilizando montantes diferentes de dois insumos (x_1 e x_2) são comparadas.

Tabela 8.1 – Utilização de insumos, geração de produto e escore de eficiência de seis DMUs hipotéticas				
DMU	x_1	x_2	q	θ
A	175	75	100	0,649
B	250	40	100	1,000
C	200	40	100	1,000
D	100	75	100	0,857
E	50	100	100	1,000
F	100	50	100	1,000

Fonte: Elaboração própria.

A quinta coluna da tabela apresenta os escores de eficiência já calculados. As unidades B, C, E e F apresentaram escore 1. Já as unidades A e D ficaram abaixo de 1, demonstrando assim que, quando comparadas com as outras DMUs da amostra, são ineficientes.

Tome o caso da DMU A – os pesos escolhidos para ela pelo modelo DEA foram os seguintes: $\omega_{1A} = 0,00649$, $\vartheta_{1A} = 0,00108$ e $\vartheta_{2A} = 0,01081$. Observe que então:

$$\theta_A = \frac{\omega_{1A}q_A}{\vartheta_{1A}x_{1A} + \vartheta_{2A}x_{2A}} = \frac{0,00649 \times 100}{0,00108 \times 175 + 0,0108 \times 75} = 0,649$$

O denominador soma uma unidade como o requerido na restrição normalizadora. O que deve ser ressaltado aqui é que os pesos escolhidos para a DMU A foram os melhores possíveis sem que alguma das restrições de consistência fosse violada. Por exemplo, considere a aplicação dos pesos escolhidos para a unidade A sobre os valores de insumos e produto da unidade F:

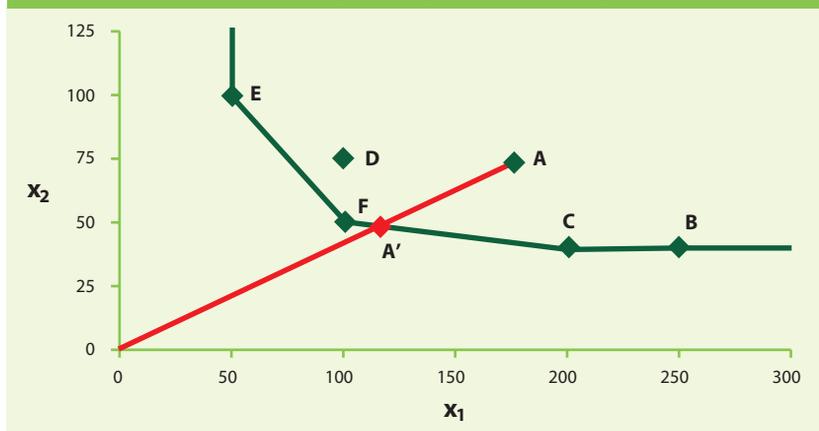
$$\frac{\omega_{1A}q_F}{\vartheta_{1A}x_{1F} + \vartheta_{2A}x_{2F}} = \frac{0,00649 \times 100}{0,00108 \times 100 + 0,0108 \times 50} = 1,000$$

Isso quer dizer que se os pesos escolhidos para A fossem levemente melhorados, isto é, se os pesos do denominador fossem diminuídos ou se o peso do numerador fosse aumentado, tais pesos modificados produziram um valor maior que 1 quando aplicados aos dados da unidade F, o que por sua vez violaria a restrição de consistência. Resumindo, não é possível melhorar os pesos de A, e com esses pesos o escore de eficiência atingido é 0,649.

O Gráfico 8.1 plota no eixo horizontal os montantes utilizados do insumo 1 e no eixo vertical os do insumo 2. Como nesse exemplo todas as unidades geram a mesma quantidade de produto, é possível se desenhar uma isoquanta.⁶

6 Apenas a título de lembrança, isoquanta é o lócus de todas as combinações de insumos capazes de produzir a mesma quantidade de produto.

Gráfico 8.1 – Combinações de insumos utilizadas para a produção de 100 unidades de produto pelas seis DMUs hipotéticas



Fonte: Elaboração própria.

Observe que as unidades que se encontram sobre a isoquanta são exatamente aquelas apontadas pelo DEA como eficientes. Já a unidade A, que como foi visto acima, é ineficiente, está à direita da isoquanta. Isso quer dizer que ela está utilizando mais insumos do que deveria. Uma interpretação do escore de eficiência é quanto uma unidade poderia reduzir proporcionalmente a sua utilização de insumos de forma a atingir a fronteira de eficiência, que nesse exemplo é a isoquanta.

Essa redução, no caso da DMU A, corresponderia a 69,4%, isto é, se essa unidade fosse eficiente, utilizaria apenas 69,4% dos insumos que ela utiliza atualmente para produzir 100 unidades de produto. Essa redução dos insumos é representada pelo ponto A no gráfico.

Um caso mais complicado é o da DMU B. Ela está sobre a fronteira de eficiência, mas é claramente ineficiente, uma vez que, para produzir a mesma quantidade de produto que a unidade C, ela gasta o mesmo montante do insumo x_2 (40 unidades), consumindo, contudo, mais do insumo x_1 (250 unidades contra 200 unidades de C). É como se ela se escondesse atrás de C, e nenhuma redução proporcional dos seus insumos fosse possível de forma que a produção de 100 unidades do produto continuasse viável. Essa situação também demonstra que o fato de a DMU estar sobre a fronteira de eficiência é condição necessária, mas não suficiente, para que seja considerada eficiente.

Nesse ponto é possível fazer uma definição mais precisa de eficiência no âmbito do modelo DEA. Uma DMU será *CCR-eficiente* caso obtenha um escore (θ_j) igual a 1 e caso nenhum dos seus pesos, tanto de insumos (ϑ_{m_i}) quanto de produtos (ω_{s_j}) seja igual a zero.

No exemplo da Tabela 8.1, ao calcularem-se os pesos atribuídos pelo modelo DEA a cada uma das DMUs (Tabela 8.2) comprova-se que, realmente, o peso atribuído ao insumo x_1 no caso da DMU B é zero.

Tabela 8.2 – Peso atribuído pelo modelo DEA a cada uma das variáveis em cada DMU			
DMU	ϑ_1	ϑ_2	ω_1
A	0,00108	0,01081	0,00649
B	0,00000	0,02500	0,01000
C	0,00167	0,01667	0,01000
D	0,00571	0,00571	0,00857
E	0,00667	0,00667	0,01000
F	0,00667	0,00667	0,01000

Fonte: Elaboração própria.

Como B gasta a mesma quantidade de x_2 que C e mais de x_1 , o modelo atribui peso zero a essa última variável, pois qualquer peso positivo faria com que a restrição de consistência fosse violada quando os pesos escolhidos para B fossem utilizados nos registros de C.

Esse exemplo acima se atém ao gasto de insumos para atingir determinada quantidade de produto. A eficiência está sendo, então, analisada pela ótica do insumo, ou seja, a questão aqui é quanto se poderia reduzir o gasto com insumos para se atingir a mesma quantidade produzida.

8.4 Modelo CCR com ótica nos produtos

O conceito de eficiência também pode ser definido pela ótica do produto. Nesse caso, a questão é: quais os máximos de produtos podem ser gerados, para determinada utilização de insumos?

Para exemplificar, suponha que em uma turma de 10 alunos, todos eles tenham estudado 10 horas para uma prova de duas questões.

O professor atribui notas de zero a cinco para cada questão, notas essas descritas nas colunas 3 e 4 da Tabela 8.3.

Tabela 8.3 – Notas dos alunos por questão, pesos atribuídos e nota final							
Aluno	Horas de Estudo	Questão 1	Questão 2	v_{Horas}	$w_{\text{Questão 1}}$	$w_{\text{Questão 2}}$	Nota Final
1	10	4,0	4,0	0,100	0,172	0,078	100 %
2	10	1,8	5,0	0,100	0,078	0,172	100 %
3	10	0,9	3,0	0,167	0,000	0,333	60 %
4	10	3,9	2,2	0,119	0,204	0,093	84 %
5	10	0,2	3,4	0,147	0,000	0,294	68 %
6	10	3,0	0,9	0,167	0,333	0,000	60 %
7	10	1,3	2,5	0,188	0,147	0,324	53 %
8	10	4,1	2,4	0,112	0,193	0,088	89 %
9	10	5,0	1,8	0,100	0,200	0,000	100 %
10	10	2,4	3,5	0,127	0,099	0,218	79 %

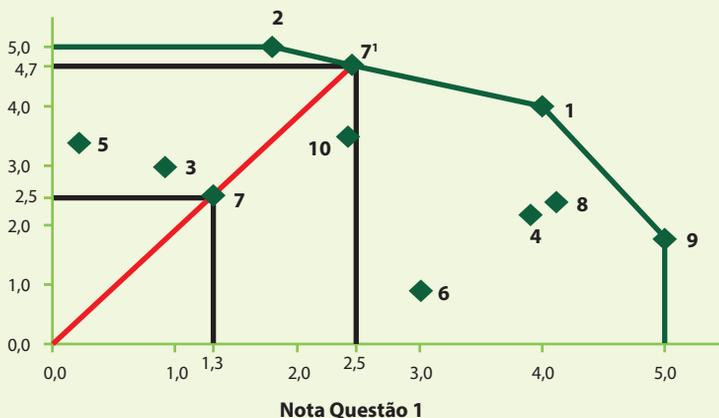
Fonte: Elaboração própria.

O professor decidiu atribuir notas finais da prova utilizando a metodologia DEA. Tomando as horas de estudo como insumo⁷ e a nota de cada questão como produtos, o professor resolveu atribuir para cada aluno os pesos mais favoráveis possíveis para cada uma das questões, mas de forma que tais pesos, quando aplicados às questões de qualquer outro aluno, não produzissem uma nota superior a 100% (restrição de consistência).

Foi, então, rodado um DEA voltado para produto, e obtiveram-se os pesos e as notas finais das provas conforme o descrito nas colunas finais da Tabela 8.3. Essas notas, plotadas em um gráfico, apareceriam da seguinte forma:

⁷ Supondo que todos os alunos tenham utilizado as 10 horas disponíveis para estudar.

Gráfico 8.2 – Notas das questões 1 e 2



Fonte: Elaboração própria.

O Gráfico 8.2 mostra a posição de cada aluno em termos de combinações de suas notas em cada questão. Observe que os alunos 1, 2 e 9 estão na fronteira e obtiveram nota máxima na prova. Já o aluno 7 tirou nota 5,3.⁸ Geometricamente, isso quer dizer que a combinação de notas do aluno 7 é apenas 53% do que poderia ser, ou seja, esse aluno consegue 1,3 e 2,5 pontos, na primeira e na segunda questão, respectivamente, quando poderia ter obtido 2,5 e 4,7, posicionando-se, então, sobre a fronteira.

A formulação DEA deste exemplo adota a ótica do produto, como já mencionado acima. O problema linear correspondente a essa formulação pode ser expresso da seguinte maneira:

$$\left\{ \begin{array}{l} \text{Min : } \eta_i = \vartheta_{1i}x_{1i} + \vartheta_{2i}x_{2i} + \dots + \vartheta_{mi}x_{mi} \\ \text{s.t. : } \omega_{1i}q_{1i} + \omega_{2i}q_{2i} + \dots + \omega_{si}q_{si} = 1 \\ \omega_{1i}q_{1j} + \omega_{2i}q_{2j} + \dots + \omega_{si}q_{sj} \leq \vartheta_{1i}x_{1j} + \vartheta_{2i}x_{2j} + \dots + \vartheta_{mi}x_{mj}, \forall j = 1, 2, \dots, n \\ \omega_{1i}, \omega_{2i}, \dots, \omega_{si} \geq 0 \text{ e } \omega_i \neq \bar{0} \\ \vartheta_{1i}, \vartheta_{2i}, \dots, \vartheta_{mi} \geq 0 \text{ e } \vartheta_i \neq \bar{0} \end{array} \right. \quad (8.6)$$

8 Considerando-se o valor máximo da prova igual a 10.

A função objetivo, contida na primeira equação, mostra que o que está sendo minimizado, nesse caso, é o denominador do problema fractal. Já a segunda equação, que é a restrição normalizadora, assegura que o numerador daquele problema esteja sendo igualado a um. As restrições de consistência garantem que o numerador será sempre menor ou igual ao denominador. Como o numerador é sempre igual a um, o denominador deverá sempre ser maior ou igual à unidade.

A solução da versão do problema orientada para produtos fornecerá sempre um resultado maior ou igual à unidade. A interpretação desse número será relacionada à capacidade de expansão proporcional dos produtos, dada uma determinada utilização de insumos. Assim, por exemplo, se determinada DMU atingir um escore de eficiência de 1,2, isso quer dizer que, pela quantidade de insumos que está utilizando, ela poderia estar produzindo 20% a mais de cada produto.

A título de exemplo, suponha uma atividade na qual as DMUs utilizam um insumo (x) para produzir dois produtos diferentes, como o descrito na Tabela 8.4.

Tabela 8.4 – Quantidade de insumo utilizada e quantidades de produtos geradas por 10 DMUs fictícias				
DMU	x	y_1	y_2	η
A	64	20	15	1,000
B	67	19	17	1,030
C	92	22	21	1,188
D	56	4	24	1,000
E	75	26	9	1,000
F	81	25	13	1,073
G	79	28	3	1,000
H	75	3	29	1,108
I	66	6	25	1,079
J	63	18	17	1,000

Fonte: Elaboração própria.

Neste exemplo, não é mais possível plotar um gráfico referenciando cada DMU à mesma fronteira, visto que cada DMU tem

Ou, mais simplesmente, em notação matricial:

$$\left\{ \begin{array}{l} \text{Min: } \theta_i \\ \text{ } \\ \text{s.t.: } \theta_i \bar{x}_i - X\bar{\lambda} \geq \bar{0}_m \\ \quad Q\bar{\lambda} - \bar{q}_i \geq \bar{0}_s \\ \quad \bar{\lambda} \geq \bar{0}_n \end{array} \right. \quad (8.8)$$

Onde θ_i é o escore de eficiência da DMU i , \bar{x}_i é o vetor ($m \times 1$) de insumos utilizados pela DMU i , \bar{q}_i é o vetor ($s \times 1$) de produtos gerados pela DMU i , $\bar{\lambda}$ é o vetor ($n \times 1$) de pesos para as combinações lineares, X é uma matriz ($m \times n$) formada pelos vetores de insumos transpostos de todas as DMUs da amostra, Q é uma matriz ($s \times n$) formada pelos vetores transpostos de produtos de todas as DMUs da amostra e $\bar{0}_j$ é o vetor nulo com a dimensão j apropriada.

Nessa formulação, o vetor $\bar{\lambda}$ contém os pesos para fazer as melhores combinações lineares com os vetores de insumos e de produtos de todas as DMUs da amostra. Então, o objetivo aqui é encontrar o menor θ possível, ou seja, promover a maior redução proporcional possível na utilização de insumos da DMU i , mas de forma que ela ainda possa ser expressa como uma combinação linear dos vetores das outras DMUs.

Mais uma vez podemos utilizar um exemplo para esclarecer esse ponto. Suponha dois agricultores que produzem a mesma quantidade de bananas. O primeiro utiliza 100 quilos de adubo e 10 litros de pesticida, enquanto o segundo utiliza, respectivamente, 80 quilos e 20 litros de cada produto. Se um terceiro agricultor produzir a mesma quantidade da fruta, mas utilizando-se de 99 quilos de adubo e 16,5 litros de pesticida, poder-se-ia afirmar que ele seria ineficiente e que tal nível de ineficiência seria de 10%.

Aqui tem-se a suposição: se é possível produzir determinada quantidade de bananas a partir da utilização de 100 quilos de adubo e 10 litros de pesticida, ou de 80 quilos de adubo e 20 litros de pesticida, então também é possível se produzir essa quantidade utilizando-se 90 quilos de adubo e 15 litros de pesticida, já que:

$$0,5 \times (100; 10) + 0,5 \times (80; 20) = (90; 15)$$

O terceiro agricultor, ao gastar 99 quilos e 16,5 litros está, portanto, gastando 10% a mais do que deveria. Nesse caso teríamos $\bar{\lambda} = (0,5; 0,5; 0)$, onde cada elemento do vetor refere-se ao peso do agricultor correspondente na combinação linear.

A grande vantagem computacional da forma envelopada do problema está no fato de que ela contém menos restrições, especialmente quando as restrições de não negatividade, que não dificultam o cômputo do problema, são desconsideradas. Por exemplo, o problema especificado pela equação 9.5 tem, já descontadas as restrições de não negatividade, $(n + 1)$ restrições, enquanto a forma envelopada, descrita pela equação 9.7, possui $(s + m)$ restrições.

Ora, para que se possa calcular eficazmente a eficiência relativa das DMUs utilizando-se o modelo DEA, faz-se necessário que o número de DMUs contido na amostra seja expressivamente maior que a soma do número de insumos e de produtos do problema, isto é, $n > s + m$.

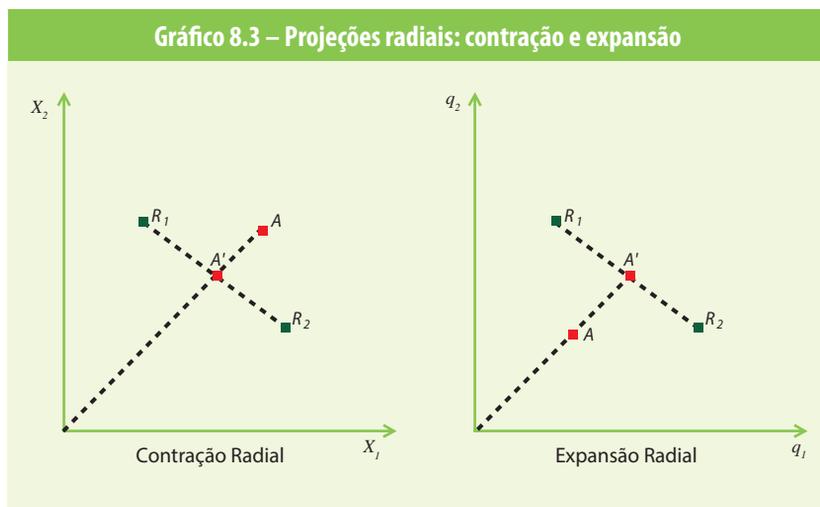
Na versão com ótica dos produtos, o problema na forma envelopada busca a maior expansão proporcional possível dos produtos de determinada DMU até que eles atinjam a maior combinação linear possível dos vetores de produtos existentes na amostra. Matematicamente:

$$\left\{ \begin{array}{l} \text{Max: } \eta_i \\ \text{ } \\ \text{s.t.: } \bar{x}_i - X\bar{\mu} \geq \bar{0}_m \\ \quad Q\bar{\mu} - \eta\bar{q}_i \geq \bar{0}_s \\ \quad \bar{\mu} \geq \bar{0}_n \end{array} \right. \quad (8.9)$$

Observe que, no caso da DMU i estar sobre a fronteira de eficiência, o seu conjunto de pesos para a melhor combinação linear será um vetor todo composto por zeros, exceto no elemento correspondente à própria DMU i , o qual será igual a 1. Isso quer dizer que a melhor combinação linear que se pode obter com os vetores da amostra e que é proporcional aos vetores da DMU i são os próprios vetores dessa DMU.

8.6 Conjuntos de referência

Como foi visto na seção anterior, a ineficiência de determinada DMU pode ser estimada pela magnitude da projeção necessária para que ela alcance a referida fronteira. Essa projeção, que pode ser uma contração dos insumos ou uma expansão dos produtos, é realizada de forma proporcional. Por esse motivo é denominada projeção radial, pois sempre pode ser expressa como um raio partindo da origem.



Fonte: Elaboração própria.

No painel esquerdo do Gráfico 8.3, uma contração radial do ponto A é demonstrada, e no painel direito, uma expansão. Observe que ambas se realizam na direção da origem, embora a expansão se dê no sentido contrário. Em ambos os casos, os resultados das projeções (ponto A') estão sobre a fronteira, que é definida como uma combinação linear dos pontos R_1 e R_2 .

Esses pontos que dão origem ao trecho da fronteira sobre o qual a projeção de um determinado ponto recai formam o conjunto de referência da DMU em questão. Por exemplo, no Gráfico 8.1 pode-se depreender que as DMUs C e F formam o conjunto de referência da DMU A , assim como as DMUs E e F formam o conjunto de referência da DMU D .

É claro que, na maioria dos casos, essa identificação visual das DMUs formadoras de determinado conjunto de referência nem sempre é possível. Então, para identificá-las, deve-se confiar nos valores calculados para o vetor $\bar{\lambda}$.

A Tabela 8.5 é baseada na Tabela 8.1, à qual se acrescentam as colunas contendo os valores calculados para o vetor de pesos das combinações lineares ($\bar{\lambda}$) para cada DMU.

Tabela 8.5 – Quantidade de insumo utilizada e quantidades de produtos geradas por DMUs hipotéticas										
DMU	x_1	x_2	q	θ	λ_A	λ_B	λ_C	λ_D	λ_E	λ_F
A	175	75	100	0,649	–	–	0,14	–	–	0,87
B	250	40	100	1,000	–	1,00	–	–	–	–
C	200	40	100	1,000	–	–	1,00	–	–	–
D	100	75	100	0,857	–	–	–	–	0,29	0,71
E	50	100	100	1,000	–	–	–	–	1,00	–
F	100	50	100	1,000	–	–	–	–	–	1,00

Fonte: Elaboração própria.

Alguns fatos são notáveis nessa tabela. O primeiro deles é que as DMUs com score igual a 1 também possuem um único λ positivo, que por sua vez também é igual a 1. Isso porque se a DMU está na fronteira de eficiência, a combinação linear que a projeta sobre a fronteira é ela própria.

Depois, para aquelas DMUs que não estão na fronteira, os λ 's positivos são aqueles correspondentes às DMUs que estão no seu conjunto de referência. Por exemplo, no caso da DMU A, $\lambda_C = 0,14$ e $\lambda_F = 0,87$, como foi constatado graficamente, as DMUs C e F são justamente aquelas que formam o conjunto de referência de A. Fato similar ocorre com a DMU D, para a qual $\lambda_E = 0,29$ e $\lambda_F = 0,71$. As outras DMUs que estão na fronteira têm seus conjuntos de referência formados por si próprias.

Então, mesmo em situações nas quais não se podem detectar graficamente os conjuntos de referência de cada DMU, tal análise é possível por meio dos valores dos λ 's. Em suma, o conjunto de referência de cada DMU será formado por aquelas DMUs cujos λ 's são positivos.

8.7 Modelos com retornos variáveis de escala

Uma das limitações mais importantes do modelo CCR é a suposição de retornos constantes de escala. Esse tipo de modelo não consegue incorporar situações nas quais as DMU estão sujeitas a variações na escala de produção.

Felizmente, Banker, Charnes e Cooper (1984) conseguiram superar essa dificuldade acrescentando uma restrição adicional ao problema envelopado. Esse modelo ficou conhecido na literatura como BCC,¹⁰ e pode ser expresso matematicamente, em sua versão voltada para insumos, da seguinte maneira:

$$\left\{ \begin{array}{l} \text{Min: } \theta_i \\ \quad \{\theta, \bar{\lambda}\} \\ \text{s.t.: } \theta_i \bar{x}_i - X \bar{\lambda} \geq \bar{0}_m \\ \quad Q \bar{\lambda} - \bar{q}_i \geq \bar{0}_s \\ \quad \bar{e}_n \bar{\lambda} = 1 \\ \quad \bar{\lambda} \geq \bar{0}_n \end{array} \right. \quad (8.10)$$

Onde \bar{e}_n é um vetor de dimensão n , exclusivamente composto por valores unitários (1). Observando a descrição do problema na equação 8.10, pode-se observar que há uma restrição a mais, a qual requer que a soma dos λ 's resulte em um, uma vez que:

$$\bar{e}_n \bar{\lambda} = \lambda_1 + \lambda_2 + \dots + \lambda_n = 1$$

Essa restrição, na prática, impede que uma DMU seja comparada com outras muito diferentes dela. Isso porque, ao forçar que os pesos das combinações lineares somem 1, ela impossibilita que DMUs com vetores de insumo ou de produtos muito maiores ou muito menores possam estar no conjunto de referência da DMU investigada.

Observe o exemplo hipotético apresentado na Tabela 8.6. As DMUs lá representadas geram um único produto a partir de dois insumos.

¹⁰ Às vezes também denominado de modelo com retornos variáveis de escala RVE, ou VRS, em inglês.

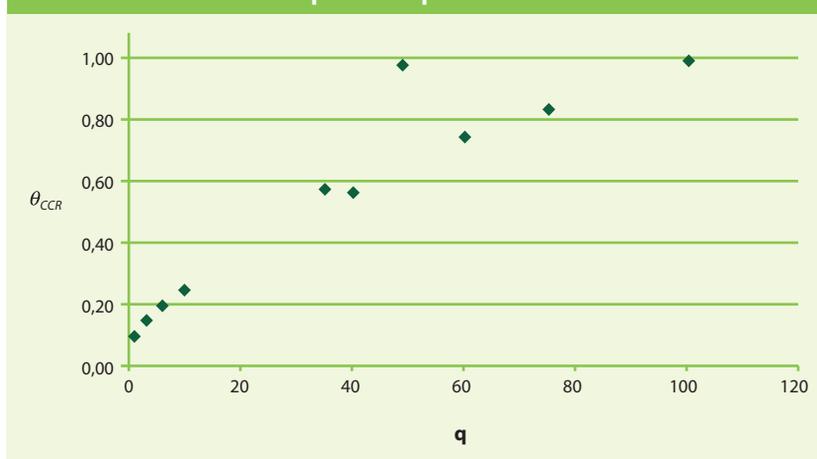
Tabela 8.6 – Utilização de insumos, geração de produto e escores de eficiência de dez DMUs hipotéticas

DMU	x_1	x_2	q	θ_{CCR}	θ_{BCC}
A	1	1	1	0,10	1,00
B	2	2	3	0,15	1,00
C	3	3	6	0,20	0,33
D	4	4	10	0,25	0,30
E	5	5	49	0,98	1,00
F	6	6	35	0,58	0,59
G	7	7	40	0,57	0,58
H	8	8	60	0,75	0,75
I	9	9	75	0,83	0,84
J	10	10	100	1,00	1,00

Fonte: Elaboração própria.

A quinta coluna da tabela apresenta os escores de eficiência calculados pelo modelo CCR, isto é, com retornos constantes de escala, voltado para produto. A única DMU que aparece com escore 1 é a DMU J – todas as outras apresentam algum grau de ineficiência relativa. Além disso, quando esses escores são plotados contra o nível de produção das DMUs, uma clara relação positiva emerge.

Gráfico 8.4 – Níveis de produção e escores de eficiência CCR das dez DMUs hipotéticas apresentadas na Tabela 8.6



Fonte: Elaboração própria.

Esse padrão pode significar que a tecnologia empregada por essas DMUs apresenta retornos crescentes de escala, isto é, a produtividade aumenta com a quantidade produzida. De fato, quando o escore de eficiência é calculado pelo modelo com retornos variáveis de escala (BCC), os resultados são bem diferentes, como se pode observar na sexta coluna da Tabela 8.6. Nesse caso, as DMUs A, B e E, além da J, também apresentam escores de 100%.

Outro ponto interessante que emerge na comparação dos modelos CCR e BCC é a possibilidade da decomposição da eficiência relativa da DMU em eficiência técnica e eficiência de escala. Essa decomposição é realizada dividindo-se θ_{CCR} por θ_{BCC} , então:

$$\theta_{scale} = \frac{\theta_{CCR}}{\theta_{BCC}} \quad (8.11)$$

Primeiramente observe que o escore de eficiência calculado com retornos variáveis de escala nunca será inferior àquele calculado com retornos constantes de escala, ou seja, sempre ocorrerá que $\theta_{BCC} \geq \theta_{CCR}$. Isso porque, no modelo BCC, as DMUs têm maior flexibilidade na avaliação de suas produções. Logo, $0 \leq \theta_{scale} \leq 1$. Esse escore será tanto maior quanto mais apropriada for a escala de operação da DMU.

Então, quando o cociente entre os dois escores é obtido, é possível avaliar quanto da ineficiência da DMU é devida a sua incapacidade técnica e quanto é devida ao fato de ela não estar produzindo na escala apropriada. A Tabela 8.7 apresenta os escores CCR e BBC das DMUs, bem como os seus níveis de eficiência de escala.

Tabela 8.7 – Escores de eficiência CCR, BCC e de escalas das DMUs apresentadas na Tabela 8.6

DMU	θ_{CCR}	θ_{BCC}	θ_{scale}
A	0,10	1,00	0,10
B	0,15	1,00	0,15
C	0,20	0,33	0,61
D	0,25	0,30	0,83
E	0,98	1,00	0,98
F	0,58	0,59	0,98
G	0,57	0,58	0,98

(continua)

(continuação)

Tabela 8.7 – Escores de eficiência CCR, BCC e de escalas das DMUs apresentadas na Tabela 8.6			
H	0,75	0,75	1,00
I	0,83	0,84	0,99
J	1,00	1,00	1,00

Fonte: Elaboração própria.

Tome o caso da DMU B: quando essa DMU é avaliada com retornos variáveis de escala, ela se mostra 100% eficiente. Portanto, toda a sua ineficiência vem da sua escala inadequada de produção. Já a DMU H apresenta os dois escores, CCR e BCC, idênticos. Isso quer dizer que a avaliação com retornos constantes ou variáveis de escala não faz diferença, nesse caso. Então a sua ineficiência não se deve à escala, podendo ser totalmente atribuída a problemas técnicos de produção. Observe que a eficiência medida pelo modelo BCC passa a ser considerada o parâmetro de eficiência técnica, pois mede a relação entre a produção efetiva¹¹ da DMU e a sua produção potencial dada a escala de produção.

8.8 Painel de dados em DEA

Se a produtividade, em termos de quilômetros rodados por litro de gasolina, de um carro fabricado na década de 1960 for comparada com aquela de um automóvel atual, será possível ter-se uma ideia dessa evolução.

A produtividade das empresas, equipamentos, governos etc. também varia de forma notável com o passar do tempo, e seria interessante desenvolver métodos capazes de aferir tais modificações.

Além disso, foi visto na seção anterior que outros fatores que não a variação da eficiência técnica podem afetar a produtividade. Naquela ocasião, estudou-se a influência da escala da produção sobre a eficiência das DMUs. Assim, se uma DMU em particular mantiver o mesmo nível de eficiência técnica, mas variar a sua escala de operação, sua produtividade total pode ser alterada.

11 Essa é a definição de eficiência pela ótica do produto.

Por exemplo, compare as DMUs B e J da Tabela 8.6. A primeira, apesar de ser tecnicamente eficiente (ver o escore BCC), tem uma produtividade muito mais baixa que a segunda, pois enquanto a DMU B utiliza 2 unidades de cada insumo para produzir 3 unidades de produto, a DMU J utiliza 10 unidades de cada insumo para produzir 100 unidades de produto. Isso significa que na primeira DMU cada unidade gasta dos insumos produz, em média, 1,5 unidades de produto, enquanto na outra DMU cada unidade gasta de insumos gera uma produção de 10 produtos.

A escala não é, contudo, além da eficiência técnica, o fator determinante das variações de produtividade das DMUs. Modificações na tecnologia e no escopo de produção também podem alterar a produtividade, como será visto mais à frente.

Para mensurar essas variações de produtividade, bem como decompô-las em suas possíveis fontes com a metodologia DEA, faz-se necessária a utilização de dados em painel, que são aqueles nos quais se encontram registros de várias DMUs por vários períodos de tempo.

A utilização de dados em painel com DEA difere um pouco da aplicação tradicional em regressões econométricas. Em primeiro lugar, com DEA só se pode utilizar dois períodos por vez. Então, um painel de dados típico para utilização com a metodologia DEA consiste em um conjunto de observações referentes a n DMUs, nos períodos s (anterior) e t (posterior), embora mais períodos possam ser utilizados, dois a dois, em várias rodadas.

Vale ainda observar que, no caso de análise de dados de painel com DEA, a utilização de painéis não balanceados, ou seja, daqueles painéis que não contam com o mesmo número de registros em cada ano, é muito mais flexível. Pode-se inclusive ter um painel no qual nenhuma das empresas existentes no período s apareça no período t .

Para que a utilização do painel DEA possa decompor as fontes de modificação da produtividade das DMUs, é necessário que se possam avaliar os esquemas produtivos de um período em relação à tecnologia existente em outro período.

8.9 Tecnologias alternativas de avaliação

Nos modelos DEA apresentados até aqui, cada DMU que compõe a amostra é avaliada por sua eficiência relativa, mas também é parte integrante da comparação que permite avaliar as outras. Isso quer dizer que cada DMU pode estabelecer a fronteira de eficiência, ou, ainda, que cada DMU faz parte da tecnologia de avaliação.

Pode-se, no entanto, pensar em uma situação um pouco diferente. Alguém pode querer, por exemplo, comparar esquemas produtivos ocorridos no passado com a tecnologia existente atualmente. Para realizar esse experimento, os esquemas produtivos atuais deveriam ser utilizados para demarcar a fronteira de eficiência e algum (ou alguns) esquema produtivo do passado seria avaliado em relação a essa fronteira.

Para definir matematicamente esse experimento, suponha que Q_s , Q_t , X_s e X_t sejam as matrizes de produtos e de insumos nos períodos s e t , respectivamente. Então, o problema de avaliação das DMUs do período s em relação às possibilidades tecnológicas do período t , possibilidades essas delimitadas pelas DMUs do período t , pode ser especificado da seguinte forma:

$$d_p^t(q_{is}, x_{is})^{-1} = \begin{cases} \text{Max: } \theta \\ \quad \{\theta, \bar{\lambda}\} \\ \text{s.t. : } \bar{x}_{is} - X_t \bar{\lambda} \geq \bar{0} \\ \quad Q_t \bar{\lambda} - \theta \bar{q}_{is} \geq \bar{0} \\ \quad \bar{e} \bar{\lambda} = 1 \\ \quad \bar{\lambda} \geq \bar{0} \end{cases} \quad (8.12)$$

Nesse caso, $d_p^t(q_{is}, x_{is})^{-1}$ significa que o esquema produtivo da DMU i , no período s , expresso por (q_{is}, x_{is}) , está sendo avaliado pela ótica do produto,¹² com rendimentos variáveis de escala, e utilizando-se como referência as DMUs do período t . A exponenciação a -1 ocorre pois, como foi visto anteriormente, o resultado do problema pela perspectiva do produto fornecerá um valor maior ou igual a 1, e para que seja compatibilizado com o conceito de eficiência relativa

12 Se fosse pela ótica dos insumos seria d_i^t .

deve ser tomado ao inverso. Nesse exercício, o que se procura é uma combinação linear dos vetores das DMUs do período t e uma expansão radial da DMU do período s que está sendo analisada.

Para obter-se avaliação semelhante, mas utilizando retornos constantes de escala, o problema deveria ser formulado como se segue:¹³

$$d_p^{*t}(q_{is}, x_{is})^{-1} = \begin{cases} \text{Max: } \theta \\ \{\theta, \bar{\lambda}\} \\ \text{s.t.: } \bar{x}_{is} - X_t \bar{\lambda} \geq \bar{0} \\ Q_t \bar{\lambda} - \theta \bar{q}_{is} \geq \bar{0} \\ \bar{\lambda} \geq \bar{0} \end{cases} \quad (8.13)$$

A partir dessa ideia, podem-se realizar várias combinações. Por exemplo, alguém poderia querer calcular $d_p^{*t}(q_{is}, x_{it})$. Essa expressão significa a eficiência, pela ótica do produto, avaliada pela tecnologia (DMUs) do período t , de uma DMU fictícia, a qual utiliza os insumos da DMU i no período t para produzir os produtos dessa mesma DMU referentes ao período anterior (s). Os motivos para se realizar um exercício como esse ficarão mais claros adiante. Esse problema poderia ser matematicamente representado assim:

$$d_p^{*t}(q_{is}, x_{it})^{-1} = \begin{cases} \text{Max: } \theta \\ \{\theta, \bar{\lambda}\} \\ \text{s.t.: } \bar{x}_{it} - X_t \bar{\lambda} \geq \bar{0} \\ Q_t \bar{\lambda} - \theta \bar{q}_{is} \geq \bar{0} \\ \bar{\lambda} \geq \bar{0} \end{cases} \quad (8.14)$$

Na prática, é necessária a manipulação dos dados para que a DMU fictícia seja criada e avaliada contra as DMUs verdadeiras do período t .

8.10 Utilizando o painel de dados para a decomposição da variação da produtividade

Uma vez que as tecnologias alternativas de avaliação foram descritas, é possível avançar na decomposição da variação da produtividade em seus diversos fatores. O primeiro, e mais simples

13 O símbolo d^* será utilizado para denotar retornos constantes de escala daqui para frente.

fator contribuinte para a variação da produtividade, é a *variação da eficiência técnica* das DMUs.

É claro que tudo mais constante, um acréscimo da eficiência técnica (medida no modelo BCC), causará elevação da produtividade da DMU. Então é possível medir-se a variação da produtividade ocasionada pela variação da eficiência técnica da seguinte forma:

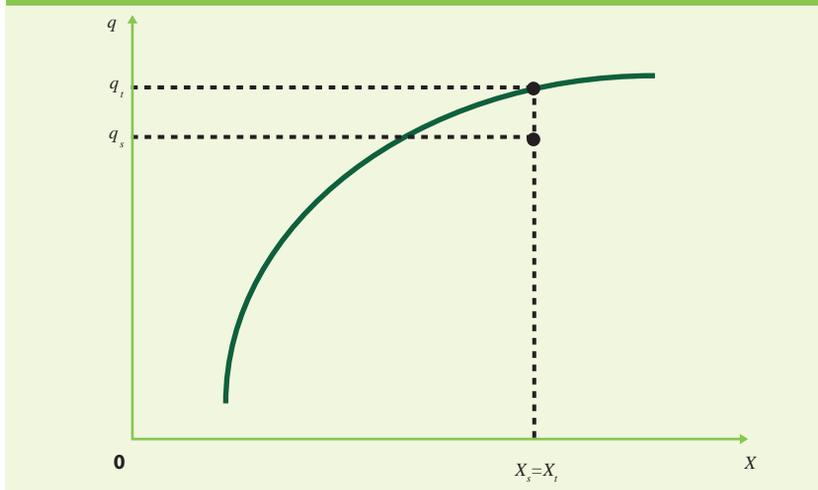
$$met_p^{s,t}(x_s, q_s, x_t, q_t) = \frac{d_p^t(x_t, q_t)}{d_p^s(x_s, q_s)} \quad (8.15)$$

Nesse caso, $met_p^{s,t}$ é o fator multiplicativo da mudança de produtividade causado *exclusivamente* por variações na eficiência técnica. Por exemplo, se a eficiência técnica de determinada DMU no período s fosse 0,60 e no período t se elevasse para 0,72, a produtividade total dessa DMU aumentaria em 20%, pois $met_p^{s,t} = 0,72 / 0,60 = 1,20$.

Para a realização desse cálculo na prática, o que precisa ser feito é o cálculo de dois DEAs: o primeiro, com os dados das DMUs no primeiro período; o segundo, com os registros do segundo período. Depois disso, basta dividir os resultados do segundo pelos resultados do primeiro. Note que cada DMU terá um valor distinto, pois cada uma delas terá sofrido a sua própria modificação em termos de eficiência técnica.

No Gráfico 8.5 a mesma DMU utiliza a mesma quantidade de insumo (x) em dois períodos consecutivos, s e t . A quantidade de produto varia, porém, entre esses dois instantes, pois a DMU produz uma quantidade maior no período t do que no período s . Notadamente a produtividade dessa DMU cresceu: no período s , tal produtividade foi de q_s/x_s , enquanto no período t foi de q_t/x_t , e como $x_s = x_t$ e $q_s < q_t$, então $q_s/x_s < q_t/x_t$.

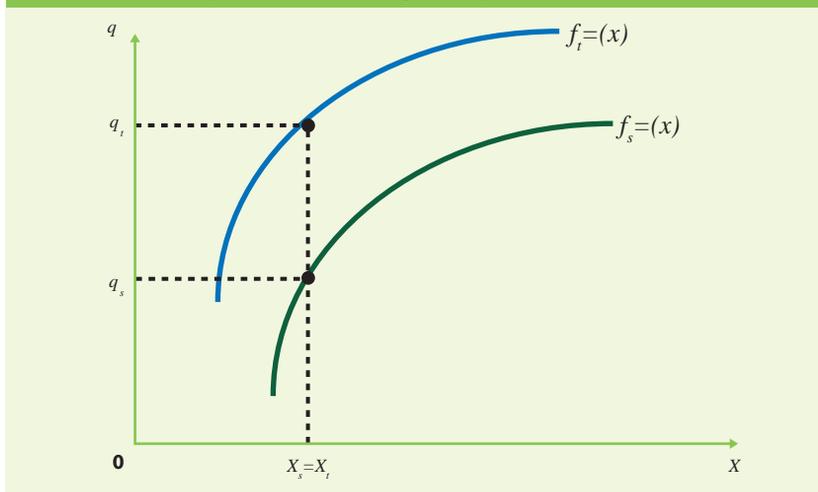
Gráfico 8.5 – Como variações na eficiência técnica podem afetar variações na produtividade



Fonte: Elaboração própria.

No entanto, existem outras possíveis fontes de variação da produtividade. Imagine, por exemplo, que ocorra uma modificação da tecnologia de produção de forma que a produção se torne mais fácil. Nesse caso, mesmo que as DMUs não melhorem sua eficiência técnica ou que não alterem sua escala de produção, elas se tornarão mais produtivas. O Gráfico 8.6 apresenta uma situação como essa, em que a quantidade de insumo também não varia de um período para outro, mas a quantidade de produto cresce. Com isso, naturalmente, a produtividade da DMU em questão também se eleva.

Gráfico 8.6 – Como variações na tecnologia de produção podem afetar variações na produtividade



Fonte: Elaboração própria.

Nessa situação, houve um deslocamento da fronteira de produção ocasionado por algum avanço tecnológico. A fronteira em verde é a que vigorava no período s , enquanto a em azul é aquela que prevalece no período t . Observe que, nos dois períodos, a eficiência técnica da DMU não se altera, é 100%, uma vez que em ambos os instantes ela está com um esquema de produção que se encontra sobre a respectiva fronteira de eficiência.

Para discernir o quanto o avanço da fronteira elevou a produtividade, utiliza-se a seguinte fórmula:

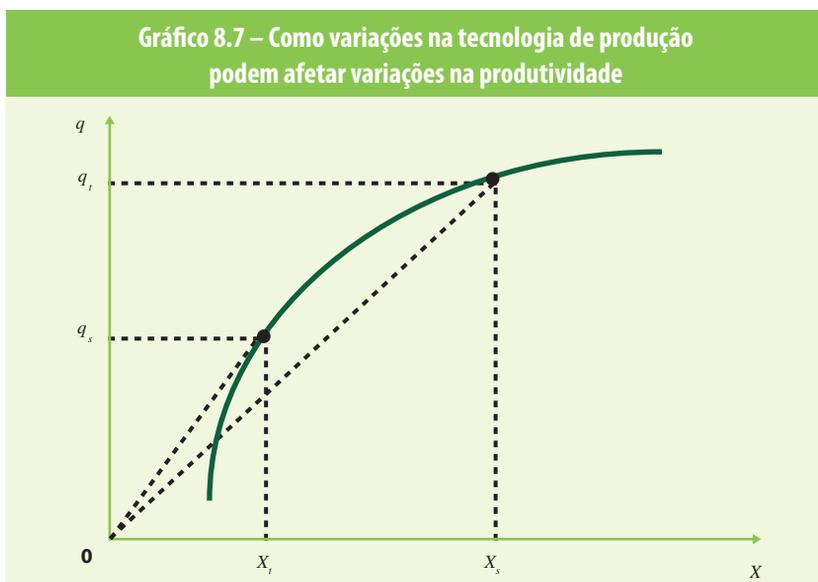
$$mt_p^{s,t}(x_s, q_s, x_t, q_t) = \sqrt{\frac{d_p^s(x_s, q_s) d_p^s(x_t, q_t)}{d_p^t(x_s, q_s) d_p^t(x_t, q_t)}} \quad (8.16)$$

Para o cálculo desse índice, denominado de índice de *Malmquist para modificações na tecnologia*, são necessárias, então, quatro rodadas de DEA diferentes. Duas delas já foram realizadas para o cálculo do *met* na equação 8.15. Faltaria calcular $d_p^t(x_s, q_s)$, que são as eficiências das DMUs do período s avaliadas com a tecnologia do período t e $d_p^s(x_t, q_t)$, que são as eficiências das DMUs do período t avaliadas com a tecnologia do período s .

Outra fonte de variação da produtividade são as variações na escala de produção. Como já foi visto, se uma DMU se aproxima de uma escala mais favorável de produção, sua produtividade tende a aumentar.

Como se pode observar no Gráfico 8.7, quando a DMU ali representada diminui a utilização de insumo do período s (X_s) para o período t (X_t), sua produção também cai de q_s para q_t . Essa queda de produção, porém, é menos proporcional do que a queda na utilização de insumos.

Esse fato pode ser comprovado pelo aumento da inclinação da reta que parte da origem e vai até o ponto que descreve o esquema produtivo da DMU.¹⁴



Fonte: Elaboração própria.

Assim, a mudança na escala de produção é o suficiente para alterar a produtividade da DMU em um contexto de retornos variáveis de escala, sem que haja necessidade de variações nem da eficiência técnica, nem na tecnologia de produção em si.

