APOSENTADORIA E MERCADO DE TRABALHO: UMA ANÁLISE USANDO REGRESSÃO DESCONTÍNUA

Gustavo Marcos Szniter Mentlik Naercio Menezes-Filho Bruno Kawaoka Komatsu

DECOMPOSIÇÃO QUANTÍLICA INCONDICIONAL DOS DIFERENCIAIS DE DESEMPENHO ENTRE ALUNOS DE ESCOLAS PRIVADAS E PÚBLICAS PROFISSIONALIZANTES

Ronaldo A. Arraes Francisca Zilania Mariano

ANÁLISE DA EFICIÊNCIA TÉCNICA E DA PRODUTIVIDADE DOS SERVIÇOS DE ÁGUA E ESGOTOS NO BRASIL DE 2006 A 2013

Felipe Ponciano da Cruz Ronaldo Seroa da Motta Alexandre Marinho

A DURAÇÃO DAS EXPORTAÇÕES BRASILEIRAS: EVIDÊNCIA E DETERMINANTES

Hélio de Sousa Ramos Filho Rennan Kertlly de Medeiros

INTRA-INDUSTRY TRADE: ANALYSIS OF THE SMOOTH ADJUSTMENT HYPOTHESIS WITH RESPECT TO BRAZIL

Priscilla Belle Oliveira Pinto Cláudio R. Foffano Vasconcelos Ricardo da Silva Freguglia

MEDINDO A EFICIÊNCIA PRODUTIVA DO TRANSPORTE POR FERROVIAS BRASILEIRAS: UMA APLICAÇÃO DOS MODELOS DEA E TOBIT

Francisco Gildemir Ferreira da Silva José Kleber Macambira Carlos Henrique Rocha



pesquisa e planejamento econômico • ppe

ipea

Governo Federal

Ministério da Economia Ministro Paulo Guedes



Fundação pública vinculada ao Ministério da Economia, o Ipea fornece suporte técnico e institucional às ações governamentais – possibilitando a formulação de inúmeras políticas públicas e programas de desenvolvimento brasileiros – e disponibiliza, para a sociedade, pesquisas e estudos realizados por seus técnicos.

Presidente

Carlos von Doellinger

Diretor de Desenvolvimento Institucional

Manoel Rodrigues Junior

Diretora de Estudos e Políticas do Estado, das Instituições e da Democracia

Flávia de Holanda Schmidt

Diretor de Estudos e Políticas Macroeconômicas

José Ronaldo de Castro Souza Júnior

Diretor de Estudos e Políticas Regionais, Urbanas e Ambientais

Nilo Luiz Saccaro Júnior

Diretor de Estudos e Políticas Setoriais de Inovação e Infraestrutura

André Tortato Rauen

Diretora de Estudos e Políticas Sociais

Lenita Maria Turchi

Diretor de Estudos e Relações Econômicas e Políticas Internacionais

Ivan Tiago Machado Oliveira

Assessora-chefe de Imprensa e Comunicação

Mylena Fiori

Ouvidoria: http://www.ipea.gov.br/ouvidoria URL: http://www.ipea.gov.br

Pesquisa e Planejamento Econômico (PPE)

Publicação quadrimestral com análises teóricas e empíricas sobre uma ampla gama de temas relacionados à economia brasileira. Estabelecida em 1971 sob o título Pesquisa e Planejamento, PPE é publicada em abril, agosto e dezembro.

Corpo Editorial

Editor

Mauricio Cortez Reis

Coeditores

Marco A.F.H. Cavalcanti José Gustavo Feres Danilo Santa Cruz Coelho

Membros

Lauro Ramos (Ipea)

Eduardo Fiuza (Ipea)

Alexandre Xavier Ywata de Carvalho (Ipea)

Daniel da Mata (Ipea)

Carlos Viana de Carvalho (PUC-RJ)

Eduardo Rios Neto (CEDEPLAR-UFMG)

José Raimundo Carvalho (CAEN-UFC)

Marcelo Portugal (UFRGS)

Marco Bonomo (Insper)

Mônica Viegas de Andrade (CEDEPLAR-UFMG)

Rafael Coutinho Costa Lima (UFPE)

Renata Narita (FEA-USP)

Ricardo Paes de Barros (Insper)

Roberto G. Ellery (UnB)

Sergio Firpo (Insper)

Vladimir Ponczek (EESP-FGV/SP)

Secretária Executiva

Thamires Fernandes da Silva

© Instituto de Pesquisa Econômica Aplicada – ipea 2020

Pesquisa e Planejamento Econômico v. 1 – n.1 – jun. 1971.