14 Se o ângulo da reta que parte da origem e chega ao ponto (x_t, q_t) for denominado α , então tem-se que $\tan(\alpha) = q_t / x_t$. Se $\tan(\beta) = q_s / x_s$, e como $\alpha > \beta$, pode-se inferir que $q_t / x_t > q_s / x_s$.

A fórmula para o cálculo do efeito da modificação de escala sobre a produtividade de determinada DMU é um pouco mais trabalhosa, como se pode observar na equação 8.17.

$$mee_p^{s,t}(x_s, q_s, x_t, q_t) = \sqrt{\frac{d_s^*(x_t, q_t) d_s(x_s, q_t) d_t^*(x_t, q_s) d_t(x_s, q_s)}{d_s(x_t, q_t) d_s^*(x_s, q_t) d_t(x_t, q_s) d_t^*(x_s, q_s)}} \quad (8.17)$$

O *mee* e o índice de Malmquist para eficiência de escala e envolve, além do cálculo de DEAs com retornos constantes de escala (CCR), o cálculo de DEAs com DMUs artificialmente criadas, nas quais a produção do período posterior deve ser associada à utilização de insumos do período anterior (x_s, q_t) e vice-versa (x_t, q_s).

O último fator que participa da composição da variação da produtividade é a variação nas economias de escopo. As composições dos produtos ou dos insumos de determinada DMU podem afetar a sua produtividade. Naturalmente, essas composições podem variar de um período para outro. A fórmula para capturar o quanto a produtividade das DMUs variou em função das modificações das economias de escopo ($mee_{esc}_p^{s,t}$) é a seguinte:

$$mee_{esc}_p^{s,t} = \sqrt{\frac{d_s^*(x_s, q_t) d_s(x_s, q_s) d_t^*(x_t, q_t) d_t(x_t, q_s)}{d_s(x_s, q_t) d_s^*(x_s, q_s) d_t(x_t, q_t) d_t^*(x_t, q_s)}} \quad (8.18)$$

As modificações na produtividade total das DMUs podem ser então calculadas como o produto das diversas fontes decompostas nesta seção, ou seja:

$$mpt_p^{s,t} = met_p^{s,t} \times mt_p^{s,t} \times mee_p^{s,t} \times mee_{esc}_p^{s,t} \quad (8.19)$$

Nesse caso, $mpt_p^{s,t}$ é a modificação da produtividade total de cada DMU entre os períodos s e t , com ótica nos produtos. Não é necessário, contudo, calcular todas as fontes de variação da produtividade para obter a variação total da produtividade. Isso porque, quando se toma o produto das equações 8.15, 8.16, 8.17 e 8.18, é possível obter uma expressão relativamente simples para ela, como o exposto na equação 8.19.

$$mpt_p^{s,t}(x_s, q_s, x_t, q_t) = \sqrt{\frac{d_s^*(x_t, q_t) d_t^*(x_t, q_t)}{d_s^*(x_s, q_s) d_t^*(x_s, q_s)}} \quad (8.20)$$

A Tabela 8.8 apresenta os valores de insumos e produtos para dez DMUs hipotéticas, em dois períodos, s (anterior) e t (posterior). Nela também constam tanto a modificação na produtividade total de cada DMU (MPT) quanto a sua decomposição em modificação da eficiência técnica (MEF), modificação da tecnologia (MT), modificação nas economias de escala (MEE) e modificação na economia de escopo ($MEESC$) de cada uma delas. Tais valores foram obtidos a partir do cálculo do painel com DEA, conforme discutido nos parágrafos anteriores. Uma boa forma de fixar os conceitos abordados é tentar reproduzir os resultados desta tabela.

Tabela 8.8 – Decomposição das modificações de produtividade de dez DMUs hipotéticas entre dois períodos

DMU	x_{1s}	x_{2s}	q_{1s}	q_{2s}	x_{1t}	x_{2t}	q_{1t}	q_{2t}	MPT	MET	MT	MEE	MEESC
A	110	270	31	8	111	277	33	9	1,05	0,95	1,11	0,99	1,00
B	88	247	32	18	87	258	30	16	0,94	0,84	1,12	1,00	1,00
C	84	176	27	17	81	182	35	21	1,23	1,00	1,23	1,01	0,99
D	59	232	24	18	59	240	24	17	0,97	1,00	0,96	1,01	1,00
E	139	218	36	18	138	222	40	18	1,06	1,00	1,09	0,99	0,99
F	86	178	27	18	86	185	27	21	1,13	1,00	1,15	0,98	1,00
G	79	211	30	18	81	201	32	18	1,05	0,91	1,16	0,99	1,00
H	99	131	25	13	98	135	27	14	1,05	0,90	1,18	0,99	1,00
I	58	55	17	6	58	55	21	7	1,20	1,00	1,20	1,00	1,00
J	80	235	27	11	79	245	31	14	1,16	1,02	1,14	1,00	1,00

Fonte: Elaboração própria.

Tome o caso da DMU A para fixar o funcionamento dessa decomposição com o painel DEA. Essa DMU utilizou mais insumos e produziu mais produtos no período t que no período s . O valor encontrado para a sua MPT , 1,05, mostra que a sua produtividade total cresceu cerca de 5% entre os dois períodos. Esse crescimento se deveu ao fato de que a tecnologia de produção da DMU sofreu melhorias de 11% ($MT = 1,11$), o que compensou as perdas de eficiência técnica e de economias de escala ocorridas entre os dois períodos, que foram, respectivamente, de 5% e 1% ($MET = 0,95$ e $MEE = 0,99$). No quesito economias de escopo, essa DMU não apresentou ganhos ou perdas em termos de produtividade, uma vez que $MEESC = 1,00$.

É válido ressaltar que as variações de produtividade são diferentes para cada DMU, bem como os fatores que formam essas variações. Isso decorre do fato de que cada DMU está operando com características de escala e de escopo distintas, além, é claro, de possuir níveis diferentes de eficiência técnica.

8.11 Variáveis ambientais e variáveis não discricionárias

A avaliação realizada por meio do modelo DEA tenta incluir todas as variáveis de insumo e de produto das DMUs que, ao mesmo tempo, são relevantes e estão sob o seu controle. No entanto, existem vários casos nos quais variáveis importantes com relação à produtividade da DMU não estão sobre seu controle.

Por exemplo, considere uma avaliação que leve em conta o sistema de saúde municipal como seu objeto de estudo. Poder-se-ia utilizar como insumos o número de médicos e de enfermeiros, o orçamento destinado à saúde, a quantidade de equipamentos hospitalares disponíveis etc. Já como produtos, poder-se-ia utilizar o número de pacientes ambulatoriais atendidos, o número e a duração das internações, e o número de cirurgias realizadas.

Um município seria mais eficiente em seu sistema de saúde se atendesse a mais pacientes utilizando menos recursos. Naturalmente uma análise como essa deixa de fora aspectos importantes do problema. O primeiro é o das variáveis ambientais.

Variáveis ambientais são parâmetros exógenos às DMUs compreendidas na análise. Por exemplo, no caso do sistema municipal de saúde, a renda da população atendida, sua faixa etária e a

densidade populacional afetarão diretamente o número de pacientes atendidos pelo sistema público de saúde, pois pacientes de maior nível de renda tenderão a procurar hospitais privados, na maioria dos casos. No entanto, se a população atendida tem maior proporção de pessoas idosas, a procura pelos hospitais municipais deverá ser maior.

Em todo caso, é possível observar que a quantidade de pacientes atendidos não depende somente da eficiência da rede municipal de saúde de cada município, mas também das condições externas.

Um problema semelhante ocorre quando se encontram presentes no problema as variáveis não discricionárias. Tais variáveis são parecidas com as variáveis ambientais, no sentido de que não estão sob o controle imediato dos gestores, pelo menos não em curto prazo.

Suponha que, ainda estudando a eficiência dos sistemas municipais de saúde, sejam detectadas em alguns municípios compras exageradas de equipamentos hospitalares em gestões anteriores. Todo aquele equipamento, bem como as instalações físicas das unidades de atendimento, não pode ser modificado pelos atuais gestores em curto prazo. No entanto, quando o DEA for aplicado para a avaliação, devem estar presentes como variáveis de insumo, uma vez que ampliam a capacidade de atendimento à população.

Portanto, se a demanda pelo sistema de saúde for reduzida, esses equipamentos poderão ficar ociosos e reduzir a eficiência dos hospitais (equipamentos são insumos), sem que o gestor nada possa fazer.

Nesse caso, tais variáveis são consideradas não discricionárias, e deverão ter um tratamento diferenciado no cálculo do DEA para avaliação, de modo a não se penalizar os gestores atuais por ineficiências geradas, no passado, por outros gestores. O problema de programação linear a ser utilizado nesse caso seria:

$$\left\{ \begin{array}{l} \text{Min} : \theta_i \\ \quad \{ \theta, \lambda \} \\ \text{s.t.} : \theta_i \bar{x}_i^D - X^D \bar{\lambda} \geq \bar{0} \\ \quad \bar{x}_i^{ND} - X^{ND} \bar{\lambda} \geq \bar{0} \\ \quad Q \bar{\lambda} - \bar{q}_i \geq \bar{0} \\ \quad \bar{\lambda} \geq \bar{0} \end{array} \right. \quad (8.21)$$

onde o sobrescrito D refere-se aos insumos discricionários (sob controle dos gestores) e o sobrescrito ND refere-se àqueles insumos não discricionários. Observe que, nessa versão do problema, sob a ótica do insumo com retornos constantes de escala, a contração possível dos insumos só é avaliada com respeito àqueles que são considerados discricionários. Para os insumos não discricionários, somente se requer que eles possam ser representados por combinações lineares dos valores das outras DMUs. Esse método pode ser não só utilizado para variáveis não discricionárias, mas também para variáveis ambientais.

No caso das variáveis ambientais, outras soluções foram propostas para lidar com o problema causado por elas na avaliação. Por exemplo, Banker e Morey (1986) sugerem uma abordagem na qual as DMUs são ordenadas de acordo com o valor da variável ambiental, e cada uma delas é comparada somente com aquelas outras unidades que possuem valores menos vantajosos da variável ambiental.

Um problema dessa abordagem é que a existência de mais de uma variável ambiental pode inviabilizá-la, uma vez que o nível de comparabilidade das DMUs poderia ser reduzido demais. Além disso, a aplicação desse método requer que os efeitos das variáveis ambientais sejam julgados *a priori* como positivos ou negativos, decisão que nem sempre é possível.

Em situações nas quais a amostra pode ser dividida em subamostras de acordo com alguma variável categórica (e.g. administração pública e privada), Charnes, Cooper e Rhodes (1981) propõem o seguinte procedimento para análise comparativa de eficiência segundo essa variável, em particular:

1. Divida a amostra nas diversas subamostras compatíveis e rode um DEA para cada subamostra;
2. Projete cada ponto na fronteira obtida para a respectiva subamostra;
3. Rode mais um DEA, desta vez com as projeções (misturadas), e calcule as médias de eficiências de cada categoria.

Esse procedimento permitirá a comparação da eficiência teórica de cada categoria e não apenas do desempenho real das DMUs existentes. O que é comparado nesse caso é o desempenho de

cada categoria se as DMUs pertencentes a cada uma delas estivesse operando com eficiência máxima. Esse método é muito utilizado para a comparação de empresas públicas com empresas privadas.

Outro método utilizado para a incorporação de variáveis ambientais no modelo DEA é a sua utilização direta como insumos e produtos. Nesse caso, as variáveis ambientais com influência positiva na produção devem ser incluídas como *inputs* não discricionários no modelo expresso na equação 8.21 e as variáveis com influência negativa sobre a produtividade deverão ser incluídas como *outputs* não discricionários.

A lógica desse processo é: se uma variável ambiental dá vantagens produtivas às DMUs, tais vantagens aumentam os produtos potenciais das DMUs com valores maiores das variáveis ambientais consideradas. Assim, essas vantagens equivalem a aumentos na utilização de insumos. Já as variáveis ambientais que apresentam influência negativa na produtividade reduzem os produtos potenciais das DMUs, e considerá-las como produtos não discricionários fornece uma compensação àquelas DMUs nas quais as variáveis ambientais negativas são mais significantes.

8.12 Modelo de dois estágios

Quando se estuda a literatura recente sobre a inclusão de variáveis ambientais em modelos de avaliação de eficiência, a forma mais encontrada para lidar com o problema é a utilização de métodos nos quais a análise DEA é combinada com a regressão linear. Nesse método chamado de semiparamétrico, ou método de dois estágios, primeiro os coeficientes de eficiência são calculados utilizando-se o DEA e, depois, esses escores são utilizados como variáveis endógenas em uma regressão, na qual as variáveis explicativas são as ambientais e não discricionárias.

Dessa forma, o método de dois estágios procura estabelecer a influência das variáveis ambientais sobre a eficiência das DMUs, permitindo assim, ao avaliador, filtrar tais influências com o objetivo de obter um escore de eficiência verdadeiro.

A versão mais simples do método poderia ser descrita, então, da seguinte forma:

1. Rode o DEA só com as variáveis (*inputs* e *outputs*) controláveis e obtenha os escores de eficiência (θ_j);

2. Regrida os escores de eficiência obtidos contra as variáveis ambientais:

$$\theta_i = z_i \beta + \mu_i \quad (8.22)$$

onde z_i é o vetor de variáveis ambientais associado a DMU i , β é o vetor de coeficientes a ser estimado e μ_i é o erro aleatório também associado a DMU i . Obtenha estimativas para β e μ_i :

3. Ajuste os escores de eficiência utilizando o vetor médio de variáveis ambientais (z_m):

$$\varphi_i = z_m b + e_i \quad (8.23)$$

onde b é a estimativa para o vetor de coeficientes β , e e é a estimativa para o erro idiossincrático μ , ambos encontrados no passo 2.

A ideia por trás do método é a de que, se as variáveis ambientais estão tendo peso sobre a eficiência relativa das DMUs, deve-se primeiro estimar esse grau de influência e depois compará-las, de forma equilibrada, tomando-se a média para isso, de modo que as vantagens e desvantagens exógenas oriundas das variáveis ambientais possam ser consideradas.

Repare que esse método, a princípio, pode ser utilizado com variáveis ambientais e não discricionárias, com variáveis contínuas ou categóricas, e mesmo que não se tenha ideia *a priori* sobre o efeito da variável em questão sobre a eficiência.

Existem, no entanto, vários problemas atrelados a esse tipo de análise, os quais necessitam ser levados em conta. O primeiro deles é que não há garantias de que os φ obtidos no passo 3 estarão entre 0 e 1, como a teoria requer.

Para evitar a obtenção de escores de eficiência ajustados maiores do que 1 ou menores do que 0, outras especificações funcionais podem ser utilizadas na regressão do passo 2, tais como a regressão logística, os modelos *Tobit* ou modelos de regressão censurada, uma vez que essas especificações asseguram que os valores projetados pelo modelo estarão dentro de intervalos predeterminados.¹⁵

15 Para detalhes sobre esses modelos, ver Greene (2011), capítulo 19.

Outro cuidado a ser tomado no emprego do método de dois estágios é o teste sobre a correlação entre as variáveis de insumo e produto utilizadas no primeiro estágio do processo e as variáveis ambientais. Se a correlação for alta, os estimadores do segundo estágio serão viesados.

O problema mais sério para a aplicação do método semiparamétrico vem do fato de que os escores de eficiência estimados no primeiro estágio serão correlacionados entre si, uma vez que são combinações lineares uns dos outros. A correlação que se origina desse processo é intrincada, sendo difícil estabelecer uma estrutura adequada para sua estimação.

Com o objetivo de sobrepujar esse problema, Simar e Wilson (2007) propuseram uma abordagem alternativa para o método de dois estágios, a qual envolve uma série de regressões utilizando-se *bootstrap* para os parâmetros de segundo estágio.

O objetivo é tentar, por meio de várias regressões simuladas, separar o efeito das variáveis ambientais sobre os escores de eficiência daquele oriundo da correlação entre os próprios escores.

Dentre os vários algoritmos sugeridos, o mais simples deles pode ser descrito assim:

1. Utilizando o DEA só com as variáveis (*inputs* e *outputs*) controláveis, obtenha os escores de eficiência (θ);
2. Retire as DMUs que se mostrarem eficientes ($\theta_i = 1$) no passo 1;
3. Obtenha os escores inversos fazendo: $\eta_i = 1/\theta_i$;
4. Utilize o modelo Tobit (ou outro modelo equivalente) para regredir os escores de eficiência obtidos contra as variáveis ambientais, mas utilizando somente os registros das DMUs não eficientes, de acordo com a seguinte equação:

$$\eta_i = z_i \gamma + \varepsilon_i \quad (8.24)$$

onde z_i é o vetor de variáveis ambientais associado a DMU i , γ é o vetor de coeficientes a ser estimado e ε_i é o erro aleatório também associado a DMU i . Obtenha estimativas para γ , ε e para o desvio padrão de ε (s_ε);

5. Produza resíduos artificiais (e_i) gerados a partir de uma distribuição normal truncada, com truncamento à esquerda¹⁶ em $(1 - z_i \gamma)$ e com desvio padrão igual a s_{e_i} , que foi estimado no passo 4¹⁷;

6. Compute a variável η_i^* segundo a seguinte equação:

$$\eta_i^* = z_i g + e_i \quad (8.25)$$

onde g é a estimativa para o parâmetro γ conforme a equação (8.24);

7. Estime mais uma vez utilizando o modelo Tobit e as DMUs citadas anteriormente não eficientes, só que agora utilizando os valores obtidos da equação 8.25 como variável endógena, e as variáveis exógenas como variáveis explicativas:

$$\eta_i^* = z_i \gamma^* + \omega_i \quad (8.26)$$

onde, agora, ω_i é o erro aleatório.

8. Obtenha as estimativas para γ^* e para o desvio padrão do erro ω (g^* e s_{ω});

9. Repita os passos 5, 6, 7 e 8 L ¹⁸ vezes, de modo a obter a matriz:

$$G = \left[g_{\ell}^*, s_{\omega, \ell} \right]_{\ell=1}^L \quad (8.27)$$

10. Calcule as médias e variâncias de cada coluna G para construir intervalos de confiança para os parâmetros;

11. Calcule a média de cada coluna para obter a estimativa dos efeitos das variáveis ambientais sobre a eficiência das DMUs.

No final desse processo, o efeito da correlação entre as variáveis endógenas sobre os estimadores será minimizado.

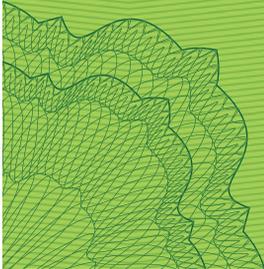
16 Pois se $e_i < 1 - z_i \gamma \Rightarrow z_i \gamma + e_i < 1 \Rightarrow \eta_i < 1$, fato que contrariaria a teoria.

17 No capítulo 11 são detalhados os procedimentos para a produção desses resíduos com a distribuição requerida.

18 Em geral, utiliza-se $L = 1000$.

REFERÊNCIAS

- BANKER, R. D.; CHARNES, A.; COOPER, W. W. Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science*, v. 30, n. 9, p. 1078-1092, 1984.
- BANKER, R. D.; MOREY, R. C. The use of categorical variables in data envelopment analysis. *Management Science*, v. 32, n. 12, p. 1613-1627, dez. 1986.
- BOUERI, R. An evaluation of the efficiency of brazilian municipalities in the provision of public goods using data envelopment analysis. In: BOUERI, R.; SABOYA, M. (Ed.). *Aspects of the fiscal development*. Brasília: IPEA, 2007. p. 139-149.
- CHARNES, A.; COOPER, W. W.; RHODES, E. Measuring the efficiency of decision makers units. *European Journal of Operational Research*, v. 2, p. 429- 444, 1978.
- _____. Evaluating program and managerial efficiency: an application of data envelopment analysis to program follow through. *Management Science*, v. 27, n. 6, p. 668-697, 1981.
- COELLI, T. J.; PRASADA RAO, D. S.; O'DONNELL, C. J.; BATTESE, G. E. *An introduction to efficiency and productivity analysis*. 2. ed. Nova York: Springer, 2005.
- COOPER, W. W.; SEIFORD, L. M.; TONE, K. *Introduction to data envelopment analysis and its uses*. New York: Springer, 2006.
- FARREL, M. J. The measurement of productive efficiency. *Journal of the Royal Statistical Society, Series A (General)*, v. 120, n. 3, p. 253-290, 1957.
- GRASS, S. I. *Linear programming: methods and applications*. 5. ed. New York: Dover Publications, 2003.
- GREENE, W. H. *Econometric Analysis*. 7. ed. New York: Prentice Hall, 2011.
- SIMAR, L.; WILSON, P. W. Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of Econometrics*, v. 136, p. 31-64, 2007.
- TECLES, P. L.; TABAK, B. M. Determinants of bank efficiency: the case of Brazil. *European Journal of Operational Research*, v. 207, p. 1587-1598, 2010.



Capítulo 9

Modelos paramétricos: *Fronteira Estocástica*

Rogério Boueri

9.1 Introdução

O objetivo deste capítulo é apresentar a metodologia de Fronteira Estocástica (FE) como outra ferramenta para a mensuração da eficiência relativa de unidades produtoras. Também serão estabelecidas as diferenças, vantagens e desvantagens desse tipo de abordagem em relação à Análise Envoltória de Dados (DEA¹), de modo a habilitar o leitor a escolher a metodologia mais adequada ao problema que se tenha em mãos.

Conceitualmente, a principal diferenciação entre as duas metodologias é que, no caso da DEA, o modelo de cálculo da eficiência relativa é não paramétrico, enquanto a FE é paramétrica. Isso quer dizer que no caso da DEA não é necessário estabelecer forma funcional que descreva a fronteira de eficiência. Já no caso da fronteira estocástica tal especificação é fundamental.

O caráter estocástico da FE também costumava ser citado como um ponto de diferenciação entre as duas metodologias, mas essa distinção vem desaparecendo à medida que processos estocásticos têm sido incorporados à análise DEA.

No que diz respeito à FE, essa metodologia pode ser utilizada de diversas maneiras para a estimação da eficiência relativa de unidades produtivas. Por exemplo, se houver diversos insumos envolvidos na produção, mas apenas um produto, a FE pode ser utilizada para a estimação da fronteira de produção. No caso contrário, em que se dispõe de apenas um insumo, em geral financeiro, mas vários produtos são produzidos, a FE estimará uma fronteira eficiente de custos.

1 Análise Envoltória de Dados é a tradução literal do termo em inglês *Data Envelopment Analysis* (DEA).

Há ainda o caso em que o esquema produtivo utiliza diversos insumos para a produção de vários produtos. Nessa situação, o modelo FE com a utilização da função translogarítmica² reversa permitirá a estimação da eficiência relativa das unidades por meio do cálculo da função distância, como será estudado na seção 9.8 deste capítulo.

Além disso, a metodologia FE também permite a utilização de painel de dados para que se possa estimar a evolução temporal da eficiência e da produtividade das unidades comparadas.

Uma das grandes vantagens da FE em relação à DEA está na incorporação de variáveis ambientais ao modelo de análise. Por se tratar de uma regressão, a metodologia FE permite a introdução de tais variáveis de maneira bem mais direta e natural do que os procedimentos necessários para a sua utilização na DEA.

Curiosamente, a principal desvantagem da FE também advém do fato de esse método ser uma regressão. Isso porque a implementação do método, a interpretação dos resultados e os testes de significância necessários podem complicar a utilização da FE. Em suma, a utilização e interpretação da DEA são bem mais simples.

Este capítulo se desenvolverá dentro da seguinte estrutura: na seção 9.2 será discutido o modelo com dados cross-section da FE, ou seja, aquele modelo no qual existem dados para diversas unidades produtoras, mas para apenas um período. O processo de estimação será apresentado na seção 9.3, bem como as limitações de tal modelo.

A seguir, na seção 9.4 será apresentada a adaptação do modelo para a estimação da função custo. A partir da seção 9.5, as técnicas de painel de dados serão discutidas. Primeiramente, será apresentado o modelo com eficiência temporalmente fixa, no qual se supõe que a eficiência de cada unidade permaneça constante em cada período de análise, embora varie entre as unidades.

Na seção 9.6 será introduzido o modelo com variação temporal da eficiência. Nesse caso, além de variar de unidade para unidade, permite-se que a eficiência também evolua com o passar dos períodos. A discussão sobre variáveis ambientais se dará na seção 9.7.

A seção seguinte discutirá a utilização da função translog, em

2 Usualmente denominada de função translog.

uma estrutura de painel de dados, para a estimação da eficiência em um contexto produtivo de múltiplos insumos e produtos. A seção 9.9 fecha o capítulo com a decomposição dos fatores influentes na evolução da produtividade das unidades produtivas, mais uma vez fazendo uso da translog e de dados em painel.

9.2 Modelo teórico

Suponha um processo produtivo descrito pela seguinte função de produção:

$$q_i = f(x_{1,i}, \dots, x_{m,i}, \beta_1, \dots, \beta_n) \quad (9.1)$$

no qual q é a quantidade produzida do único produto pela unidade i , cada x é a quantidade empregada de um insumo por essa unidade, $\{\beta_1, \dots, \beta_n\}$ é um conjunto de parâmetros que define a tecnologia produtiva e f é a função de produção. Observe que tanto os parâmetros definidores da tecnologia quanto a função de produção são comuns a todas as unidades produtoras.

Adicionalmente, suponha que cada unidade produtora possa apresentar seu próprio nível de eficiência técnica, o qual deve estar entre zero e um ($0 \leq TE_i \leq 1$). Se $TE_i = 1$, a unidade i é dita ser 100% eficiente. Caso $TE_i = 0$, a unidade i tem eficiência nula, isto é, está utilizando insumos e não produz nenhuma unidade de produto. O mais comum, no entanto é que a eficiência técnica da maioria das unidades esteja entre zero e um, denotando certo grau de ineficiência. Pode-se então deduzir que a produção de cada unidade é dada por:

$$q_i = f(x_{1,i}, \dots, x_{m,i}, \beta_1, \dots, \beta_n) \times TE_i \quad (9.2)$$

Denominado *fronteira determinística*, esse modelo afirma que a produção efetiva de cada unidade depende da tecnologia produtiva disponível e do nível de eficiência da unidade. Observe que a quantidade efetivamente produzida será sempre menor ou igual ao montante que é possibilitado pela tecnologia, uma vez que TE_i é sempre menor que 1.

Para se obter o *modelo estocástico* a partir da equação 9.2, faz-se necessária a adição de um componente aleatório à função de produção, assim:

$$q_i = f(x_{1,i}, \dots, x_{m,i}, \beta_1, \dots, \beta_n) \times TE_i \times e^{v_i} \quad (9.3)$$

Nesse modelo, a variável aleatória v_i representa o choque aleatório associado à unidade produtiva i . Aqui, a quantidade efetivamente produzida depende da tecnologia e do nível de eficiência da unidade produtora, como no caso da fronteira determinística, mas também de uma variável randômica.

Então, a eficiência técnica de uma unidade pode ser definida como:

$$TE_i = \frac{q_i}{f(x_{1,i}, \dots, x_{m,i}, \beta_1, \dots, \beta_n) \times e^{v_i}} \quad (9.4)$$

Na expressão 9.4, o numerador expressa quanto a unidade efetivamente produziu, enquanto o denominador mostra o montante máximo de produção que poderia ser atingido por essa unidade, dados: (i) a tecnologia disponível (f e β 's); (ii) a utilização de insumos pela unidade ($x_{j,i}$); e (iii) o choque recebido pela unidade (v_i). Note que o choque aleatório pode ser positivo ou negativo. No primeiro caso, ele amplia a capacidade produtiva da unidade e, no segundo, ele a diminui.

Para facilitar a discussão, suponha que a função de produção seja do tipo Cobb-Douglas. Então, a equação 9.3 poderia ser expressa assim:

$$q_i = \beta_0 x_{1,i}^{\beta_1} x_{2,i}^{\beta_2} \dots x_{m,i}^{\beta_m} \times TE_i \times e^{v_i} \quad (9.5)$$

Como foi visto, a eficiência técnica da unidade, TE_i está entre zero e um. Pode-se então modelá-la conforme uma função exponencial da seguinte forma, na qual o componente de eficiência $\mu_i \geq 0$:

$$TE_i = e^{-\mu_i} \quad (9.6)$$

Assim, quando $\mu_i \rightarrow 0$, $TE_i \rightarrow 1$, e quando $\mu_i \rightarrow \infty$, $TE_i \rightarrow 0$. Portanto, fica assegurado, com essa modelagem, que $TE_i \in [0,1]$. Quando são tomados os logaritmos da equação 9.6 – e já se levando em consideração a modelagem da eficiência técnica –, obtém-se a seguinte equação:

$$\ln q_i = \beta_0 + \beta_1 \ln x_{1,i} + \beta_2 \ln x_{2,i} + \dots + \beta_m \ln x_{m,i} + v_i - \mu_i \quad (9.7)$$

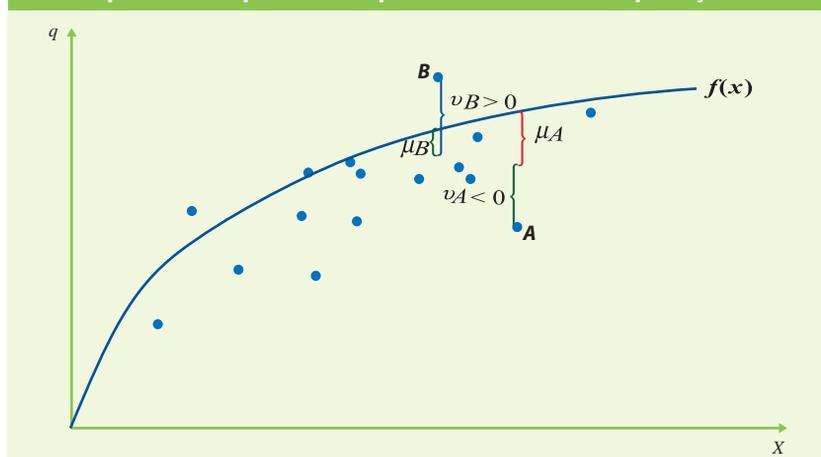
Em geral, para que se possa estimar o modelo, supõe-se que o choque aleatório v_i e o componente de eficiência μ_i são não correlacionados, ou seja, $E[v_i \mu_i] = 0$. Essa suposição quer dizer que

o choque aleatório recebido por cada unidade não tem relação com o seu nível de eficiência. Assim, as melhores unidades em termos de eficiência teriam a mesma probabilidade de receber um choque positivo do que qualquer uma das outras.

No Gráfico 9.1 pode-se observar o funcionamento do processo gerador descrito na equação 9.7 para o caso de um insumo (x), representado no eixo horizontal, e um produto (q), no eixo vertical. A linha azul representa a função de produção.

Como seria de se esperar, os pontos representativos das unidades produtoras não se encontram exatamente sobre a fronteira de produção, dado que cada uma delas sofre choques e tem um determinado nível de eficiência. Tome, por exemplo, o caso da unidade **A**. Ela se encontra bem aquém da fronteira por dois motivos: em primeiro lugar, essa unidade apresenta um nível de ineficiência representado por μ_A , que já provoca uma produção menor que a ideal; e essa unidade também sofreu um choque negativo, v_A , que reduz ainda mais a sua produção.

Gráfico 9.1 – Combinações de insumos utilizados e produtos gerados por unidades produtoras hipotéticas e a fronteira de produção



Fonte: Elaboração própria.

Note que, apesar de a maioria dos pontos se encontrar abaixo da fronteira de produção, existem alguns que estão acima dela. É o caso, por exemplo, da unidade **B**. Isso ocorre porque, apesar de essa

unidade não ser totalmente eficiente, o que pode ser constatado pelo fato de que $\mu_B \neq 0$, ela foi favorecida por um choque positivo bastante extenso (v_B). Aqui a magnitude do choque recebido pela unidade **B** mais do que compensou a sua limitação de eficiência.

9.3 Estimação da Fronteira Estocástica – fronteira de produção

O problema da estimação da FE se resume a dois principais objetivos interdependentes: (i) estabelecer a fronteira, e (ii) diferenciar o choque aleatório (v) do componente de eficiência (μ) em cada unidade produtiva.

Dois fatos curiosos podem ser citados no que diz respeito ao desenvolvimento da técnica de estimação da FE. Em primeiro lugar, essa técnica foi desenvolvida de maneira simultânea e independente por dois grupos de pesquisadores. Enquanto Aigner, Lovell e Schimdt (1977) trabalhavam nos Estados Unidos, Meeusen e van der Broeck (1977) se debruçavam sobre o assunto na Europa. Os artigos produzidos pelos dois grupos de pesquisadores foram bastante similares e apareceram com pouco mais de três meses de diferença.

A princípio se pode tentar estimar o modelo utilizando-se mínimos quadrados ordinários, mas nesse caso o modelo MQO proverá estimador viesado para o intercepto. Quando esse modelo é estimado por MQO, a equação de estimação seria:

$$\ln q_i = \beta_{MQO} + \beta_1 \ln x_{1,i} + \beta_2 \ln x_{2,i} + \dots + \beta_m \ln x_{m,i} + \varepsilon_i \quad (9.8)$$

Como ε_i tem que ter média zero e $(-\mu_i)$ é sempre negativo, a estimação proferiria:

1. $\varepsilon_i = v_i + \mu_i - E[\mu_i]$
2. $\beta_{MQO} = \beta_0 - E[\mu_i]$.

Uma vez que $-E[\mu_i] < 0$, β_{MQO} sempre será menor que o verdadeiro intercepto, β_0 , o modelo MQO sempre diminuirá a estimativa do intercepto, o que equivale a aumentar a eficiência de todas as firmas.

O estimador de MQO pode ser utilizado, no entanto, para um teste prévio sobre a relevância de aplicação do modelo FE sobre os dados que se tem à mão. Dentre os diversos testes recomendados, o de Coelli (1995) revela-se o de mais fácil aplicação.

Esse teste consiste em realizar a estimação MQO com os dados e obter os resíduos estimados. De posse desses resíduos, são calculados seus segundo e terceiro momentos (m_2 e m_3), respectivamente. Calcula-se então a seguinte estatística:

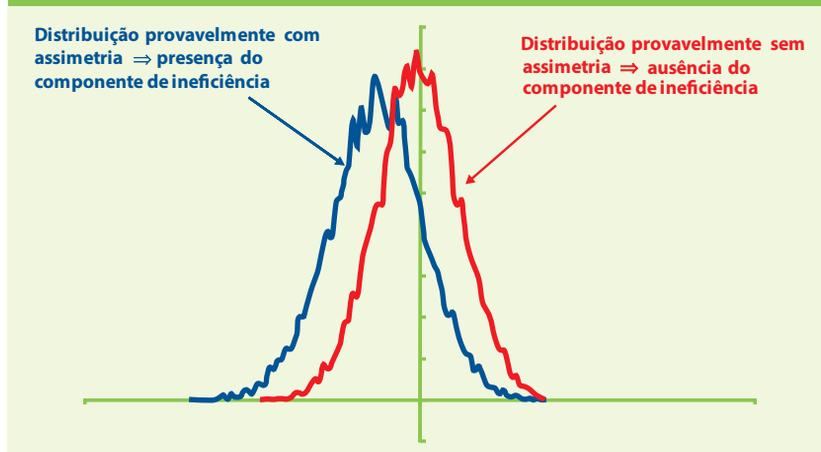
$$C_{95} = \frac{m_3 n^{1/2}}{(6m_2^3)^{1/2}} \quad (9.9)$$

Coelli (1995) mostra que a estatística C_{95} tem distribuição assintótica Normal Padrão, o que facilita bastante a aplicação. Nesse teste, as hipóteses são:

$$\begin{cases} H_0 : C_{95} < 0 \\ H_1 : C_{95} \geq 0 \end{cases}$$

A lógica do teste é a seguinte: como o terceiro momento da distribuição (m_3) representa a sua assimetria, caso os resíduos sejam simétricos tem-se que $m_3 = 0$. Nesse caso, C_{95} também seria igual a zero e a hipótese alternativa seria aceita. Se isso ocorrer, está-se aceitando que toda a variação de produção em relação à fronteira seria resultado apenas dos choques aleatórios, não havendo assim nenhum componente de eficiência nos dados. Nesse caso, o estimador de MQO seria não viesado, inclusive para o intercepto.

Gráfico 9.2 – Histogramas hipotéticos provenientes de distribuições com e sem assimetria



Se, no entanto, a hipótese nula for aceita, isto é, se houver indícios de que há diferenças no nível de eficiência das diversas unidades, far-se-á necessária a estimação dos parâmetros por máxima verossimilhança. Nesse caso, suposições sobre as distribuições estatísticas tanto do erro aleatório (v) quanto do componente de eficiência (μ) deverão ser formuladas.

As distribuições mais comumente utilizadas são a Normal para v e a Seminormal³ para μ . Existem, contudo, diversas outras possibilidades de distribuição para v e μ , como os modelos Normal-Exponencial e Normal-Gama. Na prática, porém, os resultados produzidos pelos diferentes modelos tendem a ser bastante aproximados.

Supondo então o erro aleatório Normal (i) e o componente de eficiência com distribuição Seminormal (ii), outras hipóteses são necessárias para que se possa proceder à estimação por máxima verossimilhança:

$$\left\{ \begin{array}{l} (i) \quad v_i \sim iid N(0, \sigma_v^2); \\ (ii) \quad \mu_i \sim iid N^+(0, \sigma_\mu^2); \\ (iii) \quad E[v_i \mu_i] = 0; \\ (iv) \quad E[v_i x_{m,i}] = 0, \forall m; \\ (v) \quad E[\mu_i x_{m,i}] = 0, \forall m \end{array} \right.$$

A hipótese (iii) requer que não haja correlação entre o erro aleatório e o componente de eficiência em cada unidade. Já a hipótese (iv) impõe que não haja relação entre o choque aleatório e o montante de insumos utilizados.

A hipótese (v) é mais delicada, pois se os produtores têm noção sobre o próprio nível de eficiência, sua escolha de insumos pode ser influenciada e, portanto, ter relação com o componente de eficiência. Por exemplo, o fato de um gerente saber que sua fábrica opera com um nível de eficiência muito baixo pode induzi-lo a utilizar insumos mais baratos.

3 Uma distribuição Seminormal é definida como se segue: se x for uma variável aleatória com distribuição Normal Padrão, então a variável $z = |x|$ terá distribuição Seminormal. Observe que z sempre apresentará valores positivos, sendo zero o seu mínimo.

A estimação por máxima verossimilhança fornece então uma série de parâmetros,⁴ dentre eles os estimadores para os β 's (inclusive para o intercepto). Outros estimadores importantes obtidos no processo são:

$$\hat{\sigma}^2 = \hat{\sigma}_\mu^2 + \hat{\sigma}_v^2 \quad (9.10)$$

$$\hat{\gamma} = \frac{\hat{\sigma}_\mu}{\hat{\sigma}_v} \quad (9.11)$$

Lembrando que $\hat{\sigma}_\mu$ e $\hat{\sigma}_v$ são, respectivamente os estimadores para os desvios padrão do componente de eficiência e do choque aleatório.

Observe também que, se $\hat{\gamma} \rightarrow 0$, então $\hat{\sigma}_\mu \rightarrow 0$. De qualquer maneira, $\hat{\gamma} \rightarrow 0$ significa que a variância do erro aleatório predomina sobre a variância do componente de eficiência. Então, o primeiro monopolizará o poder de explicação do modelo. Assim, quando $\hat{\gamma} \rightarrow 0$, o modelo poderá ser tratado como aquele no qual o componente de eficiência não é importante, podendo ser estimado por MQO.

Por outro lado, se $\hat{\gamma} \rightarrow \infty$, então $\hat{\sigma}_v \rightarrow 0$. Nesse caso, o choque aleatório é dominado pelo componente de ineficiência e o modelo a ser estimado é o da fronteira determinística, expresso na equação 9.2.

De posse das estimativas de $\hat{\sigma}^2$ e de $\hat{\gamma}$ e utilizando as equações 9.10 e 9.11 obtêm-se as estimativas para $\hat{\sigma}_\mu$ e $\hat{\sigma}_v$. Tais estimativas são determinantes para se ponderar a importância do efeito aleatório e do componente de eficiência sobre a produção, e também permitem calcular o valor esperado do componente de eficiência, dado por:

$$E[\mu] = \hat{\sigma}_\mu \sqrt{\frac{2}{\pi}} \quad (9.12)$$

4 Detalhes do processo de estimação podem ser obtidos em Kumbhakar e Knox Lovell (2000), p. 74 a 78.

Resta ainda obter as estimativas das eficiências individuais, o que pode ser realizado com a aplicação da seguinte fórmula:

$$\hat{\mu}_i = \frac{\hat{\sigma}_\mu \hat{\sigma}_v}{\hat{\sigma}} \left[\frac{\varphi\left(\frac{\hat{\lambda}\hat{\varepsilon}_i}{\hat{\sigma}}\right)}{1 - \Phi\left(\frac{\hat{\lambda}\hat{\varepsilon}_i}{\hat{\sigma}}\right)} - \frac{\hat{\lambda}\hat{\varepsilon}_i}{\hat{\sigma}} \right] \quad (9.13)$$

Nessa fórmula, φ e Φ são as funções de densidade e de probabilidade Normal Padrão, respectivamente.⁵ Pode-se então recuperar o nível de eficiência efetivo por meio de:

$$TE_i = e^{-\hat{\mu}_i} \quad (9.14)$$

Embora essas estimativas não sejam consistentes no sentido estatístico, elas são as melhores que podem ser obtidas caso se disponha somente de dados cross-section. Na seção 9.5 serão apresentadas técnicas para estimação com dados em painel, nas quais os estimadores obtidos são muito mais flexíveis e robustos.

9.4 Estimação da Fronteira Estocástica – fronteira de custo

Em muitas ocasiões, vários produtos são gerados a partir de um único insumo financeiro. Assim, por exemplo, municípios podem prover serviços de educação, de saúde, de urbanismo e outros utilizando como insumo as suas verbas orçamentárias, ou hospitais podem prover internações, exames, cirurgias e serviços ambulatoriais, e para tanto despendem recursos financeiros.

Nesses casos utiliza-se a estimação da fronteira de custos para avaliar a eficiência comparativa de cada unidade. Isto porque o que interessa é comparar o total despendido com os produtos gerados, a fim de se obter estimativas de quanto poderia ter sido economizado. Observe que esse enfoque é inverso ao da fronteira da produção, pois naquele caso busca-se estimar quanto a produção poderia ser expandida dada a utilização de insumos por parte das unidades produtivas.

⁵ Os detalhes práticos da implementação dessas fórmulas no R serão vistos no capítulo 10.

A função custo aqui estudada é uma versão simplificada da formulação mais usual, em que o custo é uma função dos preços dos insumos⁶ e das quantidades geradas de produtos. Na formulação abordada neste capítulo, os preços dos insumos serão considerados uniformes entre as unidades produtoras, podendo então ser parametrizados. Nesse caso, o custo mínimo em que as unidades produtoras poderão incorrer depende da quantidade produzida de cada item. Pode-se escrever então que:

$$C = C(q_1, q_2, \dots, q_s) \quad (9.15)$$

Naturalmente, o custo mínimo geralmente é superado, pois além de alguma ineficiência associada às unidades produtoras, elas estão sujeitas a choques aleatórios que podem ampliar ou reduzir seus custos efetivos (CE). A derivação teórica da ineficiência é muito similar à utilizada para a fronteira de produção:

$$TE_i = \frac{C(q_{1,i}, q_{2,i}, \dots, q_{s,i}) \times e^{v_i}}{CE_i} \quad (9.16)$$

O numerador da expressão 9.16 representa o quanto a unidade produtora i poderia ter gastado, dado que produziu o vetor $\{q_{1,i}, q_{2,i}, \dots, q_{s,i}\}$ de produtos e recebeu o choque aleatório v . Já o denominador denota quanto essa unidade realmente gastou. Quanto menor essa fração, maior será a diferença entre o que a unidade gastou e o quanto poderia ter gastado.

Em termos econométricos, é necessário escolher uma forma funcional para que os custos possam ser estimados e comparados. A mais indicada é a função translog, a ser discutida na seção 9.7. Ela permite que eventuais interações dos diversos produtos possam ser vislumbradas na estimação. Contudo, para simplificar, a discussão dessa seção será focada na função Cobb-Douglas, que permite uma exposição mais clara do tópico. Nesse caso:

$$\ln CE_i = \alpha_0 + \alpha_1 \ln q_{1,i} + \alpha_2 \ln q_{2,i} + \dots + \alpha_s \ln q_{s,i} + v_i + \mu_i \quad (9.17)$$

6 Uma versão da função custo mais completa pode ser utilizada para o cálculo da eficiência alocativa das unidades. Ver Kumbhakar e Knox Lovell (2000).

Note que nessa formulação o componente de eficiência entra com sinal positivo, uma vez que quanto maior o seu valor, menor será a eficiência⁷ e maior será o custo efetivo da unidade estudada. O restante dos procedimentos de estimação e de interpretação dos resultados é igual aos desenvolvidos na seção 9.3.

9.5 Dados em painel: eficiência temporalmente constante

A utilização de dados em painel traz uma série de vantagens à estimação da FE. Em primeiro lugar, quando se dispõe de dados em painel não é necessário fazer hipóteses sobre as distribuições do componente de eficiência. Além disso, a hipótese de que o componente de eficiência e os regressores sejam não correlacionados pode ser relaxada. Por fim, diferentemente da estimação com dados cross-section, as estimativas para as eficiências técnicas individuais são consistentes.