Rio de Janeiro: Instituto de Pesquisa Econômica Aplicada, 1990 — v. — quadrimestral.

Título anterior: Pesquisa e Planejamento v. 1, n. 1 e 2, 1971

Periodicidade anterior: semestral de 1971–1975.

1. Economia — Pesquisa — Periódicos. 2. Planejamento Econômico — Brasil. I. Brasil. Instituto de Pesquisa Econômica Aplicada.

ISSN - 0 100-0551

CDD 330.05 33(81) (05)

As publicações do Ipea estão disponíveis para download gratuito nos formatos PDF (todas) e EPUB (livros e periódicos). Acesse: http://www.ipea.gov.br/portal/publicacoes

As opiniões emitidas nesta publicação são de exclusiva e inteira responsabilidade dos autores, não exprimindo, necessariamente, o ponto de vista do Instituto de Pesquisa Econômica Aplicada ou do Ministério da Economia.

É permitida a reprodução deste texto e dos dados nele contidos, desde que citada a fonte. Reproduções para fins comerciais são proibidas.

SUMÁRIO

APOSENTADORIA E MERCADO DE TRABALHO: UMA ANÁLISE USANDO REGRESSÃO DESCONTÍNUA 5

Gustavo Marcos Szniter Mentlik Naercio Menezes-Filho Bruno Kawaoka Komatsu

DECOMPOSIÇÃO QUANTÍLICA INCONDICIONAL DOS DIFERENCIAIS DE DESEMPENHO ENTRE ALUNOS DE ESCOLAS PRIVADAS E PÚBLICAS PROFISSIONALIZANTES 29

Ronaldo A. Arraes Francisca Zilania Mariano

ANÁLISE DA EFICIÊNCIA TÉCNICA E DA PRODUTIVIDADE DOS SERVIÇOS DE ÁGUA E ESGOTOS NO BRASIL DE 2006 A 2013 81

Felipe Ponciano da Cruz Ronaldo Seroa da Motta Alexandre Marinho

A DURAÇÃO DAS EXPORTAÇÕES BRASILEIRAS: EVIDÊNCIA E DETERMINANTES 107

Hélio de Sousa Ramos Filho Rennan Kertlly de Medeiros

COMÉRCIO INTRA-INDÚSTRIA: ANÁLISE DA HIPÓTESE DE AJUSTE SUAVIZADO PARA O BRASIL 127

Priscilla Belle Oliveira Pinto Cláudio R. Foffano Vasconcelos Ricardo da Silva Freguglia

MEDINDO A EFICIÊNCIA PRODUTIVA DO TRANSPORTE POR FERROVIAS BRASILEIRAS: UMA APLICAÇÃO DOS MODELOS DEA E TOBIT 153

Francisco Gildemir Ferreira da Silva José Kleber Macambira Carlos Henrique Rocha

CONTENTS

RETIREMENT AND THE LABOR MARKET: AN ANALYSIS USING REGRESSION DISCONTINUITY DESIGN 5

Gustavo Marcos Szniter Mentlik Naercio Menezes-Filho Bruno Kawaoka Komatsu

UNCONDITIONAL QUANTILE DECOMPOSITION OF PERFORMANCE DIFFERENTIALS BETWEEN STUDENTS FROM PRIVATE AND VOCATIONAL PUBLIC SCHOOLS 31

Ronaldo A. Arraes Francisca Zilania Mariano

ANALYSIS OF TECHNICAL EFFICIENCY AND PRODUCTIVITY OF WATER AND SEWAGE SERVICES IN BRAZIL, 2006 TO 2013 83

Felipe Ponciano da Cruz Ronaldo Seroa da Motta Alexandre Marinho

THE DURATION OF BRAZILIAN EXPORTS: EVIDENCE AND DETERMINANTS 109

Hélio de Sousa Ramos Filho Rennan Kertlly de Medeiros

INTRA-INDUSTRY TRADE: ANALYSIS OF THE SMOOTH ADJUSTMENT HYPOTHESIS WITH RESPECT TO BRAZIL 129

Priscilla Belle Oliveira Pinto Cláudio R. Foffano Vasconcelos Ricardo da Silva Freguglia

MEASURING THE PRODUCTIVE EFFICIENCY OF BRAZILIAN RAILWAYS TRANSPORT: AN APPLICATION OF DEA AND TOBIT MODELS 155

Francisco Gildemir Ferreira da Silva José Kleber Macambira Carlos Henrique Rocha

APOSENTADORIA E MERCADO DE TRABALHO: UMA ANÁLISE USANDO REGRESSÃO DESCONTÍNUA

Gustavo Marcos Szniter Mentlik¹ Naercio Menezes-Filho² Bruno Kawaoka Komatsu³

A previdência social é um seguro que repõe a renda do contribuinte quando ele perde a capacidade para o trabalho. No entanto, as regras em vigor permitem que um grupo de indivíduos se aposente relativamente cedo, ainda com capacidade laborativa, e podem impactar as decisões dos indivíduos no mercado de trabalho, aumentando a propensão à inatividade. Para se ter uma ideia desse impacto, exploramos a descontinuidade das regras sobre aposentadoria por idade com regressões descontínuas (RDD). Os resultados encontrados indicam que a aposentadoria tem um efeito significativo sobre o mercado de trabalho urbano no Brasil, porém não influenciam as condições de trabalho daqueles que permanecem ocupados.

Palavras-chave: aposentadoria; mercado de trabalho.

RETIREMENT AND THE LABOR MARKET: AN ANALYSIS USING REGRESSION DISCONTINUITY DESIGN

Social security is an insurance that replaces the taxpayer's income when there is no capacity to work. However, the rules in effect allow a group of individuals to retire relatively early, when they still have work capacity, and might impact the decisions of individuals in labor market, increasing the propensity to inactivity. To get an idea of this impact, we explore the discontinuity of the rules on age retirement with regression discontinuity design (RDD). Our results indicate that retirement has a significant effect on the urban labor market in Brazil, but does not influence the work conditions of those who remain occupied.

Keywords: retirement; labor market.

JEL: H75; I38; J26.

1 INTRODUÇÃO

Nos últimos anos, o debate público brasileiro se concentrou sobre o funcionamento e as possíveis mudanças na previdência social, desde o lançamento da Proposta de Emenda Constitucional (PEC) nº 287, de 2016, e de uma nova proposta mais recente, a PEC nº 6, de 2019. Entre diversas preocupações a respeito de possíveis efeitos dessas propostas estão a sustentabilidade do sistema, o tempo de contribuição, a possibilidade de pessoas não se aposentarem, além de questões sobre a desigualdade de benefícios e as diferenças abruptas entre indivíduos semelhantes geradas pelos

^{1.} Formado na Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo (FEA/USP). *E-mail*: <qustavo.mentlik@qmail.com>.

^{2.} Professor titular da Cátedra Ruth Cardoso no Insper e professor associado da FEA/USP. *E-mail*: <naercioamf@insper. edu.br>.

^{3.} Assessor de pesquisas da Cátedra Ruth Cardoso no Insper. E-mail:

 cbrunokk@uol.com.br>.

critérios de transição. Apesar disso, pouco foi discutido a respeito de alterações possíveis no conjunto de incentivos e restrições e suas possíveis consequências sobre o comportamento dos indivíduos no mercado de trabalho.

No Brasil, as políticas de reposição da renda por perda da capacidade laborativa, como aquelas consolidadas pela previdência social, são baseadas na invalidez constatada ou presumida pela idade avançada (Camarano, Kanso e Fernandes, 2013). O recebimento do benefício previdenciário, além disso, parece exercer efeitos diretos sobre a participação e as condições de trabalho dos indivíduos (Leme e Málaga, 2001; Schwarzer e Paiva, 2003; Queiroz e Ramalho, 2009).