A estimação de dados em painel pode ser feita por dois modelos distintos: a estimação com efeitos fixos e a estimação com efeitos aleatórios. Aqui nos deteremos na estimação com efeitos fixos, por sua simplicidade e aplicabilidade.⁸

Duas são as suposições mais usuais na estimação da FE em painéis com efeitos fixos. A primeira é a de que a eficiência de cada unidade se mantém constante ao longo do período de cobertura do painel. A segunda permite que cada unidade possa evoluir (ou involuir) em sua eficiência técnica própria.

O primeiro desses modelos, denominado modelo com eficiência temporalmente fixa, será descrito nessa seção, enquanto o segundo, o modelo com variação temporal da eficiência, será o tema da próxima seção.

Mais uma vez supondo-se uma função Cobb-Douglas, pode-se utilizar a seguinte equação para estimar um painel com eficiência temporalmente fixa em um contexto da fronteira de produção:

$$\ln q_{i,t} = \beta_0 + \beta_1 \ln x_{1,i,t} + \beta_2 \ln x_{2,i,t} + \dots + \beta_m \ln x_{m,i,t} + v_{i,t} - \mu_i \quad (9.18)$$

7 Ver fórmula 9.7.

8 O leitor interessado poderá encontrar uma ampla discussão sobre os modelos de efeitos aleatórios em Greene (1993).

O subscrito t denota que agora se trata de um painel com T períodos. Assim, $q_{i,t}$ representa a quantidade de produtos gerada na unidade i , no período t . Já $x_{m,i,t}$ denota a quantidade utilizada do insumo m , pela unidade i , no período t . O choque $v_{i,t}$ é aquele que foi recebido pela unidade i , no período t . Repare que o componente de eficiência μ_i não varia no tempo, apenas por unidade produtiva, sendo isso o que caracteriza o modelo temporalmente fixo.

O modelo com efeitos fixos é estimado por mínimos quadrados ordinários. Como não é possível, a princípio, discernir entre os dois tipos de variáveis não observáveis (μ_i e $v_{i,t}$), o que se obtém são estimadores para os efeitos fixos que agregam o componente de eficiência em seu bojo. Então, a estimação MQO com efeitos fixos produziria:

$$\ln q_{i,t} = \hat{\beta}_{0,i} + \hat{\beta}_1 \ln x_{1,i,t} + \hat{\beta}_2 \ln x_{2,i,t} + \dots + \hat{\beta}_m \ln x_{m,i,t} + \hat{v}_{i,t} \quad (9.19)$$

Repare que $\hat{\beta}_{0,i} = \hat{\beta}_0 - \hat{\mu}_i$, ou seja, os estimadores dos efeitos fixos obtidos constituem-se, na verdade, em uma mistura do estimador do intercepto com o estimador do componente de eficiência. É válido ressaltar que, da maneira como está sendo apresentada essa estimação, a constante da regressão não deve entrar como um dos regressores – é necessário rodar o painel com efeitos fixos sem a constante.

Com a regressão, são obtidos então os n estimadores para os efeitos fixos ($\hat{\beta}_{0,i}$) e os m estimadores associados a cada um dos insumos utilizados. Obtidas as estimativas para tais parâmetros, procede-se à seguinte normalização, com o objetivo de se obter uma estimativa para o intercepto ($\hat{\beta}_0$):

$$\hat{\beta}_0 = \underset{i}{Max} \{ \hat{\beta}_{0,i} \} \quad (9.20)$$

O processo de estimação contido na equação 9.19 prescreve que o maior efeito fixo individual seja considerado como o intercepto. Isso quer dizer que à unidade produtora com o maior efeito fixo será imputada eficiência de 100%, pois nesse caso:

$$\hat{\beta}_{0,i} = \hat{\beta}_0 \Rightarrow \hat{\mu}_i = 0 \Rightarrow TE_i = e^{-0} = 1$$

Para todas as demais unidades ocorrerá que:

$$\hat{\mu}_i = \hat{\beta}_0 - \hat{\beta}_{0,i} \quad (9.21)$$

Com $\hat{\mu}_i > 0 \Rightarrow TE_i = e^{-\hat{\mu}_i} < 1$.

De posse dos componentes de eficiência individuais é possível calcular as eficiências técnicas de cada unidade produtora, utilizando novamente a equação 9.14. Como foi mencionado no início da seção, os estimadores obtidos são consistentes e o seu viés sobre-estimador da eficiência tende a desaparecer quando a amostra cresce.

O viés do estimador provém do fato de que, ao se escolher a unidade de maior efeito fixo como eficiente, provavelmente se está atribuindo essa característica indevidamente. Se por acaso a amostra em questão só contiver unidades bastante ineficientes, a menos pior será considerada eficiente. No entanto, à medida que a amostra aumenta, torna-se mais provável que unidades realmente eficientes no sentido absoluto sejam incluídas, aumentando a possibilidade de que o viés seja afastado.

Para consolidar o processo de estimação observe o seguinte exemplo, descrito em detalhes no próximo capítulo. Existem 12 unidades produtivas ($n = 12$) que realizam a produção por cinco períodos consecutivos ($T = 5$), utilizando dois insumos (x_1 e x_2) para produzir um único produto (q).

A Tabela 9.1 apresenta os resultados obtidos para os efeitos fixos de cada unidade considerada.

Tabela 9.1 – Estimação da eficiência técnica para 12 DMUs hipotéticas

DMU	$\hat{\beta}_{0,i}$	μ_i	TE_i
1	0,528	0,352	0,703
2	0,880	0,000	1,000
3	0,540	0,340	0,712
4	0,510	0,369	0,691
5	0,463	0,416	0,659
6	0,598	0,281	0,755
7	0,588	0,291	0,747
8	0,213	0,667	0,513
9	0,200	0,680	0,507
10	-0,493	1,372	0,254
11	-0,137	1,016	0,362
12	-0,907	1,786	0,168

Fonte: Elaboração própria.

A primeira coluna da tabela numera as unidades produtoras e a segunda apresenta o resultado das estimações dos efeitos fixos. O valor máximo obtido foi 0,880, correspondente ao efeito da segunda unidade, e que, seguindo a equação 9.20, passa a ser o estimador para β_0 .

A terceira coluna apresenta os valores de μ para cada unidade, obtidos de acordo com a equação 9.21. Os valores da última coluna foram obtidos de acordo com a equação 9.14 e refletem a eficiência técnica para cada unidade. Observe que a segunda unidade com o maior efeito fixo estimado possui eficiência de 100%.

9.6 Dados em painel: variação temporal da eficiência

O modelo anterior supõe que cada unidade produtora possua um determinado nível de eficiência e que esse seja constante ao longo do tempo. No entanto, quando se consideram painéis mais longos, ou mais espaçados temporalmente,⁹ ou de setores sujeitos a variações tecnológicas repentinas, a suposição de eficiência constante no tempo passa a ser bastante questionável.

Para essas situações é possível utilizar o modelo com variação temporal da eficiência, que busca captar a evolução desse parâmetro em cada unidade produtora da amostra, para grupos de unidades. Nesse caso, supõe-se algum processo de evolução da eficiência, de modo que esta possa variar temporalmente de acordo com o seguinte modelo:

$$\ln q_{i,t} = \beta_0 + \beta_1 \ln x_{1,i,t} + \beta_2 \ln x_{2,i,t} + \dots + \beta_m \ln x_{m,i,t} + v_{i,t} - \mu_{i,t} \quad (9.22)$$

A diferença entre esse modelo e aquele com eficiência temporalmente invariante, expresso na equação 9.18, está no componente de ineficiência μ , para o qual agora se permite que varie temporalmente. O modelo de estimação por MQO em painel segue a seguinte equação:

$$\ln q_{i,t} = \hat{\beta}_{0,i,t} + \hat{\beta}_1 \ln x_{1,i,t} + \hat{\beta}_2 \ln x_{2,i,t} + \dots + \hat{\beta}_m \ln x_{m,i,t} + \hat{v}_{i,t} \quad (9.23)$$

Note que agora o estimador de β_0 varia não só entre as unidades produtoras, mas também com o tempo, já que $\hat{\beta}_{0,i,t} = \hat{\beta}_0 - \hat{\mu}_{i,t}$. Isso causa um problema para a estimação do modelo, pois mesmo ao

9 Painéis quinquenais ou decenais, por exemplo.

se supor um painel balanceado, o número de observações será de $n \times T$ e os parâmetros a serem estimados totalizam $n \times T + m$. Nessas condições, a estimação do modelo não é viável, uma vez que existiriam mais parâmetros para estimar do que observações e não há, portanto, graus de liberdade suficientes.

Para tornar viável a estimação desse modelo, faz-se necessária a designação de uma estrutura na evolução para β_{it} . Uma formulação muito utilizada para essa estrutura é:

$$\beta_{0,i,t} = \Omega_{0,i} + \Omega_{1,i}t + \Omega_{2,i}t^2 \quad (9.24)$$

A equação a ser estimada, então, seria:

$$\ln q_{i,t} = \hat{\Omega}_{0,i} + \hat{\Omega}_{1,i}t + \hat{\Omega}_{2,i}t^2 + \hat{\beta}_1 \ln x_{1,i,t} + \hat{\beta}_2 \ln x_{2,i,t} + \dots + \hat{\beta}_m \ln x_{m,i,t} + \hat{v}_{i,t} \quad (9.25)$$

Na equação 9.25 o número de parâmetros estimados seria $m+3n$, enquanto o número de observações é de $n \times T$. Assim, se o número de DMUs for maior que 3 e o número de períodos superior a 2, o problema de suficiência dos graus de liberdade será resolvido.

Nessa equação, os Ω s são parâmetros que governam a evolução temporal da eficiência de cada unidade. A introdução de um contador temporal quadrático possibilita não só a aferição do sentido, mas também a da velocidade de evolução, permitindo ainda o teste de uma série de hipóteses sobre a evolução da eficiência, como mostra o Quadro 9.1.

Quadro 9.1 – Hipóteses passíveis de serem testadas a partir da estimação dos parâmetros Ω

Hipóteses	Significado
$\Omega_{1,i} = 0$ e $\Omega_{2,i} = 0$	Eficiência temporalmente fixa na unidade i .
$\Omega_{1,i} = \Omega_1$ e $\Omega_{2,i} = \Omega_2$ $\forall i = 1, 2, \dots, n$	Eficiência temporalmente variante, mas com variação uniforme entre as unidades.
$\Omega_{1,i} < 0$	Tendência de diminuição da eficiência na unidade i .
$\Omega_{1,j} > 0$ e $\Omega_{2,i} < 0$	Tendência de diminuição da eficiência na unidade i , mas com ganhos cada vez menores.

Outra possibilidade é a de dividir a amostra em dois grupos e testar a diferença entre os coeficientes Ω s. Supondo que a amostra fosse dividida entre prestadores públicos e privados, seria estimada a seguinte equação:

$$\ln q_{i,t} = S\hat{\Omega}_{S0} + S\hat{\Omega}_{S1}t + S\hat{\Omega}_{S2}t^2 + P\hat{\Omega}_{P0} + P\hat{\Omega}_{P1}t + P\hat{\Omega}_{P2}t^2 + \hat{\beta}_1 \ln x_{1,i,t} + \dots + \hat{\beta}_m \ln x_{m,i,t} + \hat{v}_{i,t} \quad (9.26)$$

Nela, S assume valor 1 se o registro for de uma empresa pública e 0, caso contrário; P assume valor 1 ou 0, dependendo se a empresa é privada ou não. Aqui o teste de significância poderia aferir não só se existe diferença na eficiência das empresas públicas e privadas, mas também qual a distinção na velocidade de evolução desse parâmetro.

Esse teste pode ser útil na avaliação a respeito da prestação de diversos serviços públicos, como água, saneamento, distribuição de energia. Nesses setores, poucas evidências são encontradas em favor dos serviços privados no período posterior à privatização deles.

Uma das explicações para isso é a de que esses serviços são sujeitos a processos de *learning by doing* e, portanto, levaria algum tempo para que as empresas privadas aprendessem a gerir o processo de forma efetiva. No entanto, se isso é verdade, então esse tipo de empresa, a despeito do nível de eficiência (Ω_ρ) semelhante ao das públicas, deveria apresentar coeficientes de evolução da eficiência (Ω_1 e Ω_2) superiores.

9.7 Variáveis ambientais

Como já mencionado, a introdução de variáveis ambientais se dá de forma mais direta e natural na análise da FE do que quando se utiliza a metodologia DEA. Isso porque a FE é por si só uma regressão e, portanto, passível de lidar com variáveis de controle. A Fronteira Estocástica dispensa, então, a utilização de um segundo estágio, pois as influências exógenas (variáveis ambientais) são incorporadas já na primeira regressão.

No caso da análise de regressão tradicional são consideradas variáveis de controle aquelas que têm influência sobre a variável endógena, mas que não são diretamente objeto das hipóteses a serem testadas na regressão. Por exemplo, se um pesquisador quiser estimar a elasticidade do preço da demanda por energia elétrica, ele deverá

incluir variáveis exógenas na regressão, como renda dos consumidores, temperatura ambiente e outros, mesmo que ele não tenha interesse em investigar seus efeitos sobre o consumo de energia.

Quando a análise é realizada por meio da FE, as variáveis de controle são denominadas variáveis ambientais. A principal característica delas é que podem influenciar diretamente a eficiência das unidades sem, contudo, estar sob o seu controle. Assim, se uma escola está localizada em uma região pobre e atende as crianças de lá, é provável que ela apresente resultados acadêmicos inferiores àqueles observados em unidades de localidades mais ricas, simplesmente porque os alunos dessas últimas possuem vantagens iniciais que não dependem do controle dos educadores.

Existem duas formas principais de incorporação das variáveis ambientais na regressão da FE. A primeira delas é utilizada quando se supõe que as variáveis ambientais afetam diretamente a eficiência das unidades estudadas. Já no segundo caso, a hipótese é de que as variáveis ambientais possam deslocar a fronteira de eficiência.

Um exemplo do segundo caso pode ser observado em um contexto de estudo da produtividade agrícola. O regime pluviométrico, que é uma variável ambiental, desloca a fronteira de possibilidades de produção dos agricultores. Assim, um empreendimento agrícola situado em uma região com regime de chuvas mais favorável terá as suas possibilidades de produção estendidas, enquanto regiões que sofrem secas verão as suas fronteiras recuarem.

Já a influência direta sobre a eficiência se dá quando a variável ambiental afeta a habilidade da unidade produtora. Por exemplo, uma situação na qual se avalia a eficiência de escolas públicas primárias. Geralmente esse tipo de avaliação é realizado utilizando-se o número de professores, o de equipamentos didáticos e outros como variáveis de insumo, e a nota dos estudantes em algum teste padronizado como variável de produto (ou de resultado).

Acontece que o desempenho dos alunos é positivamente correlacionado com o nível de escolaridade dos seus pais. A teoria é de que a maior escolaridade dos pais tende a proporcionar um convívio mais educativo em casa, o que por sua vez favoreceria o aprendizado na escola. Portanto, crianças com convívio familiar

menos enriquecedor em termos culturais teriam as suas habilidades de aprendizado prejudicadas.

Embora a distinção entre a natureza das variáveis ambientais possa ser tênue, o modelo de estimação é substancialmente diferente. O caso mais simples é aquele em que as variáveis ambientais deslocam a fronteira de eficiência. Nessa situação, as variáveis ambientais entram diretamente na regressão como variáveis de controle. Se tomarmos a função Cobb-Douglas em um painel com eficiência temporalmente fixa:

$$\ln q_{i,t} = \beta_0 + \beta_1 \ln x_{1,i,t} + \dots + \beta_m \ln x_{m,i,t} + \phi_1 \ln z_{1,i,t} + \dots + \phi_q \ln z_{q,i,t} + v_{i,t} - \mu_i \quad (9.27)$$

Nessa equação, $z_{q,i,t}$ é o valor que a q -ésima variável ambiental assumiu no período t e na unidade i , e $\tau_{q,i}$ o seu coeficiente. Aqui a estimação é realizada de acordo com a técnica desenvolvida na seção 9.5, ou seja, por mínimos quadrados ordinários.

No caso da incorporação das variáveis ambientais com efeito direto sobre a eficiência das unidades produtoras, ter-se-ia o seguinte modelo:

$$\ln q_{i,t} = \beta_0 + \beta_1 \ln x_{1,i,t} + \dots + \beta_m \ln x_{m,i,t} + v_{i,t} - \mu_i \quad (9.28)$$

$$\mu_i = \phi_1 z_{1,i} + \dots + \phi_q z_{q,i} + \xi_i \quad (9.29)$$

Aqui, as variáveis ambientais afetam diretamente o resultado da unidade e a estimação deve ser por máxima verossimilhança, uma vez que a equação se torna não linear, pois ao aglutinar 9.28 com 9.29 chega-se a:

$$\ln q_{i,t} = \beta_0 + \beta_1 \ln x_{1,i,t} + \dots + \beta_m \ln x_{m,i,t} + v_{i,t} - \phi_1 z_{1,i} - \dots - \phi_q z_{q,i} - \xi_i \quad (9.30)$$

Felizmente, as rotinas para todas essas estimações já estão escritas no *software* R, como será visto na próxima seção.

9.8 Função translogarítmica inversa (translog)

A função translogarítmica inversa, ou mais simplesmente função translog, foi sugerida por Christensen, Jorgenson e Lau (1975) como uma função de utilidade que não apresentava nem homotetia, nem

aditividade, com o objetivo de obter funções de demanda mais gerais. Desde então ela vem sendo extensivamente utilizada na literatura empírica para a estimação de diversos fenômenos.

No caso da estimação da FE ela é particularmente útil, uma vez que a partir dessa forma funcional é possível proceder tanto à estimação da fronteira de produção quanto à da fronteira de custos. E mais ainda, é possível também estimar a função distância, que permite avaliação das eficiências das unidades produtoras em um cenário de múltiplos insumos e produtos.

Uma das vantagens da função translog – em sua versão de custos, de produção ou distância – é a de permitir a interação das diversas variáveis, não sendo os seus efeitos independentes dos efeitos das demais.

Para entender esse ponto, observe a fórmula da função Cobb-Douglas expressa, por exemplo, na equação 9.7. Nela, aumentos de x_1 repercutem sobre q de forma completamente independente das outras variáveis. Assim, qualquer que seja o nível de x_2 , uma unidade a mais de x_1 elevará o produto q em β_1 unidades. Compare com o modelo translog:

$$\ln q_i = \beta_0 + \beta_1 \ln x_{1,i} + \dots + \beta_m \ln x_{m,i} + \frac{1}{2} \beta_{11} \ln x_{1,i} \ln x_{1,i} + \dots \quad (9.31)$$

$$+ \frac{1}{2} \beta_{m1} \ln x_{1,i} \ln x_{m,i} + \dots + \frac{1}{2} \beta_{mm} \ln x_{m,i} \ln x_{m,i} + v_i - \mu_i$$

a qual pode ser expressa em uma formulação simplificada como:

$$\ln q_i = \beta_0 + \sum_{j=1}^m \beta_j \ln x_{1,j} + \frac{1}{2} \sum_{j=1}^m \sum_{k \neq j}^m \beta_{jk} \ln x_{1,j} \ln x_{1,k} + v_i - \mu_i \quad (9.31')$$

Nessa formulação, o efeito líquido de um aumento de uma unidade de, digamos, x_1 sobre q depende dos níveis de todas as outras variáveis de input, uma vez que:

$$\frac{\partial \ln q_i}{\partial \ln x_{1,i}} = \beta_1 + \beta_{11} \ln x_{1,i} + \frac{1}{2} \sum_{k \neq j}^m \beta_{1k} \ln x_{k,i} \quad (9.32)$$

Portanto, o impacto de x_1 sobre q depende não só de seu coeficiente, mas também de todos os coeficientes $\beta_{1k'}$ bem como dos valores absolutos de todas as outras variáveis de insumo.

Como se pode notar pela expressão 9.30, a desvantagem na utilização da função translog é o grande número de parâmetros a serem estimados. Tendo em mente uma situação com três insumos e um produto, a estimação com a translog necessitaria estimar dez betas distintos.¹⁰ Se o número de insumos for 5, deverão ser estimados 21 betas. Isso quer dizer que, a menos que se disponha de muitos registros, a utilização da função translog poderá ser inviável.

A versão expressa nas equações 9.31 e 9.32 refere-se à fronteira de produção, situação na qual se tem um só produto e um número m de insumos utilizados na sua produção. A função translog também se presta à estimação da fronteira de custos. Nesse caso, como foi observado na seção 9.4., a variável endógena é o montante despendido e as variáveis explicativas são as quantidades produzidas. Então, em sua versão translog, a função para estimação da fronteira de custos pode ser descrita como:

$$\ln CE_i = \alpha_0 + \sum_{j=1}^s \alpha_j \ln q_{j,i} + \frac{1}{2} \sum_{j=1}^s \sum_{k=1}^s \alpha_{jk} \ln q_{j,i} \ln q_{k,i} + v_i + \mu_i \quad (9.33)$$

Mais uma vez, esta forma funcional capta a interação das variáveis exógenas – nesse caso, os montantes produzidos –, permitindo, inclusive, a análise das economias de escopo. Para tanto, devem-se analisar o sinal e a magnitude dos parâmetros cruzados α_{jk} . Para que haja economias de escopo entre dois produtos – digamos, o produto 2 e o produto 5 –, é necessário que $\alpha_{25} < 0$. Isso quer dizer que ao se aumentar a produção desses dois produtos, os custos subiram menos que proporcionalmente.

Como mencionado anteriormente, além das estimações da fronteira de produção e da de custo, a função translog também pode ser utilizada para a estimação da FE em situações nas quais estão presentes múltiplos insumos e múltiplos produtos. Essa aplicação é bastante útil, pois pode-se dizer que na maioria dos processos produtivos reais existem múltiplos produtos e insumos.

Para realizar essa estimação é necessária a especificação da função distância, que traz como seus argumentos tanto insumos quanto produtos, e que permite avaliar qual a distância entre um

10 $\beta_0, \beta_1, \beta_2, \beta_3, \beta_{11}, \beta_{22}, \beta_{33}, \beta_{12}, \beta_{13}, \beta_{23}$.

determinado esquema produtivo e a fronteira de eficiência.

Na ótica do produto, função distância pode ser expressa da seguinte forma:

$$\begin{aligned} \ln DP_i = & \alpha_0 + \sum_{j=1}^s \alpha_j \ln q_{j,i} + \frac{1}{2} \sum_{j=1}^s \sum_{k=1}^s \alpha_{jk} \ln q_{j,i} \ln q_{k,i} + \sum_{j=1}^m \beta_j \ln x_{j,i} \\ & + \frac{1}{2} \sum_{j=1}^m \sum_{k=1}^m \beta_{jk} \ln x_{j,i} \ln x_{k,i} + \frac{1}{2} \sum_{j=1}^s \sum_{k=1}^m \delta_{jk} \ln q_{j,i} \ln x_{k,i} \end{aligned} \quad (9.34)$$

Manipulando-se a equação 9.33 obtém-se:

$$\begin{aligned} -\ln q_{1,i} = & \alpha_0 + \sum_{j \neq 1}^s \alpha_j \ln q_{j,i}^* + \frac{1}{2} \sum_{j \neq 1}^s \sum_{k \neq 1}^s \alpha_{jk} \ln q_{j,i}^* \ln q_{k,i}^* \\ & + \sum_{j=1}^m \beta_j \ln x_{j,i} + \frac{1}{2} \sum_{j=1}^m \sum_{k=1}^m \beta_{jk} \ln x_{j,i} \ln x_{k,i} \\ & + \frac{1}{2} \sum_{j \neq 1}^s \sum_{k=1}^m \delta_{jk} \ln q_{j,i} \ln x_{k,i}^* - \ln DP_i \end{aligned} \quad (9.35)$$

Aqui $\ln q_{j,i}^* = \ln (q_{j,i}/q_{1,i})$ e analogamente $\ln x_{k,i}^* = \ln (x_{k,i}/x_{1,i})$. Pode-se escolher qualquer um dos produtos (na equação 9.34), ou qualquer um dos insumos (na equação 9.35) para fazer o papel de variável endógena, contanto que a mesma variável seja utilizada como denominador do lado direito da equação. O próximo passo é substituir o termo do logaritmo da distância pelo resíduo estocástico e pelo componente de ineficiência para se obter uma variação estimável da equação. Assim:

$$\begin{aligned} -\ln q_{1,i} = & \alpha_0 + \sum_{j \neq 1}^s \alpha_j \ln q_{j,i}^* + \frac{1}{2} \sum_{j \neq 1}^s \sum_{k \neq 1}^s \alpha_{jk} \ln q_{j,i}^* \ln q_{k,i}^* \\ & + \sum_{j=1}^m \beta_j \ln x_{j,i} + \frac{1}{2} \sum_{j=1}^m \sum_{k=1}^m \beta_{jk} \ln x_{j,i} \ln x_{k,i} \\ & + \frac{1}{2} \sum_{j \neq 1}^s \sum_{k=1}^m \delta_{jk} \ln q_{j,i} \ln x_{k,i}^* + v_i + \mu_i \end{aligned} \quad (9.36)$$

A eficiência técnica é, então, obtida de acordo com a equação 9.6.

A versão da função de distância voltada para insumos pode ser obtida de forma análoga:

$$\begin{aligned}
-\ln x_{1,i} &= \varphi_0 + \sum_{j \neq 1}^m \varphi_j \ln x_{j,i}^* + \frac{1}{2} \sum_{j \neq 1}^m \sum_{k \neq 1}^m \varphi_{jk} \ln x_{j,i}^* \ln x_{k,i}^* \\
&+ \sum_{j=1}^s \rho_j \ln q_{j,i} + \frac{1}{2} \sum_{j=1}^s \sum_{k=1}^s \rho_{jk} \ln q_{j,i} \ln q_{k,i} \\
&+ \frac{1}{2} \sum_{j \neq 1}^m \sum_{k=1}^s \psi_{jk} \ln x_{j,i} \ln q_{k,i} + v_i - \mu_i
\end{aligned} \tag{9.37}$$

Observe que os sinais do componente de eficiência nas duas formulações são trocados. Na versão voltada para produto (equação 9.35), quanto maior μ , menor a eficiência; mas também menor, em termos absolutos, a produção de q_j . Já no caso do foco em insumos, quanto maior μ , mais negativo será $-\ln x_{1,i}$, indicando um valor absoluto maior e, portanto, maior uso do insumo.

9.9 Decomposição da evolução da eficiência com a FE

A função translog também pode ser utilizada para a estimação da evolução da produtividade e para a decomposição dos fatores influentes sobre esse fenômeno. Para tanto, deve haver a disponibilidade de dados em painel.

Quando esse for o caso, algumas modificações na função de estimação serão necessárias para que ela possa acomodar a dimensão temporal. A equação 9.38 pode ser utilizada para a estimação da fronteira de produção com múltiplos insumos, mas com um único produto. O termo t é o contador de períodos que incorpora a dimensão temporal.

$$\begin{aligned}
q_{it} &= \beta_0 + \sum_n \beta_n \ln x_{nit} + \beta_t t + \frac{1}{2} \sum_n \sum_k \beta_{nk} \ln x_{nit} \ln x_{kit} \\
&+ \frac{1}{2} \beta_{tt} t^2 + \sum_n \beta_{nt} t \ln x_{nit} + v_{it} - u_{it}
\end{aligned} \tag{9.38}$$

Os parâmetros aqui são estimados de acordo com o exposto na seção 9.6 (modelo com variação temporal da eficiência). A partir da obtenção dos estimadores dos parâmetros, é possível decompor as fontes de variação da produtividade e calcular essa variação entre dois períodos de forma semelhante ao processo exposto no capítulo anterior.

Para se obter o índice de variação da tecnologia no período t ,

deve-se utilizar a seguinte equação:

$$tc_{i,t} = \hat{\beta}_t + \hat{\beta}_n t + \sum_n \hat{\beta}_{nt} \ln x_{nit} \quad (9.39)$$

onde os $\hat{\beta}$ s correspondem aos estimadores para os parâmetros da equação 9.37. Valores para $tc_{i,t}$ maiores que um indicam que a tecnologia contribuiu para o aumento da produtividade no período estudado.

A variação da eficiência técnica pode ser obtida a partir da estrutura estipulada na equação 9.24, tomando-se a sua derivada em relação ao tempo:

$$tec_{i,t} = \hat{\Omega}_{2,i} + 2t\hat{\Omega}_{3,i} \quad (9.40)$$

Analogamente, $tec_{i,t}$ mede o quanto o aumento da eficiência técnica contribuiu para a evolução da produtividade entre dois períodos consecutivos. A produtividade também pode ser afetada por modificações na escala de produção, pois quanto mais o nível de operação de uma unidade produtiva se aproxima da escala ótima de produção, maior será a sua produtividade, tudo o mais constante. Para capturar esse efeito, utiliza-se a seguinte equação:

$$\varepsilon_{i,t} = \sum_m (\hat{\beta}_m + \sum_k \hat{\beta}_{mk} \ln x_{kit} + \hat{\beta}_{mt} t) \quad (9.41)$$

Para o cálculo da variação total da produtividade, é necessário ainda o cálculo da elasticidade dos insumos em relação ao produto:

$$\varepsilon_{m,i,t} = \hat{\beta}_m + \sum_k \hat{\beta}_{mk} \ln x_{kit} + \hat{\beta}_{mt} t \quad (9.42)$$

De posse de todos esses índices, calcula-se a variação da produtividade total da unidade por meio da seguinte equação:

$$pc_{i,t} = tc_{i,t} + tec_{i,t} + (\varepsilon_{i,t} - 1) \sum_k \left(\frac{\varepsilon_{k,i,t}}{\varepsilon_{i,t}} \right) (x_{k,i,t} - x_{k,i,t-1}) \quad (9.43)$$

O índice $pc_{i,t}$ fornece então a variação total da produtividade no período.

REFERÊNCIAS

AIGNER, D.; LOVELL, C. A. K.; SCHMIDT, P. Formulation and estimation

of stochastic frontier production function models. *Journal of Econometrics*, v. 6, p. 21-37, 1977.

BAUER, P. W. Decomposing TFP growth in the presence of cost inefficiency, nonconstant returns to scale, and technological progress. *Journal of Productivity Analysis*, v. 1, p. 287-299, 1990.

COELLI, T. J. Estimators and hypothesis tests for a stochastic frontier function: a Monte Carlo analysis. *Journal of Productivity Analysis*, v. 6, p. 247-268, 1995.

COELLI, T. J.; PRASADA RAO, D. S.; O'DONNELL, C. J.; BATTESE, G. E. *An introduction to efficiency and productivity analysis*. 2. ed. Nova York: Springer, 2005.

CHRISTENSEN, L. R.; JORGENSON, D. W.; LAU, L. J. Transcendental logarithmic utility functions. *American Economic Review*, v. 65, p. 367-383, 1975.

GREENE, W. H. The econometric approach to efficiency analysis. In: FRIED, H. O.; SCHMIDT, S. S.; LOVELL, C. A. K. (Ed.). *The measurement of productive efficiency: techniques and applications*. Oxford University Press, 1993.

JONDROW, J.; LOVELL, C. A. K.; METEROV, I. S.; SCHMIDT, P. On the estimation of technical inefficiency in the stochastic frontier of production function model. *Journal of Econometrics*, v. 19, p. 233-238, 1982.

KUMBHAKAR, S. C.; LOVELL, C. A. K. *Stochastic frontier analysis*. Cambridge University Press, 2000.

MEEUSEN, W.; VAN DER BROECK, J. Efficiency estimation from Cobb-Douglas production function with composed error. *International Economic Review*, v. 18, p. 435-444, 1977.

Capítulo 10

R: Um *software* livre para mensuração da eficiência

Luis Felipe Vital Nunes Pereira

10.1 Introdução: o desafio computacional da mensuração da eficiência

A metodologia Análise Envoltória de Dados (DEA¹), apresentada no capítulo 8, há alguns anos exigia conhecimento prévio de técnicas de Programação Linear e implementação utilizando linguagens de programação como Fortran e C. As aplicações também estavam restritas a problemas relativamente pequenos, devido a limitações de hardware.

Já a metodologia de Fronteira Estocástica, apresentada no capítulo 9, consiste em problemas computacionalmente mais intensivos e exigem a utilização de um arcabouço de estatística e econometria. Estas também eram utilizadas apenas por pesquisadores com acesso a computadores de grande porte e conhecimento de programação.

A facilidade no acesso a ferramentas computacionais permitiu grandes avanços na mensuração da eficiência nas últimas décadas. Hoje, esses problemas podem ser resolvidos por estudantes e gestores públicos em seus computadores pessoais.

Você vai perceber ao longo deste capítulo que os problemas são resolvidos em segundos, mesmo quando utilizamos bases de dados grandes. Até mesmo os problemas que exigem estimações recursivas, como os utilizados na detecção de *outliers* apresentados na Parte V deste livro, possuem tempo de processamento bastante razoável e não podem ser vistos como uma restrição para a mensuração da eficiência.

O objetivo deste capítulo é apresentar ao leitor o R, capacitando-o a instalar, encontrar ajuda, importar dados, obter estatísticas e utilizar rotinas para cálculo do DEA e da Fronteira Estocástica. Esta publicação optou pelo *software* R para os exercícios práticos de forma a tornar possível o acesso de gestores e estudantes.

1 Análise Envoltória de Dados é a tradução literal do termo em inglês *Data Envelopment Analysis* (DEA).

10.1.1 *Softwares* para mensuração da eficiência

Existe grande variedade de *softwares* no mercado, capazes de realizar mensuração de eficiência utilizando DEA e Fronteira Estocástica. Especificamente para o DEA, os mais populares são o DEA-Solver, OnFront, IDEAS, Frontier Analyst e DEA-P. No entanto, é possível aplicar DEA utilizando rotinas e pacotes em *softwares* estatísticos como R, SAS, Stata e Shazam, ou ainda implementar a programação linear em *softwares* de programação matemática como GAMS e AMPL, com ajuda de um solver como o C-Plex, por exemplo.

A implementação mais popular de Fronteira Estocástica, o Frontier, foi originalmente desenvolvida em Fortran 77 e era executado em modo de comando com uso de arquivos auxiliares de dados. Hoje, o mesmo Frontier passou por mudanças que o deixaram mais robusto e eficiente e pode ser chamado por meio de *softwares* estatísticos. Alguns *softwares* econométricos, como o Stata, também calculam Fronteiras Estocásticas.

Nesta publicação, as aplicações de DEA e Fronteira Estocástica serão realizadas a partir do *software* R utilizando os pacotes Benchmarking e Frontier. O Benchmarking se baseia em outros pacotes para R que resolvem problemas de programação linear, enquanto o Frontier é baseado em um código binário programado em Fortran, mas chamado a partir da interface do R.

10.1.2 Os gestores públicos e a infraestrutura de TI

Conforme apresentado na seção anterior, existe uma variedade de *softwares* capazes de mensurar eficiência com as técnicas apresentadas neste livro. No entanto, esses *softwares* nem sempre estão à disposição do gestor público e o processo de aquisição pode ser longo e custoso.

A aquisição de alguns dos *softwares* citados pode custar à administração pública alguns milhares de reais. Ainda existem dificuldades muito específicas como a existência de um único fornecedor para o *software*, mudanças de versões, descontinuidade de suporte ao *software*, prazos de entrega etc.

Nesse sentido, a ação do governo federal para adoção de *softwares* livres tem ganhado espaço. *Software* livre é aquele disponibilizado, gratuitamente ou comercializado, com as premissas de liberdade de

instalação, plena utilização, acesso ao código fonte, possibilidade de modificações/aperfeiçoamentos para necessidades específicas, distribuição da forma original ou modificada, com ou sem custos (RIBEIRO, 2004).

O Ministério do Planejamento criou, em 2003, um grupo de trabalho com o objetivo de formular orientações para a migração para *software* livre. A opção pelo *software* livre é uma decisão estratégica que reduz custos, amplia a concorrência, gera empregos e desenvolve o conhecimento e a inteligência no país. A redução nos custos de propriedade de *software* já pode ser observada em diversas áreas do governo. Do ponto de vista legal, a adoção do *software* livre pela administração pública também é amparada pelos princípios constitucionais de impessoalidade, eficiência e razoabilidade (BRASIL, 2005).

10.2 R: Um *software* livre

O R é uma linguagem e um ambiente para computação estatística popular entre estatísticos e economistas e disponível gratuitamente no site: <http://www.r-project.org/> para os sistemas operacionais Windows, OS X e Linux, nas versões 32 bits e 64 bits.

10.2.1 O Projeto R

O R é uma implementação de uma linguagem de programação mais antiga, denominada S, que permite uma grande variedade de recursos estatísticos e econométricos. Entre as funcionalidades mais conhecidas estão a modelagem linear, não linear, econometria de séries de tempo e dados em painel, além de programação linear e técnicas de *cluster*.

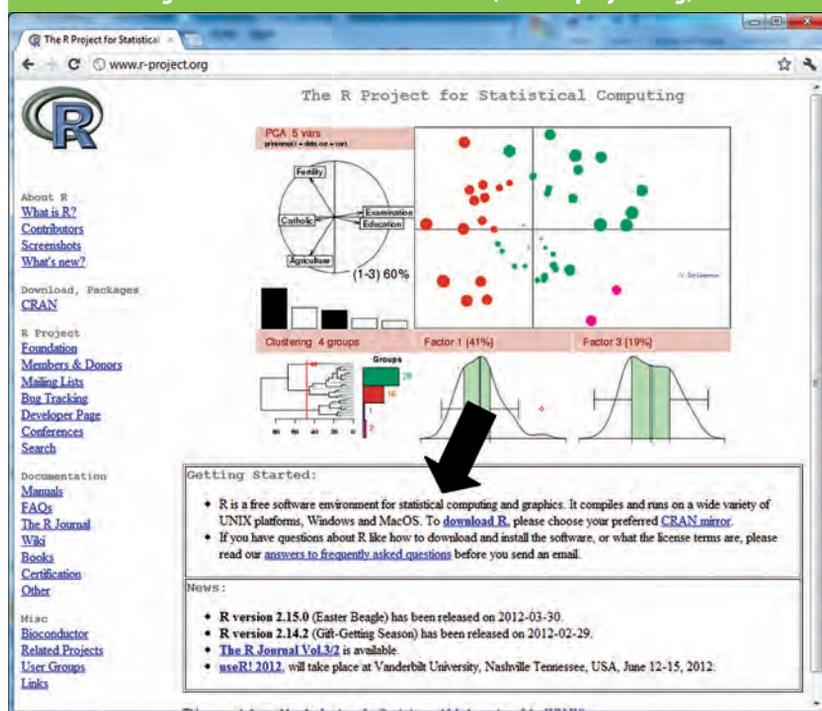
Um grupo de estatísticos trabalha constantemente no esforço de aprimorar a ferramenta, e frequentemente são lançadas atualizações com novas funcionalidades e correções de eventuais problemas (*bugs* e erros) nas versões atuais.

O fato de ser um *software* livre e estar disponível gratuitamente pela internet levou diversos pesquisadores a adotar o R como ferramenta principal para divulgação de seus algoritmos. Como consequência, o R é capaz de disponibilizar funcionalidades antes dos principais *softwares* pagos.

10.2.2 Visão geral do *software*

Para instalar o R, visite o site www.r-project.org e clique em *download*. O link levará a uma página de repositórios (*mirrors*), que são servidores ao redor do mundo que armazenam arquivos do R. Escolha o servidor mais próximo da sua localidade (existem diversas opções de servidores no Brasil).

Figura 10.1 – Site do R na Internet (www.r-project.org)

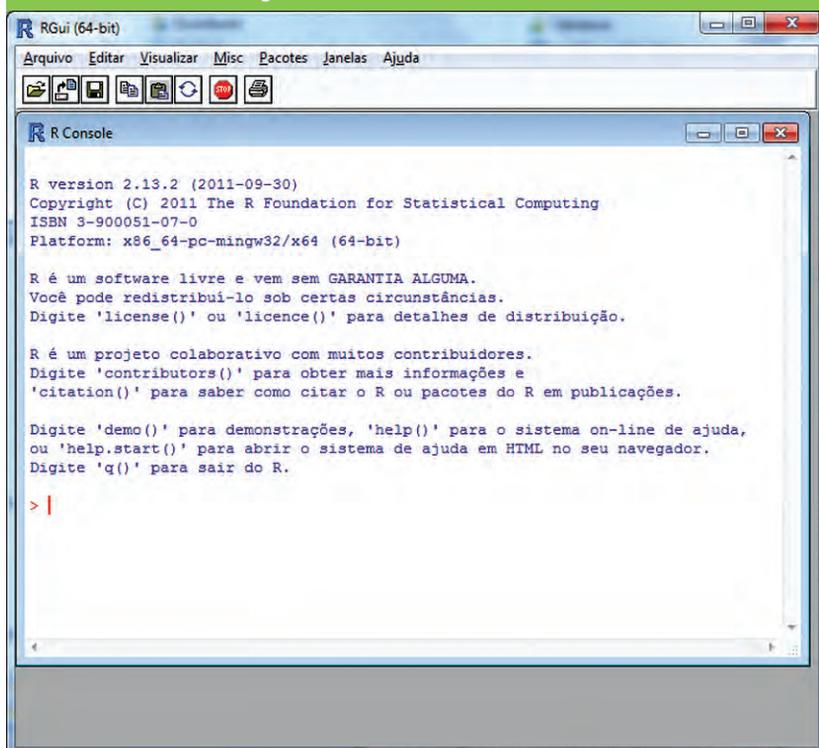


Fonte: Elaboração própria.

Após a escolha do servidor, faça o *download* do arquivo referente ao seu sistema operacional e em seguida clique em *install R for the first time*.² Agora basta seguir os passos da instalação. A Figura 10.2 apresenta a tela de abertura do R.

² Do inglês: "Instale o R pela primeira vez".

Figura 10.2 – Tela de abertura do R



Fonte: Elaboração própria.

Você notará que muito pouco pode ser feito no R por meio de menus. A maior parte dos comandos é digitada na tela após o sinal “>”, que indica que o R está pronto para receber comandos.

Neste livro, os comandos para o R serão apresentados em fonte courier negrito, antecidos pelo sinal “>”. A saída referente a este comando é apresentada em courier.

Exemplo: Obtendo detalhes sobre a licença do R.

```

> license()

This software is distributed under the terms of the GNU General
Public License, either Version 2, June 1991 or Version 3, June 2007.
The terms of version 2 of the license are in a file called COPYING
which you should have received with this software and which can be
displayed by RShowDoc("COPYING").

Copies of both versions 2 and 3 of the license can be found at
http://www.R-project.org/Licenses/.

A small number of files (the API header files listed in R_DOC_DIR/
COPYRIGHTS) are distributed under the Lesser GNU General Public
License version 2.1. This can be displayed by RShowDoc("COPYING.
LIB"), or obtained at the URI given.

'Share and Enjoy.'
```

Um tutorial bem completo sobre o R está disponível em <http://cran.r-project.org/doc/manuals/R-intro.pdf>. Na próxima seção, faremos um tutorial passo a passo, para que no final do capítulo você seja capaz de aplicar as ferramentas de DEA e Fronteira Estocástica.

10.2.3 Comandos básicos

Inicialmente vamos utilizar operações básicas do R para familiarização com o *software*. Na tela de comando, você pode inserir expressões matemáticas para que o R calcule.

```

> 2+2
[1] 4
> 5*(3+2)^2
[1] 125
>
```

Além de calcular expressões matemáticas, você pode desejar guardar o resultado de alguns cálculos. Para isso, será necessário criar *variáveis* e atribuir valores a elas. A atribuição de valores no R é feita com a expressão "<-" ou com "=".³

3 No R a expressão "=" é usada para atribuição de valores, enquanto a expressão "==" é utilizada para fazer a comparação (lógica) entre dois valores.

```

> a<- 2+2
> b<-5*(3+2)^2
> c<- b/a
> c
[1] 31.25

```

Outra forma de armazenar valores no R é por meio de vetores, que são sequências ordenadas de valores, que podem ser usadas em expressões algébricas. O R utiliza a função `c()` para concatenar valores em um vetor. Um exemplo de função algébrica é a soma de dois vetores:

```

> d <- c(3, 4, 7)
> d
[1] 3 4 7
> e<- c(2, 5, -1)
> e
[1] 2 5 -1
> d+e
[1] 5 9 6

```

Vetores podem ser combinados de forma a montar uma matriz com o comando `cbind()`, utilizado para agrupar os vetores coluna a coluna.⁴

```

> f <- cbind(d,e)
> f
      d e
[1,] 3 2
[2,] 4 5
[3,] 7 -1

```

Os elementos de uma matriz podem ser apresentados usando colchetes após o nome da matriz. O primeiro parâmetro dentro do colchete se refere à linha e o segundo, à coluna.