Alguns estudos mostram a associação negativa entre a aposentadoria e a participação no mercado de trabalho, com interesse especial sobre as aposentadorias na modalidade por tempo de contribuição entre indivíduos com idade relativamente baixa. Schwarzer e Paiva (2003), por exemplo, concluíram que homens e mulheres que recebem renda de aposentadoria têm taxa de participação no mercado de trabalho expressivamente inferior à taxa daqueles que não recebem nenhum benefício da previdência. Estima-se que, caso os beneficiários de aposentadorias precoces deixassem de receber o benefício, sua taxa de ocupação aumentaria de 39% para 87% (Paiva, Rangel e Caetano, 2016).

Por sua vez, parte significativa da população aposentada permanece em atividade econômica, e alguns fatores parecem favorecer essa situação. No geral, quanto mais jovem, mais escolarizado ou menor for o valor do benefício de segurado, maior a probabilidade de ele continuar ofertando sua força de trabalho após adquirir o benefício (Liberato, 2003). Além disso, há o argumento de que o valor dos benefícios previdenciários seria tão baixo que os trabalhadores teriam que permanecer no mercado de trabalho para conseguirem manter condições mínimas de consumo.

Neste artigo, procuramos contribuir para o debate sobre os efeitos da aposentadoria sobre o mercado de trabalho. Nesse sentido, utilizamos a descontinuidade da regra atual de aposentadoria por idade do Regime Geral de Previdência Social (RGPS) e do Regime Próprio de Previdência Social (RPPS) para estimar os impactos da aposentadoria sobre o comportamento no mercado de trabalho, além de discutir possíveis efeitos da PEC nº 287/2016. A principal contribuição desta pesquisa é a utilização do método de regressão descontínua (RDD) no contexto das regras de aposentadorias, que permite uma inferência sobre o efeito causal do benefício. Um estudo com especificação semelhante foi realizado para o sistema de aposentaria alemão (Eibich, 2014) e não há ainda estudo semelhante para o caso brasileiro.

Os efeitos da aposentaria sobre o mercado de trabalho ganham relevância especialmente quando temos em vista os possíveis problemas a serem enfrentados pelo RGPS no futuro. O sistema brasileiro de previdência social adota o regime de financiamento de repartição simples (comumente chamado de *pay as you go*),

em que os segurados em atividade econômica contribuem para o pagamento dos benefícios dos segurados inativos. Esse modelo se apoia em um esquema de transferências intergeracionais, de modo que sua sustentação atuarial é bastante dependente da dinâmica demográfica.

Como apresentado por diversos autores, como Camarano (2014), o Brasil passará até 2060 por uma drástica transição demográfica, com uma diminuição do grupo etário jovem e adulto (potencialmente produtivo) e substancial crescimento da população idosa. Esta saltará dos 22 milhões de pessoas, em 2013, para em torno de 73,5 milhões, em 2060 – 50 milhões de idosos a mais. Visto por outro ângulo, a relação do número de pessoas em idade ativa (15 a 59 anos) por idoso (60 anos ou mais) deve diminuir drasticamente, de cerca de 6,5, em 2010, para 1,9, em 2050 (Tafner, 2012), indicando que o sistema previdenciário brasileiro deverá sustentar mais idosos beneficiários com menos contribuintes ativos nas próximas décadas O contexto atual de transição da estrutura demográfica no país, portanto, tende a fragilizar as bases do RGPS.

Outro fator que pode afetar a sustentabilidade do sistema é a situação fiscal do RGPS. Em 2015, havia 32,6 milhões de benefícios ativos, o que representou uma despesa de R\$ 436 bilhões (ou 7,38% do produto interno bruto – PIB), segundo o Anuário Estatístico da Previdência Social – AEPS (Brasil, 2016b). No entanto, havia cerca de 70 milhões de contribuintes pessoas físicas e uma arrecadação líquida de R\$ 350 bilhões (5,93% do PIB). O *deficit* do RGPS, em 2015, portanto, ficou na casa de R\$ 85 bilhões (1,45% do PIB); por sua vez, o *deficit* previsto para o RGPS, em 2016, é de, aproximadamente, R\$ 150 bilhões, conforme o Resultado Fiscal do Governo Central, do Tesouro Nacional.⁴

A perspectiva de envelhecimento da população, devido à maior esperança de vida dos idosos, traz a expectativa de maiores gastos da previdência social no futuro. Considerando a manutenção de níveis elevados de informalidade no mercado de trabalho brasileiro e o menor número de contribuintes em idade ativa nas próximas décadas, é provável que o RGPS enfrente dificuldades fiscais nos próximos cinquenta anos (Tafner, Botelho e Erbisti, 2014).