⁴ No manual de referência do R, encontrado em <http://cran.r-project.org/manuals.html>, existe uma lista completa de comandos para o R.

```

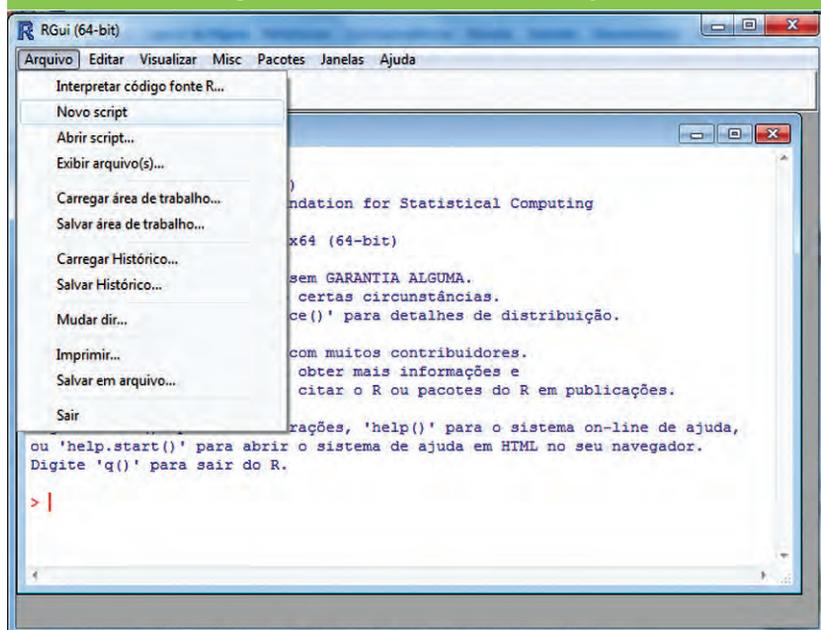
> f[3,2]
e
-1
> f[2,1]
d
4
-

```

Outro objeto que será importante para as nossas aplicações em R é o *data frame*, que é uma lista de vetores de mesmo tamanho. Quando importamos dados no R usando o comando `read.table()`, os dados são armazenados em *data frames*, como veremos a seguir.

Agora que você já está familiarizado com comandos básicos e é capaz de escrever seqüências de comandos, talvez seja necessário você guardar essas seqüências. Isso pode ser feito utilizando os *scripts*.

Figura 10.3 – Abrindo o editor de *scripts*



Fonte: Elaboração própria.

A janela *Editor R* pode ser utilizada como um editor de textos que armazenará a sequência de comandos que será executada pelo R. Os comandos podem ser executados selecionando a linha e clicando em “Executar linha ou seleção”, ou pressionando as teclas Ctrl + R.

Para utilizar o R, você precisará definir um diretório de trabalho. Este é o local onde ficarão armazenados os arquivos e os dados que você utilizará na sua sessão. Os nossos exemplos todos serão referentes ao diretório “C:\Eficiencia”. Para criar um diretório de trabalho no Windows, abra o Windows Explorer, clique com o botão direito do mouse “c:” e em seguida clique em “Nova Pasta”. Renomeie a pasta criada com o nome “Eficiencia”.

Ao abrir o R, você pode verificar qual o diretório de trabalho ao qual o R está conectado com o comando:

```
> getwd()  
[1] "C:/Windows/system32"  
>
```

O R está utilizando o diretório “c:\Windows\system32”. Será necessário alterar o diretório de trabalho para “C:\Eficiencia”. Porém, note que no R é necessário substituir a barra invertida “\” por uma barra normal “/”.

```
> setwd("c:/Eficiencia")  
>
```

Você pode utilizar novamente o comando `getwd()` para se certificar de que o diretório de trabalho foi alterado:

```
> getwd()  
[1] "c:/Eficiencia"  
>
```

Nosso próximo passo será importar os dados do arquivo “exemplo1.csv” para o R com o comando `read.table()`.⁵ No comando do R, será necessário informar o arquivo que contém os dados, se o arquivo tem cabeçalho (*header*), o separador entre as colunas (no nosso caso “;”) e o separador de decimais, que no nosso caso é uma vírgula.

```
> dados<-read.table("exemplo1.csv",header=TRUE,,sep=";",dec = ",")
>
```

Note que o comando acima cria um objeto chamado “dados”, e atribuímos a ele os valores lidos no arquivo exemplo1.csv. Esse objeto é um *data frame*. Para verificar se os dados foram importados corretamente, você pode usar o comando `list()` para apresentar os dados na tela de comandos do R, ou o comando `fix()` para apresentar em uma janela separada. Perceba que, no caso da janela separada, será necessário fechar a janela antes de voltar a trabalhar no R.

```
> list(dados)
[[1]]
  DMU  x1  x2  q
1   A 175  75 100
2   B 250  40 100
3   C 200  40 100
4   D 100  75 100
5   E  50 100 100
6   F 100  50 100
>
```

5 Note que o R não lê arquivos no formato padrão do Excel (.xls e .xlsx) diretamente, é necessário salvar o arquivo no formato .csv antes de fazer a importação. Isso pode ser feito no Excel utilizando a opção “Salvar como”.

Para obter estatísticas básicas sobre o objeto “dados” utilize o comando `summary()`:

```
> summary(dados)
  DMU      x1      x2      q
A:1  Min.   : 50.0  Min.   : 40.00  Min.   :100
B:1  1st Qu.:100.0  1st Qu.: 42.50  1st Qu.:100
C:1  Median :137.5  Median : 62.50  Median :100
D:1  Mean   :145.8  Mean   : 63.33  Mean   :100
E:1  3rd Qu.:193.8  3rd Qu.: 75.00  3rd Qu.:100
F:1  Max.   :250.0  Max.   :100.00  Max.   :100
>
```

Muitas funcionalidades do R não estão disponíveis logo após a instalação, sendo necessário instalar pacotes. Existem pacotes para diversas metodologias estatísticas e econométricas como Regressão em Painel, Programação Linear, Redes Neurais etc. Para instalar pacotes clique em “Pacotes - > Instalar Pacotes”. Nesse passo será necessário definir o *mirror*, que é um dos servidores espalhados pelo mundo que armazenam arquivos do R para *download*. Escolha o *mirror* mais próximo da sua localidade. Em seguida encontre o pacote que você procura e clique em ok.

O procedimento descrito copiará o pacote do *mirror* para o seu computador, no entanto é necessário carregar (instalar) para que ele funcione corretamente. Clique em “Pacotes - > Carregar pacote...”. Encontre o pacote que você precisa carregar e clique em ok.

Os procedimentos de instalar e carregar pacotes também podem ser feitos pela linha de comando usando as funções `install.packages()` e `library()`, respectivamente.

10.2.4 Pacotes para mensuração da eficiência

Frontier

O *software* denominado Frontier foi criado por Tim Coelli para ser utilizado nas estimações com máxima verossimilhança das fronteiras estocásticas de produção e função custo. Originalmente

o *software* foi escrito em Fortran e funcionava em modo comando com o auxílio de arquivos-texto contendo dados.

O Frontier foi adaptado na forma de um pacote para o R, em uma versão mais estável, com melhoria nos algoritmos numéricos e com a possibilidade de rodar em diversos sistemas operacionais que rodam o R. A sua utilização também passou a ser mais simples, já que não mais necessita de arquivos adicionais de dados; basta carregar os dados no R, conforme já descrito na seção anterior.

O Frontier possui implementações das fronteiras estocásticas com variação temporal da eficiência (BATTESE; COELLI, 1992) e o modelo no qual as firmas são influenciadas por variáveis ambientais (BATTESE; COELLI, 1995). Ele pode ser obtido nos repositórios do R, ou utilizando o comando *install.packages("frontier")*.

Benchmarking

O pacote Benchmarking foi desenvolvido por Bogetoft e Otto (2011) e contém métodos para análise de fronteiras. No caso do DEA, os métodos programados incluem diferentes hipóteses de retorno de escala, podendo ser orientados a insumo, produto e outras opções. O pacote Benchmarking também inclui algumas funcionalidades que utilizam o pacote FEAR, produzido por Paul W. Wilson,⁶ para fazer detecção de *outliers* e *bootstrap* nos dados.

O Benchmarking está disponível nos repositórios do R e pode ser instalado utilizando o comando *install.packages("Benchmarking")*. O FEAR não está disponível nos repositórios R e possui algumas restrições de licença para uso profissional ou no governo. O pacote pode ser obtido no site: <<http://www.clemson.edu/economics/faculty/wilson/Software/FEAR/fear.html>>, que também possui um roteiro de instalação.

6 Informações sobre o pacote FEAR podem ser obtidas no site: <<http://www.clemson.edu/economics/faculty/wilson/Software/FEAR/fear.html>>.

10.3 Aplicações de DEA com o pacote Benchmarking

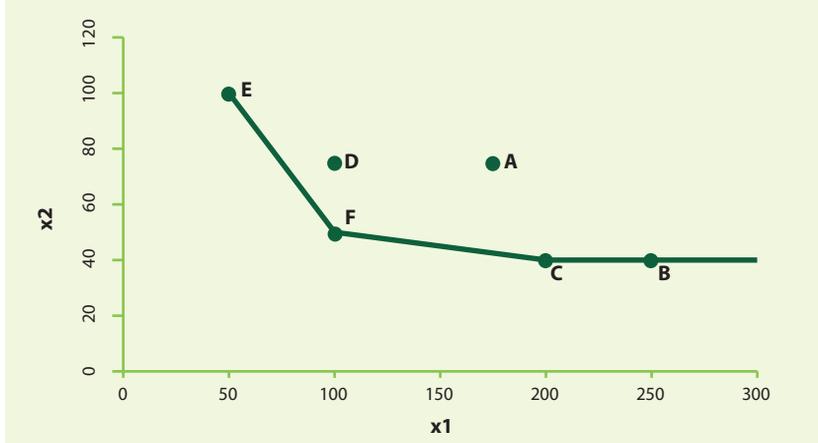
Nesta seção nós exploraremos as principais funções do pacote Benchmarking. Para iniciar os trabalhos, escreva o *script* a seguir no editor do R e execute selecionando o texto e clicando em “executar linha ou seleção”:

```
library(Benchmarking)
setwd("C:/Eficiencia")
dados<-read.table("exemplo1.csv",header=TRUE,,sep=";",dec = ",",)
x1 <-dados$x1
x2 <-dados$x2
x<-cbind(x1,x2)
y<- dados$y
dmu<-dados$DMU
```

10.3.1 DEA orientado a insumos

Essa sequência de comandos carrega o pacote Benchmarking, define o diretório de trabalho como “C:/Eficiencia”, em seguida faz a leitura dos dados do arquivo exemplo1.csv, armazena os dados dos insumos nas variáveis x1 e x2, concatena estes em uma matriz x, armazena os dados de produto no vetor y e, finalmente, armazena os “rótulos” das DMUs na variável DMU.

A função `dea.plot()` plota gráficos de tecnologias DEA, podendo ser aplicada na forma de isoquantas (2 insumos), curva de transformação (2 produtos), ou função de produção (1 insumo e 1 produto). Neste último caso, quando há mais de um insumo, a função de produção é plotada usando uma média ponderada de insumos.

Gráfico 10.1 – Isoquanta: `dea.plot.isoquant(x1,x2)`

Fonte: Elaboração própria.

Conforme vimos no capítulo 8, as DMUs D e A não são eficientes. Note que só foi possível plotar a isoquanta corretamente porque na nossa base de dados os produtos têm o mesmo valor em todas as DMUs.

Para estimar a fronteira DEA utilizaremos a função `dea()`. Para aplicar o DEA-CCR orientado a insumo aos dados já carregados, digite:

```
> dea(x,y, RTS="crs", ORIENTATION = "in", SLACK=TRUE)
[1] 0.6486 1.0000 1.0000 0.8571 1.0000 1.0000
```

O resultado mostra que a DMU A tem eficiência de 64,86%. Como o modelo é orientado a insumo, isso implica que seria possível produzir o mesmo nível de produto com apenas 64,86% dos insumos utilizados por essa DMU. Note que a DMU D também não é eficiente pelo DEA.

Os resultados da estimação DEA podem ser armazenados em um objeto e, nesse caso, basta fazer:

```
> ex_ccr <- dea(x,y, RTS="crs", ORIENTATION = "in", SLACK=TRUE)
>
```

O objeto chamado “ex_ccr” recebe várias informações a respeito do cálculo das eficiências. No capítulo 8, vimos que a DMU B, apesar de ter eficiência 100% e estar na fronteira, não é CCR-Eficiente. Isso ocorre porque a DMU B produz a mesma quantidade que a DMU C, mas gasta mais insumo x1. A informação da eficiência CCR pode ser obtida com os “slacks”.

```
> ex_ccr$slack  
[1] FALSE TRUE FALSE FALSE FALSE FALSE
```

Note que a única DMU com *slack* “TRUE” é a segunda, que corresponde à DMU B. Também está armazenado no objeto “ex_ccr” o conjunto de referência para cada DMU não eficiente, bem como o peso atribuído a cada uma delas:

```
> peers(ex_ccr)  
[1,]    3    6  
[2,]    3  NA  
[3,]    3  NA  
[4,]    5    6  
[5,]    5  NA  
[6,]    6  NA  
>
```

Perceba que as DMUs eficientes têm como referência elas mesmas, enquanto as não eficientes utilizam duas outras DMUs⁷. As proporções das DMUs de referências são obtidas a partir dos *lambdas*.

```
> lambda(ex_ccr)
      L3      L5      L6
[1,] 0.1351351 0.0000000 0.8648649
[2,] 1.0000000 0.0000000 0.0000000
[3,] 1.0000000 0.0000000 0.0000000
[4,] 0.0000000 0.2857143 0.7142857
[5,] 0.0000000 1.0000000 0.0000000
[6,] 0.0000000 0.0000000 1.0000000
>
```

Na primeira linha, os rótulos L3, L5, L6 correspondem às DMUs C, E e F, que se encontram na fronteira. A linha [1,] nos diz que a DMU A tem como conjunto de referência as DMUs C e F, nas proporções 0,14 e 0,86, respectivamente. A linha [4,] mostra que a DMU D tem como referência as DMUs E e F nas proporções 0,29 e 0,71, respectivamente.

10.3.2 DEA orientado a produtos

Execute o *script* a seguir para carregar os dados referentes ao exemplo 2.

```
library(Benchmarking)
setwd("C:/Eficiencia")
dados2<-read.table("exemplo2.csv",header=TRUE,,sep=";",dec = ",")
estudo <-dados2$estudo
q1 <-dados2$q1
q2<-dados2$q2
y<- cbind(q1,q2)
aluno<-dados2$aluno
```

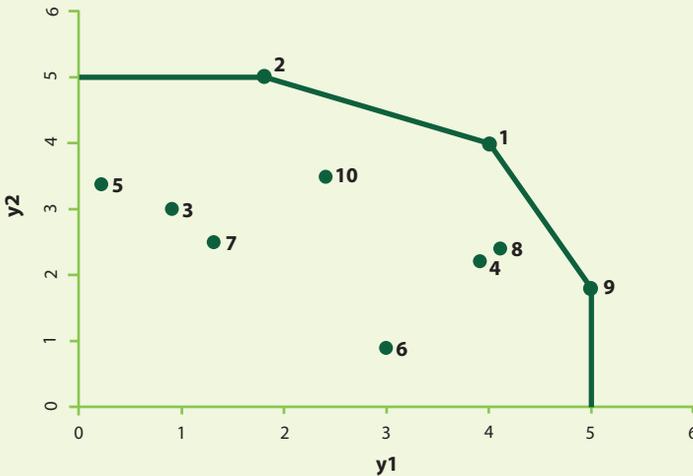
⁷ Com exceção da DMU B, posto que ela só tem como referência a DMU C.

Os dados são fictícios de uma turma de alunos, e apresentam o número de horas estudadas e as notas na questão 1 (q1) e na questão 2 (q2). O professor decidiu atribuir as notas finais do curso utilizando DEA pela ótica do produto. O comando `list()` apresenta os dados carregados:

```
> list(dados2)
[[1]]
  aluno estudo  q1  q2
1     1     10 4.0 4.0
2     2     10 1.8 5.0
3     3     10 0.9 3.0
4     4     10 3.9 2.2
5     5     10 0.2 3.4
6     6     10 3.0 0.9
7     7     10 1.3 2.5
8     8     10 4.1 2.4
9     9     10 5.0 1.8
10    10     10 2.4 3.5
>
```

Como todos os alunos têm o mesmo nível de insumo, ou seja, estudaram o mesmo número de horas, é possível plotar a curva de transformação.

Gráfico 10.2 – Curva de transformação: `dea.plot(q1,q2, RTS="crs",txt=aluno, ORIENTATION="out")`



Fonte: Elaboração própria.

Para obter as notas finais é necessário rodar o DEA-CCR com o comando:

```
> dea(estudo, y,RTS="crs", ORIENTATION="out", SLACK = TRUE)
[1] 1.000 1.000 1.667 1.187 1.471 1.667 1.882 1.121 1.000 1.267
```

Observe que as notas (eficiências) calculadas estão acima de 1 em alguns casos. Isso é porque estão sendo apresentadas as eficiências de Farrell. Para obter as eficiências convencionais que estudamos, primeiramente vamos guardar o resultado do DEA em um objeto denominado "ex_ccr2".

```
> ex_ccr2<- dea(estudo, y,RTS="crs", ORIENTATION="out", SLACK = TRUE)
[1] 1.000 1.000 1.667 1.187 1.471 1.667 1.882 1.121 1.000 1.267
```

Os resultados indicam o quanto as notas poderiam ser aumentadas, utilizando o mesmo nível de produto. No caso do aluno 3, este poderia ter obtido notas 67% superiores às que obteve. Para alcançar as notas na forma de eficiência, basta obter o recíproco da eficiência de Farrell. Para isso devemos executar:

```
> 1/eff(ex_ccr2)
[1] 1.0000000 1.0000000 0.6000000 0.8421875 0.6800000 0.6000000
0.5312500 0.8921875 1.0000000 0.7890625
```

No exemplo orientado a insumos, descobrimos que uma DMU não era CCR-Eficiente, pois possuía *slack*. Para fazer a mesma verificação utilizamos novamente o comando *slack*:

```
> ccr2$slack
[1] FALSE FALSE TRUE FALSE TRUE TRUE FALSE FALSE FALSE FALSE
```

Os alunos 3, 5 e 6 possuem *slacks*. Se analisarmos a Gráfico 10.1, a expansão radial dos alunos 3 e 5 até a fronteira levará a um ponto “escondido” atrás do aluno 2. A mesma expansão radial para o aluno 6 levará a um ponto também escondido atrás do aluno 9. Isso indica que ainda que as notas desses alunos fossem aumentadas, elas não seriam CCR-Eficientes, pois estariam dominadas por outros alunos capazes de obter notas maiores com o mesmo número de horas de estudo.

Para encontrarmos o conjunto de referência para os alunos que não estão na fronteira, ou seja, não tiveram nota final 10, usaremos novamente o comando `peers()`:

```
> peers(ex_ccr2)
      peer1 peer2
[1,]     1  NA
[2,]     2  NA
[3,]     2  NA
[4,]     1    9
[5,]     2  NA
[6,]     9  NA
[7,]     1    2
[8,]     1    9
[9,]     9  NA
[10,]    1    2
>
```

Note que apenas os alunos com nota máxima (1, 2 e 9) são referência para o cálculo da nota dos demais.

10.3.3 Modelo com retornos variáveis de escala

Até o momento trabalhamos apenas com o DEA-CCR, que assume retornos constantes de escala. Agora passaremos a utilizar também o DEA-BCC com retornos variáveis de escala e analisaremos como calcular os ganhos de escala.

Inicialmente, rode o *script* a seguir para carregar os dados do exemplo 3:

```
library(Benchmarking)
setwd("C:/Eficiencia")
dados3<-read.table("exemplo3.csv",header=TRUE,,sep=";",dec = ",")
x1 <-dados3$x1
x2 <-dados3$x2
q<-dados3$q
x<- cbind(x1,x2)
DMU<-dados3$DMU
```

O *script* lê os dados do arquivo exemplo3.csv e os grava no objeto dados3. Em seguida, define os vetores de insumo x1 e x2, o vetor de produtos q, cria a matriz de insumos x e o vetor de rótulos DMU. Vamos calcular a eficiência orientada a insumos com as hipóteses de retornos constantes de escala ("crs") e retornos variáveis de escala ("vrs").

```
> ex_ccr3<- dea(x, q,RTS="crs", ORIENTATION="in", SLACK = TRUE)
> eff(ex_ccr3)
[1] 0.1000000 0.1500000 0.2000000 0.2500000 0.9800000 0.5833333
0.5714286 0.7500000 0.8333333 1.0000000
> ex_bcc3<- dea(x, q,RTS="vrs", ORIENTATION="in", SLACK = TRUE)
> eff(ex_bcc3)
[1] 1.0000000 0.5833333 0.4722222 0.4375000 1.0000000 0.6388889
0.6071429 0.7598039 0.8387800 1.0000000
>
```

Esses códigos calcularam as fronteiras DEA com as duas hipóteses de retorno, armazenaram os resultados nos objetos "ex_ccr3" e "ex_bcc3" e as eficiências foram obtidas usando o comando `eff()`. Note que a eficiência BCC, medida por retornos variáveis, é sempre maior que a eficiência CCR.

As eficiências de escala podem ser obtidas dividindo as eficiências ccr pelas bcc:

```
> scale <- eff(ex_ccr3)/eff(ex_bcc3)
> scale
[1,] 0.1000000 0.2571429 0.4235294 0.5714286 0.9800000 0.9130435
0.9411765 0.9870968 0.9935065 1.0000000
>
```

Para obter a Tabela 8.7 basta combinar as colunas com o comando `cbind()`.

```
> cbind(eff(ex_ccr3), eff(ex_bcc3), scale)
              scale
[1,] 0.1000000 1.0000000 0.1000000
[2,] 0.1500000 0.5833333 0.2571429
[3,] 0.2000000 0.4722222 0.4235294
[4,] 0.2500000 0.4375000 0.5714286
[5,] 0.9800000 1.0000000 0.9800000
[6,] 0.5833333 0.6388889 0.9130435
[7,] 0.5714286 0.6071429 0.9411765
[8,] 0.7500000 0.7598039 0.9870968
[9,] 0.8333333 0.8387800 0.9935065
[10,] 1.0000000 1.0000000 1.0000000
>
```

O resultado mostra que a DMU 1 tem eficiência de 10% sob a hipótese de retornos constantes e eficiência de 100% sob a hipótese de retornos variáveis. Isso indica que o problema de eficiência observado deve-se à escala utilizada.

10.3.4 Painel de dados em DEA

Conforme vimos no capítulo 8, é possível trabalhar com painel de dados em DEA. Esse tipo de painel difere bastante no conceito de painel de dados utilizado em econometria, tendo como principal característica o fato de permitir apenas dois períodos e não ser necessário que as mesmas DMUs estejam em ambos os

períodos analisados. Para trabalhar com DEA em painel de dados, inicialmente rode o *script* a seguir:

```
library(Benchmarking)
setwd("C:/Eficiencia")
paine1<-read.table("exemplo4.csv",header=TRUE,,sep=";",dec = ",",")
x1s <-paine1$x1s
x2s <-paine1$x2s
q1s <-paine1$q1s
q2s <-paine1$q2s
x1t <-paine1$x1t
x2t <-paine1$x2t
q1t <-paine1$q1t
q2t <-paine1$q2t
xs <- cbind(x1s, x2s)
qs <- cbind(q1s, q2s)
xt <- cbind(x1t, x2t)
qt <- cbind(q1t, q2t)
DMU <-paine1$DMU
```

O *script* carrega os dados do arquivo "exemplo4.csv", em seguida cria os vetores com os insumos e produtos nos períodos t e s. Finalmente, o código cria matrizes de insumo e de produto para cada um dos períodos e cria um vetor de rótulos. Para verificar se os dados foram carregados corretamente, você pode usar o comando `list()`.

```
> list(paine1)
[[1]]
   DMU x1s x2s q1s q2s x1t x2t q1t q2t
1    A 110 270  31   8 111 277  33   9
2    B  88 247  32  18  87 258  30  16
3    C  84 176  27  17  81 182  35  21
4    D  59 232  24  18  59 240  24  17
5    E 139 218  36  18 138 222  40  18
6    F  86 178  27  18  86 185  27  21
7    G  79 211  30  18  81 201  32  18
8    H  99 131  25  13  98 135  27  14
9    I  58  55  17   6  58  55  21   7
10   J  80 235  27  11  79 245  31  14
>
```

Nosso primeiro passo será calcular a variação da eficiência técnica. Para isso, vamos rodar o DEA BCC orientado a produto para o período t e para o período s da seguinte forma, e em seguida calcularemos o fator multiplicativo da mudança de produtividade:

```
> MetNum<-dea(xt,qt,RTS="vrs",XREF=xt,YREF=qt, ORIENTATION="out")
> MetDen<-dea(xs,qs,RTS="vrs",XREF=xs,YREF=qs, ORIENTATION="out")
> Met <-eff(MetNum)^-1 /eff(MetDen)^-1
> Met
[1] 0.9540213 0.8444444 1.0027831 1.0000000 1.0000000 1.0000000
0.9142857 0.9046365 1.0000000 1.0205761
>
```

Note que o comando que utilizamos para o DEA é um pouco diferente dos anteriores. Os parâmetros XREF e YREF são empregados com a função DEA para informar quais vetores de insumos e produtos estamos usando para delimitar a tecnologia. No primeiro caso, estamos utilizando os insumos e produtos no período t , com a tecnologia também do período t , e armazenando os resultados na variável MetNum. Em seguida fazemos o mesmo para o período s e armazenamos em MetDen. Finalmente, calculamos a variação da eficiência técnica dividindo as eficiências do período t pelas eficiências do período s e armazenamos na variável Met.

Nosso próximo passo é calcular o índice de Malmquist para modificações na tecnologia. Para isso executamos o seguinte código:

```
> MtNum1<-dea(xs,qs,RTS="vrs",XREF=xs,YREF=qs, ORIENTATION="out")
> MtNum2<-dea(xt,qt,RTS="vrs",XREF=xs,YREF=qs, ORIENTATION="out")
> MtDen1<-dea(xs,qs,RTS="vrs",XREF=xt,YREF=qt, ORIENTATION="out")
> MtDen2<-dea(xt,qt,RTS="vrs",XREF=xt,YREF=qt, ORIENTATION="out")
> Mt<-sqrt(eff(MtNum1)^-1*eff(MtNum2)^-1/eff(MtDen1)^-1/
eff(MtDen2)^-1)
> Mt
[1] 1.1132251 1.1154470 1.2281388 0.9597857 1.0867494 1.1543532
1.1601028 1.1757321 1.2004901 1.1374313
>
```

A primeira linha calcula as eficiências do período s com a tecnologia do período s . A segunda calcula as eficiências do período t com a tecnologia do período s e assim por diante. Finalmente, para calcular o índice para modificações na tecnologia utilizamos a função `sqrt()` para a raiz quadrada e a fórmula 8.16 do capítulo 8.

Para calcular o índice de Malmquist para eficiências de escala, precisamos calcular as eficiências com outras variações de insumos e produto, conforme o código:

```
> MeeNum1<-dea(xt,qt,RTS="crs",XREF=xs,YREF=qs,ORIENTATION="out")
> MeeNum2<-dea(xs,qt,RTS="vrs",XREF=xs,YREF=qs,ORIENTATION="out")
> MeeNum3<-dea(xt,qs,RTS="crs",XREF=xt,YREF=qt,ORIENTATION="out")
> MeeNum4<-dea(xs,qs,RTS="vrs",XREF=xt,YREF=qt,ORIENTATION="out")
> MeeDen1<-dea(xt,qt,RTS="vrs",XREF=xs,YREF=qs,ORIENTATION="out")
> MeeDen2<-dea(xs,qt,RTS="crs",XREF=xs,YREF=qs,ORIENTATION="out")
> MeeDen3<-dea(xt,qs,RTS="vrs",XREF=xt,YREF=qt,ORIENTATION="out")
> MeeDen4<-dea(xs,qs,RTS="crs",XREF=xt,YREF=qt,ORIENTATION="out")
> Mee = sqrt( (eff(MeeNum1)^-1 * eff(MeeNum2)^-1 * eff(MeeNum3)^-1
* eff(MeeNum4)^-1) / (eff(MeeDen1)^-1 / eff(MeeDen2)^-1 / eff(MeeDen3)^-1 /
eff(MeeDen4)^-1) )
> Mee
[1] 0.9905393 1.0022893 1.0071782 1.0087588 0.9916549 0.9831050
0.9933662 0.9892550 1.0000000 0.9970614
>
```

Note que para determinar esse índice calculamos eficiências alternando insumos dos períodos com produtos do período t , usando tecnologia de ambos os períodos, além de usarmos eficiências CCR e BCC no cálculo.

O índice para modificação das economias de escopo é calculado de forma semelhante, seguindo a fórmula 8.18:

```
> MeescNum1<-dea(xs,qt,RTS="crs",XREF=xs,YREF=qs,ORIENTATION="out")
> MeescNum2<-dea(xs,qs,RTS="vrs",XREF=xs,YREF=qs,ORIENTATION="out")
> MeescNum3<-dea(xt,qt,RTS="crs",XREF=xt,YREF=qt,ORIENTATION="out")
> MeescNum4<-dea(xt,qs,RTS="vrs",XREF=xt,YREF=qt,ORIENTATION="out")
> MeescDen1<-dea(xs,qt,RTS="vrs",XREF=xs,YREF=qs,ORIENTATION="out")
> MeescDen2<-dea(xs,qs,RTS="crs",XREF=xs,YREF=qs,ORIENTATION="out")
> MeescDen3<-dea(xt,qt,RTS="vrs",XREF=xt,YREF=qt,ORIENTATION="out")
> MeescDen4<-dea(xt,qs,RTS="crs",XREF=xt,YREF=qt,ORIENTATION="out")
> Meesc = sqrt( (eff(MeescNum1)^-1*eff(MeescNum2)^-1*eff(MeescNum3)^-1
*eff(MeescNum4)^-1/eff(MeescDen1)^-1/eff(MeescDen2)^-1/
eff(MeescDen3)^-1/eff(MeescDen4)^-1)
> Meesc
[1] 1.0000000 1.0000000 0.9943353 1.0000000 0.9854438 1.0000000
1.0000000 0.9975504 1.0000000 1.0000000
>
```

O último índice a ser calculado é a modificação da produtividade total. Existem duas formas de calcular esse índice, sendo a primeira: pelo produto dos índices de economia de escopo, eficiência de escala, modificação na tecnologia e variação da eficiência técnica. Neste caso:

```
> Mpt1 = Met * Mt * Mee * Meesc
> Mpt1
[1] 1.0519929 0.9440894 1.2333706 0.9681923 1.0619934 1.1348504
1.0536292 1.0496043 1.2004901 1.1574241
>
```

A segunda forma é pela utilização da fórmula 8.17, que deriva do produto dos índices, mas são realizados alguns cancelamentos. Neste caso:

```
MptNum1<-dea(xt,qt,RTS="crs",XREF=xs,YREF=qs,ORIENTATION="out")
> MptNum2<-dea(xt,qt,RTS="crs",XREF=xt,YREF=qt,ORIENTATION="out")
> MptDen1<-dea(xs,qs,RTS="crs",XREF=xs,YREF=qs,ORIENTATION="out")
> MptDen2<-dea(xs,qs,RTS="crs",XREF=xt,YREF=qt,ORIENTATION="out")
> Mpt2<-sqrt( (eff(MptNum1)^-1 * eff(MptNum2)^-1 / eff(MptDen1)^-1 /
eff(MptDen2)^-1 )
> Mpt2
[1] 1.0519929 0.9440894 1.2333706 0.9681923 1.0619934 1.1348504
1.0536292 1.0496043 1.2004901 1.1574241
>
```

Para obter os dados da Tabela 8.8 basta concatenar os vetores usando a fórmula `cbind()`. Também arredondaremos os valores na segunda casa decimal com a fórmula `round()`.

```
> tabela<- cbind(Mpt1, Mpt2, Met, Mt, Mee, Meesc)
> round(tabela,2)
      Mpt1 Mpt2  Met   Mt  Mee Meesc
[1,] 1.05 1.05 0.95 1.11 0.99 1.00
[2,] 0.94 0.94 0.84 1.12 1.00 1.00
[3,] 1.23 1.23 1.00 1.23 1.01 0.99
[4,] 0.97 0.97 1.00 0.96 1.01 1.00
[5,] 1.06 1.06 1.00 1.09 0.99 0.99
[6,] 1.13 1.13 1.00 1.15 0.98 1.00
[7,] 1.05 1.05 0.91 1.16 0.99 1.00
[8,] 1.05 1.05 0.90 1.18 0.99 1.00
[9,] 1.20 1.20 1.00 1.20 1.00 1.00
[10,] 1.16 1.16 1.02 1.14 1.00 1.00
>
```

10.3.5 Modelo de dois estágios

Conforme vimos no capítulo 8, o modelo de dois estágios pode ser utilizado em modelos com variáveis ambientais, com um método semiparamétrico. Nesta seção vamos utilizar uma base com dados de

fazendeiros produtores de madeira na Noruega. Essa base de dados pode ser carregada a partir do pacote Benchmarking.

```
> library(Benchmarking)
Carregando pacotes exigidos: lpSolveAPI
Carregando pacotes exigidos: ucminf
> data(norWood2004)
> d<-norWood2004
> d
[[1]]
      firm      x      z1      m      y z3 z6
1       1 600000 848470 123109 163366 52  0
2       2 254000 2486106 30962  14728 45  0
...
112    112 127500 415650 81159  55212 38  0
113    113 302200 806046 85073   9362 50  0
```

O *data frame* “norWood2004” possui 113 observações das variáveis:

- firm: Nome da firma
- m: Custo variável
- x: Terras (valor da área cultivada e da terra)
- y: Lucro
- z1: Renda secundária
- z3: Idade do proprietário
- z6: Variável binária que indica a existência de planejamento de longo prazo (1= existe planejamento, 0 = não existe planejamento).

Vamos seguir o algoritmo criado por Simar e Wilson (2007), apresentado no capítulo 8.

1) Utilizando o DEA, só com variáveis (*inputs* e *outputs* controláveis), obtenha os escores de eficiência θ_i :

```
> insumos<- cbind(d$x, d$m)
> produtos <- d$y
> estagiol<-dea(insumos,produtos, RTS="vrs", ORIENTATION = "in",
  SLACK=TRUE)
> d$theta<-eff(estagiol)
```

2) Retire as DMUs que se mostrarem eficientes ($\theta_i = 1$) no passo 1:

```
> d<-subset(d, theta!=1)
```

3) Obtenha os escores inversos fazendo ($\eta_i = 1/\theta_i$):

```
> d$eta=1/d$theta
```

4) Utilize o modelo Tobit (ou outro modelo equivalente) para regredir os escores de eficiência obtidos contra as variáveis ambientais, mas utilizando somente os registros das DMUs não eficientes, de acordo com a seguinte equação: ($\eta_i = z_i \gamma + \varepsilon_i$), onde z_i é o vetor de variáveis ambientais associado a DMU i , γ é o vetor de coeficientes a ser estimado e ε_i é o erro aleatório também associado a DMU i . Obtenha estimativas para γ , ε e para o desvio padrão de ε (S_ε);

```
> install.packages("AER")
> library(AER)
> d$zz1<-d$z1/1e6
> etobit<-tobit(eta~zz1+z3+z6, left=1, right=Inf, data=d)
> d$e<-residuals(etobit)
> s_e<-sqrt(var(d$e))
> d$predito<-fitted(etobit)
```

5) Produza resíduos artificiais (e_i) gerados a partir de uma distribuição normal truncada, com truncamento à esquerda⁸ em $(1 - z_i; \gamma)$ e com desvio padrão igual a s_{e_i} , que foi estimado no passo 4:

```
> install.packages("truncnorm")
> library(truncnorm)
> d$e_artif<-rtruncnorm(1, a=1-d$predito, b=Inf, mean = 0, sd = s_e)
```

6) Compute a variável η_i^* segundo a equação: $\eta_i^* = z_i g + e_i$, onde g é o estimador para o parâmetro γ :

```
> d$eta_est<-d$predito+d$e_artif
```

7) Estime mais uma vez utilizando o modelo Tobit e as DMUs acima não eficientes, só que agora utilizando os valores obtidos da equação 8.24 como variável endógena, e as variáveis exógenas como variáveis explicativas $\eta_i^* = z_i \gamma^* + \omega_i$, onde agora ω_i é o erro aleatório.

```
> etobit_novo<-tobit(eta_est~zz1+z3+z6, left=1, right=Inf, data=d)
```

8) Obtenha as estimativas para γ^* e para o desvio padrão do erro ω , (g^* e s_ω).

```
> gamma_est<-coef(etobit_novo)
> s_w<-sqrt(var(residuals(etobit_novo)))
```

8 Pois se $e_i < 1 - z_i \gamma \Rightarrow z_i \gamma + e_i < 1 \Rightarrow \eta_i < 1$, fato que contrariaria a teoria.

9) Repita os passos 5, 6, 7 e 8 L^9 vezes, de modo a obter a matriz $G = \left[g_{\ell}^*, s_{\omega, \ell} \right]_{\ell=1}^L$

Inicialmente vamos definir a matriz G e atribuir os valores dos coeficientes e do desvio padrão a ela:

```
> G<-cbind(rep(0,100),rep(0,100),rep(0,100),rep(0,100),rep(0,100))
> G[1,]<-cbind(t(gamma_est), s_w)
Para preencher toda a matriz G, executaremos o loop:
L=100
G<-cbind(rep(0,L),rep(0,L),rep(0,L),rep(0,L),rep(0,L))
for (i in 1:L){
d$e_artif<-rtruncnorm(1, a=1-d$predito, b=Inf, mean = 0, sd = s_e)
d$eta_est<-d$predito+d$e_artif
etobit_novo<-tobit(eta_est~zz1+z3+z6, left=1, right=Inf, data=d)
gamma_est<-coef(etobit_novo)
s_w<-sqrt(var(residuals(etobit_novo)))
G[i,]<-cbind(t(gamma_est), s_w)
}
```

9 Em geral, utiliza-se $L = 1000$.

10) Calcule as médias e variâncias de cada coluna G para construir intervalos de confiança para os parâmetros.

```
> medias<-rep(0,5)
> desvios<-rep(0,5)
> intervalo<-data.frame(cbind(rep(0,5),rep(0,5)))
> names(intervalo)<-cbind("Inferior", "Superior")
> rownames(intervalo)<-rbind("Intercepto", "zz1", "z3", "z6",
"Desvio do Erro")
> for(i in 1:5){
+ medias[i]<-mean(G[,i])
+ desvios[i]<-sqrt(var(G[,i]))
+ erro <- qnorm(0.95)*desvios[i]/sqrt(L)
+ intervalo[i,1]<- medias[i]-erro
+ intervalo[i,2]<- medias[i]+erro
+ }
> intervalo
```

	Inferior	Superior
Intercepto	6.99628833	7.62079517
zz1	-0.41871938	-0.15974810
z3	-0.02796693	-0.01527041
z6	0.30049021	0.53936585
Desvio do Erro	3.39650445	3.48331000

```
>
```

11) Calcule a média de cada coluna para obter a estimativa dos efeitos das variáveis ambientais sobre a eficiência das DMUs.

```
> names(medias)<-cbind("Intercepto", "zz1", "z3", "z6", "Desvio do
Erro")
> medias
```

	zz1	z3	z6	Desvio do Erro
Intercepto	7.30854175	-0.28923374	-0.02161867	0.41992803
	3.43990722			

10.3.6 Exercícios

- 1 - Utilizando os dados do arquivo exercicio1.csv:
 - a) Calcule as eficiências DEA orientadas a produto com retornos constantes de Escala.
 - b) Identifique os conjuntos de referência para cada DMU.
 - c) Identifique se existem DMUs eficientes que não são CCR-Eficientes.
- 2 - Utilizando os dados do arquivo exercicio2.csv:
 - a) Calcule as eficiências DEA orientadas a insumo com retornos constantes de Escala.
 - b) Identifique os conjuntos de referência para cada DMU.
 - c) Identifique se existem DMUs eficientes que não são CCR-Eficientes.
 - d) Calcule as eficiências DEA com retornos variáveis de escala.
 - e) Calcule as eficiências de escala.
- 3 - O artigo de Charnes et al. (1981) utiliza a base de dados de um programa federal nos Estados Unidos que fornece assistência a estudantes desfavorecidos em diversas localidades americanas. As DMUs são 70 escolas nos Estados Unidos e as variáveis utilizadas são:
 - Y1: Nota de Leitura
 - Y2: Nota de Matemática
 - Y3: Nota de autoestima
 - X1: Nível de educação da mãe
 - X2: Ocupação do membro da família
 - X3: Visitas dos pais à escola
 - X4: Tempo gasto com as crianças em assuntos relacionados à escola
 - X5: Número de Professores

Pft: Variável que identifica se a escola está no programa

Não é necessário carregar os dados de um arquivo .csv. Esta base de dados faz parte do pacote Benchmarking. Para carregar, rode o *script* a seguir:

```
library(Benchmarking)
data(charnes1981)
x1<- charnes1981$x1
x2<- charnes1981$x2
x3<- charnes1981$x3
x4<- charnes1981$x4
x5<- charnes1981$x5
y1<- charnes1981$y1
y2<- charnes1981$y2
y3<- charnes1981$y3
pft<- charnes1981$pft
x<-cbind(x1,x2,x3,x4,x5)
y<-cbind(y1,y2,y3)
```

- Calcule as eficiências DEA com retornos constantes de Escala.
- Identifique os conjuntos de referência para cada DMU.
- Identifique se existem DMUs eficientes que não são CCR-Eficientes.
- Calcule as eficiências DEA com retornos variáveis de escala.
- Calcule as eficiências de escala.

10.4 Aplicações de fronteira estocástica com o pacote Frontier

Nesta seção nós exploraremos as principais funções do pacote Frontier. Para iniciar os trabalhos, execute o *script* a seguir a partir do editor do R:

```
library(frontier)
setwd("C:/Eficiencia")
dados<-read.table("exemplo5.csv",header=TRUE,,sep=";",dec = ",")
list(dados)
```

10.4.1 Fronteiras *cross section*

Esta sequência de comandos carrega o pacote Frontier, define o diretório de trabalho como "C:/Eficiencia", em seguida faz a leitura dos dados do arquivo exemplo5.csv e os armazena no objeto "dados". Para verificar se os dados foram carregados corretamente, utilizamos a função `list()` para apresentá-los na tela.

Vamos começar estimando uma fronteira de produção com dados *cross section*, ou seja, temos dados de diversas DMUs em um mesmo instante de tempo. A fronteira será estimada por máxima verossimilhança utilizando a função `sfa()`.

A sintaxe da função `sfa()` é bastante diferente do comando `dea()` que utilizamos na seção anterior. Nas fronteiras estocásticas precisamos informar qual equação estamos estimando, em qual objeto estão os dados e se é uma fronteira de produção ou de custo, ou seja, se o termo de ineficiência aumenta ou diminui a variável dependente. O comando utilizado no R seguirá a seguinte forma geral:

```
sfa( [equação], [dados], [produção ou custo]).
```

No exemplo a seguir vamos estimar uma fronteira de produção tipo Cobb Douglas. Para isso, precisaremos fazer a estimação em logaritmos, usando a função `log()`.

```
> ex_cross <- sfa(log(q)~log(x1) + log(x2), data = dados,
ineffDecrease = TRUE)
>
```

A equação estimada equivale à forma funcional $q = Ax_1^\alpha x_2^\beta$. Os dados de q , x_1 e x_2 estão armazenados no *data frame* "dados" e como se trata de uma função de produção, o termo de ineficiência diminui a variável dependente (ineffDecrease = True). Os resultados foram gravados no objeto `ex_cross`. Para visualizar os resultados da estimação, digite:

```
> summary(ex_cross)
Error Components Frontier (see Battese & Coelli 1992)
Inefficiency decreases the endogenous variable (as in a production
function)
The dependent variable is logged
Iterative ML estimation terminated after 9 iterations:
log likelihood values and parameters of two successive iterations
are within the tolerance limit

final maximum likelihood estimates
              Estimate Std. Error z value Pr(>|z|)
(Intercept) 0.057149   1.083779  0.0527  0.95795
log(x1)      0.354749   0.184126  1.9267  0.05402 .
log(x2)      0.846055   0.166188  5.0909 3.563e-07 ***
sigmaSq      0.443650   0.162405  2.7318  0.00630 **
gamma        0.859115   0.153627  5.5922 2.242e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
log likelihood value: -29.49461

cross-sectional data
total number of observations = 50
mean efficiency: 0.6463099
```

Observe que o coeficiente estimado para o intercepto não é significativamente diferente de zero, dado que o p-valor obtido a partir de uma distribuição-t sob a hipótese nula de que o coeficiente é zero é igual 0,95795. Os demais coeficientes estimados são todos estatisticamente significativos ao nível de 10%. O Frontier também informa o valor da log-verossimilhança (-29,49461) e calcula as estatísticas `sigmaSq` e `gamma`, conforme as equações 8.10 e 8.9 do capítulo 8. De acordo com o resultado, a soma da variância do termo de eficiência e do termo de choque aleatório é 0,443650, sendo

estatisticamente diferente de zero, com nível de significância de 1%. Já a razão entre os desvios padrão do componente de choque e do componente de eficiência é 0,859115, sendo estatisticamente diferente de zero, com nível de significância inferior a 0,1%.

A eficiência média estimada a partir da fronteira de produção é 0,6463. Para obter as eficiências de cada DMU digite:

```
> efficiencias (cross)
  efficiency
1  0.7696421
2  0.8958975
3  0.7161633
...
48 0.3716450
49 0.8741082
50 0.7038094
>
```

10.4.2 Fronteiras em painel

Na seção anterior estimamos a fronteira a partir de dados *cross section*, formados pelas informações de 50 DMUs no mesmo instante de tempo. No entanto, é possível fazermos uma extensão para um painel de dados, ou seja, com dados de várias DMUs em vários instantes de tempo.