Quanto às regras atuais do RGPS em vigência, o Brasil é um dos únicos treze países do mundo que concede benefícios de aposentadoria sem idade mínima. Esse requisito está presente tanto em países desenvolvidos, como Alemanha, França e Estados Unidos, quanto em países subdesenvolvidos mais próximos ao Brasil, como África do Sul, Argentina e México. A aposentadoria por tempo de contribuição (ATC) que vigora atualmente não exige idade mínima e tem como requisito apenas o tempo mínimo de contribuição de 35 anos para os homens e 30 para as mulheres. A suposição adotada é que esse tipo de aposentadoria permite que trabalhadores se aposentem antes de perder aptidão laborativa.

^{4.} Informações disponíveis em: https://bit.ly/2SdsjOq>..

De acordo com cálculos realizados a partir de dados do Ministério do Trabalho e Previdência Social (Brasil, 2018), a idade média com que os trabalhadores se aposentam por tempo de contribuição foi de 53 anos para as mulheres e 55,7 para os homens em 2016 (idades aquém dos 60 anos, que define os indivíduos como idosos). Esses valores fazem com que a idade média de aposentadoria no Brasil seja consideravelmente menor do que a idade média com que os trabalhadores costumam se aposentar nos países da Organização para a Cooperação e o Desenvolvimento Econômico (OCDE) — entre os homens, por exemplo, a média fica em 59,4 anos para o Brasil e 64,2 anos para a média da OCDE, de acordo com o mesmo estudo.

Pretende-se analisar, aqui, os efeitos do atual conjunto de regras do RGPS sobre o mercado de trabalho brasileiro e o impacto da possível medida que estabelece uma idade mínima de 65 anos para todas as categorias de aposentadoria (parte da PEC nº 287/2016). A questão de interesse principal é o impacto de aposentadorias na ocupação, na atividade e no salário dos idosos. Esse é o cerne deste estudo porque o conjunto atual de regras da previdência social pode exercer influência sobre as decisões dos indivíduos no mercado de trabalho. Dado que o objetivo que o próprio RGPS se coloca é substituir a renda do segurado quando da perda de sua capacidade de trabalho, o regime não deveria se constituir como um mecanismo que incentiva os trabalhadores a se afastarem do mercado de trabalho.

Este artigo aponta para o fato de que as aposentadorias consideradas precoces são destinadas, majoritariamente, à população concentrada na parte superior da distribuição de renda e contribuem para afastar esses indivíduos do mercado de trabalho. Isso ocorre na medida em que a condição de aposentado tem impacto significativo na taxa de ocupação e atividade dos indivíduos de áreas urbanas, de forma que aqueles que recebem benefícios previdenciários têm uma probabilidade cerca de 50% maior de permanecerem desocupados e inativos em comparação com os não aposentados.

O método de RDD é utilizado com o intuito de indicar e mensurar o efeito da condição de aposentado sobre variáveis relacionadas ao mercado de trabalho, contribuição que ainda não está presente na literatura brasileira de aposentadorias. A regressão de primeiro estágio demonstra o efeito da regra sobre a condição de aposentado, e as outras regressões de segundo estágio apresentam o efeito do *status* de aposentado sobre ocupação, atividade e salário dos indivíduos.

Além desta introdução, o artigo é organizado da seguinte forma: a seção 2 trata das atuais regras da previdência e seu contexto institucional, e a seção 3 apresenta uma breve descrição sobre a base de dados utilizada e a metodologia adotada, discutindo-se as hipóteses necessárias para a implementação das regressões e dos resultados descritivos. Na seção 4, os resultados obtidos e a questão central abordada são expostos, comparando-se com a literatura existente. Por fim, na seção 5, as conclusões são explicitadas buscando responder às questões de pesquisa colocadas inicialmente.

2 CONTEXTO INSTITUCIONAL

O RGPS tem suas políticas elaboradas pelo Ministério da Previdência Social (MPS) e realizadas pelo Instituto Nacional do Seguro Social (INSS), autarquia federal a ele vinculada. A previdência social é um seguro que garante a renda do contribuinte e de sua família e exerce a função de repor a renda do segurado quando ele perde a capacidade para o trabalho.