Escreva o *script* a seguir no editor do R:

```
library(frontier)
setwd("C:/Eficiencia")
dados2<-read.table("exemplo6.csv",header=TRUE,,sep=";",dec = ",")
list(dados2)
```

Essa sequência de comandos carrega o pacote Frontier, define o diretório de trabalho para "C:/Eficiencia", carrega os dados do arquivo exemplo6.csv no *data frame* "dados2" e utiliza a função list() para apresentar os dados na tela.

Com os dados em tela, você verificará que temos um total de 60 observações que foram obtidas de 12 DMUs em cinco períodos de tempo.

É necessário informar ao R que estamos trabalhando com um painel de dados, definindo qual variável indica o instante de tempo e qual variável identifica a DMU. Para isso utilizamos o comando `plm.data()`:

```
> dados2_painel <-plm.data(dados2,c("DMU","TIME"))
> dados2_painel
  DMU TIME      x1      x2      q
1    1   1  5.374113  5.472774  5.997639
13   1   2  6.697868  4.775656 10.953230
25   1   3  6.250772  3.470539 12.848809
37   1   4  4.760019  5.100358 17.527600
49   1   5  5.407279  6.244785 11.576708
2    2   1  5.301897  5.575616 15.117391
14   2   2  6.780184  7.235184 19.357388
26   2   3  8.731247  8.464446 36.432030
38   2   4  8.674794  9.440677 16.710627
50   2   5  9.930518 11.257279 59.266039
...
12  12   1  5.146620  5.919449  3.051909
24  12   2  6.652940  7.590464  3.504011
36  12   3  5.549627  7.461824  4.614911
48  12   4  4.443723  8.293831  1.848757
60  12   5  4.388440  7.817372  3.973715
```

Note que o comando `plm.data()` reorganizou os dados em um novo *data frame*, agora ordenado pelas DMUs. Para estimar a fronteira de produção com painel de dados utilize:

```
> painel <- sfa(log(q) ~ log(x1) + log(x2), data = dados2_painel,
ineffDecrease = TRUE)
> summary(painel)
Error Components Frontier (see Battese & Coelli 1992)
Inefficiency decreases the endogenous variable (as in a production
function)
The dependent variable is logged
Iterative ML estimation terminated after 11 iterations:
log likelihood values and parameters of two successive iterations
are within the tolerance limit

final maximum likelihood estimates

              Estimate Std. Error z value Pr(>|z|)
(Intercept)  0.43953    0.45668  0.9624  0.335830
log(x1)      0.74149    0.23769  3.1195  0.001811 **
log(x2)      0.58052    0.19580  2.9649  0.003028 **
sigmaSq      0.60933    0.23732  2.5676  0.010241 *
gamma        0.72316    0.11824  6.1162  9.584e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
log likelihood value: -41.68798

panel data
number of cross-sections = 12
number of time periods = 5
total number of observations = 60
thus there are 0 observations not in the panel

mean efficiency: 0.659408
>
```

Os resultados da estimação foram guardados no objeto “painel”. Um resumo da estimação pode ser obtido utilizando a função `summary()`. Note que todos os coeficientes, à exceção do intercepto, são significativos ao nível de 5%. O resultado também mostra que o painel foi especificado corretamente, com doze *cross sections* (DMUs) e cinco períodos de tempo. O valor da log verossimilhança e a eficiência média também são apresentados. Para obter as eficiências individuais utilize novamente o comando `efficiencies()`.

```
> efficiencies (painel)
  efficiency
1  0.7808186
2  0.8890732
3  0.7914690
4  0.7527375
5  0.7532368
6  0.7931134
7  0.8258926
8  0.6636512
9  0.6216153
10 0.3193467
11 0.5003255
12 0.2216167
>
```

10.4.3 Exercícios

1 - O pacote Frontier contém dados de provedores de serviços de telecomunicações de 21 países no ano de 1990. Para carregar o *data frame* você deve digitar:

```
> data(telecom)
>
```

Um *data frame* denominado telecom será carregado, contendo as seguintes informações:

```
Country: Nome do país.
Output: índice de produto.
Mainlines: linhas de transmissão (mil km).
Employees: número de empregados (10.000 pessoas).
```

- a) Os dados apresentados formam um modelo *cross section* ou painel? Justifique.
- b) Estime uma fronteira utilizando o índice de produto como produto, e como insumo: linhas de transmissão e número de empregados, usando uma forma Cobb-Douglas.
- c) Quais coeficientes são significativos ao nível de 10%?
- d) Qual a eficiência média dos países?
- e) Identifique os países com maior e menor eficiência.

2 - O pacote Frontier contém os dados coletados de 43 pequenos produtores de arroz das Filipinas no período de 1990 e 1997. Para carregar os dados você deve digitar:

```
> data(riceProdPhil)
>
```

Esse comando carregará o *data frame* “riceProdPhil” que contém as seguintes variáveis:

YEARDUM: Período (1= 1990, ..., 8 = 1997).

FMERCODE: Fazendeiro/DMU (1, ..., 43).

PROD *Output*: Produção de arroz em toneladas.

AREA: Área plantada (hectares).

LABOR: Mão de obra (homens/dia).

NPK: Fertilizantes (Kg).

OTHER: Outros insumos (índice de Laspeyres index = 100 para DMU 17 em 1991).

PRICE: Preço do kg de arroz (pesos por kg)

AREAP: Custo de aluguel da terra (pesos por hectare).

LABORP: Custo da mão de obra (pesos por homens/dia).

NPKP: Preço do fertilizante (pesos por kg).

OTHERP: Preço de outros insumos (preço implícito do índice).

AGE: Idade do chefe de família (anos).

EDYRS: Grau de instrução do chefe de família (anos de estudo).

HHSIZE: Número de pessoas na unidade familiar.

NADULT: Número de adultos na unidade familiar.

BANRAT: Percentual da área classificada como montanhosa.

Lembre-se de que esses dados compõem um painel, logo você deverá usar o comando `plm.data()` para criar um *data frame* em formato de painel.

- a) Estime uma fronteira de produção usando uma forma Cobb-Douglas, tendo a Produção de Arroz como produto e como insumo: mão de obra, área plantada, fertilizantes e o índice de outros insumos;
- b) Identifique quais coeficientes são significativos ao nível de 10%;
- c) Identifique a eficiência média e a menor eficiência entre todas as DMUs; e
- d) Apresente as características da DMU de menor eficiência. Quais destas você acredita que explicam a menor eficiência? Justifique.

REFERÊNCIAS

BATTESE, G. E.; COELLI, T. Frontier production functions, technical efficiency and panel data: with application to paddy farmers in India. *Journal of Productivity Analysis*, v. 3, p. 153-169, 1992.

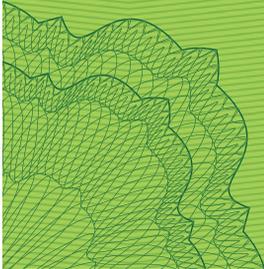
_____. A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics*, v. 20, p. 325-332, 1995.

BOGETOFT, P.; OTTO, L. *Benchmarking, with DEA, SFA and R*. Nova York: Springer, 2011.

BRASIL. *Guia livre: referência para migração para software livre no Governo Federal*, versão 1.0. Brasília: Ministério do Planejamento, 2005.

CHARNES, A.; COOPER, W. W.; RHODES, E. Evaluating program and managerial efficiency: an application of data envelopment analysis to program follow through. *Management Science*, v. 27, n. 6, p. 668-697, 1981.

RIBEIRO, D. D. C. *Software livre na administração pública: estudo de caso sobre adoção do SAMBA na auditoria geral do Estado de Minas Gerais*. 2004. 88f. Monografia (Especialização em Administração em Redes Linux) – Universidade Federal de Lavras, 2004.



Capítulo 11

Federalismo e Descentralização

Janete Duarte
Sérgio Ricardo de Brito Gadelha
Plínio Portela de Oliveira
Luis Felipe V. N. Pereira

11.1 Introdução

A provisão de serviços públicos de qualidade é uma demanda crescente da sociedade. Para atendê-la, a responsabilidade no fornecimento desses serviços é repartida entre os diferentes níveis de governo (União, estados e municípios), levando em consideração o diferencial de eficiência de cada ente na provisão desses serviços.

Assim, onde há ganhos de escala no fornecimento de serviços públicos, o fornecimento é competência da União. Por sua vez, onde a eficiência na provisão está relacionada ao melhor conhecimento das preferências e necessidades da população local, devido à proximidade que tem com o cidadão, cabe aos municípios a tomada das melhores decisões. Onde o *meio-termo* é o mais adequado, a responsabilidade se volta para os estados.

Esse tema é conhecido na economia como descentralização, ou seja, processo de transferência de poder e responsabilidade de gestão da União para os estados e municípios, e é estudado na literatura sobre federalismo fiscal. Nessa visão, a descentralização deveria melhorar a qualidade do gasto por meio de uma gestão mais assertiva e eficiente do custo e benefício de cada bem e de cada serviço público prestado.

11.1.1 Federalismo fiscal

O federalismo fiscal estuda a divisão de responsabilidades entre os diferentes níveis de governo. Essas responsabilidades incluem a oferta de serviços públicos (educação, saúde, segurança etc.) e a arrecadação de tributos.

No caso brasileiro, a maior parte das responsabilidades está definida na Constituição Federal. O artigo 22 estabelece que compete privativamente à União legislar sobre serviços monetários, comércio exterior, seguridade social, defesa territorial, dentre outros. Aos municípios compete legislar sobre o transporte coletivo, ordenamento territorial, assistência à União na prestação de serviços de saúde etc. As competências dos estados são entendidas como residuais, ou seja, tudo aquilo que não foi listado como responsabilidade da União, ou dos municípios, caberia a princípio aos estados.

Quanto à tributação, a teoria sobre federalismo fiscal nos diz que a divisão das responsabilidades deve seguir alguns critérios (MENDES, 2005, p. 430), a saber:

1. Mobilidade da base tributária: Se for possível deslocar a base tributária, o agente buscará as localidades com alíquotas menores. Nesse caso, o maior nível de governo (federal) deve tributar, para impedir evasões ou comportamentos oportunistas. Outro caso extremo é a tributação sobre imóveis, que fica a cargo do menor nível de governo (municipal), pois não há possibilidade de o agente deslocar a base tributária para obter benefícios.
2. Economia de escala na administração do tributo: Alguns tributos podem ser demasiadamente custosos para a arrecadação em pequena escala. Nesse caso, níveis mais altos de governo passariam a ter ganhos de escala na arrecadação.

No Brasil, esses critérios explicam, por exemplo, o fato de o Imposto Predial e Territorial Urbano (IPTU) ser municipal (mobilidade da base) e o Imposto de Renda ser federal (economia de escala¹).

Definidas as competências tributárias e as responsabilidades na provisão de serviços públicos, surge o questionamento: "O que acontece quando os recursos de um dos níveis de governo não são suficientes para custear os serviços públicos?"

1 Argumentos de mobilidade da base também explicam o fato de o Imposto de Renda ser federal.

11.2 As transferências intergovernamentais

Nesse caso, existe a figura da transferência vertical (entre níveis de governo). Aqueles níveis com maior poder arrecadatório fazem transferências para os de menor poder, de forma a permitir que haja recursos para a prestação de serviços. Existem três tipos principais de transferências: condicionais ou incondicionais, com contrapartida ou sem contrapartida, e limitada ou ilimitada.

11.2.1 Condicionais ou incondicionais

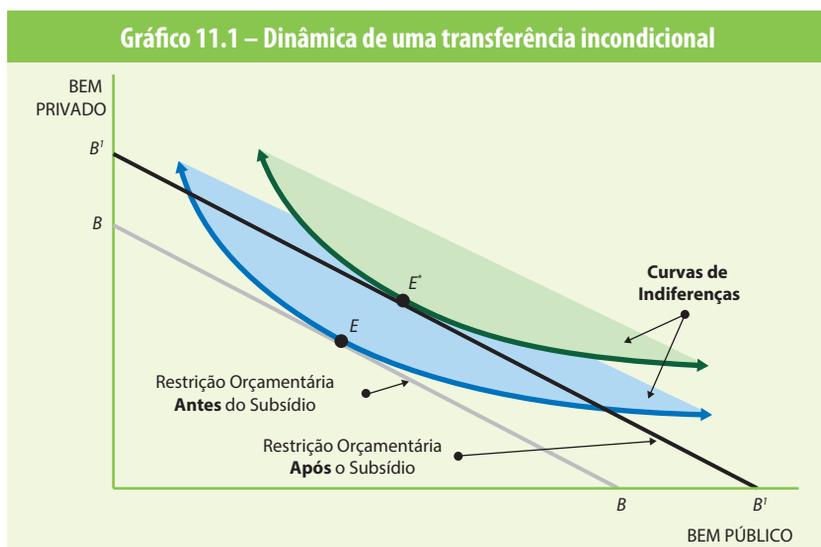
Uma transferência condicional é aquela em que o doador dos recursos determina o tipo de gasto em que o recurso orçamentário deve ser aplicado. As verbas federais transferidas aos estados e municípios para financiar a manutenção de escolas públicas são exemplos de transferências condicionais (MENDES, 2005, p. 435).

Já a transferência incondicional – como o próprio nome indica – consiste no aporte de recursos sem a obrigação (contrapartida) de se aplicar a verba em despesa específica. O Fundo de Participação dos Estados (FPE) e o Fundo de Participação dos Municípios (FPM) se enquadram nesse tipo de transferência (MENDES, 2005, p. 435).

O Gráfico 11.1, a seguir, descreve a dinâmica de uma transferência incondicional, a qual desloca a restrição orçamentária para a direita e o novo ponto de equilíbrio passa a ser E^* . Note que a variação na produção de bens públicos (distância entre as projeções de E e E^* no eixo horizontal) é menor que a transferência (distância entre B e B') (STIGLITZ, 2000, p. 744). Isso ocorre porque parte dos recursos recebidos pode ser utilizada tanto na expansão dos gastos quanto na redução dos impostos locais, o que permitiria um aumento de bens privados. Essa transferência visa minorar o problema de desequilíbrio vertical, uma vez que o governo federal (ou governo central), que pode arrecadar a maior parte dos tributos da forma mais eficiente, repassa parte dessa arrecadação aos estados e municípios (MENDES, 2005, p. 435-436).

As transferências incondicionais são utilizadas como instrumento que busca solucionar o desequilíbrio vertical entre os níveis de governo. Essas transferências também exercem o papel de redistribuição de recursos fiscais entre as regiões ricas e aquelas

mais pobres, com o propósito de reduzir o hiato na capacidade fiscal dessas regiões (MENDES, 2005, p. 435-436).

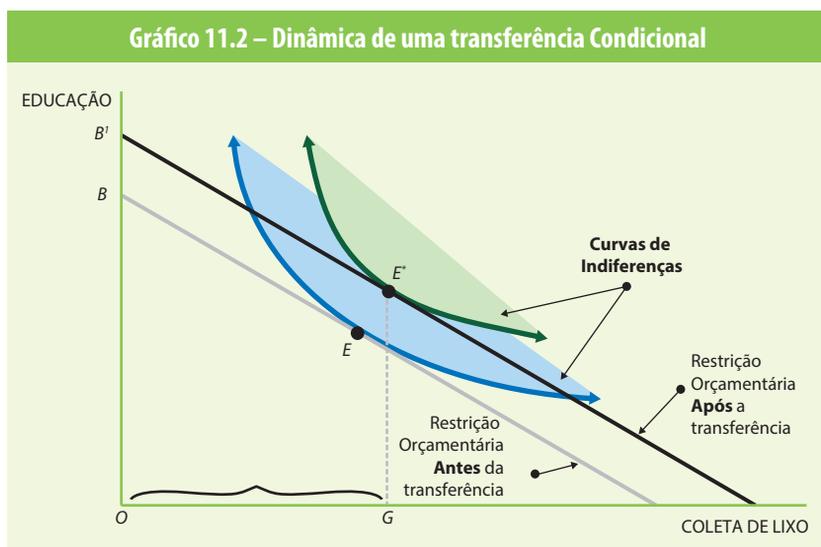


Fonte: Baseado em Stiglitz (2000, p. 744).

No caso da transferência condicional, observa-se que essa modalidade deve ser aplicada quando o governo federal (ou governo central) estiver interessado em garantir um padrão mínimo de prestação de um dado serviço público local – por exemplo, educação. Ou seja, trata-se de garantir financiamento federal para provisão de bens públicos de interesse das parcelas da população com necessidades especiais (pobres, idosos etc.), que podem ser ofertados de forma mais eficiente por governos locais (MENDES, 2005, p. 436-437).

Importante ressaltar que a categoria de gasto a qual essa transferência se destina – por exemplo, gasto em saúde, educação, coleta de lixo etc. – pode ser irrelevante, sempre que o valor da transferência for inferior ao ótimo desejado pela comunidade. O recurso do governo local que seria gasto em um bem público específico para o qual houve transferência pode ser redirecionado para outro, de forma que o caráter condicional da transferência fique sem efeitos (STIGLITZ, 2000, p. 745).

O Gráfico 11.2, a seguir, apresenta a dinâmica de uma transferência condicional.



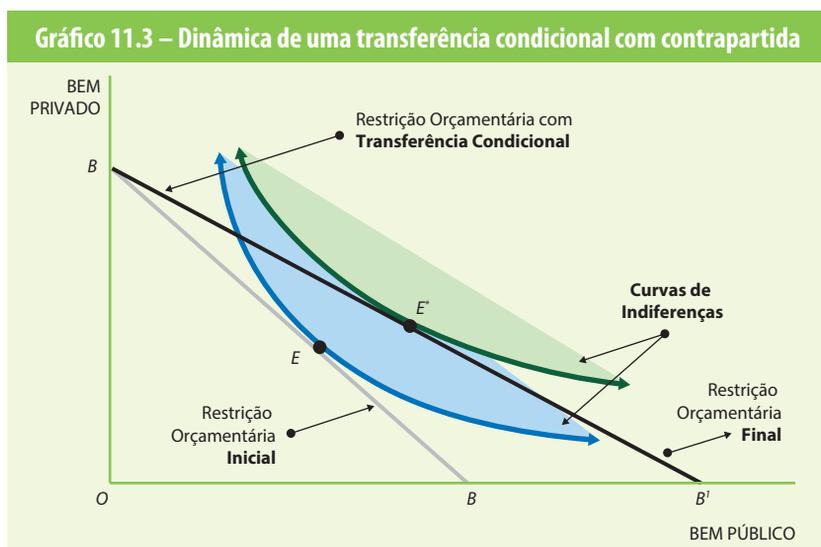
Fonte: Baseado em Stiglitz (2000, p. 745).

11.2.2 Com ou sem contrapartida

Uma transferência condicional é classificada como *com contrapartida* se o valor transferido pelo governo federal (ou governo central) for proporcional ao montante que o receptor aplica no gasto incentivado (MENDES, 2005). Por exemplo, para cada “x” reais gastos pelo governo local existe uma transferência de “y” reais como contrapartida do governo federal (ou governo central).

De modo específico, a nova restrição orçamentária, com o subsídio para os gastos governamentais locais, sofre uma rotação ao redor do ponto B , pois a contrapartida altera os preços relativos. Uma vez que os bens públicos passem a ser relativamente menos caros, a comunidade irá escolher consumir mais desse bem, em detrimento do bem privado, de modo que o equilíbrio muda do ponto E para o ponto E^* (STIGLITZ, 2000, p. 746-747).

O Gráfico 11.3, a seguir, apresenta a dinâmica de uma transferência condicional com contrapartida.



Fonte: Baseado em Stiglitz (2000, p. 747).

11.2.3 Limitada ou ilimitada

Um problema no desenho de transferências é a mensuração do valor total a ser transferido. No caso das transferências com contrapartida, por exemplo, municípios de maior renda podem ser excessivamente favorecidos, pois possuem mais recursos para utilizar na contrapartida. Para tentar controlar o volume de transferência em determinados programas, existe a modalidade de transferência limitada, ou seja, o doador estipula um valor máximo para a transferência (MENDES, 2005, p. 435). Por exemplo, o governo federal (ou central) poderia estabelecer que transferirá para os estados a quantia de R\$ 0,50 para cada R\$ 1,00 que aquele estado aplicar em determinada política pública, e estipular que essa contrapartida não deve ultrapassar um total de R\$ 1 bilhão.

Deve-se destacar que a utilização das transferências – sejam elas condicionais, incondicionais, com contrapartida ou não, limitadas ou ilimitadas – como instrumento para lidar com o desequilíbrio vertical

não é consenso entre os economistas. Os pontos de conflito passam necessariamente pela redistribuição de recursos fiscais entre regiões ricas e pobres, e pela estimativa do valor exato a ser transferido. Enfim, muito cuidado deve ser dispensado ao desenho das transferências, porque um desenho mal feito pode agravar o desequilíbrio que elas se propõem a solucionar (MENDES, 2005).

11.3 Distribuição das responsabilidades de gastos

Aparentemente, não pairam dúvidas sobre a responsabilidade do governo federal em fornecer, de maneira eficiente, bens e serviços públicos nacionais (como segurança nacional, controle de fronteiras etc.). Todavia, muitos são os questionamentos em relação à oferta de bens e serviços públicos por governos locais: seriam os municípios eficientes nessa oferta? Há justificativa teórica que corrobore tal argumento?

Sim, é possível questionar, à luz da teoria da descentralização, arranjos de responsabilidades e recursos entre os diferentes níveis de governo na provisão de bens e serviços públicos, alguns com maior probabilidade de serem mais eficientes que outros. Em um notável artigo escrito há mais de 40 anos, Tiebout² argumentou que a concorrência entre os governos locais conduziria à eficiência na prestação de bens públicos locais, assim como a competição entre as empresas privadas levaria a uma oferta eficiente de bens privados. Já Oates³ trabalhou com a ideia de que o governo local possui, pelo menos, uma vantagem em relação ao governo central na provisão de serviços públicos locais porque conhece melhor as preferências dos contribuintes.

11.3.1 Tiebout: concorrência entre governos locais revela sua preferência por bens públicos

Tiebout confiava na mobilidade como mecanismo fundamental para revelar as preferências dos indivíduos por serviços públicos. A ideia é simples: ao invés de haver um único governo que tenta adivinhar as preferências dos cidadãos para ofertar os bens públicos, pode-se pensar em diversas comunidades locais (municípios), conduzidas por governos locais, ofertando não apenas uma cesta

2 Tiebout (1956).

3 Oates (1972).

de bens públicos diferente quanto cardápios diferentes entre bens públicos e tributação. É óbvio que, nessas circunstâncias, cada cidadão escolheria viver na localidade que atendesse melhor às suas necessidades. Ao agir racionalmente e optar viver em uma dada comunidade, o cidadão estaria, em última análise, revelando suas preferências pela cesta de bens públicos.

Assim, para explicar a importância da descentralização fiscal na provisão eficiente de serviços públicos, Tiebout faz uma analogia entre o comportamento dos governos locais e a Teoria da Firma sob Concorrência Perfeita. Na ausência de falhas de mercado, uma economia competitiva seria eficiente no sentido de Pareto,⁴ quando, ao agir em benefício próprio, toma decisões que conduzem à eficiência na oferta de bens e serviços ao menor custo possível. Analogamente, a concorrência entre os governos locais assegura a eficiência na oferta de serviços públicos locais, assim como a concorrência entre as empresas garante a eficiência no fornecimento de bens privados.

11.3.2 Oates: cooperação entre níveis de governo aumenta eficiência do setor público

Oates busca mostrar que o federalismo fiscal, formado por uma divisão de tarefas entre diferentes níveis de governos, seria uma opção intermediária à adoção de uma administração pública totalmente centralizada ou amplamente descentralizada. Para tal, explora dois aspectos: (i) o federalismo fiscal reúne vantagens comuns de cada um dos casos extremos, ao passo que minimiza seus problemas; (ii) o federalismo fiscal induz a população a uma maior participação e fiscalização nas decisões de governos locais.

Sobre o primeiro aspecto, um setor público amplamente descentralizado teria sérias dificuldades para implementar políticas macroeconômicas, como promover estabilidade de preços ou programas de distribuição de renda. Aqui, a ausência de um governo central na condução dessas políticas voltadas para o coletivo inviabilizaria a oferta de bens e serviços que beneficiassem o país como um todo, uma vez que o foco de cada municipalidade seria a implementação das políticas

4 Neste sentido, a eficiência é considerada ótima se não for possível melhorar a situação de alguém sem prejudicar outrem.

locais. Já o governo extremamente centralizado não conseguiria atender eficientemente às diferentes preferências locais, pois o custo de se obter essas informações seria muito elevado. Já os governos locais, dada a proximidade com os indivíduos, levariam vantagem no atendimento dessas demandas, pois tendem a conhecer melhor as preferências dos seus cidadãos a um baixo custo.

Quanto ao segundo aspecto, Oates vincula a maior participação dos cidadãos à relação *direta* entre o que os contribuintes pagam de impostos àquela municipalidade e o que eles recebem dela em serviços públicos. Ou seja, os contribuintes tendem a ficar mais atentos à utilização do valor arrecadado em impostos, pois a relação custo-benefício é direta. Certamente, em um governo centralizado tornar-se-ia mais difícil a associação direta entre o pagamento de impostos e o retorno em serviços públicos.

Em suma, existe fundamento teórico para explicar por que alguns serviços públicos são ofertados pelos governos locais e outros, não. No Brasil, temos na educação e na saúde aplicações claras dessa teoria. No caso da educação, o ensino fundamental fica a cargo dos municípios, pois a dispersão geográfica dificultaria uma administração centralizada e a proximidade da administração escolar com os pais permite melhor atendimento da demanda local.

No caso da saúde, como veremos na próxima seção, o fornecimento de serviços é concentrado nos governos locais. Atividades ambulatoriais e medidas preventivas de educação para a saúde ficam a cargo dos municípios, enquanto procedimentos hospitalares mais complexos podem ficar a cargo dos estados ou do governo federal.

Para que essa divisão de responsabilidades funcione adequadamente, é necessário que cada um dos entes tenha recursos suficientes para fornecer os serviços. Assim, existe uma estrutura de transferências para saúde e educação no Brasil, a ser explorada a seguir.

11.4 Financiamento da saúde e da educação no Brasil

11.4.1 Financiamento da saúde

Como sabemos, uma das funções do Estado, a distributiva, visa promover uma divisão mais equânime da renda na sociedade. Uma

das formas de atuação da função distributiva é por meio da seguridade social, que compreende ações voltadas à saúde, previdência social e assistência social.

Essas ações são de responsabilidade conjunta dos três níveis de governo e cada uma delas conta com orçamento específico. Os recursos podem vir de tributos especiais, denominados contribuições sociais, ou do orçamento dos entes da Federação.

Anualmente o Congresso define as prioridades relativas a cada área e a proporção dos recursos destinados à saúde, à previdência e à assistência social. No caso específico da saúde, a Constituição Federal estabelece montantes mínimos a serem destinados obrigatoriamente para ações e serviços sociais, sob a forma de percentuais que deverão ser aplicados em algumas categorias de receitas.

No caso da União, enquanto não existe Lei Complementar definindo a base de cálculo e o percentual aplicável, vigora a exigência de se aplicar em saúde, a cada ano, pelo menos o mesmo valor efetivamente empenhado em ações e serviços públicos da área no ano imediatamente anterior, corrigido pela variação nominal do PIB do ano anterior.

Para os estados, aplica-se o percentual mínimo de 12% a uma base de cálculo formada por impostos e receitas de transferência, deduzindo-se as transferências feitas a municípios.

Já para os municípios, o percentual mínimo é de 15% e a base é composta por impostos e transferências da União e do estado. O Distrito Federal é um caso excepcional, onde são aplicados o mínimo de 12% a uma parte da base de cálculo e 15% ao restante.

Definido o montante mínimo de recursos que deve obrigatoriamente ser gasto em saúde, ainda é preciso definir o que de fato são *serviços de saúde*. Existe controvérsia sobre este tema, desde a promulgação da Emenda Constitucional nº 29 de 2000. O conceito que prevalece hoje é o previsto na Resolução do CNS nº 322/2003, definido a partir de diretrizes que tentam limitar o que pode e o que não pode ser incluído como gasto em saúde para fins de cumprimento dos mínimos legais.

A gestão dos recursos de saúde é feita pelo ministério ou secretaria responsável, em cada ente da Federação, auxiliados pelo Conselho de Saúde, que reúne representantes do governo, prestadores de serviço, profissionais da área e usuários dos serviços de saúde. A cada

quatro anos também é realizada a Conferência de Saúde, que reúne todos os segmentos sociais interessados na saúde pública a fim de propor diretrizes para o setor.

A gestão financeira é claramente separada e individualizada com relação às demais funções orçamentárias. O Fundo de Saúde, que todo ente deve criar, recebe os recursos mínimos constitucionalmente exigidos e dele são sacados os recursos necessários para ações e serviços de saúde.

A direção nacional do SUS (centralizada na União) define princípios, regras, padrões de atenção à saúde e diretrizes nacionais da política de saúde. As direções estaduais fazem o controle e a avaliação das redes hierarquizadas do SUS e prestam apoio à União e aos municípios. As direções municipais do SUS são as responsáveis pela maior parte da execução das ações e dos serviços de saúde, sobretudo no campo da atenção básica.

Para que os entes subnacionais possam desempenhar suas atribuições, a legislação prevê o repasse de recursos do Fundo Nacional de Saúde aos fundos estaduais e municipais. A Lei nº 8.080, de 1990, prevê critérios técnicos para esses repasses, mas na ausência de norma que regulamente a aplicação dos critérios, os repasses são feitos exclusivamente segundo o quociente da divisão dos recursos pelo número de habitantes do ente beneficiário. Pelo menos 70% dos recursos são repassados aos municípios e o restante, aos estados. Esses repasses são regulares e automáticos.

11.4.2 Financiamento da educação

A intervenção social do Estado também se dá por meio da promoção e do incentivo à educação. Um dos princípios do modelo brasileiro é a coexistência de instituições públicas e privadas de ensino. Apesar de o ensino ser livre à iniciativa privada, é dever do Estado garantir educação básica obrigatória e gratuita dos 4 aos 17 anos, a progressiva universalização do ensino médio gratuito, educação infantil em creches e pré-escolas para crianças de até 5 anos de idade e acesso aos níveis mais elevados de ensino, pesquisa e criação científica, segundo a capacidade de cada um.

Em síntese, cabe ao Estado assegurar a oferta direta e universal dos serviços educacionais completos em todas as suas etapas.

De acordo com a Constituição Federal, os serviços educacionais públicos não podem ser custeados por taxas cobradas aos usuários. Portanto, a única maneira de custeá-los é por tributos.

A organização dos sistemas de ensino é obrigação conjunta de todas as esferas de governo, que devem cumpri-la em regime de colaboração. Os recursos para financiamento provêm dos orçamentos e de uma contribuição social denominada *salário-educação*.

Da mesma forma que ocorre com os gastos em saúde, existem recursos mínimos que devem ser destinados à educação, definidos pela Constituição Federal. São estabelecidos percentuais que devem ser aplicados a uma base de cálculo formada pela totalidade das receitas de impostos e pelas receitas líquidas de cada ente. No caso da União, esse percentual é de 18%, e para os estados e municípios, de 25%. Os recursos devem ser aplicados, necessariamente, nos sistemas públicos de ensino, escolas comunitárias, confessionais e filantrópicas, ou ainda em atividades universitárias de pesquisa e extensão.

Essa regra geral é complementada por outras de natureza temporária. Desde 2007 vigora o Fundo de Manutenção e Desenvolvimento da Educação Básica e de Valorização dos Profissionais da Educação (Fundeb), que a princípio não vincula recursos adicionais aos previstos na regra geral. Em vez disso, o objetivo é determinar a forma como os entes subnacionais deverão gastar os recursos, assegurando o desenvolvimento e a manutenção da educação básica, bem como a remuneração condigna dos trabalhadores do setor.

O Fundeb possui um mecanismo de complementação de fundos pela União, que pode acarretar em recursos excedentes aos constitucionalmente vinculados. Estados e municípios devem destinar 20% das receitas de impostos e transferências líquidas ao Fundeb. Uma parcela adicional de 5% é vinculada ao ensino, mas não se submete ao regime do fundo. Cada estado possui o seu próprio Fundeb – municípios e a União não o possuem.

A complementação federal ocorre sempre que o volume de recursos não for suficiente para assegurar, no âmbito de cada estado, determinado valor médio mínimo por aluno, estabelecido nacionalmente. A ponderação é feita de acordo com a etapa educacional e o tipo de estabelecimento de ensino em que as matrículas são realizadas.

O gasto em educação passa pela mesma dificuldade encontrada no caso da saúde: quais despesas podem ser consideradas como educação? A resposta para este questionamento é encontrada na Lei de Diretrizes e Bases da Educação, a LDB (Lei nº 9.394/1996), que define quais despesas podem ser compreendidas como dessa área.

11.5.0 papel dos consórcios na eficiência dos serviços descentralizados

A fragmentação ocorrida no Brasil, no final do século XX, gerou uma grande quantidade de municípios com número reduzido de habitantes. Estima-se que entre 1984 e 1997 o total tenha crescido em 34,3%⁵.

A estrutura descentralizada de alguns serviços públicos, principalmente da saúde, também foi afetada por esse aumento no número de municípios. A dificuldade em obter ganhos de escala, somada a problemas relacionados à gestão, levou a uma queda na qualidade dos serviços de saúde nos municípios pequenos.

Dado o elevado custo de provisão de alguns serviços de saúde, em especial os custos com infraestrutura hospitalar, a escala passa a ser fator determinante para a eficiência. Mendes (2001) cita casos em que um grupo de oito municípios tem hospitais com taxa de ocupação média de 22% e que possuem sete aparelhos de ultrassonografia, quando seriam suficientes apenas dois.

Uma forma de contornar a questão da escala em municípios é a criação de consórcios, que podem ser destinados à saúde, ao saneamento ou a outras áreas, inclusive as de infraestrutura. Dados do IBGE apontam que 35% dos municípios brasileiros se uniram em consórcios de saúde.

Todavia, existem vários aspectos na estrutura de consórcios que dificultam sua implementação. Uma legislação ainda recente tenta ampliar a segurança jurídica e a eficiência do modelo, mas ainda há alguns pontos criticados por especialistas.

No formato atual, os municípios comprometem uma parcela considerável do seu orçamento, que deve ser destinada aos consórcios. No entanto, existe a possibilidade de o município enfrentar condições

5 De acordo com Gomes e Mac Dowell (2000).

econômicas adversas, como as decorrentes de uma calamidade, e não possuir flexibilidade no orçamento para enfrentá-las.

Outra fragilidade do modelo de consórcios é o comportamento oportunista (*free rider*) criado pelo desenho atual. Determinado município pode aderir ao consórcio e, em um segundo momento, sair. Nesse caso, não precisará arcar com as despesas do consórcio e sua população poderá continuar com acesso aos serviços de saúde, ainda que em outro município.

Teixeira, Bugarin e Mac Dowell (2003) estudam o formato dos consórcios à luz da Teoria dos Jogos e concluem que o comportamento oportunista pode ser bastante reduzido, caso haja uma punição para o município que deixar o consórcio: a população deixaria de ter acesso aos serviços de saúde desse consórcio. Na prática, tal desenho de contrato dificilmente poderia ser aplicado ao Brasil, pois a proibição do acesso à saúde para a população de um determinado município é inconstitucional – as ações e serviços de saúde devem ser universais e iguais.

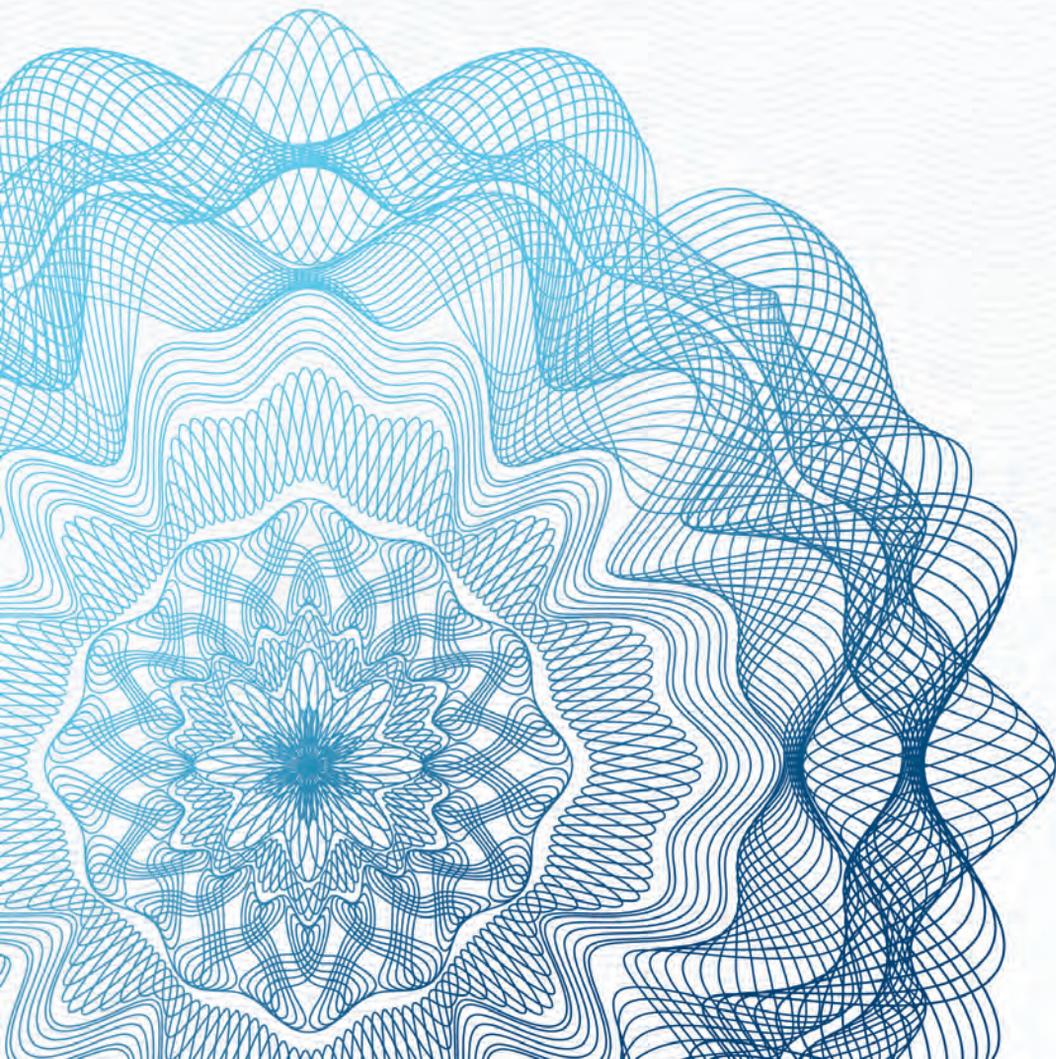
Uma alternativa sugerida por Teixeira (2006) é a inserção dos estados e da União no desenho dos contratos, atuando como uma espécie de *juiz*. Uma das propostas baseia-se na destinação direta de recursos provenientes do Fundo de Participação dos Municípios (FPM) de municípios em situação de inadimplência com o consórcio. Mas essa proposta também requer a criação de legislação específica. Aos estados caberia o apoio à formação de consórcios, destinando recursos e *expertise* para a criação deles.

REFERÊNCIAS

- ARVATE, P. R.; BIDERMAN, C. *Economia do setor público no Brasil*. Rio de Janeiro: Elsevier: Campus, 2005.
- GOMES, G. M.; MAC DOWELL, M. C. *Descentralização política, federalismo fiscal e criação de municípios: o que é mau para o econômico nem sempre é bom para o social*. Brasília: IPEA, 2000. (Texto para Discussão n. 706).
- MENDES, E. V. *Os grandes dilemas do SUS: Tomos I e II*. Salvador, BA: Casa da Qualidade Editora, 2001. (Saúde Coletiva, n. 4).
- MENDES, M. Federalismo fiscal. In: BIDERMAN, C.; ARVATE, P. (Org.). *Economia do setor público no Brasil*. 8. Tiragem. Rio de Janeiro: Elsevier, 2005.
- OATES, W. E. *Fiscal federalism*. Nova York: Harcourt Brace Jovanovich, 1972.
- TEIXEIRA, L. S. Consórcios intermunicipais: instrumento para aumentar eficiência do gasto público. In: MENDES, M. (Org.). *Gasto público eficiente: propostas para o desenvolvimento do Brasil*. Rio de Janeiro: Topbooks Editora, 2006.
- TEIXEIRA, L.; BUGARIN, M.; MAC DOWELL, M. C. Consórcios intermunicipais de saúde: uma análise à luz da teoria dos jogos. *Revista Brasileira de Econometria*, v. 57, n. 2, 2003.
- TIEBOUT, C. M. A pure theory of local expenditures. *Journal of Political Economy*, v. 64, p. 416-424, 1956.
- STIGLITZ, J. E. *Economics of the Public Sector*. 3. ed. Nova York: W. W. Norton & Co., 2000.

PARTE V

TÓPICOS AVANÇADOS EM
MENSURAÇÃO DE EFICIÊNCIA



Capítulo 12

Eficiência na provisão de educação e saúde: resenha e aplicações para os municípios brasileiros

Fabiana Rocha

Janete Duarte

Sérgio Ricardo de Brito Gadelha

José de Anchieta Semedo Neves

Plínio Portela de Oliveira

Luis Felipe Vital Nunes Pereira

12.1 Introdução

É possível dizer que nos últimos anos uma das grandes preocupações da literatura e da prática de finanças públicas esteve relacionada à avaliação e à melhoria da qualidade dos gastos públicos dos diferentes níveis de governo.

A saúde e a educação, por serem os serviços mais importantes providos pelo setor público, receberam atenção especial.

O objetivo deste capítulo é apresentar brevemente os principais trabalhos que procuram medir e comparar o produto do sistema educacional e de saúde de países, estados ou municípios com os recursos empregados, a partir da derivação de uma fronteira de produção teórica. Se um país, estado ou município está operando na fronteira, ele é considerado eficiente. Se, por sua vez, tem um desempenho abaixo da fronteira, é possível estimar a distância em relação a essa fronteira, ou seja, o seu escore de eficiência.

Além da revisão da literatura, o capítulo faz duas aplicações que buscam implementar a técnica da metafronteira na mensuração da eficiência relativa dos municípios brasileiros na provisão de educação e de saúde.

O capítulo está organizado da seguinte maneira: a seção 12.2 faz uma resenha da literatura, chamando atenção em particular para os métodos utilizados e os resultados obtidos; a 12.3 apresenta uma

abordagem mais avançada para a mensuração da eficiência, com a aplicação da técnica de metafronteira utilizando as metodologias Análise Envoltória de Dados (DEA¹), na provisão de educação, e de fronteira estocástica, na provisão de saúde.

12.2 Revisão da literatura

12.2.1 Educação

A primeira tentativa sistematizada de avaliar a eficiência dos gastos em educação pode ser atribuída a Clements (2002). Ele usou uma abordagem *Free Disposable Hull* (FDH) para avaliar o desempenho relativo dos países da Organização para Cooperação e Desenvolvimento Econômico (OCDE) na provisão de serviços de educação. Os resultados obtidos indicam que os países europeus poderiam alcançar aproximadamente os mesmos resultados em educação com cerca de 25% menos de recursos.

Gupta, Verhoeven e Tiongson (2002) analisam, também por meio da abordagem FDH, a relação entre gasto e desempenho educacional para uma amostra de 38 países africanos e para uma amostra de 85 países africanos, asiáticos e do hemisfério ocidental durante o período 1984-1995. Os resultados indicam que, na média, os países africanos são menos eficientes que os da Ásia e do hemisfério ocidental, mas que os gastos em educação na África se tornaram mais eficientes nesse período. Diante disso, maiores alocações orçamentárias para os setores sociais nos países africanos não necessariamente irão se refletir em melhoria nos seus resultados sociais, a menos que medidas específicas sejam adotadas para corrigir a ineficiência subjacente ao gasto.

Afonso e St. Aubyn (2005), por sua vez, avaliam a eficiência do gasto em educação para uma amostra de países da OCDE, utilizando os métodos FDH e DEA. Quando estimam as fronteiras, usam como indicadores de insumo tanto medidas de gasto por aluno quanto de quantidade de insumos (tempo total de aula e número de professores por aluno), uma vez que um país pode ser eficiente do ponto de vista técnico, mas parecer ineficiente se os insumos que utiliza são

1 Análise Envoltória de Dados é a tradução literal do termo em inglês *Data Envelopment Analysis* (DEA).

caros. Concluem que medidas de insumo quantitativas captam mais acuradamente a importância relativa dos insumos usados por cada país do que medidas financeiras e, ainda, que nos países classificados como menos eficientes existe um amplo espaço para atingir melhores resultados usando os mesmos recursos.

Afonso e St. Aubyn (2006) avaliam a eficiência do gasto em educação em 25 países, a maioria da OCDE, a partir da combinação de duas vertentes distintas da literatura: a de eficiência de gastos públicos e a de função de produção de educação.²

Essa junção é feita ao se estimar um modelo semiparamétrico de processo de produção de educação usando uma abordagem em dois estágios. No primeiro deles é estimado o escore de eficiência de produto de cada país, usando uma abordagem DEA que relaciona insumos educacionais a produtos. A análise DEA, contudo, incorpora somente insumos discricionários, ou seja, aqueles cujas quantidades podem ser modificadas de acordo com os desejos dos países avaliados (tempo total de aula e número de professores por aluno). Contudo, diferenças socioeconômicas (inclusive riqueza das famílias e educação dos pais) podem desempenhar um papel relevante ao determinar a heterogeneidade entre países e influenciar os resultados. É preciso, então, levar em conta a presença de variáveis ou fatores ambientais, também conhecidos como insumos não discricionários. Diante disso, em um segundo estágio, os escores de eficiência são utilizados como variáveis dependentes em uma análise de regressão. A evidência empírica obtida indica que as variáveis de *background* familiar identificadas pelos autores da vertente da função de produção de educação são de fato altamente correlacionadas com a ineficiência, ou seja, representam variáveis ambientais significantes.

Pang e Herrera (2005) também utilizam uma abordagem em dois estágios para avaliar a eficiência dos gastos em educação de 140 países no período de 1996 a 2002. No primeiro estágio, os escores

2 Esta última procura avaliar os determinantes da qualidade das escolas a partir de regressões *cross-country*, especificando e estimando modelos lineares para a relação entre qualidade da escola (medida por diversas variáveis que indicam o resultado da educação) e suas possíveis fontes (os recursos alocados para a educação, por exemplo, professores por aluno ou gastos por aluno e outros fatores que podem afetar o produto educacional, como a renda ou o nível de educação dos pais). Hanushek e Kimko (2000) e Hanushek e Luque (2003) encontram pouca ou nenhuma evidência de uma relação positiva entre os recursos alocados para o sistema educacional e o desempenho em exames padronizados. Contudo, apontam que os níveis de educação dos pais têm um efeito positivo e significativo sobre o desempenho dos filhos.

de eficiência são obtidos por meio do uso tanto da abordagem FDH quanto da abordagem DEA. No segundo estágio, os escores de eficiência são regredidos em seus possíveis determinantes, a fim de verificar regularidades empíricas da variação *cross-country* na eficiência. Uma conclusão importante do trabalho é a de que, em termos de orientação de política, o importante é diferenciar entre o nível tecnicamente eficiente e o nível de gasto ótimo ou desejável. Pode ser ainda necessário que um país aumente o seu gasto público, mesmo identificado como eficiente, se ele quer atingir uma meta estabelecida para um indicador educacional. O fundamental, então, é que ao expandirem suas escalas de operação os países o façam ao longo da fronteira eficiente.

Gimenez, Prior e Thieme (2007) calculam escores de eficiência para o sistema educacional de 31 países. Os resultados indicam que o aumento médio no desempenho acadêmico poderia ser de 10%, dos quais 6% atribuíveis aos fatores ambientais e 4%, à ineficiência do sistema propriamente dito. Quando são considerados somente os países em desenvolvimento, os fatores ambientais aparecem como essenciais para explicar as diferenças de eficiência entre eles.

Sutherland et al. (2007) avançam em relação ao resto da literatura ao fazerem a análise tanto em nível de países quanto em nível de escolas. Os insumos incluem a razão professor/aluno, a disponibilidade de computadores (*proxy* para capital) e indicadores das características dos alunos (*background* socioeconômico, se é imigrante ou não, língua falada em casa). Os escores do Programa Internacional de Avaliação de Alunos (*Program for International Student Assessment - Pisa*) são utilizados como produto.

Quando a amostra é composta de escolas, o número grande de observações permite que sejam obtidos resultados por meio da abordagem DEA e também da estimação de uma fronteira estocástica. Três resultados aparecem como particularmente interessantes. O primeiro é o de que para a escola mediana obter o mesmo nível de produto, os insumos poderiam ser reduzidos em 1/3. O segundo é o de que as diferenças nas estimativas de eficiência não são muito grandes para os diferentes tipos e tamanhos de escola. A escola pública mediana, na amostra, é somente um pouco menos eficiente do que a escola privada mediana dependente do governo e do que a escola privada mediana independente do governo. O terceiro é o

de que existem diferenças significantes nas estimativas de eficiência em certos países. Por exemplo, embora no Japão e na Irlanda o nível mediano de ineficiência seja baixo, há mais heterogeneidade na eficiência em nível de escola no primeiro do que no segundo país.

12.2.2. Saúde

O relatório original produzido por Evans et al. (2000) para a Organização Mundial da Saúde (OMS) pode ser considerado o trabalho seminal de avaliação de eficiência de sistemas de saúde. Nele é estimado um modelo de painel com efeitos fixos para 191 países, entre 1993 e 1997, para gerar uma fronteira de produção. Os efeitos fixos individuais para países são usados como indicadores de ineficiência. O país com o intercepto mais alto é caracterizado como o mais eficiente, e os desvios dos demais países em relação a ele aparecem como medidas de ineficiência.

São usados como insumos os gastos em saúde e também em educação; e como produto, os resultados de saúde medidos pela expectativa de vida ajustada (*disability adjusted life expectancy* – Dale) e uma medida composta de cuidado de saúde (Comp).³

Hollingsworth e Wildman (2003) argumentam que alternativas metodológicas podem trazer informação adicional sobre o desempenho dos países na provisão de serviços de saúde. Dessa forma, reestimam os modelos da OMS usando métodos de dados em painel, estimadores de dados em painel com constantes variantes no tempo, DEA, índices de Malmquist e fronteiras estocásticas. Concluem que o modelo usado por Evans et al. (2000) está mal especificado para os países da OCDE e que é um erro combinar os dois grupos em uma única amostra. Eles concluem ainda que a escolha entre um método não paramétrico e um método paramétrico afeta a estimação da eficiência.

3 A Dale indica o número de anos que uma pessoa espera viver em plena saúde, ou seja, levando em conta os anos prejudicados por doenças e/ou ferimentos. A Comp consiste na média ponderada dos cinco objetivos especificados para o sistema de saúde (saúde, desigualdade em saúde, nível de resposta, distribuição da resposta e financiamento justo), com os pesos baseados em pesquisa feita pela OMS para estabelecer as preferências dos indivíduos a partir de suas valorações relativas dos objetivos do sistema de saúde. A renda nacional não é considerada um determinante dos resultados em saúde, uma vez que os seus efeitos sobre a saúde ocorrem basicamente por meio da educação.

Gravelle et al. (2003) têm como objetivo avaliar a sensibilidade dos resultados a diferentes especificações. Dessa forma, os escores de eficiência são calculados usando-se um estimador de efeitos fixos e um estimador *between*, uma vez que 99,8% da variação no logaritmo da variável Dale, 98,9% da variação no logaritmo dos gastos em saúde e 99,8% da variação na escolaridade ocorrem entre os grupos de países e não dentro de cada grupo de país.

Greene (2004) critica metodologicamente o trabalho de Evans et al. (2000), ao apontar que o estimador de efeitos fixos não é capaz de distinguir ineficiência e heterogeneidade entre países. Isso porque é avaliado um painel de 191 países ao longo de cinco anos em que as diferenças econômicas e culturais são enormes, o que acaba produzindo grande heterogeneidade nos dados, confundida com ineficiência.

O autor apresenta, então, várias abordagens alternativas de análise de fronteira estocástica para dados de painel e aplica algumas delas aos dados da OMS.

Ele incorpora, por exemplo, os determinantes da eficiência, chamados de variação de país, na estimação original da fronteira estocástica. É estimada uma fronteira estocástica com efeitos aleatórios em que as covariadas invariantes no tempo explicam a média subjacente de u_i . É estimada ainda uma extensão do modelo de efeitos aleatórios, proposta por Battese e Coelli (1988, 1995), que incorpora alguma variação temporal na ineficiência. A especificação preferida corresponde a um modelo de efeitos aleatórios que inclui variação entre países na função de produção e na ineficiência. Nesse caso, os resultados são bem diferentes dos obtidos com o modelo de efeitos fixos, tanto em estimativas quanto em ordenamento dos países.

A conclusão geral é a de que os dados usados pela OMS são marcados pela heterogeneidade que ficou encoberta nos demais estudos como ineficiência. Ao usar modelos mais gerais, flexíveis e que se baseiam em diferentes indicadores de heterogeneidade, novas evidências emergem e os resultados existentes passam a ser questionados.

A maior parte dos trabalhos que avalia a eficiência dos gastos em educação analisa também a eficiência dos gastos em saúde.

Gupta, Verhoeven e Tiongson (2002) concluem que o grau de ineficiência aumenta com o nível de gasto do governo, o que sugere que os governos devem ser mais cuidadosos ao considerarem expandir seus gastos em saúde quando o nível inicial de dispêndios já é alto. A melhora no produto seria factível simplesmente pela correção das ineficiências nos gastos em saúde.

Afonso e St. Aubyn (2005) também avaliam duas especificações para mensurar a eficiência dos gastos em saúde. Na primeira especificação, o insumo é o gasto em saúde *per capita* em Paridade do Poder de Compra ou, na sigla em inglês, Purchasing Power Parity PPP; na segunda, medidas físicas de insumo (número de médicos, número de enfermeiras e número de leitos por mil habitantes) são comparadas aos resultados (produtos). Os resultados são os mesmos obtidos para a educação: nos países considerados menos eficientes é possível atingir resultados iguais usando menos recursos; e a eficiência mensurada, considerando os recursos financeiros alocados para um setor, gera resultados diferentes daqueles obtidos quando os recursos (insumos) são considerados em termos físicos.

Pang e Herrera (2005) repetem para os gastos em saúde a avaliação em dois estágios feita para estimar a eficiência dos gastos em educação. No segundo estágio, os principais resultados (caso de um insumo – um produto e múltiplos insumos – múltiplos produtos) são: (i) países com os níveis mais altos de gasto também apresentam os escores mais baixos (maior ineficiência); (ii) países em que a folha de salário representa uma parcela maior do gasto total tendem a ser mais ineficientes; (iii) países em que o financiamento público é uma parcela maior do gasto total também registram menores escores de eficiência; (iv) a taxa de urbanização é positivamente associada aos escores de eficiência; (v) o efeito da incidência de HIV/Aids é negativo, como esperado, e significativo em poucos casos; (vi) a distribuição de renda tem o efeito negativo esperado sobre os escores de eficiência; (vii) existe uma relação negativa entre alguns escores de eficiência e a razão de dependência de ajuda externa; (viii) nenhuma das variáveis institucionais é estatisticamente significativa, contrariando a expectativa de que melhores instituições, mais transparência e menos corrupção estivessem associados com escores de eficiência mais altos.

12.2.3 Evidência para o Brasil

Seguindo a literatura internacional, alguns trabalhos também foram feitos para o Brasil.

Brunet et al. (2006) buscam relacionar a utilização dos recursos dos estados brasileiros e do Distrito Federal à oferta de produtos e serviços (eficiência) e aos resultados obtidos (efetividade). Para tanto, comparam o insumo (as despesas efetuadas segundo a classificação funcional dos orçamentos estaduais) com indicadores socioeconômicos de produto e de resultado. São utilizados dois métodos distintos: o FDH e o método de ajuste pela função de Hill, que gera uma curva contínua por meio da interpolação dos pontos associados aos estados que se localizam sobre a fronteira eficiente, e analisadas 14 categorias funcionais, incluindo educação e saúde.

Marinho (2003) avalia a eficiência dos municípios do estado do Rio de Janeiro na oferta de serviços ambulatoriais e hospitalares. Realiza inicialmente uma análise não paramétrica (DEA) para determinar os escores de eficiência. Os resultados diferem pouco entre as regiões do estado, e a média geral de eficiência é de 83,07%. É feita ainda uma regressão dos escores de eficiência em variáveis que refletem o ambiente em que o gestor atua, como a população e o produto.

Mattos et al. (2009) calculam escores de eficiência para os gastos em saúde dos municípios paulistas com o objetivo de avaliar o impacto de variáveis de escala sobre a qualidade da oferta de serviços públicos de saúde. Por um lado, a literatura de federalismo fiscal recomenda que os bens e serviços públicos devam ser providos pelo nível de governo mais próximo da população, por conhecer melhor as preferências dela. Por outro lado, a oferta pulverizada pode levar a significativas deseconomias de escala.

É utilizado um procedimento em duas etapas. Na primeira, os escores de eficiência são estimados usando uma abordagem FDH. Na segunda, são estimados os efeitos das variáveis de escala sobre a eficiência e diretamente sobre os indicadores de desempenho e de gasto *per capita* em saúde. A evidência obtida indica que os municípios menores são menos eficientes na provisão de serviços públicos de saúde. Apresentam ainda maiores gastos *per capita* e pior desempenho medido pelo acesso ao sistema de saúde e pela

prevenção de doenças infecciosas. Diante disso, os autores concluem que a descentralização pode ter levado à ineficiência, aumentando gastos desnecessariamente sem promover melhoras na qualidade.

Souza, Nishijima e Rocha (2010) também olham para os municípios do estado de São Paulo, mas procurando avaliar a eficiência produtiva do setor hospitalar entre os anos de 1998 e 2003. É adotada a metodologia proposta por Battese e Coelli (1995) para painéis, que permite a variação da ineficiência ao longo do tempo, modelando-se a sua média a partir de diferentes características dos municípios. Os três autores concluem que os municípios mais eficientes são aqueles que contratam maior percentual de leitos de hospitais privados e de leitos universitários, com maior número de internações por ano e que apresentam menor tempo de internação. A ausência de significância das variáveis de gestão sugere que a qualidade da internação hospitalar independe do ciclo político e que a gestão municipal não apresenta ganhos em relação à gestão estadual.

Finalmente, Alves Júnior e Sampaio de Sousa (2011) utilizam uma metodologia de três estágios para avaliar a eficiência dos estados e do Distrito Federal na provisão de serviços educacionais. É aplicada a metodologia DEA no primeiro estágio. No segundo, seguindo a proposta de Fried et al. (1999; 2002), os excessos de insumos obtidos no primeiro estágio são estimados usando uma fronteira estocástica, em que as variáveis exógenas são explicitamente levadas em conta a fim de *separar* os componentes de ineficiência que podem ser atribuídos ao ambiente em que operam os estados, ao ruído estatístico e ao mau gerenciamento de recursos. O objetivo é controlar para que as condições externas e outros fatores possam favorecer ou prejudicar a atuação dos estados. Finalmente, no terceiro estágio os insumos corrigidos (ou pseudoinsumos) são utilizados em uma nova análise DEA, possibilitando uma comparação efetiva, uma vez que as unidades federativas são avaliadas em igualdade de condições. O resultado da substituição dos insumos originais pelos insumos corrigidos deve representar exclusivamente o grau de eficiência, refletindo as competências estaduais em gerenciar os recursos empregados no processo produtivo de educação. Ao se fazer a correção do excesso de consumo dos recursos, os resultados mudam bastante. Para o conjunto dos estados, o nível de eficiência aumenta para 0,81, o que representa um incremento de 4,4% frente à média de 0,77 encontrada no primeiro estágio.

12.3 Tópicos avançados na mensuração da eficiência: a abordagem da metafronteira

Estimar uma fronteira de produção global para todos os municípios é assumir implicitamente que eles dividem uma tecnologia de produção comum. Essa hipótese não parece ser adequada à realidade, uma vez que os municípios não só fazem escolhas a partir de combinações de insumo e produto diferentes, como também têm conjuntos de tecnologia distintos, decorrentes dos seus diferentes estoques de capital físico, humano e financeiro, da infraestrutura econômica, da dotação de recursos etc.

Estimar fronteiras individuais em vez de uma fronteira comum para a amostra completa de municípios não é adequado, porque se cada um deles tem sua fronteira de produção própria, os escores de eficiência técnica dos diferentes municípios não são diretamente comparáveis.

Essas dificuldades podem ser resolvidas com a aplicação da técnica de metafronteira.⁴ Para fazer as comparações de eficiência entre grupos de municípios, mensura-se a eficiência relativamente a uma metafronteira comum, definida como a fronteira de um conjunto tecnológico não restrito. A metafronteira é o envelope das fronteiras de grupo que, por sua vez, são fronteiras de conjuntos de tecnologia restritos, em que as restrições derivam de falta de infraestrutura econômica e/ou de outras características do ambiente de produção, como discutido anteriormente. Assim, as eficiências relativas à metafronteira podem ser decompostas em: um componente que mede a distância do ponto de insumo-produto em relação à fronteira de grupo (a medida comum de eficiência técnica) e outro que mede a distância entre a fronteira de grupo e a metafronteira (representando a natureza restritiva do ambiente de produção).

Battese e Rao (2002) tentam comparar as eficiências técnicas de firmas em diferentes grupos com base na função de produção metafronteira estocástica. Eles assumem que existem dois tipos diferentes de mecanismos de geração para os dados: um com relação à fronteira estocástica, usando dados pertencentes àquele grupo, e outro com respeito à metafronteira, que é estimada usando-se os dados da

4 Ver Hayami (1969) e Hayami e Ruttan (1970;1971) para as formulações iniciais.

amostra inteira. O *gap* tecnológico fornece informação sobre a habilidade de as firmas em um grupo competirem com outras de diferentes grupos dentro de uma indústria, uma região ou um país. O problema com essa abordagem é que os pontos na metafronteira estimada podem se situar abaixo dos pontos nas fronteiras de grupo estimadas.

Battese, Rao e O'Donnell (2004) resolvem o problema especificando um único processo de geração de dados que explica os desvios entre os produtos observados e as fronteiras de grupo, ou seja, os modelos de fronteira para firmas nos diferentes grupos. Definem a metafronteira como uma função que envelopa os componentes determinísticos de um conjunto de funções de produção estocásticas para firmas que operam sob diferentes tecnologias (fronteiras de grupo). Contudo, eles só consideram a estimação da metafronteira usando um tipo de metodologia de fronteira estocástica.

Finalmente, O'Donnell, Rao e Battese (2008) usam tanto DEA quanto abordagens alternativas de fronteira estocástica para estimar metafronteiras e fronteiras de grupo, assim como para decompor as diferenças nos desempenhos entre as firmas em eficiência técnica e efeitos de *gap* tecnológico.

Como ilustração da técnica, será aplicada a seguir a metodologia da metafronteira, utilizando as abordagens DEA e Fronteira Estocástica, aos dados disponíveis de educação e saúde.

12.3.1 Educação: aplicação da técnica de metafronteira com abordagem DEA

Para estimar a metafronteira e as fronteiras de grupo será utilizada a metodologia DEA na versão orientada pelo insumo, em que o produto é mantido constante e procura-se a maior redução proporcional no uso de insumos. Estima-se, portanto, uma função metafronteira de custo, que é o envelope das fronteiras de custo específicas para os grupos de municípios.

A estimação seguirá a metodologia proposta por O'Donnell, Rao e Battese (2008), cujo procedimento envolve os seguintes passos:

1. Classificar os S municípios em S_1, S_2, \dots, S_k grupos;
2. Estimar as eficiências $(\alpha_s^{m, S_1}, \alpha_s^{m, S_2}, \dots, \alpha_s^{m, S_k})$ para cada um dos municípios classificados em cada um dos agrupamentos

S_1, S_2, \dots, S_k . Essas eficiências de custo para os grupos individuais serão denominadas CE^k (eficiência de custo);

3. Aplicar o DEA à amostra completa para obter as eficiências com relação à metafronteira (α_s^m). Os escores de eficiência com relação à metafronteira serão denominados CE^* ;
4. As estimativas ($\alpha_s^m / \alpha_s^{m,S_1}, \alpha_s^m / \alpha_s^{m,S_2}, \dots, \alpha_s^m / \alpha_s^{m,S_k}$) são denominadas razões de *gap* tecnológico por Battese et al. (2004) e razões de metatecnologia por O'Donnell, Rao e Battese (2008).

A razão de metatecnologia basicamente avalia o tamanho do *gap* tecnológico para determinado grupo cuja tecnologia correntemente adotada nos municípios está mais restrita que aquela disponível a todos os grupos, representada pela função metafronteira de custo. A um dado nível de produto, a razão de metatecnologia é definida pelo menor custo possível dentro da metafronteira dividido pelo menor custo no grupo específico. Assim, quanto maior o valor médio da razão de metatecnologia para um grupo, mais avançada é a tecnologia de produção que ele adota.

Na prática, é conveniente escrever a eficiência técnica com relação à metafronteira de custo usando a seguinte decomposição:

$$\hat{CE}_i^* = \hat{CE}_i^k \times \hat{MTR}_i^k$$

Onde o primeiro elemento é a eficiência técnica convencional ($\alpha_s^{m,S_1}, \alpha_s^{m,S_2}, \dots, \alpha_s^{m,S_k}$) medindo o desvio do custo efetivo do município da fronteira de custo específica do grupo e o segundo elemento mede o desvio da fronteira específica de grupo da função custo metafronteira. O escore de eficiência da metafronteira de custo reflete quão bem o município se comporta em relação ao desempenho previsto dos pares com as melhores práticas que exploram a melhor tecnologia disponível para todos os grupos, a fim de produzir um dado produto.

Como medida de produto é utilizado o Índice de Desenvolvimento da Educação Básica (Ideb), calculado com base no desempenho dos alunos em testes padronizados (Prova Brasil) e nas taxas de aprovação deles. Os testes são aplicados na quarta e na oitava séries (quinto e nono anos) do ensino fundamental.

O Ideb é apresentado em uma escala de 0 a 10. Uma vez que este índice é uma medida clara de resultado e utilizado como diretriz na política educacional, ele aparece como medida natural de desempenho na área de educação. Os dados são referentes ao Ideb 2011, o último ano para o qual estão disponíveis os resultados contabilizados.⁵

Como insumos, são utilizados os gastos em educação fundamental, obtidos do Fundo Nacional de Desenvolvimento da Educação (FNDE), pelo Sistema de Informações sobre Orçamentos Públicos em Educação (Siope)⁶, em termos *per capita*, e a escolaridade média das mães de alunos da rede pública de ensino fundamental estimada a partir dos microdados do Censo IBGE 2010. Como visto anteriormente, a escolaridade aparece como insumo não discricionário importante para explicar os resultados em educação.

Os municípios são agrupados de acordo com o tamanho de suas populações, segundo o estabelecido na Lei de Responsabilidade Fiscal: até 50 mil habitantes, entre 50 mil e 100 mil habitantes, e acima de 100 mil habitantes. Uma vez que o terceiro grupo engloba ainda municípios muito heterogêneos, ele foi dividido em dois subgrupos: municípios com até 500 mil habitantes e municípios com mais de 500 mil habitantes.

Uma das críticas ao método DEA, e aos métodos não paramétricos em geral, decorre do fato de que a presença de *outliers* pode afetar consideravelmente o cálculo da eficiência. Assim, é preciso, antes de tudo, verificar se observações atípicas estão presentes e, caso estejam, descartá-las. Optou-se pelo método da nuvem de dados (*data cloud method*).⁷

5 Consideramos a média aritmética simples dos valores do Ideb 4ª série e do Ideb 8ª série. Evidentemente, para aqueles municípios onde havia apenas um dos dois números, atribuímos seu valor à variável.

6 Média anual da despesa municipal por aluno, segundo o Siope, para o período 2008 a 2011, a preços constantes (médios de 2010), com deflacionamento feito pelo IPCA mensal médio.

7 A matriz combinada $[XY]$, onde X é a matriz com os insumos e Y é a matriz com os produtos para todos os municípios, contém, então, todas as observações. Essas observações podem ser vistas como uma nuvem de pontos, em que cada ponto representa um município. Se removendo um município dos dados o volume da nuvem de dados fica muito menor, tem-se a indicação de que esse município é um *outlier*. Assim, o método consiste em verificar como o volume da nuvem de dados muda quando uma ou mais observações são removidas. Para maiores detalhes, ver Bogetoft e Otto (2011), capítulo 5 e apêndice, onde é descrito sucintamente o método proposto por Wilson, P. W. (1993) e Wilson, P. W. (2010).

As estimativas para as eficiências técnicas e as razões de metatecnologia quando se assume a hipótese de retornos constantes de escala são resumidas na Tabela 12.1.

Os valores médios das eficiências dos grupos variam entre 0,503 para os municípios menores (até 50.000 habitantes) e 0,812 para os maiores (mais de 500.000 habitantes).

O contrário ocorre com relação às razões de metatecnologia. Os municípios com população acima de 500.000 habitantes possuem os escores de eficiência mais altos entre todos os grupos, mas a razão de metatecnologia é a menor entre os grupos. Os municípios pequenos operam sob tecnologias superiores, mas à custa de escores de eficiência médios mais baixos. Seus custos de produção se situam acima e longe da sua própria fronteira de custo.

A eficiência de custo média relativa à metafronteira dos grupos é igual a 0,50, 0,53, 0,47 e 0,38, respectivamente, o que sugere que os municípios na amostra devem tentar adotar a tecnologia potencial disponível para todos os municípios, de forma a deslocarem suas fronteiras de custo para baixo.

População	Quant. municípios	Eficiência técnica (grupos)	Razões de metatecnologia	Eficiência técnica (metafronteira)
Até 50.000	4341	0,503	0,998	0,5020
50.001-100.000	312	0,567	0,942	0,5337
100.001-500.000	236	0,682	0,686	0,4684
Mais de 500.000	32	0,812	0,467	0,3790

Fonte: Elaboração própria.

As estimativas para as eficiências técnicas e as razões de metatecnologia sob a hipótese de retornos variáveis de escala são resumidas na Tabela 12.2. Os resultados seguem o mesmo padrão dos obtidos usando-se a hipótese de retornos constantes de escala, ainda que os escores de eficiência obtidos sejam mais altos.

**Tabela 12.2 – Estimativas das eficiências técnicas
DEA retornos variáveis de escala**

Habitantes	Quant. municípios	Eficiência técnica (grupos)	Razões de metatecnologia	Eficiência técnica (metafronteira)
Até 50.000	4341	0,564	0,947	0,534
50.001-100.000	312	0,655	0,864	0,566
100.001-500.000	236	0,719	0,680	0,489
Mais de 500.000	32	0,868	0,455	0,395

Fonte: Elaboração própria.

12.3.2 Saúde: aplicação da técnica de metafronteira com abordagem de Fronteira Estocástica

O procedimento de estimação seguirá a abordagem de fronteira estocástica proposta por O'Donnell, Rao e Battese (2008)⁸.

Uma vez que dados de insumos e produtos para as amostras aleatórias dos diferentes grupos de municípios estão disponíveis, é possível estimar a metafronteira e as fronteiras de grupo.

Um modelo de fronteira estocástica para o grupo-k é:

$$y_i = f(x_{1i}, x_{2i}, \dots, x_{Ni}; \beta^k) \times e^{v_i^k - u_i^k} \quad (12.1)$$

Onde x_{ni} é a quantidade do n-ésimo insumo do i-ésimo município; β^k é um vetor de parâmetros desconhecidos associado com o k-ésimo grupo. Os v_i^k s representam ruído estatístico e assume-se que são independentemente e identicamente distribuídos como variáveis aleatórias $N(0, \sigma_{vk}^2)$. Os u_i^k s representam ineficiência e são definidos pela truncagem (em zero) de distribuições $N(\mu_i^k, \sigma_k^2)$, onde os μ_i^k s são definidos por algum modelo de ineficiência apropriada (por exemplo, o modelo de Battese e Coelli – 1995)⁹.

8 Esta seção se baseia em O'Donnell, Rao e Battese (2008). Eles utilizam, contudo, dados em painel, ou seja, observações para diferentes firmas em diferentes períodos de tempo, e aqui serão utilizados dados de corte transversal.

9 Para maiores detalhes sobre os modelos de fronteira estocástica, ver capítulo específico sobre o assunto neste livro.

Se o expoente da função de produção fronteira é linear no vetor de parâmetros (β^k), então o modelo pode ser escrito como:

$$y_i = f(x_{1i}, x_{2i}, \dots, x_{Ni}; \beta^k) \times e^{v_i^k - u_i^k} \equiv e^{x_i' \beta^k + v_i^k - u_i^k} \quad (12.2)$$

Onde x_i é agora um vetor de transformações dos insumos para o i -ésimo município. Dados de insumos e produtos dos municípios no k -ésimo grupo são então usados para estimar os parâmetros desconhecidos desta fronteira.

Depois de obtidas as estimativas, torna-se possível obter a eficiência técnica para o i -ésimo município com relação à fronteira do grupo- k usando o resultado:

$$TE_i^k = \frac{y_i}{e^{x_i' \beta^k + v_i^k}} = e^{-u_i^k} \quad (12.3)$$

Uma função de produção metafronteira determinística é definida como:

$$y_i^* = f(x_{1i}, x_{2i}, \dots, x_{Ni}; \beta) \equiv e^{x_i' \beta} \quad (12.4)$$

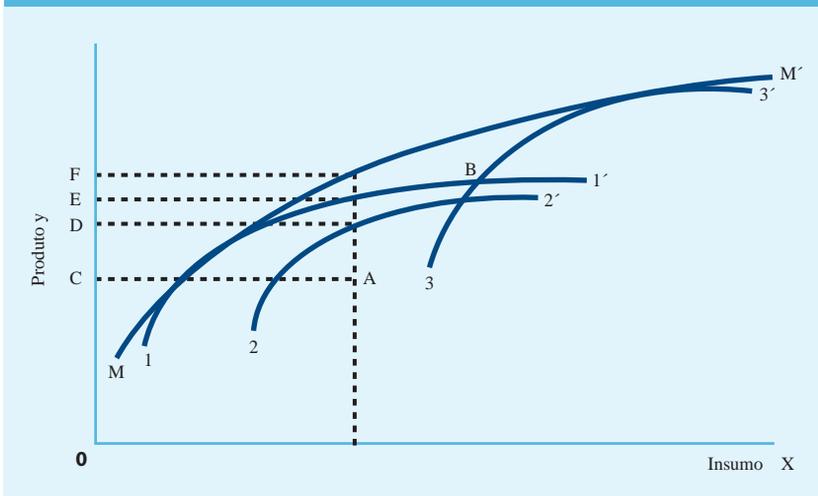
Onde y_i^* é o produto da metafronteira e β é um vetor de parâmetros da metafronteira satisfazendo as restrições $x_i' \beta \geq x_i' \beta^k$ para todo $k=1,2,\dots,K$.

É importante comentar duas características do modelo dado pelas equações (12.1) a (12.4).

Primeiro, as restrições dadas por (12.4) apontam que a função metafronteira não pode se situar abaixo de qualquer das fronteiras de grupo. Segundo, o modelo é caracterizado por um único processo de geração de dados¹⁰.

10 Isso contrasta com o modelo de metafronteira estocástica de Battese e Rao (2002), que assume um mecanismo gerador de dados diferente para a metafronteira e para as diferentes fronteiras de grupo. A metafronteira estocástica definida por eles pode ser estimada usando-se os insumos e os produtos de todas as firmas em todos os grupos e em todos os períodos de tempo, mas não há garantia de que a metafronteira estimada seja o envelope das fronteiras de grupo estimadas.

Gráfico 12.1 – Fronteiras intragrupo e metafronteira



Fonte: Elaboração própria.

Para obter uma função metafronteira estimada que seja o envelope das fronteiras de grupo estimadas é preciso resolver o seguinte problema de otimização:

$$\min_{\beta} \sum_{i=1}^L [\ln f(x_{1i}, x_{2i}, \dots, x_{Ni}; \beta) - \ln f(x_{1i}, x_{2i}, \dots, x_{Ni}; \hat{\beta}^k)] \quad (12.5)$$

Tal que $\ln f(x_{1i}, x_{2i}, \dots, x_{Ni}; \beta) \geq \ln f(x_{1i}, x_{2i}, \dots, x_{Ni}; \hat{\beta}^k)$ para todo i ; onde $\hat{\beta}^k$ é o vetor de coeficientes estimados associados com a fronteira estocástica do grupo- k .

Uma vez que os vetores de coeficientes estimados são fixos para o problema acima, uma forma equivalente da programação linear definida em (12.5) é:

$$\min_{\beta} \sum_{i=1}^T [\ln f(x_{1i}, x_{2i}, \dots, x_{Ni}; \beta)] \quad (12.6)$$

Tal que $\ln f(x_{1i}, x_{2i}, \dots, x_{Ni}; \beta) \geq \ln f(x_{1i}, x_{2i}, \dots, x_{Ni}; \hat{\beta}^k)$ para todo i .

Caso a função $f(\cdot)$ seja log-linear nos parâmetros, o problema de programação linear se torna:

$$\min_{\beta} \bar{x}\beta \text{ tal que } x'_i\beta \geq x'_i\beta^k \text{ para todo } i, \quad (12.7)$$

Onde \bar{x} é a média aritmética dos vetores x_i para todas as firmas.

Desvios padrão para os estimadores dos parâmetros da metafronteira podem ser obtidos usando-se métodos de simulação e *bootstrapping*.

Para obter as estimativas das razões de metatecnologia e eficiências técnicas com relação à metafronteira deve ser feita a seguinte decomposição da equação (12.2):

$$y_i = e^{-u_i^k} \times \frac{e^{x'_i\beta^k}}{e^{x'_i\beta}} \times e^{x'_i\beta + v_i^k} \quad (12.8)$$

O primeiro termo do lado direito é a eficiência técnica do i -ésimo município com relação à fronteira do grupo- k , definida pela equação (12.3). O segundo termo do lado direito é a razão metatecnologia para o i -ésimo município no k -ésimo grupo:

$$MTR_i^k = \frac{e^{x'_i\beta^k}}{e^{x'_i\beta}} \quad (12.9)$$

Para estimar a razão de metatecnologia basta substituir as estimativas de β e β^k na equação (12.9).¹¹

Finalmente, a eficiência técnica do i -ésimo município com relação à metafronteira é:

$$TE_i = \frac{y_i}{e^{x'_i\beta + v_i^k}} \quad (12.10)$$

11 A garantia de que as razões de metatecnologia se situem no intervalo unitário é dada pelas restrições no problema de programação linear dadas pela equação (12.7).

A eficiência técnica relativa à metafronteira é então a razão entre o produto observado e o produto da fronteira, ajustado pelo erro aleatório correspondente. Na prática, a eficiência técnica com relação à metafronteira é prevista usando-se a seguinte decomposição:

$$\hat{TE}_i = \hat{TE}_i^k \times \hat{MTR}_i^k \quad (12.11)$$

Onde \hat{TE}_i^k e \hat{MTR}_i^k são os previsores relativos às equações (12.3) e (12.9).

A eficiência técnica medida com referência à metafronteira (representando o estado de conhecimento existente) pode ser decomposta no produto entre a eficiência técnica medida com referência à fronteira do grupo-k e a razão de metatecnologia para o grupo k (representando o estado de conhecimento existente e o ambiente físico, social e econômico que caracteriza o grupo k).

Para implementação da técnica, é utilizado o Índice de Desempenho do SUS (IDSUS, divulgado pelo Ministério da Saúde em 2012: <<http://portal.saude.gov.br/portal/saude/area.cfm?id_area=1080>>) como produto. O IDSUS é um indicador síntese que procura medir o desempenho do sistema de saúde a partir de uma série de indicadores que refletem tanto acesso quanto efetividade.

Como insumos, são utilizados os gastos *per capita* em saúde. Os gastos em saúde feitos pelos municípios têm como fonte a Secretaria do Tesouro Nacional, de maneira específica, as despesas na função saúde obtidas no banco de dados *Finanças do Brasil: Dados Contábeis dos Municípios (Finbra)*, a preços médios de 2010 e deflacionamento pelo IPCA mensal médio de cada ano. Já os dados de população e de educação (escolaridade das pessoas com ao menos 25 anos de idade) têm como fonte o Censo 2010 divulgado pelo Instituto Brasileiro de Geografia e Estatística (IBGE).¹²

12 No tocante à variável de educação, buscou-se reconstruir a variável *anos de estudo*, atribuindo-lhe os seguintes valores conforme o valor assumido pela variável V6400 nos microdados do Censo 2010, nível de instrução: (i) sem instrução até fundamental incompleto => 0 anos de estudo; (ii) fundamental completo até médio incompleto => 9 anos de estudo; (iii) médio completo e superior incompleto => 12 anos de estudo; (iv) superior completo => 15 anos de estudo.

Existe certo consenso na literatura de que tanto os gastos com saúde quanto a escolaridade afetam os resultados em saúde. Assim, a função de produção assumirá a seguinte forma (HOLLINGSWORTH; WILDMAN, 2003; GREENE, 2004):

$$Saúde_i = f(Educação_i, Gasto_i) + v_i - u_i$$

A forma funcional da função de produção precisaria, contudo, ser determinada. Inicialmente será especificado um modelo translog:

$$\begin{aligned} \text{Logsaúde} = & \alpha + \beta_1 \times \log \text{gasto} + \beta_2 \times \log \text{educ} + \beta_{11} \times [(\log^2 \text{educ})/2] + \\ & \beta_{32} \times [(\log^2 \text{gasto})/2] + \beta_{12} \times [(\log \text{educ}) \times (\log \text{gasto})] + v - u \end{aligned}$$

As estimativas foram obtidas a partir da adaptação para o programa R do código Shazam fornecido por O'Donnell.

Os municípios, como no caso do gasto em educação, são agrupados de acordo com o número de habitantes. A opção por esse tipo de agrupamento parece particularmente adequada no caso da saúde, pois existem várias evidências na literatura brasileira de que escala é um fator importante para explicar a (in)eficiência dos municípios. Sampaio de Souza, Criabari-Neto e Stosic (2005), por exemplo, usam como variável de escala a densidade demográfica e observam que cidades com baixa concentração populacional gastam relativamente mais. A baixa concentração populacional leva a um aumento dos custos médios de provisão dos serviços públicos, o que conduz a uma incapacidade de exploração das economias de escala associadas à produção daqueles serviços e, em última instância, à utilização subótima dos recursos.

As estimativas dos modelos de fronteira definidos para os agrupamentos e para a metafronteira foram inicialmente obtidas assumindo-se a forma funcional translog, como discutido anteriormente. Para alguns agrupamentos, observa-se que a maioria das estimativas dos parâmetros das fronteiras de grupo é estatisticamente significantes pelo menos ao nível de 10%. Contudo, para garantir a robustez dos resultados, as estimações das fronteiras de grupo foram refeitas utilizando-se outras duas formas funcionais para a função de produção de saúde. A primeira é a formulação da OMS que elimina da função translog, por questões de parcimônia, o termo do quadrado do gasto em saúde e o termo cruzado. As eficiências médias obtidas para cada um dos grupos, usando-se as

três formas funcionais diferentes, foram muito similares, o que trouxe confiança de que os resultados independem do formato específico assumido para a função de produção. A segunda é uma formulação Cobb-Douglas, em que são considerados somente os gastos em saúde *per capita* e o nível de escolaridade. São apresentados somente os resultados para a Cobb-Douglas, mas os demais resultados podem ser obtidos diretamente com os autores.

Os resultados da estimação das fronteiras de grupos são apresentados na Tabela 12.3.

Como é possível observar, há uma associação positiva tanto dos gastos em saúde quanto do nível de educação com os resultados em saúde, conforme o esperado.

Tabela 12.3 – Estimativas das fronteiras de grupo			
Variáveis	Até 50 mil habitantes	Mais de 50 mil a 100 mil habitantes	Mais de 100 mil habitantes
Intercepto	0,88428*** (0,02746)	1,21493*** (0,11818)	1,08106** (0,13182)
Gastos	0,15484*** (0,004759)	0,05367** (0,0237)	0,09722*** (0,02104)
Educação	0,06069*** (0,006658)	0,18829*** (0,02689)	0,08199 (0,05398)
Lambda	1,83388*** (0,100644)	3,20755*** (0,67256)	1,22529*** (0,4467)
No.Obs.	4919	324	280

Fonte: Elaboração própria.

Obs: ***, ** e * indicam significâncias estatísticas aos níveis de 1%, 5% e 10%, respectivamente. Desvios padrão entre parêntesis.

Uma vez estimadas e analisadas as fronteiras de produção dos agrupamentos, a próxima questão é saber se os municípios dos diferentes grupos operam sob um único tipo de tecnologia. Se todos os municípios dividem a mesma tecnologia, então é desnecessário analisar os dados utilizando um modelo de metafronteira.

É importante chamar a atenção, contudo, para o fato de que no caso de municípios fala-se em tecnologia e utiliza-se o conceito de

razão de *gap* tecnológico de Battese e Rao (2002), Battese, Rao e O'Donnell (2004) e O'Donnell, Rao e Battese (2008), apesar de as unidades de análise não serem firmas privadas. Fala-se em diferentes tecnologias somente a fim de facilitar a comparação com o trabalho deles, mas o que simplesmente está sendo sugerido é que o efeito da hipótese sob as condições ambientais (população) pode ser importante para diferenciar os municípios.

Para isso é estimada uma fronteira estocástica usando-se os dados de todos os municípios conjuntamente, sem levar em conta as possíveis diferenças tecnológicas – modelo *pooled* de Battese e Rao (2002). A diferenciação entre as tecnologias é testada formalmente com a utilização do teste de Wald. Verifica-se inicialmente se as fronteiras de cada grupo são iguais entre si; em seguida, se as fronteiras dos grupos são iguais à do modelo *pooled*. As hipóteses nulas de igualdade entre as fronteiras são rejeitadas ao nível de 10%, levando à conclusão de que a metodologia de metafronteira é apropriada.

As estimativas para as eficiências técnicas e as razões de metatecnologia são resumidas na Tabela 12.4.

Tabela 12.4 – Estimativas das eficiências técnicas e das razões de metatecnologia					
	Quant. municípios	Média	Desvio padrão	Mínimo	Máximo
Eficiências intra-grupo					
Até 50 mil habitantes	4919	0,884	0,062	0,531	0,987
Mais de 50 mil a 100 mil habitantes	324	0,841	0,096	0,591	0,974
Mais de 100 mil habitantes	280	0,901	0,043	0,746	0,966
Razões de meta-tecnologia					
Até 50 mil habitantes	4915	0,963	0,011	0,873	1,000
Mais de 50 mil a 100 mil habitantes	324	0,918	0,028	0,826	1,000
Mais de 100 mil habitantes	280	0,862	0,013	0,830	0,897
Eficiências meta-fronteira					
Até 50 mil habitantes	4911	0,851	0,061	0,505	0,954
Mais de 50 mil a 100 mil habitantes	324	0,772	0,091	0,524	0,942
Mais de 100 mil habitantes	280	0,777	0,038	0,626	0,853

Fonte: Elaboração própria.

Os resultados indicam que um município pequeno qualquer (com população menor do que 50 mil habitantes) é 88% eficiente quando medido relativamente à fronteira de municípios pequenos. Isso significa que o produto médio nesse município corresponde a 88% do que seria possível obter usando os mesmos níveis de insumo e tecnologia disponíveis para os municípios pequenos. A razão de metatecnologia média para os municípios pequenos é 0,96, indicando que o produto máximo que poderia ser obtido usando-se os insumos de um município pequeno e a tecnologia dos municípios pequenos é 96% do produto máximo que poderia ser obtido usando-se os mesmos insumos e a tecnologia representada pela metafronteira. Finalmente, a eficiência dos municípios pequenos, medida em relação à metafronteira, é 85%.

Para os outros grupos, a interpretação dos resultados é similar.

Como observado por O'Donnell, Rao e Battese (2008), as estimativas de eficiência técnica são calculadas para utilização, em última instância, em programas de melhoria de desempenho a partir de mudanças no gerenciamento dos municípios. As razões de metatecnologia (*gap* entre as fronteiras de grupo e a metafronteira), por sua vez, podem ser utilizadas em programas que buscam mudar as características do ambiente em que a produção ocorre (por exemplo, infraestrutura, financiamento etc.).

REFERÊNCIAS

AFONSO, A.; ST. AUBYN, M. Non-parametric approaches to education and health efficiency in OECD countries. *Journal of Applied Economics*, v. 8, n. 2, p. 227-246, nov. 2005.

_____. Cross-country efficiency of secondary education provision: a semi-parametric analysis with non-discretionary inputs. *Economic Modeling*, v. 23, n. 3, p. 476-491, 2006.

ALVES JR. J. A.; SAMPAIO DE SOUSA; M. C. Eficiência na aplicação de recursos educacionais em presença de condicionantes exógenos e efeitos aleatórios: uma abordagem em três estágios. In: ENCONTRO BRASILEIRO DE ECONOMETRIA, 33., 2011, Foz do Iguaçu. *Anais... Foz do Iguaçu: Sociedade Brasileira de Econometria*, 2011. 1 CD-ROM.

BATTESE, G.; COELLI, T. Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data. *Journal of Econometrics*, v. 38, p. 387-399, 1988.

_____. A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics*, v. 20, n. 2, p. 325-332, 1995.

BATTESE, G. E.; RAO, D. S. P. Technology gap, efficiency, and a stochastic metafrontier function. *International Journal of Business and Economics*, v. 1, n. 2. p. 87-93, 2002.

BATTESE, G. E.; RAO, D. S. P.; O'DONNELL, C. J. A metafrontier production function for estimation of technical efficiencies and technology gaps for firms operating under different technologies. *Journal of Productivity Analysis*, v. 21, p. 91-103, 2004.

BOGETOFT, P.; OTTO, L. *Benchmarking, with DEA, SFA and R*. Nova York: Springer, 2011.

BRUNET, J. F. G.; BORGES, C. B.; BERTÊ, A. M. A.; BUSSATO, L. M. *Estados comparados por funções do orçamento: uma avaliação da eficiência e efetividade dos gastos públicos estaduais*. Prêmio IPEA-Caixa, Menção Honrosa I, 2006.

CLEMENTS, B. How efficient is public spending in Europe? *European Review of Economics and Finance*, v. 1, p. 3-26, 2002.