A partir da promulgação da Constituição Federal de 1988, a previdência social solidificou sua base institucional e, de acordo com Rangel *et al.* (2009), avançou no sentido da universalização, da participação social e da diminuição das desigualdades econômicas. O texto constitucional assegura a cobertura previdenciária de caráter contributivo a idade, invalidez, morte, acidente no trabalho, reclusão e proteção à família. Além disso, garante a manutenção do valor real dos benefícios e estabelece o salário mínimo como piso para os benefícios previdenciários.

A Constituição Federal também conferiu grande movimento de ampliação da cobertura previdenciária, principalmente devido à possibilidade de que mais de um benefício rural fosse concedido por unidade familiar. Isso ocorreu pela modificação dos critérios de elegibilidade da aposentadoria para os trabalhadores rurais, que devem comprovar efetivo exercício de atividade rural no período de requerimento do benefício para se enquadrarem na categoria.

Atualmente, existem quatro tipos de aposentadoria: por idade, por tempo de contribuição, por invalidez e aposentadoria especial. Conforme o art. 201 da Constituição Federal (Brasil, 1988), para que trabalhadores urbanos possam se aposentar por idade, exigem-se 65 anos para os homens e 60 para as mulheres (inciso II, § 7º), além de um período de contribuição mínimo de quinze anos – 180 contribuições mensais. Os trabalhadores rurais podem requerer esse tipo de aposentadoria com cinco anos de idade a menos: a partir dos 60 para os homens e 55 para as mulheres. Também devem ser comprovados 180 meses de atividade rural. Para os chamados segurados especiais, categoria em que estão incluídos o trabalhador rural e seus familiares que produzem em regime de economia familiar (produtor rural pessoa física sem empregados), o pescador artesanal e o índio, a idade mínima também é reduzida em cinco anos para ambos os sexos.

A ATC pode ser integral ou proporcional e não exige idade mínima. A aposentadoria integral é um benefício para aqueles que comprovarem tempo total de contribuição durante 35 anos para os homens e 30 anos para as mulheres. Para os professores que realizam exclusivamente atividade nos ensinos básico, fundamental, médio ou técnico, é exigido tempo mínimo de contribuição no magistério de 25 anos para as mulheres e 30 anos para os homens.

Para evitar que os contribuintes se aposentem cedo, foi instituído, em 1999, por meio da Lei nº 9.876, o fator previdenciário, que é aplicado sobre o valor dos benefícios daqueles que optam pela ATC. O fator é resultado de uma fórmula que leva em conta o tempo de contribuição, a idade do segurado e a expectativa de sobrevida média do brasileiro na data da aposentadoria, adicionando-se um prêmio com o aumento da idade em que se requer o benefício. Concretamente, trata-se de um desconto no valor da aposentadoria para aqueles que decidem se aposentar mais cedo, ou seja, é um desestímulo às aposentadorias precoces. Assim, esse fator tem implícito um incentivo para que os contribuintes permaneçam em atividade, mesmo depois de atingidas as condições de elegibilidade para a aposentadoria (Liberato, 2003).

Como alternativa ao fator previdenciário, foi criada, em 2015, pela Lei nº 13.183, a fórmula 85/95, que soma idade ao tempo de contribuição para garantir a aposentadoria integral, uma opção de cálculo que permite não aplicar o fator. A soma desses anos para as mulheres deve ser de 85 e, para os homens, 95. Para os professores, o período de contribuição somado à idade deve resultar em 80 anos para as mulheres e 90 para os homens. No entanto, para se enquadrar nessa regra, é necessário ter contribuído, pelo menos, trinta anos para o INSS, assim sendo uma mulher pode se aposentar aos 55 anos e um homem, a partir de 65. Essa fórmula é progressiva, como uma tentativa de acompanhar as mudanças na demografia do país, tendo em vista que as pessoas estão vivendo mais. A métrica 85/95 é válida até 2018 e vai aumentando progressivamente até chegar a 90/100, em 2027.