EVANS, D. B.; TANDOM, A.; MURRAY, C. J. L.; LAUER, J. A. The comparative efficiency of national health systems in producing health: an analysis of 191 countries. Geneva: World Health Organization, 2000. (GPE Discussion Paper Series, n. 29).

FRIED, H. O.; SCHMIDT, S. S.; YAISAWARNG, S. Incorporating the operating environment into a nonparametric measure of technical efficiency. *Journal of Productivity Analysis*, v. 12, p. 249-267, 1999.

FRIED, H. O.; LOVELL, C. A. K.; SCHMIDT, S. S.; YAISAWARNG, S. Accounting for environmental effects and statistical noise in Data Envelopment Analysis. *Journal of Productivity Analysis*, v. 17, p. 154-174, 2002.

GIMENEZ, V.; PRIOR, D.; THIEME, C. Technical efficiency, managerial efficiency and objective setting in the educational system: an international comparison. *Journal of the Operational Research Society*, v. 58, p. 996-1007, 2007.

GRAVELLE, H.; JACOBS, R.; JONES, M. A.; STREET, A. Comparing the efficiency of national health systems: a sensitivity analysis of the WHO approach. *Applied Health Economics and Health Policy*, v. 2, n. 3, p. 1-7, 2003.

GREENE, W. H. Distinguishing between heterogeneity and inefficiency: stochastic frontier analysis of the World Health Organization's panel data on national health care systems. *Health Economics*, v. 13, n. 10, p. 959-980, out. 2004.

GUPTA, S.; VERHOEVEN, M.; TIONGSON, E. R. The effectiveness of government spending on education and health care in developing and transition economies. *European Journal of Political Economy*, v. 18, p. 717-737, 2002.

HANUSHEK, E.; KINKO, D. Schooling, labor force quality, and economic growth. *American Economic Review*, v. 90, n. 5, p. 1184-1208, 2000.

HANUSHEK, E.; LUQUE, J. Efficiency and equity in schools around the world. *Economics of Education Review*, v. 22, p. 481-502, 2003.

HAYAMI, Y. Sources of agricultural productivity gap among selected countries. *American Journal of Agricultural Economics*, v. 51, p. 564-575, 1969.

HAYAMI, Y.; RUTTAN, V. W. Agricultural productivity differences among countries. *American Economic Review*, v. 40, p. 895-911, 1970.

_____. *Agricultural development: an international perspective*. Baltimore: Johns Hopkins University Press, 1971.

HOLLIHGSWORTH, B.; WILDMAN, J. The efficiency of health production: re-estimating the WHO panel data using parametric and non-parametric approaches to provide additional information. *Health Economics*, v. 12, p. 493-504, 2003.

MARINHO, A. Avaliação da eficiência técnica nos serviços de saúde nos municípios do Estado do Rio de Janeiro. *Revista Brasileira de Economia*, v. 57, n. 3, p. 515-534, 2003.

MATTOS, E.; ROCHA, F.; NOVAES, L.; AVARTE, P.; ORELLANO, V. Economias de escala na oferta de serviços públicos de saúde: um estudo para os municípios paulistas. *Economia*, v. 10, n. 2, p. 357-386, 2009.

O'DONNELL, C. J.; RAO, D. S. P.; BATTESE, G. E. Metafrontier frameworks for the study of firm-level efficiencies and technology ratios. *Empirical Economics*, v. 34, p. 231-255, 2008.

PANG, G.; HERRERA, S. Efficiency of public spending in developing countries: an efficiency frontier approach. *World Bank Policy Research Working Paper*, n. 3645, 2005.

SAMPAIO DE SOUZA, M. C.; CRIABARI-NETO, F.; STOSIC, B. D. Explaining DEA technical efficiency scores in an outlier corrected environment: the case of public services in Brazilian municipalities. *Brazilian Review of Econometrics*, v. 25, n. 2, p. 287-313, 2005.

SOUZA, I. V.; NISHIJIMA, M.; ROCHA, F. Eficiência do setor hospitalar nos municípios paulistas. *Economia Aplicada*, v. 14, n. 1, p. 51-66, 2010.

SUTHERLAND, D.; PRICE, R.; JOUMARD, L.; NICQ, C. Performance indicators for public spending efficiency in primary and secondary education. *OECD Economics Department Working Paper*, n. 546, 2007.

Wilson, P.W. (1993), Detecting outliers in deterministic nonparametric frontier models with multiple outputs, *Journal of Business and Economic Statistics* 11, 319-323.

Wilson, P.W. (2010), Detecting outliers in deterministic nonparametric frontier models with multiple outputs: Correction, unpublished working paper, Department of Economics, Clemson University, Clemson, South Carolina 29634 (available on-line at <http://www.clemson.edu/economics/faculty/wilson/Papers/ap-corrected.pdf>).

Capítulo 13

Detecção de *outliers* em modelos não paramétricos: o método *Jackstrap* ampliado

Maria da Conceição Sampaio de Sousa
Borko D. Stosic

13.1 Introdução

Métodos não paramétricos como a abordagem Análise Envoltória de Dados (DEA¹) são particularmente apropriados para avaliar a eficiência das organizações públicas sem fins lucrativos por várias razões. Não somente eles produzem um indicador agregado da eficiência relativa dessas organizações, construído a partir de múltiplos insumos e múltiplos produtos, mas também porque esse tipo de metodologia não determina *a priori* os pesos atrelados aos diferentes insumos e produtos, que entram no cálculo dos escores de eficiência. Por fim, a análise não paramétrica permite também levar em conta fatores externos, que estão fora do controle dos responsáveis pelas organizações, e dispensa o uso de preços no cálculo do índice agregado.

No entanto, o uso de medidas não paramétricas *naive*² pode comprometer a credibilidade dos escores de eficiência, porque nesses modelos determinísticos todas as observações são factíveis com probabilidade um. As fronteiras de eficiência assim computadas são muito sensíveis ao tamanho da amostra e à dimensão do modelo (número de insumos e produtos), mas principalmente à existência de observações discrepantes referentes às DMUs (*Decision Making Units*) que exibem desempenho com baixa probabilidade de ocorrência, doravante mencionadas como *outliers*.

Esse ponto é particularmente importante em abordagens que se baseiam em técnicas de programação linear para mensurar a

1 Charnes, Cooper and Rhodes (1978); Banker, Charnes e Cooper (1984); Färe, Grosskopf e Lovell (1985). Análise Envoltória de Dados é a tradução literal do termo em inglês *Data Envelopment Analysis* (DEA).

2 A literatura refere-se ao DEA *naive* como a aplicação direta da metodologia DEA a determinada base de dados.

eficiência, como é o caso da metodologia DEA (TIMMER, 1971). Como nessa metodologia as fronteiras de eficiência são construídas com base em observações extremas, a presença de *outliers* reduz, artificialmente, os escores DEA para as demais DMUs da amostra, subestimando, assim, seus níveis de desempenho. Ademais, essa subestimação altera não somente a magnitude, mas também modifica a ordenação dos escores de eficiência, o que compromete a comparação entre DMUs. Embora alguns desses *outliers* correspondam, de fato, a observações atípicas e boas práticas de gestão, em muitos casos eles resultam de erros de medida e de outras discrepâncias estatísticas, não consideradas na versão convencional (*naive*) do modelo DEA.

Em pequenas bases de dados, essas informações errôneas podem ser mais facilmente detectadas. Não ocorre o mesmo, porém, em grandes bases de dados, nas quais a inspeção manual das informações não é viável. Portanto, para garantir que os escores de eficiência DEA sejam efetivamente úteis no processo decisório, é necessário incluir na abordagem DEA o tratamento das discrepâncias supracitadas. Isso inclui (i) o cômputo de medidas que permitam a identificação automática de *outliers* que possam ser implementadas em bases de qualquer tamanho; e (ii) a fixação de regras de inclusão/exclusão de *outliers*, quando do cálculo das medidas DEA de eficiência.

Vários autores propuseram técnicas para detectar observações influentes no cômputo da eficiência não paramétrica. Wilson (1993; 1995) utilizou métodos descritivos para detectar observações influentes no cômputo da eficiência não paramétrica. Seaver e Triantis (1992; 1995) combinaram procedimentos da lógica *fuzzy*, em que a análise de eficiência é usada junto com a programação paramétrica *fuzzy*, com medidas robustas de distância, para a detecção de *outliers*. Andersen e Peterson (1993) recorreram ao modelo de supereficiência para detectar observações discrepantes, que recebem escores superiores à unidade, mediante a exclusão dessas DMUs da coluna avaliada no programa linear.

Mais recentemente, a Fronteira de Ordem-m (CAZALS; FLORENS; SIMAR, 2002; SIMAR, 2003) e a Fronteira de Eficiência Robusta (CHERCHYE; KUOSMANENE; POST, 2000) ampliaram a lista de abordagens para identificar e corrigir observações discrepantes.

A primeira baseia-se no conceito de insumo esperado mínimo (produto esperado máximo) e gera fronteiras de graus variados de robustez. A Fronteira de Eficiência Robusta decompõe a fronteira DEA original em diferentes conjuntos de referência, sendo a fronteira mensurada em relação a esses conjuntos. Note-se, porém, que todas essas abordagens, além de se basearem em inspeção manual dos dados, utilizam regras *ad-hoc* de exclusão de *outliers*. Esses aspectos são particularmente restritivos quando se trabalha com grandes bases de dados.

Para sanar esses problemas e, assim, obter escores de eficiência mais robustos, Sampaio de Sousa e Stosic (2005) desenvolveram uma metodologia denominada *Jackstrap*, que combina as técnicas de amostragem *Jackknife* e *Bootstrap* para calcular uma medida que mostra o efeito da remoção de uma DMU sobre os escores de eficiência das demais componentes da amostra. Baseando-se nessa medida – doravante mencionada como alavancagem (*leverage*) – podem-se também distinguir as observações influentes, cuja remoção afeta substancialmente o desempenho das demais, daquelas que, embora apresentem altas alavancagens, têm pouco (ou nenhum) impacto sobre o desempenho de seus pares.

O objetivo deste capítulo é apresentar o método *Jackstrap* Ampliado, que vai além dos limites metodológicos da abordagem *Jackstrap*, no sentido de usar as informações sobre alavancagens para reduzir a probabilidade de selecionar os *outliers* no processo estocástico de reamostragem. Não se trata aqui de descartá-los, mas somente de impedir que essas informações atípicas distorçam os escores de eficiência das demais DMUs. Essa abordagem compreende duas fases. Na fase 1, computam-se as alavancagens para cada DMU; na fase 2, essas medidas são utilizadas para excluir as observações discrepantes e/ou selecioná-las com menor probabilidade. Por fim, procede-se ao cálculo dos escores robustos de eficiência.

Este capítulo organiza-se da seguinte maneira. A seção 13.2 faz uma breve descrição da metodologia *Jackstrap*. A seção 13.3 detalha o método *Jackstrap* Ampliado e o cálculo dos índices de eficiência DEA robustos. Finalmente, a seção 13.4 contém a conclusão do capítulo.

13.2 O método *Jackstrap*

A abordagem *Jackstrap* baseia-se nos escores de eficiência DEA. Busca-se aqui não somente mensurar a influência de uma dada DMU no cômputo desses escores para as demais observações, mas também decidir como incluí-la na amostra, de forma a minimizar a distorção por ela causada nas medidas de eficiência DEA.

Para tal, o procedimento *Jackstrap* combina as técnicas de reamostragem *Jackknife* (determinística) e *Bootstrap* (estocástica) para extrair o impacto da remoção de uma dada DMU sobre o cálculo dos escores de eficiência DEA para o resto da amostra. Esse impacto corresponde à alavancagem (*leverage*) dessa DMU. Note-se que o uso do termo alavancagem para caracterizar observações discrepantes foi emprestado da análise econométrica (BELSLEY, 1980) e é usado aqui de forma análoga, isto é, refere-se à influência de observações com valores atípicos para os insumos ou produtos que compõem a base de dados analisada. A ideia subjacente é a de que as observações ditas *outliers* apresentam altas alavancagens, bem superiores à média da amostra. Por essa razão, as observações devem ter menor probabilidade de serem selecionadas quando da composição da amostra a partir da qual os escores DEA serão computados. A essência da abordagem *Jackstrap* é, pois, *reduzir estocasticamente o impacto de poucas observações muito influentes sobre os escores de eficiência finais. Na seção subsequente, explicaremos em detalhes o método Jackstrap Ampliado para detectar e selecionar observações discrepantes, permitindo, assim, o cômputo de escores de eficiência DEA robustos.*

13.3 Detecção e inclusão de *outliers*

A abordagem *Jackstrap Ampliado* compreende duas fases. A primeira etapa corresponde ao método proposto por Sampaio de Sousa e Stosic (2005), no qual as medidas de influência, doravante mencionadas como alavancagens, são computadas (subseção 13.3.1). A segunda fase utiliza as alavancagens obtidas na fase 1 para identificar os *outliers* e selecioná-los para formarem as amostras aleatórias que servirão de base para o cálculo dos escores robustos de eficiência (subseções 13.3.2 e 13.3.3).

13.3.1 Fase 1: cálculo das alavancagens

Vamos, de início, definir o conceito de alavancagem ℓ_j , da j -ésima DMU. Essa medida corresponde ao desvio padrão das medidas de eficiência após a sua remoção em relação às medidas antes da sua retirada. Se computada por meio da técnica de reamostragem *Jackknife* ℓ_j é dada pela seguinte expressão, na qual n é o número de DMUs na amostra:

$$\ell_j = \sqrt{\frac{\sum_{k=1; k \neq j}^n (\theta_{kj}^* - \theta_k)^2}{n-1}} \quad (13.1)$$

Nesse procedimento, os escores de eficiência $\{\theta_k \mid k = 1, \dots, n\}$ são obtidos utilizando-se o conjunto de dados original; em seguida, remove-se uma DMU por vez e recalculam-se os escores $\{\theta_{k,j}^* \mid k = 1, \dots, n; k \neq j\}$. Então $\theta_{k,j}^*$ é o escore de eficiência da DMU k quando este é calculado sem a presença da DMU j .

O problema do uso do método *Jackknife* nesse contexto advém do fato de ele ser computacionalmente intensivo. Mais especificamente, para retirar cada DMU e realizar $(n-1)$ cálculos de DEA, exige-se a resolução de $n(n-1)$ problemas de programação linear, dificultando a sua execução em grandes bancos de dados, mesmo com as atuais disponibilidades computacionais. Para contornar esse problema, Sampaio de Sousa e Stosic (2005) propuseram o método *Jackstrap*, que combina a técnica *Jackknife* com o método de reamostragem *Bootstrap*, de acordo com os seguintes passos:

1. Seleciona-se aleatoriamente uma amostra de L DMUs que comporte entre 10% e 20% do total de DMUs, e calcula-se um subconjunto de alavancagens $\tilde{\ell}_j$ utilizando a técnica *Jackknife*, onde o índice j assume L valores (selecionados aleatoriamente) do conjunto $\{1, \dots, n\}$.
2. Repete-se B vezes o passo 1, acumulando os valores das alavancagens calculadas com base nas diversas amostras aleatórias. Para B suficientemente grande, cada DMU deve ser selecionada n vezes, onde $n_j \approx BL/n$.

3. Calcula-se a alavancagem média para cada DMU como:

$$\tilde{\ell}_k = \frac{\sum_{b=1}^{n_k} \tilde{\ell}_{kb}}{n_k} \quad (13.2)$$

4. A alavancagem média global escreve-se como:

$$\tilde{\ell} = \frac{\sum_{k=1}^K \tilde{\ell}_k}{K} \quad (13.3)$$

As alavancagens ℓ_k , variam entre 0 e 1; $\ell_k = 0$ significa que a remoção de uma dada DMU não tem efeito sobre os escores de eficiência das outras DMUs, enquanto $\ell_j = 1$ implica que a retirada da DMU observada altera todos os outros escores de zero para um. Completa-se, então, a primeira fase da abordagem *Jackstrap*.

Sampaio de Sousa e Stosic (2004) investigaram o impacto do tamanho das subamostras aleatórias L e B sobre as alavancagens calculadas e mostraram que à medida que L aumenta, ℓ_k diminui. Esses autores evidenciaram, ainda, que a ordenação das alavancagens é preservada no espectro dos diferentes valores de L.

Ressalte-se, também, que o fato de as DMUs serem avaliadas dentro de inúmeras subamostras aleatórias contribui para tornar o método *Jackstrap* relativamente insensível ao efeito máscara (*masking effect*).³ Isso porque a inspeção aleatória da base de dados revela não somente as observações discrepantes que compõem as várias fronteiras de eficiências, mas também aquelas que influenciam observações no interior dessa fronteira. Por essa razão, nossa abordagem permite identificar também os *outliers locais*.

13.3.2 Fase 2: seleção de outliers

Na segunda fase do procedimento *Jackstrap*, utilizam-se as alavancagens obtidas para identificar os *outliers* e decidir

3 Este efeito refere-se ao fato de que o procedimento de detecção pode não identificar *outliers* que estão grupados. Nesse caso, um ou mais desses *outliers* podem esconder (mascarar) os demais e/ou não reconhecê-los.

como incluí-los na amostra. Uma possibilidade é remover, sistematicamente, uma a uma as observações com maiores alavancagens e comparar as sucessivas distribuições empíricas das eficiências resultantes desse processo. O impacto da remoção das observações influentes pode, então, ser usado para fixar os patamares a partir dos quais as DMUs seriam excluídas da amostra.

O problema maior desse procedimento advém do fato de que ele se baseia em regras *ad-hocs* de exclusão de *outliers*. Outra alternativa – usada por Sampaio de Sousa e Stosic (2005) – envolve a aplicação do teste de Kolmogorov–Smirnov (K–S) para quantificar as diferenças entre as distribuições pré e pós remoção da j -ésima DMU influente e, então, observar o ponto a partir do qual as diferenças entre as duas distribuições deixam de ser significantes. Essa abordagem é bastante intensiva em termos computacionais, sobretudo quando se trata de grandes bases de dados.

Por fim, é possível usar as informações sobre alavancagem, calculadas na fase 1, para reduzir a probabilidade de selecionar os *outliers* no processo estocástico de reamostragem. Não se trata aqui de descartá-los – embora isso possa ser feito, por meio de uma função específica de probabilidade, como a *heaviside* – mas somente de impedir que essas informações atípicas distorçam os escores de eficiência das demais DMUs. Em ambos os casos, é necessário adotar funções de probabilidades que penalizem aquelas com altas alavancagens.

Dentre as possíveis escolhas dessas funções, quatro são comumente utilizadas: linear, exponencial, inversa e *heaviside (step)*. A seguir, descreveremos essas funções.

A função de probabilidade linear é dada por:

$$P(\tilde{\ell}_k) = \frac{\ell_{\max} - \tilde{\ell}_k}{\ell_{\max} - \ell_{\min}} \quad (13.4)$$

onde ℓ_{\max} e ℓ_{\min} correspondem, respectivamente, às alavancagens máxima e mínima do conjunto $\{\tilde{\ell}_k; k = 1, \dots, K\}$. A probabilidade de se reter uma DMU com alavancagem ℓ_{\min} é, portanto, igual a um; aquela referente à DMU com maior alavancagem (ℓ_{\max}) é zero.

A distribuição de probabilidade inversa pode ser representada como:

$$P(\tilde{\ell}_k) = \begin{cases} 1 & , \quad \tilde{\ell}_k < \ell_0 \\ \left(\frac{\ell_0}{\tilde{\ell}_k}\right) \left(\frac{\ell_{\max} - \tilde{\ell}_k}{\ell_{\max} - \ell_0}\right) & , \quad \tilde{\ell}_k \geq \ell_0 \end{cases} \quad (13.5)$$

onde $\ell_0 > 0$ é o limite inferior para $\tilde{\ell}_k$, isto é, trata-se de um parâmetro independente (*leverage threshold*) abaixo do qual as DMUs são retidas com probabilidade igual a 1. O limite inferior ℓ_0 pode ser arbitrariamente pequeno.

A distribuição exponencial é dada por:

$$P(\tilde{\ell}_k) = \frac{e^{-\tilde{\ell}_k} - e^{-\ell_{\max}}}{e^{-\ell_{\min}} - e^{-\ell_{\max}}} \quad (13.6)$$

Finalmente, a função *heaviside* (ou *step*) escreve-se como:

$$P(\tilde{\ell}_k) = \begin{cases} 1 & , \quad \tilde{\ell}_k < \tilde{\ell} \log K \\ 0 & , \quad \tilde{\ell}_k \geq \tilde{\ell} \log K \end{cases} \quad (13.7)$$

Aqui, o ponto de corte $\tilde{\ell} \log K$ é escolhido de modo a levar em conta o tamanho da amostra, de sorte que para $K=1000$, por exemplo, rejeita-se uma DMU cuja alavancagem é superior ao triplo da média global.

13.3.3 Escores de eficiência robustos

O cômputo dos escores de eficiência robustos é feito da seguinte maneira:

1. Utiliza-se uma das funções de probabilidade descritas pelas equações 13.4 a 13.7 para definir as DMUs a serem incluídas na pseudoamostra aleatória, com reposição;
2. Definida a amostra, calcula-se o escore de eficiência, θ_{kb} , para cada municipalidade;

3. Repete-se B vezes os passos 1 e 2, acumulando os valores dos escores de eficiência, que são calculados com base nas diversas amostras aleatórias, geradas por meio da função de probabilidade escolhida;
4. Calculam-se os escores robustos de eficiência média para cada DMU como:

$$\hat{\theta}_k = \frac{\sum_{b=1}^B \theta_{kb}}{B} \quad (13.8)$$

Note-se que nossa abordagem é inspirada no método *Bootstrap*. Porém, em vez de fazermos a reamostragem nos insumos e produtos originais (ZHANG; BARTELS, 1998) ou nos escores de eficiência (SIMAR; WILSON, 1998), ela é feita com base nas alavancagens, que guiarão a seleção aleatória, com reposição das DMUs que irão compor as B amostras, a partir das quais os escores DEA serão calculados.

O procedimento acima descrito é consistente tanto com a visão instrumentalista quanto com a versão convencionalista da análise de eficiência. Na visão instrumentalista, os escores de eficiência DEA são vistos como medidas descritivas do desempenho relativo das DMUs (MCDONALD, 2009). Nessa interpretação, a fronteira de eficiência corresponde ao construto das melhores práticas observadas. A visão concorrente – também chamada de convencionalista – considera esses escores como estimadores da eficiência verdadeira das DMUs (BANKER; NATARAJAN, 2008; SIMAR; WILSON, 2007). Esse ponto é particularmente relevante para a escolha do método econométrico a ser usado na análise DEA em dois estágios, cujos escores robustos de eficiência constituem as variáveis dependentes, no segundo estágio.

13.4 Considerações finais

Neste capítulo, apresentamos o método *Jackstrap* Ampliado, que permite não somente identificar observações influentes (*outliers*) na metodologia DEA, mas também incluí-las no cômputo dos escores de eficiência, selecionando-as, porém, com menor probabilidade.

Uma das vantagens dessa abordagem advém do fato de ela ser relativamente insensível ao efeito máscara (*masking effect*).

De fato, como as DMUs são avaliadas dentro de inúmeras subamostras aleatórias, o procedimento *Jackstrap* Ampliado permite identificar *outliers* que estão grupados. Esse grupamento, que pode esconder (mascarar) alguns *outliers* e/ou não reconhecê-los, é dificilmente captado pelas metodologias de detecção de observações discrepantes.

Ademais, por ser um método completamente automático, ele é particularmente apropriado para tratar com grandes bases de dados. Por fim, ao permitir a inclusão dos *outliers* no cômputo dos escores de eficiência, em vez de removê-los por meio de regras *ad-hoc*, o procedimento *Jackstrap* Ampliado faz melhor uso da informação existente.

No próximo capítulo, ilustraremos a metodologia supracitada mediante sua aplicação aos serviços da Atenção Básica à Saúde nos municípios brasileiros.

REFERÊNCIAS

ANDERSON, P.; PETERSON, N. C. A procedure for ranking efficient units in data envelopment analysis. *Management Science*, v. 39, p. 1261-1264, 1993.

BANKER, R. D.; CHARNES, A.; COOPER, W. W. Some models for estimating technical and scale efficiencies in data envelopment analysis. *Management Science*, v. 30, p. 1078-1092, 1984.

BANKER, R. D.; NATARAJAN, R. Evaluating contextual variables affecting productivity using data envelopment analysis. *Operations Research*, v. 56, n. 1, p. 48-58, 2008.

BELSLEY, D. A.; KUH, E.; WELSCH, R. H. *Regression diagnostics: identifying influential data and sources of collinearity*. New York: John Wiley & Sons, 1980.

CAZALS, C.; FLORENS, J. P.; SIMAR, L. Non parametric frontier estimation: a robust approach. *Journal of Econometrics*, v. 106, p. 1-25, 2002.

CHARNES, A.; COOPER, W. W.; RHODES, E. Measuring efficiency of decision making units. *European Journal of Operational Research*, v. 1, p. 429-444, 1978.

CHERCHYE, L.; KUOSMANEN, T.; POST, G. T. New tools for dealing with errors in variables in DEA. *Center for Economic Studies Discussions Paper Series 00.06*, Katholieke Universiteit Leuven, Faculty of Economics and Applied Economics, 2000.

FÄRE, R.; GROSSKPOF, S.; C. K. LOVELL. *The measurement of efficiency of production*. Boston-Dordrech: Kluwer-Nijhoff Publishing, 1985.

_____. *Production frontiers*. Cambridge: Cambridge University Press, 1994.

MCDONALD, J. Using least squares and tobit in second stage DEA efficiency analyses. *European Journal of Operational Research*, v. 197, n. 2, p. 792-798, 2009.

STOSIC, B.; SAMPAIO DE SOUSA, M.C. Jackstrapping DEA scores for robust efficiency measurement. In: LAMES ENCUESTRO LATINO AMERICANO DE LA SOCIEDAD ECONOMETRICA, Santiago, Chile, Jul. 2004.

SAMPAIO DE SOUSA, M. C.; STOSIC, B. Technical efficiency of the Brazilian municipalities: correcting nonparametric frontier measurements for outliers. *Journal of Productivity Analysis*, v. 24, p. 155-179, 2005.

SEEVER, B.; TRIANTIS, K. A fuzzy clustering approach used in evaluating technical efficiency measures in manufacturing. *Journal of Productivity Analysis*, v. 3, p. 337-363, 1992.

_____. The impact of outliers and leverage points for technical efficiency measurement using high breakdown procedures. *Management Science*, v. 41, p. 937-956, 1995.

SIMAR, L. Detecting outliers in frontiers models: a simple approach. *Journal of Productivity Analysis*, v. 20, p. 391-424, 2003.

SIMAR, L.; WILSON, P. W. Sensitivity analysis of efficiency scores: how to bootstrap in nonparametric frontier models. *Management Science*, v. 44, p. 49-61, 1998.

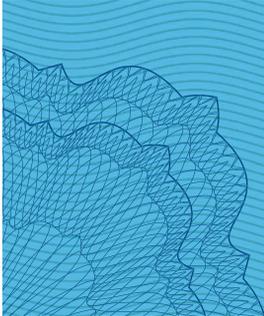
_____. Estimation and inference in two-stage, semiparametric models of production process. *Journal of Econometrics*, v. 136, n. 1, p. 31-64, 2007.

TIMMER, C. P. Using a probabilistic frontier function to measure technical efficiency. *Journal of Political Economy*, v. 79, p. 776-794, 1971.

WILSON, P. Detecting influential observations in data envelopment analysis. *Journal of Productivity Analysis*, v. 6, p. 27-45, 1993.

_____. Detecting influential observations in deterministic non-parametric frontiers models. *Journal of Business and Economic Statistics*, v. 11, p. 319-323, 1995.

ZHANG, Y.; BARTELS, R. The effect of sample size on the mean efficiency in DEA with an application to electricity distribution in Australia, Sweden and New Zealand. *Journal of Productivity Analysis*, v. 9, n. 3, p. 187-204, 1998.



Capítulo 14

Aplicação do Método *Jackstrap* na Atenção Básica à Saúde

Maria da Conceição Sampaio de Sousa
Rodrigo Hitoshi Dias

14.1 Introdução

O crescimento substancial dos sistemas de Atenção Básica à Saúde (AB) vem sendo observado em muitos países. No Brasil, essa tendência à ampliação dos serviços de AB vem ocorrendo desde os anos 1990 e manifesta-se, sobretudo, pela expansão da cobertura da Estratégia Saúde da Família (ESF), presente hoje em praticamente todos os municípios brasileiros.¹ Implantada em 1994, com o nome de Programa de Saúde da Família (PSF), essa estratégia baseia-se em equipes multiprofissionais, agrupadas em Unidades Básicas de Saúde (UBS) que são responsáveis pelo acompanhamento da saúde das famílias, geograficamente referenciadas.²

O atendimento dessas demandas ampliadas envolve um volume substancial de recursos públicos, cujo gerenciamento constitui um desafio para os gestores da ESF e, em particular, para as municipalidades, já que a descentralização as tornou responsáveis pela gestão da rede de serviços de saúde, inclusive das ações empreendidas no âmbito da ESF. Considerando-se os elevados custos de oportunidade dos fundos públicos, a questão fundamental é saber se os gestores municipais utilizam esses recursos de modo eficiente, alocando-os de forma a maximizar os resultados em saúde e bem-estar das populações. A resposta a essa questão passa pela avaliação do desempenho municipal na provisão dos serviços de saúde na esfera da atenção básica.

1 Em 2008, cerca de 93 milhões de brasileiros eram atendidos no âmbito da ESF.

2 Ver a esse respeito Santana e Carmagnani (2001).

No Brasil, poucos estudos focaram-se na análise de eficiência na provisão de serviços de saúde; a maioria deles centra-se na análise da eficiência hospitalar.³ Ao nosso conhecimento, apenas os trabalhos de Varela (2008) e Hitoshi Dias (2010) computaram escores de eficiência para a AB. O trabalho de Varela, além de utilizar uma versão *naive* da abordagem de Análise Envoltória de Dados (DEA⁴), refere-se apenas aos municípios do estado de São Paulo. Hitoshi Dias (2010) aplicou a metodologia *Jackstrap* proposta por Sampaio de Sousa e Stosic (2005) para computar escores de eficiência robustos, mediante a exclusão de *outliers*.

O objetivo deste capítulo é, pois, avaliar o desempenho dos sistemas de atenção primária à saúde nos municípios brasileiros mediante o uso do método *Jackstrap Ampliado*, descrito no capítulo anterior. A Seção 14.2 discute os insumos e produtos utilizados e a Seção 14.3 contém os resultados das fases 1 e 2 da abordagem *Jackstrap*. A Seção 14.4 analisa os escores de eficiência robustos para a Atenção Básica nos municípios brasileiros. Por fim, a Seção 14.5 reúne as conclusões do capítulo.

14.2 Insumos e produtos

Apesar de não existir um consenso sobre quais insumos e produtos estão relacionados à prestação de serviços de atenção primária, a maioria dos estudos inclui o uso de profissionais da atenção primária (em quantitativo ou horas trabalhadas), gastos/custos, medicamentos e exames laboratoriais como insumos; visitas e consultas a profissionais da atenção primária, procedimentos (e.g., imunizações, nascimentos, exames citopatológicos), bem como número de pacientes como produtos.

A literatura sobre eficiência na atenção básica⁵ constata que a maioria dos trabalhos na área inclui indicadores de estrutura – similares ao conceito de insumo na análise de eficiência – e indicadores de processo – referentes às atividades que ocorrem entre os trabalhadores de saúde e os pacientes –, porém, não levam em conta os indicadores de resultado que mensuram os impactos na saúde dos pacientes decorrentes dos cuidados prestados (DONABEDIAN, 1980).

3 Souza, Nishijima e Rocha (2010); Cesconetto, Lapa e Calvo (2008); Proite e Sampaio de Sousa (2004).

4 Análise Envoltória de Dados é a tradução literal do termo em inglês *Data Envelopment Analysis* (DEA).

5 Para uma excelente revisão da literatura, ver Amado e Dyson (2008).

Isso porque, em muitos casos, não existem informações suficientes sobre os resultados em saúde associados a esse nível de atenção.

Esse, porém, não é o caso do Brasil, que dispõe de um vasto conjunto de indicadores, coletados no âmbito do Pacto pela Saúde. A escolha de insumos e produtos considerou alguns dos indicadores que fazem parte da prioridade VI – Fortalecimento da Atenção Básica desse pacto. Os insumos selecionados representam a estrutura presente no município: quantidade de estabelecimentos de atenção primária, número de profissionais de saúde e investimentos nesse nível de atenção. Já os produtos dividem-se entre variáveis relacionadas ao processo (i.e., pré-natais, visitas domiciliares, consultas médicas) e aos resultados (i.e., famílias com perfil saúde, crianças abaixo do peso, internações por DM e AVC). Indicadores como *crianças abaixo do peso por idade* e *internações por diabetes e AVC* são claramente produtos da AB, pois se referem à situação de saúde obtida pela população. Neste trabalho, eles entram como insumo, porque o objetivo é minimizá-los.⁶ Nossa amostra é constituída de um corte seccional de 4.007 municípios para o ano de 2007. Os dados foram obtidos de diferentes sistemas de informação do Ministério da Saúde (Datasus). A Tabela 14.1 apresenta as estatísticas descritivas dos insumos e produtos utilizados no cômputo dos índices de eficiência.

Tabela 14.1 – Estatísticas descritivas de produtos e insumos, 2007

<i>Variáveis</i>	Média	Mediana	Desvio padrão	Min	Máx	Fonte
Produtos						
Número de nascidos vivos de mães com quatro ou mais consultas de pré-natal em determinado local e período	508,7	166,0	31,5	4,0	73760	Sinasc ^d
Número de visitas domiciliares por família realizadas por Agentes Comunitários de Saúde	4697,0	2169,6	160,6	0,1	269213	SIA ^e
Número de consultas médicas nas especialidades básicas	4809,1	1803,3	241,7	22,9	476032	SIA
Número de famílias com perfil saúde acompanhadas pela atenção básica	986,8	491,5	29,1	0,5	46065	Sisvan ^f

(continua)

6 Mais detalhes sobre esse tipo de procedimento podem ser encontrados na seção 11 do capítulo 9.

(continuação)

<i>Variáveis</i>	Média	Mediana	Desvio padrão	Min	Máx	Fonte
Produtos como Insumos						
Número de crianças menores de cinco anos com peso por idade abaixo do percentil 3	40,6	16,0	1,3	1,0	2289	Sisvan
Número internações por complicações do diabetes mellitus na população de 30 anos e mais e por acidente cerebral vascular (AVC) na população de 40 anos e mais	55,3	19,0	2,7	1,0	5621	SIH ^g
Insumos						
Gasto municipal na atenção básica e transferências federais para a atenção básica municipal ^a (R\$ milhões)	4,9	2,1	0,2	0,1	361	Siops ^h / FNS ⁱ
Número de Unidades Básicas de Saúde	8,5	5,0	0,2	1,0	255	CNES ^j
Número de médicos da atenção básica ^c	11,0	5,0	0,5	0,2	678	CNES
Número de enfermeiros, auxiliares de enfermagem e técnicos de enfermagem ^c	30,2	15,0	1,4	2,0	2506	CNES
Número de agentes comunitários de saúde ^c	50,91	27,00	1,61	2,00	1739	CNES

Fonte: Hitoshi Dias (2010).

Notas: a) Transferências federais para a atenção básica municipal no ano de 2006; b) médicos da atenção básica: clínicos gerais, obstetras, ginecologistas, pediatras e médicos de saúde da família; c) o número de profissionais de saúde refere-se a 40 horas semanais de trabalho em unidades de atenção primária. Por exemplo, um enfermeiro que trabalhe 20 horas por semana é considerado como "meio trabalhador"; d) Sinasc – Sistema de Informações de Nascidos Vivos; e) SIA – Sistema de Informações Ambulatoriais; f) Sisvan – Sistema de Vigilância Alimentar e Nutricional; g) SIH – Sistema de Informações Hospitalares; h) Siops – Sistema de Informações sobre Orçamento Público em Saúde; i) FNS – Fundo Nacional de Saúde; j) CNES – Cadastro Nacional de Estabelecimentos.

14.3 Resultados

As seções 14.3.1 e 14.3.2 descrevem os resultados das duas fases do modelo *Jackstrap* aplicado aos serviços de Atenção Básica dos municípios brasileiros, expostos nas Tabelas 14.2 a 14.5.

14.3.1 Fase 1: Alavancagens e influência

Para implementar a primeira fase do modelo, computamos alavancagens⁷ para 4.007 municípios utilizando o método *Jackstrap*

7 Medida que mostra o efeito da remoção de uma DMU sobre os escores de eficiência das demais componentes da amostra.

proposto por Sampaio de Sousa e Stosic (2005). Essas alavancagens foram calculadas para duas versões da metodologia DEA: DEA-CCR (retornos constantes de escala) e DEA-BCC (retornos variáveis de escala). Ambas as versões seguem a orientação produto. Fixamos o tamanho das bolhas a $L=400$ e perfizemos $B=1000$ passos de *bootstrap*.⁸ Para o cálculo das alavancagens, cada municipalidade foi selecionada, em média, 99,8 vezes ($BL/K=99,8$).

As Tabelas 14.2 e 14.3, que mostram as alavancagens para os 10 municípios mais (menos) influentes, computadas por meio das abordagens DEA-CCR e DEA-BCC, ilustram o funcionamento da primeira fase do método *Jackstrap*. A coluna *alav* lista a alavancagem média obtida por uma dada municipalidade, usando a expressão 13.2.9 do capítulo anterior; essa alavancagem será doravante mencionada como *alavancagem bruta*. A coluna *tot-hits* mostra o número total de vezes em que a municipalidade foi selecionada no procedimento *jackstrap* ($100 \approx BL/K = (1000 \times 400) / 4007$). A coluna *a-hits* denota o número de vezes em que a municipalidade selecionada obteve alavancagem não nula. Por fim, a coluna *alav-a* mostra a alavancagem ajustada ($alav-a = alav \times (a-hits/tot-hits)$), que constitui a informação relevante sobre a influência das observações. Se uma determinada observação é selecionada aproximadamente 100 vezes, porém somente em uma delas influenciou os escores das demais, então, o poder dessa observação é muito pequeno, mesmo que ela apresente uma alavancagem elevada. Esse ponto será esclarecido quando da discussão dos resultados mostrados nas tabelas supracitadas.

A Tabela 14.2 apresenta os municípios com maiores alavancagens e influentes, ordenados pelas alavancagens ajustadas, obtidas pelo método CCR. Várzea Grande aparece como o maior *outlier* da amostra. Embora sua alavancagem seja apenas a 18ª maior da amostra ($alav=0,0913$), das 78 vezes em que foi aleatoriamente escolhida, em praticamente todas elas ($a-hits = 71$) gerou alavancagens não nulas. Trata-se, pois, de um município altamente influente, que tem o poder de deslocar a fronteira de eficiência e alterar os escores das outras cidades. Ajustando sua alavancagem inicial pelo peso – ($a-hits/$

8 Ver capítulo 13 para detalhamento do processo.

total hits) = 71/78 – essa observação torna-se a mais influente da amostra. Raciocínio análogo pode ser feito para os demais *outliers* apresentados na Tabela 14.2.

Tabela 14.2 – Municípios com elevadas alavancagens e influentes – DEA-CCR e DEA-BCC												
Municípios	DEA CCR						DEA BCC					
	alav	ord	a-hits	tot hits	lev	ord	alav	ord	a-hits	tot hits	lev	ord
Várzea Grande	0,0913	18	72	78	0,0843	1	0,1355	639	71	78	0,1234	1
Coroatá	0,0557	19	94	96	0,0546	2	0,1000	832	92	96	0,0959	5
Itaquaquetuba	0,0517	20	91	99	0,0475	3	0,0931	877	91	99	0,0856	14
Tailândia	0,0481	21	100	110	0,0438	4	0,0799	992	100	110	0,0727	35
Medicilândia	0,0394	22	102	110	0,0366	5	0,1029	815	99	110	0,0926	7
São Joaquim de Bicas	0,0373	23	86	92	0,0349	6	0,0695	1115	86	92	0,0650	70
Santa Gertrudes	0,0358	26	81	85	0,0341	7	0,0463	1523	80	85	0,0436	265
Seropédica	0,0362	25	85	91	0,0338	8	0,1169	731	87	91	0,1118	2
São José do Rio Claro	0,0346	27	85	91	0,0323	9	0,0855	945	84	91	0,0790	22

Fonte: Elaboração própria.

No outro extremo, temos os municípios cuja alavancagem bruta é alta, porém, com pouca (ou nenhuma) influência sobre seus pares. A Tabela 14.3 apresenta 10 desses municípios, ordenados também pela alavancagem CCR. Vejamos, por exemplo, o caso de Lavras da Mangabeira. Essa cidade apresentou a maior alavancagem bruta da amostra, quando os escores de eficiência são computados pela variante DEA-CCR (alav=0,2760). Porém, das 105 vezes em que foi aleatoriamente selecionado, apenas uma vez esse município afetou os escores das demais observações. Quando utilizamos a alavancagem ponderada (lev), Lavras de Mangabeira passa da 1^a para a 364^a posição. Portanto, sua influência sobre os escores das outras cidades é muito pequena, visto que ela não integra a fronteira de eficiência. O mesmo ocorre com as demais cidades da Tabela 14.3.

Tabela 14.3 – Municípios com altas alavancagens e sem influência

Municípios	DEA CCR						DEA BCC					
	alav	ord	a-hits	tot hits	lev	ord	alav	ord	a-hits	tot hits	lev	ord
Lavras da Mangabeira	0,2760	1	1	105	0,0026	364	0,1669	423	18	105	0,0286	531
Carnaúba dos Dantas	0,2651	2	1	92	0,0029	347	0,1749	368	8	92	0,0152	1029
Joaquim Távora	0,2651	2	1	115	0,0023	400	0,0000	2754	0	115	0,0000	2722
Itabirito	0,2244	4	1	104	0,0022	409	0,2078	163	2	104	0,0040	1870
Petrópolis	0,2239	5	1	92	0,0024	384	0,1110	764	12	92	0,0145	1072
Oeiras	0,2239	5	1	95	0,0024	394	0,2278	80	4	95	0,0096	1353
Cidade Ocidental	0,1952	7	1	80	0,0024	383	0,0868	932	3	80	0,0033	1981
Palmares	0,1949	8	1	95	0,0021	431	0,0964	852	2	95	0,0020	2240
Ribeirão das Neves	0,1949	8	1	98	0,0020	445	0,0649	1192	26	98	0,0172	937

Fonte: Elaboração própria.

Voltaremos, agora, à análise dos maiores *outliers* – municípios com alta alavancagem e influentes – ordenados pela alavancagem ponderada obtida por meio da variante DEA-CCR, como o exposto na Tabela 14.4, na qual são mostradas as alavancagens para os 20 municípios mais influentes. Para facilitar a discussão, apresentamos também na Tabela 14.4 os produtos e insumos da AB utilizados no cômputo das alavancagens.

Note-se, em primeiro lugar, que muitos desses municípios, em vez de se constituírem em exemplos de adoção das melhores práticas, devem suas altas alavancagens a distorções e/ou erros de mensuração nos insumos e produtos declarados por essas cidades. Em geral, elas apresentam número excessivo de consultas médicas, baixas taxas de hospitalização, gastos *per capita* com AB muito baixos e pequeno número de profissionais de saúde por habitante.

Tabela 14.4 – Maiores *outliers* – indicadores selecionados

MUNICÍPIO	TXAVC	TXDM	PRENAT	PERF	CONSUL	GASTO	MED	ENF	ACS
Várzea Grande	36,4	15,4	91,8	0,11	5,95	55,1	1,1	2,5	6,6
Coroatá	29,9	27,6	87,2	0,01	5,42	203,5	3,2	6,9	28,2
Itaquaquetuba	21,6	7,4	91,4	1,00	1,78	77,9	1,3	2,3	1,1
Tailândia	12,8	7,2	77,4	0,21	2,30	140,4	0,4	3,2	15,3
Medicilândia	55,2	9,1	87,4	0,67	0,92	90,2	1,5	16,8	31,1
Sao Joaquim das Bicas	11,0	5,1	91,2	0,81	1,66	18,0	1,0	3,7	5,1
Santa Gertrudes	12,0	0	96,1	0,47	3,43	25,4	4,6	11,9	9,3
Seropédia	12,4	9,5	90,4	0,45	8,56	28,2	3,2	4,3	11,1
São José do Rio Claro	52,3	14,6	98,8	0,40	26,01	224,6	6,0	17,9	19,1
Brasil	33,4	16,7	89,3	0,53	1,92	193,5	4,3	13,3	22,4

Fonte: Hitoshi Dias (2010).

No topo da lista, o município de Várzea Grande apresentou problemas em três desses aspectos, especialmente no que se refere ao elevado número de consultas médicas da AB por habitante/ano (5,95), quando a média dos municípios brasileiros é de 1,92 consultas. De fato, quando avaliamos a quantidade mensal de consultas por médico em Várzea Grande, temos um valor de 4.741, equivalente a aproximadamente 1 consulta por médico a cada 2 minutos (considerando que cada médico trabalha 8 horas por dia e 23 dias por mês). Trata-se, pois, claramente de um caso de erro na entrada de dados. Outro exemplo extremo em relação a problemas na variável *consultas médicas* é o município de São José do Rio Claro-MT; de acordo com os dados, obteve-se uma média de 26 consultas por habitante em 2007, muito acima da média nacional, igual a 1,92.

No tocante aos insumos, Tailândia, no Pará, não declara nenhum médico trabalhando 40 horas por semana nos serviços de atenção básica, o que não é razoável para uma cidade com mais de 70 mil habitantes. Esses números, exemplos atípicos da base de dados, são provavelmente resultantes de erros de coleta/digitação. Quando não corrigidos, contribuem para reduzir artificialmente os escores de eficiência das demais observações comprometendo, assim, a credibilidade dos escores obtidos por meio da aplicação *naive* do método DEA.

14.3.2 Fase 2: Seleção de municipalidades

Passaremos agora à análise dos resultados da Fase 2 do procedimento *Jackstrap*. De posse das alavancagens, essa fase seleciona as municipalidades que entrarão no cômputo dos escores de eficiência. Para construir as pseudoamostras, utilizamos as distribuições de probabilidades, descritas na subseção 13.3.2 do capítulo anterior.

A Tabela 14.5 ilustra o procedimento de seleção. Ela mostra o número de vezes ($nboot$) em que o município foi selecionado, bem como a sua probabilidade de seleção, computada de acordo com as equações 13.4 a 13.7.⁹ Os resultados referem-se às municipalidades que apresentam as cinco maiores (menores) alavancagens ajustadas (lev).

Nenhuma das funções de probabilidade usadas para escolher os municípios que compuseram as diferentes amostras selecionou Várzea Grande, visto que esse município apresenta a maior alavancagem. Ademais, quando se usa a função *heaviside*, nenhum *outlier* dos apresentados na Tabela 14.5 integra as amostras, a partir das quais os índices de eficiência foram computados ($nboot = 0$). Nas demais funções, essas observações discrepantes foram incluídas na amostra, porém com uma menor probabilidade de seleção.

Vejamos, por exemplo, o caso de Coroatá. O número de vezes em que essa cidade foi incluída na amostra ($nboot$) variou de 63 (para ≈ 1000 passos de *bootstrap*), quando a função inversa foi usada, até 384 vezes, quando se aplicou a função linear; sua probabilidade de inclusão na amostra variou de 0,06 (função inversa) a 0,352 (função linear). De forma geral, quanto maior a influência de uma observação, menor sua probabilidade de compor as pseudoamostras. Na segunda parte da Tabela 14.5, vê-se que os municípios menos influentes entram na amostra em praticamente todos os passos de *bootstrap* (≈ 1000), sendo, pois, selecionados com probabilidade igual a 1.

Ressalte-se que tanto a função *heaviside* quanto a função inversa têm a desvantagem de exigirem pontos de corte arbitrário: $\tilde{\ell} \log K$, na função *heaviside* e ℓ_0 na função inversa. Além disso, a função *heaviside* usa esse limiar para remover as municipalidades, descartando todas as informações referentes a essas cidades.

⁹ Também do capítulo 13.

Tabela 14.5 – Número de vezes e probabilidade de seleção dos municípios

Municípios	Funções de probabilidade							
	Linear		Exponencial		Inversa		Heaviside	
	nboot	p(k)	nboot	p(k)	nboot	p(k)	nboot	p(k)
Mais influentes								
Várzea Grande	0	0,0000	0	0,0000	0	0,0000	0	0,0000
Coroatá	384	0,3523	381	0,3427	63	0,0604	0	0,0000
Itaquaquecetuba	499	0,4358	501	0,4255	99	0,0858	0	0,0000
Tailândia	517	0,4808	477	0,4703	80	0,1029	0	0,0000
Medicilândia	505	0,5661	503	0,5557	131	0,1450	0	0,0000
Menos influentes								
São Luíz do Norte	963	1,0000	1015	1,0000	1015	1,0000	1011	1,0000
Silvânia	1020	1,0000	1046	1,0000	1044	1,0000	1078	1,0000
Teresina de Goiás	1018	1,0000	990	1,0000	989	1,0000	1007	1,0000
Terezópolis de Goiás	975	1,0000	995	1,0000	995	1,0000	996	1,0000
Uruaçu	994	1,0000	1002	1,0000	1002	1,0000	1021	1,0000

Fonte: Elaboração própria.

As distribuições linear e exponencial, além de não exigirem a fixação de limiares *ad-hoc*, incluem praticamente a totalidade das informações. Quando essas distribuições foram utilizadas, por construção, somente Várzea Grande, que apresentou a maior alavancagem ponderada, foi retirada da amostra. Os demais *outliers* foram apenas selecionados com menor probabilidade e não removidos da amostra, como ocorreu na distribuição *heaviside*.

14.4 Escores robustos de eficiência

A Tabela 14.6 apresenta as estatísticas descritivas para os escores de eficiência, computados a partir das pseudoamostras geradas pelas diferentes funções de probabilidade, nas variantes DEA-CCR e DEA-BCC. Note-se, em primeiro lugar, que em presença de *outliers* a aplicação do DEA *naive* conduz à subestimação dos índices de eficiência, em ambas as variantes do método DEA. Isso ocorre apesar de o elevado número de insumos e produtos provocar um viés superior nos escores, em virtude da existência da maldição da dimensionalidade. A remoção de

Várzea Grande, o maior *outlier*, com a menor probabilidade de inclusão na amostra das demais observações discrepantes, eleva os escores de eficiência, em todos os casos examinados.

Tabela 14.6 – Estatísticas descritivas para os escores de eficiência DEA-CCR e DEA-BCC										
Métodos /funções	# Municípios				Municipalidades ineficientes					
					Escore de eficiência – estatísticas descritivas – municipalidades ineficientes					
	Total	Remov	Eficientes		Média	Mediana	Desvio padrão	Assimetria	Curtose	Mínimo
#		#	%							
DEA-CCR										
Dea	4007	0	264	6,59	0,5804	0,5451	0,2058	0,4843	-0,5823	0,1307
Linear	4006	1	272	6,79	0,6418	0,6147	0,2021	0,2546	-0,8487	0,1381
Exponencial	4006	1	272	6,80	0,6424	0,6154	0,2020	0,2518	-0,8511	0,1378
Inversa	4006	1	281	7,01	0,6801	0,6601	0,1989	0,0921	-0,9413	0,1428
Step	3890	117	624	16,04	0,7451	0,7387	0,1908	-0,2067	-0,9424	0,1646
DEA-BCC										
Dea	4007	0	533	13,30	0,6372	0,6033	0,2211	0,2600	-1,0240	0,1343
Linear	4006	1	550	13,73	0,7132	0,6964	0,2080	-0,0635	-1,1352	0,1528
Exponencial	4006	1	551	13,75	0,7143	0,6983	0,2079	-0,0690	-1,1350	0,1528
Inversa	4006	1	605	15,10	0,7744	0,7863	0,1961	-0,3883	-1,0103	0,1673
Step	3896	111	706	18,12	0,7421	0,7390	0,1992	-0,1879	-1,0615	0,1530

Fonte: Elaboração própria.

Como esperado, exceto quando usamos a distribuição *heaviside*, os escores de eficiência da variante DEA-BCC são mais elevados já que, por construção, $\theta^{CRS} \leq \theta^{VRS}$, para uma mesma amostra. Aqueles obtidos das amostras geradas pelas distribuições linear e exponencial, além de mais elevados, são muito similares, em ambas as variantes do método DEA. Essas distribuições produzem, inclusive, o mesmo número de municipalidades eficientes (272). Já a função *heaviside*, por remover 117 e 111 municipalidades, respectivamente, nas variantes DEA-CCR e DEA-BCC, não somente aumenta substancialmente o número de observações eficientes, mas também eleva a magnitude dos escores computados, particularmente no método CCR.

Com exceção dos resultados atrelados à distribuição *heaviside*, a assimetria positiva das medidas de eficiência na variante CCR indica que, em presença de *outliers*, os escores de eficiência são enviesados para cima; de fato, a inclusão de todos os 4.007 municípios praticamente dobra a assimetria, em relação às demais amostras. Quando a função *heaviside* é usada, a assimetria negativa deve-se à exclusão dos principais *outliers*. Por fim, as curtoses negativas – em ambas as variantes e para todas as funções de probabilidade utilizadas – indicam que as distribuições das eficiências, comparadas com a distribuição normal, têm caudas mais curtas e mais finas; os escores estão pouco concentrados em torno da média.

14.4.1 Tamanho do município e desempenho municipal na AB

A proliferação de pequenos municípios tornou-se uma característica do movimento de transferência do poder político aos níveis inferiores de governo. Esse fenômeno contribuiu para elevar os custos dos serviços públicos no Brasil porque municípios muito pequenos não exploram as economias de escala que caracterizam muitos dos serviços públicos e, portanto, não utilizam de maneira ótima os recursos disponíveis.¹⁰ Uma questão relevante é saber se essa argumentação valeria também para a produção de serviços de saúde no âmbito da AB. De fato, não se pode garantir que esse desmembramento de cidades tenha levado em conta a escala mínima de operações exigida para que essas municipalidades sejam capazes de assegurar aos seus habitantes os serviços básicos de saúde, sem incorrer em custos, desnecessariamente, elevados.

Para investigar esse tópico, agrupamos os escores municipais de eficiência da AB por classe de população (Tabela 14.7). Por razões de espaço, incluímos apenas os resultados para as distribuições exponencial e *heaviside*. Vê-se que os escores de eficiência aumentam com o tamanho da municipalidade, em ambas as distribuições apresentadas. Esses resultados corroboram estudos anteriores que mostraram que o nível de eficiência da AB tende a aumentar com o tamanho do município (SIDDHARTHAN; AHERN; ROSENMAN, 2000; KONTODIMOPOULOS et al., 2007) e sugerem que a produção dos cuidados primários de saúde no Brasil caracteriza-se pela existência de economias de escala.

10 Sampaio de Sousa e Souza (1999); Maia Gomes e McDowell (1997).

Tabela 14.7 – Escores de eficiência robustos por classe de população – estatísticas descritivas

Classes de população	Estatísticas descritivas				
	Media	Mediana	Desvio	Assimetria	Curtose
Exponencial					
0-9999	0,5823	0,5514	0,1958	0,4360	-0,5216
10000-49999	0,6409	0,6175	0,1865	0,3312	-0,7392
50000-99999	0,7217	0,7101	0,1826	0,0095	-1,0888
100000-499999	0,7216	0,6985	0,1647	0,1404	-0,9118
>50000	0,8114	0,8688	0,1706	-0,3544	-1,4256
Heaviside(Step)					
0-9999	0,6727	0,6502	0,1986	0,1299	-0,8493
10000-49999	0,7378	0,7247	0,1813	-0,0531	-0,9951
50000-99999	0,8223	0,8420	0,1622	-0,5162	-0,7312
100000-499999	0,8349	0,8481	0,1501	-0,4858	-0,8477
>50000	0,8941	0,9878	0,1267	-0,6540	-1,2450

Fonte: Elaboração própria.

Resta, porém, saber se a relação positiva entre tamanho da cidade e desempenho gerencial da AB deve-se efetivamente às especificidades da tecnologia da produção de cuidados básicos de saúde, ou se decorre da presença de fatores exógenos, não considerados na abordagem *Jackstrap*. Voltaremos a discutir esse ponto quando da discussão dos escores residuais de eficiência.

14.4.2 Municípios situados nos extremos da distribuição das eficiências: resultados selecionados

Nesta seção, vamos discutir brevemente as municipalidades situadas nos extremos da distribuição de eficiência. Além da população, usamos três indicadores de relações entre produtos e insumos para caracterizar essas municipalidades: (i) número de consultas pré-natal por real gasto com a atenção básica, *PRENAT/G*; (ii) número de visitas por agentes comunitários de saúde, *VISIT/ACS*; e (iii) número de consultas por médicos, *CONSULT/MED*.

A Tabela 14.8 apresenta esses indicadores para dez municipalidades selecionadas, que aparecem como eficientes em todas as amostras – linear, exponencial, inversa e *heaviside*, na variante DEA-CCR. Note-se, em primeiro lugar, que cidades de diferentes portes aparecem nessa lista. Temos capitais (Manaus), cidades de porte médio, como Santana de Parnaíba, regiões metropolitanas (Jaboatão dos Guararapes), e cidades pequenas, como Bodó e Aguiar. Esse resultado, consistente com a existência de divisibilidade, contrapõe-se àquele sugerido pelas informações da Tabela 14.7, que indicava a possibilidade de economias de escala, na produção de serviços da AB.

Tabela 14.8 – Municípios eficientes – indicadores selecionados					
Municípios	UF	Indicadores			
		POP	PRENAT/G	VISIT/ACS	CONSULT/MED
Manaus	AM	1731993	728,15	88,04	306,05
Jaboatão dos Guararapes	PE	511891	217,86	97,85	479,65
Aguiar	PB	4261	53,21	87,45	288,63
Armação dos Búzios	RJ	24730	47,20	133,16	795,16
Belém do Brejo do Cruz	PB	6050	45,49	113,77	212,90
Icém	SP	7342	34,13	132,93	601,65
Santana de Parnaíba	SP	106361	34,03	95,07	1003,20
Bodó	RN	2783	31,47	83,46	298,83
Atalaia do Norte	AM	11752	30,82	14,74	732,61
Carapebus	RJ	10604	13,60	185,31	429,51

Fonte: Elaboração própria.

Além disso, os níveis de eficiência podem estar sendo inflados para municípios cujas consultas médicas são mais curtas – incluindo-se aí aquelas de pré-natal – o que, por sua vez, não necessariamente está relacionado à qualidade ou à resolutividade em saúde. Um maior número de consultas por médico pode refletir tanto uma demanda reprimida por esse tipo de atendimento quanto o baixo número de profissionais médicos. Santana do Parnaíba, Armação dos Búzios e Atalaia do Norte detêm o maior número de consultas por médico. Manaus e Jaboatão dos Guararapes destacam-se no indicador PRENAT/G.

A Tabela 14.9 apresenta os escores de eficiência para as dez cidades mais ineficientes. Mostraremos apenas os resultados obtidos mediante o uso das funções de probabilidade exponencial (efic-exp) e *heaviside* (efic-step), ordenados pela eficiência média CCR das amostras geradas pela função exponencial.

Tabela 14.9 – Municípios ineficientes – indicadores selecionados								
Municípios	CRS		VRS		Indicadores			
	Efic - exp	Efic-step	Efic - exp	Efic-step	pop	prenat/g	visit/acs	consult/med
Rio Preto - MG	0,1378	0,1646	0,1528	0,1554	5388	0,023	20,5	64,7
São Martinho - RS	0,1496	0,1870	0,1655	0,1805	3409	0,020	23,0	57,5
Chupinguaia - RO	0,1497	0,1740	0,1528	0,1530	7456	0,017	16,3	27,8
Luciara - MS	0,1837	0,2100	0,3004	0,3072	2405	0,015	47,4	137,4
Charrua - RS	0,1838	0,2663	0,2260	0,2512	3581	0,021	10,5	255,8
Silvianópolis - MG	0,1876	0,2541	0,2046	0,2411	3803	0,036	0,0	235,9
José Raydan - MG	0,1914	0,2532	0,3647	0,3565	4146	0,038	1,3	228,2
Chiapetta - RS	0,1931	0,2772	0,2157	0,2253	4058	0,020	4,0	141,5
Antônio Prado - RS	0,2160	0,3108	0,2443	0,2933	13315	0,047	34,3	380,3
Jenipapo de Minas - MG	0,2200	0,2860	0,2482	0,2891	6905	0,037	30,5	424,0

Fonte: Elaboração própria.

Contrariamente aos municípios situados na fronteira de eficiência, aqueles de pior desempenho apresentam várias características em comum. Com exceção de Antônio Prado, eles têm menos de 10 mil habitantes, apresentam baixa densidade populacional e são, em sua grande maioria, localizados nas regiões Sudeste e Sul. Em alguns deles, a população, inclusive, está estagnada ou se reduzindo ao longo do tempo, como é o caso dos municípios gaúchos listados na Tabela 14.9: Antônio Prado, Charrua, Chiapetta e São Martinho.¹¹

Mesmo supondo-se divisibilidade na produção de serviços da AB, esses pequenos municípios não conseguem atingir a escala mínima exigida para a produção eficiente desses serviços e, portanto, trabalham com custos médios mais elevados, o que leva à redução da

11 Ver <<http://www.censo2010.ibge.gov.br/>>.

eficiência. Além disso, em muitos deles, a proporção de idosos é maior quando comparada com a média brasileira. As condições de renda e de educação são superiores à média nacional, e eles dispõem de melhores condições de infraestrutura. Seus baixos escores de eficiência devem-se mais aos seus elevados gastos com AB e suas altas taxas de internações por diabetes mellitus e AVC, mais recorrentes em população mais idosas, do que a problemas de gerenciamento nesse nível de atenção.

14.5 Considerações finais

Neste capítulo, computamos escores de eficiência DEA-CCR e DEA-BCC para a Atenção Primária a Saúde (APS), mediante o uso do método *Jackstrap* Ampliado. O uso dessa abordagem permitiu a identificação de municipalidades influentes, que têm o poder de deslocar a fronteira de eficiência e alterar os escores de eficiência das outras cidades.

Na maioria dos casos, as altas alavancagens obtidas por esses municípios, em vez de apontarem para a adoção das melhores práticas de gestão da AB, remetem-se a distorções e/ou erros de mensuração nos insumos e produtos, declarados por essas cidades. Como a presença desses *outliers* conduz à subestimação dos índices de eficiência, geramos pseudoamostras baseadas nessas alavancagens, nas quais as observações discrepantes foram selecionadas com menor probabilidade. Esse procedimento, além de permitir melhor aproveitamento das informações, contribuiu para elevar os escores de eficiência gerados pelas pseudoamostras.

Para concluir, vale lembrar que a discussão dos escores de eficiência computados neste capítulo baseia-se nos insumos e produtos gerenciáveis pelas municipalidades. Não leva, pois, em conta o fato de que parte dessas ineficiências pode resultar da presença de fatores exógenos, que estão fora de controle das municipalidades. Portanto, o uso de medidas de eficiência baseadas unicamente nos insumos e produtos gerenciáveis pelos municípios pode distorcer o desempenho observado de uma dada municipalidade, comprometendo, assim, o uso desses indicadores em políticas de saúde pública. É necessário, pois, considerar os elementos ambientais na análise da eficiência técnica para melhorar a qualidade e representatividade dos indicadores de eficiência e permitir, assim, um ordenamento inequívoco das municipalidades.

REFERÊNCIAS

AMADO, C. A.; DYSON, R. G. On comparing the performance of primary care providers. *European Journal of Operational Research*, v. 3, p. 915-932, 2008.

CESCONETTO, A.; LAPA, J. dos S.; CALVO, M. C. M. Avaliação da eficiência produtiva de hospitais do SUS de Santa Catarina, Brasil. *Caderno de Saúde Pública*, Rio de Janeiro, v. 24, n. 10, p. 2407-2417, out. 2008.

DONABEDIAN, A. *The definition of quality and approaches to its assessment*. Ann Arbor, MI: Health Administration Press, 1980.

FARRELL, M. J. The measurement of productive efficiency. *Journal of the Royal Statistical Society Series A*, 120, p. 253-281, 1957.

HITOSHI DIAS, R. *Eficiência da atenção primária à saúde nos municípios brasileiros*. 2010. 42f. Dissertação (Mestrado em Economia) – Departamento de Economia, Universidade de Brasília, Brasília, 2010.

KONTODIMOPOULOS, N; MOSCHOVAKIS, G; VASSILIS, H A; NIAKAS, D. The effect of environmental factors on technical and scale efficiency of primary health care providers in Greece. *Cost Efficiency and Resource Allocation*, v. 5, n. 14, 2007.

MAIA GOMES, G.; MACDOWELL, C. Os elos frágeis da descentralização: observações sobre as finanças dos municípios brasileiros. In: ENCONTRO NACIONAL DE ECONOMIA, 25., 1995, Recife. *Anais...* Recife: Associação Nacional dos Cursos de Pós-Graduação em Economia (ANPEC), 1995, p. 645-660.

PROITE, A.; SAMPAIO DE SOUSA, M. C. Eficiência técnica, economias de escala, estrutura da propriedade e tipo de gestão no sistema hospitalar brasileiro. In: ENCONTRO NACIONAL DE ECONOMIA, 32., 2004, João Pessoa. *Anais...* João Pessoa: Recife: Associação Nacional dos Cursos de Pós-Graduação em Economia (ANPEC), 2004.

SAMPAIO DE SOUSA, M. C.; SOUZA, F. R. Eficiência técnica e retornos de escala na produção de serviços públicos municipais: o caso do Nordeste e do Sudeste brasileiros. *Revista Brasileira de Economia*, v. 53, p. 433-461, 2000.

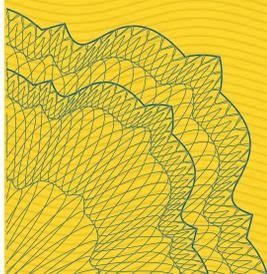
SAMPAIO DE SOUSA, M. C.; STOSIC, B. D. Technical efficiency of the Brazilian municipalities: correcting nonparametric frontier measurements for outliers. *Journal of Productivity Analysis*, v. 24, p. 155-179, 2005.

SANTANA, M. L.; CARMAGNANI, M. I Programa saúde da família no Brasil: um enfoque sobre seus pressupostos básicos, operacionalização e vantagens. *Saúde e Sociedade*, São Paulo, v. 10, jan./jul. 2001.

SIDDHARTHAN K.; AHERN M.; ROSENMAN R. Data envelopment analysis to determine efficiencies of health maintenance organizations. *Health Care Management Sciences*, v. 3, p. 23-29, 2000.

SOUZA, I. G.; NISHIJIMA, M.; ROCHA, F. Eficiência do setor hospitalar nos municípios paulistas. *Economia Aplicada*, v. 14, n. 1, jan./mar. 2010.

VARELA, P. S. *Financiamento e controladoria dos municípios paulistas no setor saúde: uma avaliação de eficiência*. 2008. 211f. Tese (Doutorado em Ciências Contábeis) – Departamento de Contabilidade e Atuária da Faculdade de Economia, Administração e Economia, Universidade de São Paulo, São Paulo, 2008.



Anexo

Bases de dados disponíveis em áreas finalísticas para a avaliação de políticas públicas

Finanças públicas nacionais

O principal *site* para o acesso aos dados de finanças públicas brasileiras, tanto no plano nacional quanto no âmbito subnacional (estados e municípios), é o da Secretaria do Tesouro Nacional.

No caso dos governos subnacionais, a STN disponibiliza bases completas com os dados fiscais de estados e municípios. A base Finbra (Finanças Municipais do Brasil), por exemplo, traz informações detalhadas sobre as contas fiscais dos municípios brasileiros de forma individualizada.

Além dos gastos por rubrica orçamentária, a base Finbra também contém os gastos municipais por funções, o que permite ao pesquisador saber quanto cada município gastou anualmente em saúde, educação etc.

A base Finbra vem sendo melhorada ao longo dos anos, e as últimas edições contêm maior número de municípios com informações completas. O pesquisador deverá possuir o *Microsoft Access* para ler os dados. Essa base pode ser obtida a partir do seguinte endereço: <<https://www.tesouro.fazenda.gov.br/pt/prefeituras-governos-estaduais/sobre>>.

Outra base muito útil, disponibilizada pela STN em seu *site*, é a planilha de orçamentos estaduais. Nela podem ser encontradas informações detalhadas sobre as finanças dos estados brasileiros, bem como os seus gastos por função. O painel se estende de 1995 a 2011, com informações para todos os estados. É uma ótima fonte de dados para a montagem de painéis balanceados. Essa base pode ser obtida a partir do seguinte endereço: <<https://www.tesouro.fazenda.gov.br/pt/prefeituras-governos-estaduais/sobre>>.

As informações sobre as transferências da União aos governos estaduais e municipais também estão disponíveis no *site*. Tais transferências são separadas por sua obrigatoriedade em constitucionais e voluntárias. No caso das transferências voluntárias, estas podem ser acessadas a partir do seguinte endereço, no qual o pesquisador poderá obter planilhas com dados anuais de transferências para cada estado, individualmente, e para os municípios agregados por estados: <<https://www.tesouro.fazenda.gov.br/pt/prefeituras-governos-estaduais/transferencias-estados-municipios/transferencias-voluntarias>>.

As informações sobre as transferências constitucionais são obtidas a partir do seguinte endereço: <<https://www.tesouro.fazenda.gov.br/pt/prefeituras-governos-estaduais/transferencias-estados-municipios/transferencias-constitucionais>>.

Nele são encontradas informações sobre o Fundo de Participação dos Estados e do Distrito Federal (FPE), Fundo de Participação dos Municípios (FPM), Fundo de Manutenção e Desenvolvimento da Educação Básica e de Valorização dos Profissionais da Educação (Fundeb) etc. No entanto, a sistemática de obtenção dos dados é diferente. Ao acessar o endereço, o pesquisador se deparará com a página exposta na Figura A.1.

Figura A.1 – Página da STN para obtenção das informações sobre transferências constitucionais da União

ESTADOS E MUNICÍPIOS
Transferências Constitucionais

Parcela das receitas federais arrecadadas pela União é repassada aos Estados, ao Distrito Federal e aos Municípios. O rateio da receita proveniente da arrecadação de impostos entre os entes federados representa um mecanismo fundamental para arremediar as desigualdades regionais, na busca incessante de promover o equilíbrio sócio-econômico entre Estados e Municípios.

Cabe ao Tesouro Nacional, em cumprimento aos dispositivos constitucionais, efetuar as transferências desses recursos aos entes federados, nos prazos legalmente estabelecidos.

Dentre as principais transferências da União para os Estados, o DF e os Municípios, previstas na Constituição, destacam-se: o Fundo de Participação dos Estados e do Distrito Federal (FPE); o Fundo de Participação dos Municípios (FPM); o Fundo de Compensação pela Exportação de Produtos Industrializados - FPEX; o Fundo de Manutenção e Desenvolvimento da Educação Básica e de Valorização dos Profissionais da Educação - Fundeb; e o Imposto sobre a Propriedade Territorial Rural - ITR.

Estados

Estado: [dropdown]
Ano: 2010 [dropdown]
Mês: [dropdown]

Transferências: FPE [dropdown]
 HTML
 XLS
 Consultar [button]

Municípios

Estado: -- Escolha um Estado -- [dropdown]
Ano: [dropdown]
Mês: [dropdown]

Transferências: [dropdown]
 HTML
 XLS
 Consultar [button]

Fonte: http://www3.tesouro.fazenda.gov.br/estados_municipios/transferencias_constitucionais.asp

A princípio, pode-se ter a impressão de que será necessário coletar as informações estaduais uma por vez, mas esse não é o caso. A configuração apresentada na Figura A.1, por exemplo, apresentará os repasses de FPE a todos os estados em todos os meses de 2010. Isso porque, ao deixar a janela em branco, requisita-se que as informações de todas as categorias constantes na janela sejam apresentadas.

Para os municípios, o funcionamento é semelhante, com a exceção de que a janela *estado* não pode ser deixada em branco. Dessa forma, só é possível coletar as informações municipais individuais de um estado de cada vez. Mas com um pouco de paciência, o pesquisador poderá realizar o processo repetidamente e obter as informações para todos os municípios, individualmente.

Para o leitor interessado em conhecer mais profundamente as séries disponíveis no site da STN, recomenda-se que acesse: <http://www3.tesouro.gov.br/series_temporais/principal.aspx#ancora_consulta>.

Bases regionais brasileiras

A principal fonte de dados regionais brasileira é o *site* do Instituto Brasileiro de Geografia e Estatística (IBGE) <www.ibge.gov.br>, no qual podem ser encontradas duas das informações mais importantes para quem se dispõe a tratar dados regionalizados, a saber: as informações sobre os Produtos Internos Brutos (PIBs) estaduais e municipais.

A obtenção dos PIBs estaduais pode ser efetuada de duas formas distintas. A primeira delas é: ao entrar na página do instituto, clicar na aba intermediária *Economia*, como o mostrado na Figura A.2. a seguir. Esse movimento abrirá opções nessa aba, dentre as quais se podem encontrar as *Contas Regionais*, como o destacado também na Figura A.2.



Fonte: <http://www.ibge.gov.br>

Na página que se abre, podem ser encontradas informações metodológicas, tabelas prontas e os dados sobre os PIBs estaduais de 1995 a 2009.

Outra forma de se obter tais informações na página do IBGE é clicando na aba *Download* (também destacada na Figura A.2) e a

partir daí seguir o seguinte caminho: *Estatística\Contas Regionais*. Procedendo dessa forma, tem-se acesso às contas regionais desde 1985. Contudo, deve-se atentar ao fato de que as séries aí obtidas têm dois padrões distintos, visto que houve uma modificação na metodologia de cálculo em 2002.

Os PIBs municipais podem ser obtidos no site do IBGE de forma semelhante. Na aba *Economia* existe uma opção denominada *Produto Interno Bruto Municipal* (ver Figura A.2). Clicando-se nessa opção, obtém-se acesso às séries de PIBs municipais de 1999 a 2009.

Outra fonte de dados importante aos estudiosos da economia regional brasileira é o site do Ipeadata (<http://www.ipeadata.gov.br/>). É válido observar que as séries contidas nesse sítio não são produzidas pelo Instituto de Pesquisa Econômica Aplicada (Ipea), sendo coletadas a partir das instituições que as produzem e pelo Instituto.

Ao entrar na página, o leitor encontrará três abas (macroeconômico, regional e social). Ao clicar na aba *Regional* (em verde), esta se desdobra em quatro opções: *Temas*, *Fontes*, *Nível Geográfico* e *Ajuda*.

A escolha da opção *Temas* provoca ainda um novo desdobramento, no qual vários assuntos são apresentados. É importante notar que todas as bases contidas nos diversos temas apresentam informações subnacionais. Por exemplo, o tema *Finanças Públicas* disponibiliza diversas séries de receitas e gastos de governos estaduais e municipais.

Outro tema bastante útil é o de *Contas Nacionais*, que, nesse caso, retrata as contas regionais. Ali se encontram séries de produto interno estadual e municipal, bem como as suas desagregações por setor econômico e os seus valores *per capita*.

Bases setoriais: educação

A base de dados educacionais mais completa e relevante é a do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). O pesquisador interessado nas informações sobre educação pode acessar especificamente a página de divulgação dos microdados do instituto (<http://portal.inep.gov.br/basicalevantamentos-acessar>).

Nessa página se encontram diversos conjuntos de informações como aquelas referentes aos resultados do Exame Nacional do Ensino

Médio (Enem), da Prova Brasil e do Exame Nacional de Desempenho de Estudantes (Enade). Os arquivos existentes referem-se aos microdados dos participantes e escolas, e necessitam da utilização de *softwares* para a sua leitura.

Na página de *Indicadores Educacionais* (<http://portal.inep.gov.br/indicadores-educacionais>) existem várias informações sobre rendimento escolar, número de vagas e outros parâmetros educacionais, também discriminados por unidades subnacionais.

Além disso, no sítio da STN ou no do Ipea, é possível obter o gasto de cada município e de cada estado na função educação, desagregado inclusive em suas diversas subfunções. Dessa forma, podem-se desenvolver diversos modelos de análise de eficiência na área de educação pelo contraste dos recursos utilizados com os resultados obtidos.

Bases setoriais: saúde

A fonte mais completa de informações sobre a saúde no Brasil é o *site* do Datasus (www.datasus.gov.br), onde podem ser encontradas informações sobre estados e municípios, dados operacionais e financeiros, bem como estatísticas socioeconômicas relacionadas à saúde.

O primeiro grande tópico no *site* se refere às informações de saúde propriamente dita, que são subdivididas em oito subtópicos. O segundo grande tópico traz informações financeiras do sistema de saúde brasileiro e contém cinco subtópicos, dentre os quais se destaca, para os objetivos desse volume, o item *Transferências a Municípios*. No Quadro A.1, o leitor pode visualizar a estrutura de tópicos informacionais contidos no Datasus.

Quadro A.1 – Estrutura das Informações do Datasus

Informações de Saúde (Tabnet)

- Indicadores de saúde
- Assistência à saúde
- Epidemiológicas e morbidade
- Rede assistencial
- Estatísticas vitais
- Demográficas e socioeconômicas
- Inquéritos e pesquisas
- Saúde suplementar

Informações Financeiras

- Recursos do SUS
- Siops
- Transferência a municípios
- Créditos a prestadores
- GAP

Outras bases úteis para avaliação de gastos governamentais

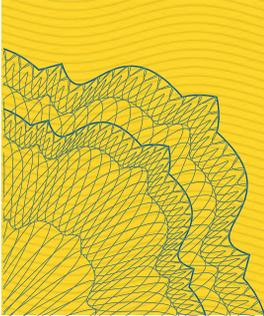
As secretarias de Fazenda da maioria dos estados e as dos grandes municípios mantêm em seus *sites* informações pormenorizadas dos gastos públicos de seus governos. Outra fonte de informações diretas sobre finanças públicas estaduais são os *sites* dos Tribunais de Contas Estaduais (TCEs), muito embora nesse caso a heterogeneidade entre os dados das diversas unidades seja acentuada.

A Secretaria da Receita Federal (SRF) pode ser útil para aqueles pesquisadores interessados na distribuição estadual de arrecadação de tributos federais. Em seu *site*, na página <<http://www.receita.fazenda.gov.br/Arrecadacao/default.htm>>, podem ser encontrados, além de informações sobre arrecadação federal em geral, dados

regionalizados por unidade da Federação sobre a arrecadação de cada tributo federal entre os anos de 1994 e 2012.

Os dados de arrecadação por município não estão disponíveis por uma questão de manutenção do sigilo fiscal, pois algumas empresas sediadas em municípios pequenos teriam seus dados operacionais praticamente descortinados.

Outra restrição aos dados de arrecadação regional da SRF é a de que geralmente os impostos são apropriados no estado no qual a sede da empresa se localiza. Assim, o Imposto de Renda Pessoa Jurídica (IRPJ) relativo, digamos, ao Banco do Brasil, será todo registrado como tendo sido arrecadado no Distrito Federal, quando as operações que o geraram foram realizadas por todo o território nacional.



Nota sobre os autores

Alex Pereira Benício é graduado e mestre em Economia pela Universidade de Brasília (UnB). Ingressou na Secretaria do Tesouro Nacional em 1998, lotado na Coordenação-Geral de Estudos Econômico-Fiscais, onde foi gerente de Planejamento Fiscal e atualmente é o coordenador.

Ana Carolina Giuberti é graduada em Ciências Econômicas e doutora em Economia do Desenvolvimento pela Universidade de São Paulo. Professora do Departamento de Economia da Universidade Federal do Espírito Santo (UFES) desde 2005, atua nas áreas de macroeconomia e economia do setor público.

Borko Stosic é doutor em Física Teórica pela Universidade de Belgrado e possui pós-doutorado pela Universidade de Boston. É professor associado da Universidade Federal Rural de Pernambuco (UFRPE) e pesquisador do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). Desenvolve pesquisas em Métodos Computacionalmente Intensivos, Física Estatística, Pesquisa Operacional.

Enlison Mattos é professor adjunto (associado) da Escola de Economia da Fundação Getulio Vargas (EESP-FGV). Possui bacharelado em Economia pela Universidade de São Paulo – campus Ribeirão Preto (FEARP-USP) e mestrado em economia pela Universidade de São Paulo (FEA-USP). É Ph.D. em Economia pela University of Illinois – Estados Unidos. É professor da FGV desde 2006. Desenvolve pesquisas sobre finanças públicas e possui, nessa área, publicações científicas em diversas revistas acadêmicas internacionais e nacionais.

Fabiana Magalhães Almeida Rodopoulos é graduada e mestre em Ciências Econômicas pela Universidade de Brasília (UnB). Ingressou na Secretaria do Tesouro Nacional em 2001, onde atualmente é Coordenadora-Geral de Estudos Econômico-Fiscais.

Fabiana Rocha é graduada em Ciências Econômicas pela Universidade Federal de Minas Gerais (UFMG), mestre em Economia pela Universidade de São Paulo (USP) e doutora em Economia pela University of Illinois at Urbana-Champaign. É professora titular do Departamento de Economia da Universidade de São Paulo, atuando nas áreas de macroeconomia e economia do setor público.

Fabiano Silvio Colbano é graduado e mestre em Ciências Econômicas pela Universidade de São Paulo (USP). Ingressou na Secretaria do Tesouro Nacional em 2007 como analista na Gerência de Análise Econômica e Cenários da Coordenação-Geral de Planejamento Estratégico da Dívida Pública Federal, e foi gerente de Pesquisa e Desenvolvimento desta mesma coordenação-geral. Atualmente, é gerente de Planejamento Fiscal da Coordenação-Geral de Estudos Econômico-Fiscais.

Felipe Palmeira Bardella é engenheiro eletrônico graduado pelo Instituto Tecnológico de Aeronáutica (ITA) e mestre em Economia pela Universidade de São Paulo (FEA/USP). Ingressou na Secretaria do Tesouro Nacional em 2005, onde ocupou o cargo de Coordenador de Estudos Econômico-Fiscais no período de 2008-2013. Atualmente cursa o programa de doutorado em Economia no Department of Economics at Michigan State University.

Ian Lienert é graduado em Matemática e Economia pela Universidade de Canterbury, Nova Zelândia, tendo iniciado sua carreira no Tesouro Neozelandês. Durante 1976-89, trabalhou no Departamento de Economia da OCDE. Em 1989, integrou-se à equipe do FMI e trabalhou no Departamento de Finanças Públicas no período 1996-2010, onde se especializou em finanças públicas e gestão do orçamento. Desde 2010, é consultor independente em gestão pública financeira. Suas publicações e interesses de pesquisa incluem comparação de sistemas de leis orçamentárias, transparência fiscal e o papel do legislativo em processos orçamentários.

Janete Duarte é graduada e mestre em Economia pela Universidade Federal de Minas Gerais (FACE/CEDEPLAR – UFMG). Ingressou na Secretaria do Tesouro Nacional em 2001. Atualmente, é gerente de Estudos Econômico-Fiscais da Coordenação-Geral de Estudos Econômico-Fiscais.

José de Anchieta Semedo Neves é graduado em Economia pela Universidade Federal do Espírito Santo (UFES) e mestre em Economia pela Universidade Federal do Rio Grande do Sul (UFRGS). Ingressou na Secretaria do Tesouro Nacional em 2009 como analista e foi gerente de projeto de Estudos Econômico-Fiscais da Coordenação-Geral de Estudos Econômico-Fiscais e está cursando programa de doutorado na Universitat Autònoma de Barcelona.

Luis Felipe Vital Nunes Pereira é graduado em Engenharia Elétrica pela USP (Poli-USP) e é doutor em Economia pela Universidade Católica de Brasília, com período sanduíche na Duke University. Ingressou na Secretaria do Tesouro Nacional em 2007. Atualmente, é gerente de Análise e Monitoramento da Coordenação-Geral de Análise Econômico-Fiscal de Projetos de Investimentos Públicos.

Marc Robinson é consultor internacional da *Public Financial Management* (PFM), especializado em elaboração de orçamento, e está à frente da Consultoria *PFM Results* desde 2009. É membro do Painel Consultivo sobre Orçamento e Despesa Pública da OCDE, tendo trabalhado como consultor em temas relacionados à reforma orçamentária em mais de 25 países, em nome do Banco Mundial, da OCDE, do FMI etc. Dentre suas publicações, destacam-se as relacionadas ao desempenho orçamentário, tais como: *Linking Funding and Results* (Palgrave/IMF, 2007) and *Performance Budgeting* (World Bank, 2011).

Maria da Conceição Sampaio de Sousa é licenciada e mestre em Ciências Econômicas pela Universidade Católica de Louvain e doutora em Economia pela Universidade Livre de Bruxelas, Bélgica. Professora titular aposentada da Universidade de Brasília (UnB), é atualmente professora titular da Universidade Federal da Paraíba (UFPB), tendo lecionado na Universidade Tsukuba, no Japão, na Universidade de São Paulo (USP) e na Universidade Federal de Pernambuco (UFPE). Pesquisadora 1A do CNPq, concentra suas pesquisas em Economia do Setor Público e Análise de Eficiência.

Mauricio Dias Leister é graduado e doutor em Ciências Econômicas pela Universidade de São Paulo (USP) e mestre em Ciências Econômicas pela Universidade Estadual de Campinas (Unicamp). Ingressou na Secretaria do Tesouro Nacional em 2011, onde atualmente é gerente da Gerência de Pesquisa e Desenvolvimento da Coordenação-Geral de Planejamento Estratégico da Dívida Pública Federal.

Plínio Portela de Oliveira é analista de finanças e controle da Secretaria do Tesouro Nacional. Graduou-se em Economia pela Universidade Federal de Pernambuco (UFPE) e atualmente cursa graduação de Matemática pela Universidade de Brasília (UnB).

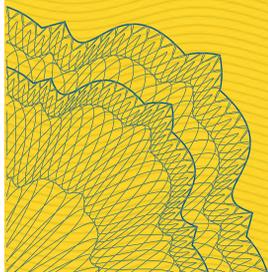
Rafael Terra é professor adjunto da Universidade de Brasília desde 2012. É bacharel em Economia pela Universidade de São Paulo (FEA-USP), mestre em Economia Aplicada pela Universidade de São Paulo (FEARP-USP) e doutor em Economia pela Escola de Economia da Fundação Getúlio Vargas (EESP-FGV) com período sanduíche na University of Illinois. Desenvolve trabalhos em Economia do Setor Público e Microeconomia Aplicada. Tem publicações em revistas e anais de congressos nacionais e internacionais. Trabalhou como consultor para diversos organismos internacionais, como PNUD, Unesco e BID.

Rodrigo Hitoshi Dias formou-se em Psicologia pela Bates College-EUA e é mestre em Economia do Setor Público pela Universidade de Brasília. Trabalhou na Harvard Initiative for Global Health, Ministério da Saúde e Ministério Público Federal, e, atualmente, é analista técnico de políticas sociais no Ministério da Saúde.

Rogério Boueri possui graduação em Ciências Econômicas pela Universidade de Brasília (UnB), mestrado em Economia pela Fundação Getulio Vargas (FGV-RJ) e doutorado em Economia pela Universidade Maryland College Park. Atualmente é economista do Instituto de Pesquisa Econômica Aplicada (IPEA). Tem experiência na área de Economia, com ênfase em Análise de Eficiência e Produtividade e Finanças Públicas.

Sérgio Ricardo de Brito Gadelha é doutor em Economia pela Universidade Católica de Brasília. Ingressou na Secretaria do Tesouro Nacional em 2003. Atualmente, exerce suas atividades na Gerência de Estudos Econômico-Fiscais da Coordenação-Geral de Estudos Econômico-Fiscais.

Sobre o Projeto Gráfico



O projeto gráfico do livro **Avaliação da Qualidade do Gasto Público e Mensuração da Eficiência** foi desenvolvido pelo diretor de arte Humberto Cunha e busca representar o universo visual do tema por meio de grafismos e ilustrações que utilizam a técnica *guilloche*.

A *guilloche* foi muito utilizada na arquitetura grega, na romana e nas construções neoclássicas. Apesar de o alemão Hans Schwanhardt também ser citado, os créditos pelo desenvolvimento dessa técnica são dados ao engenheiro francês Guillot. Foi ele o inventor de uma máquina que originalmente criava padrões decorativos com linhas finas e repetidas, que se entrelaçavam formando um padrão em torno de um espaço central, gerando grafismos visualmente agradáveis.

Nos dias de hoje, ilustrações em *guilloche* são utilizadas em diversos segmentos. Por serem dificilmente copiadas em um período que antecedeu à invenção de *scanners* e afins, foi adotada como linguagem gráfica das cédulas de dinheiro ao redor do mundo. Isso fez com que essa técnica fosse popularmente conhecida mesmo quase 500 anos depois do seu desenvolvimento e a tornou amplamente associada, visualmente, a temas relacionados à economia.



Com muitos países esforçando-se para alcançar as metas fiscais ou as de dívida pública, há considerável interesse na melhoria da eficiência do gasto público. Ao priorizar o gasto, o governo necessita de ferramentas para identificar ineficiências do gasto público e eliminar o desperdício. Os contribuintes desejam potencializar o valor do seu dinheiro nas políticas públicas existentes e observar as economias no orçamento público advindas do aprimoramento da eficiência do gasto, sem comprometer a qualidade dos serviços.

É possível definir e mensurar a eficiência do gasto no setor público? Este livro fornece uma resposta afirmativa para essa questão. A obra discute o conceito de eficiência do gasto público, que é menos claro que no setor privado, uma vez que os produtos do governo, os custos dos fatores e o progresso tecnológico não são fáceis de mensurar. A produção de bens e de serviços públicos não é feita apenas para o seu retorno econômico, mas também para alcançar objetivos sociais e ambientais. Apesar desses desafios, o livro apresenta várias técnicas, inclusive programas de computação, para medir a eficiência dos gastos públicos. Considerações teóricas são importantes e também são examinadas, porém o livro é empírico, uma vez que examina como a temática de eficiência do gasto público tem sido abordada no Brasil e no âmbito internacional.

Este volume, preparado pelo Ministério da Fazenda em colaboração com o Banco Mundial, ajuda a esclarecer como diversos países podem detectar programas públicos ineficientes e de baixa prioridade. As ferramentas analíticas simples e avançadas descritas no livro são aplicáveis em todos os níveis de governo, tanto federal quanto estadual e municipal, bem como em setores específicos, como educação e saúde. Fortalecendo a compreensão das múltiplas facetas da eficiência do gasto público, este livro beneficiará tanto executores quanto pesquisadores no Brasil e no exterior.

Otaviano Canuto - Vice-presidente, Banco Mundial



Ministério da
Fazenda