A aposentadoria por invalidez é um benefício concedido aos trabalhadores que, por doença ou acidente, se tornaram incapazes de realizar uma atividade laborativa, sem possibilidade de reabilitação em outra função. No caso de doença, o trabalhador tem de contribuir para a previdência social por, no mínimo, doze meses para receber o benefício e, em caso de acidente, esse prazo de carência não é exigido. Esse tipo de aposentadoria tem como valor mensal 100% do salário de benefício para aqueles que contribuíram com o INSS. Os trabalhadores rurais têm direito a um salário mínimo caso não tenham contribuído espontaneamente. Já a aposentadoria especial é destinada aos trabalhadores que tenham trabalhado em condições prejudiciais à saúde ou à integridade física, de forma contínua e ininterrupta. Nesse caso, o trabalhador pode aposentar-se após cumprir 25, 20 ou 15 anos de contribuição, conforme o agente nocivo. Também é necessário ter trabalhado efetivamente por, no mínimo, 180 meses.

A proposta de reforma da previdência social que tramita como PEC nº 287/2016 é a mais abrangente dos últimos anos. Ela propunha inicialmente idade mínima de 65 anos para todos os contribuintes, incluindo mulheres e trabalhadores rurais. O tempo mínimo de contribuição passaria a ser de 25 anos – inclusive para os trabalhadores rurais, que atualmente não são obrigados a

contribuir para o INSS. Com relação aos servidores públicos e aos professores, o projeto prevê o fim das diferenças entre o regime da previdência geral e público. Nesse sentido, a PEC nº 287/2016 envolvia uma regra de transição para homens com mais de 50 anos e mulheres com mais de 45 anos. Os primeiros poderiam se aposentar pelas regras atuais, porém pagando um pedágio de 50% sobre o tempo que faltava para a aposentadoria. Outro ponto importante daquela proposta residia no fato de que o piso do valor do benefício continuaria sendo o salário mínimo (de R\$ 998,00 em 2019).

A PEC nº 287/2016 gerou uma série de controvérsias no debate público e de resistências e modificações dessa proposição no Congresso Nacional (Fernandes *et al.*, 2019). Apesar dessa PEC já ter tramitado na Comissão de Constituição e Justiça da Câmara dos Deputados, em 2016, e na Comissão Especial da Câmara dos Deputados, em 2017, o governo eleito em 2018 lançou uma nova proposta, planejando uma economia nos gastos públicos de mais de R\$ 1 trilhão em dez anos (Martello, 2019). Assim, a PEC nº 6/2019 foi apresentada ao Congresso, no dia 20 de fevereiro de 2019, sugerindo alterações semelhantes às da PEC nº 287/2016.

A PEC nº 6/2019 também impõe uma idade mínima para a aposentadoria para trabalhadores do RGPS e do RPPS, contudo diferenciada por sexo: de 65 anos para os homens e de 62 para as mulheres. Para trabalhadores do setor privado, essa proposta coloca um período de vinte anos de contribuição, já para trabalhadores do setor público, há um tempo mínimo de 25 anos de contribuição, dez anos no serviço público e cinco anos no cargo. Para os trabalhadores rurais, a idade mínima seria de 60 anos para homens e mulheres, com período de contribuição mínima de vinte anos. Professores da educação básica terão uma idade mínima de 60 anos e tempo de contribuição de trinta anos, e, para aqueles da rede pública, há ainda a exigência de dez anos de serviço público e de cinco anos no cargo.

3 BASE DE DADOS E METODOLOGIA

Neste estudo, analisamos os efeitos da aposentadoria sobre variáveis do mercado de trabalho. Utilizamos a Pesquisa Nacional por Amostra de Domicílios (PNAD) de 2014 como base de dados para as estimações, além de considerarmos as variáveis relacionadas a ocupação, atividade, salários do trabalho principal, horas trabalhadas e formalização. A atividade se refere ao *status* de atividade econômica da pessoa, ou seja, à participação no mercado de trabalho como ocupado ou desocupado.

A PNAD é uma pesquisa amostral domiciliar de abrangência nacional realizada anualmente pelo Instituto Brasileiro de Geografia e Estatística (IBGE), cuja amostra é composta por cerca de 120 mil domicílios e 360 mil pessoas. Suas questões abrangem temas como educação, mercado de trabalho, rendimentos, migração e fecundidade. Entre as variáveis de resultado analisadas, é importante observar que

a ocupação e a atividade são determinadas pela permanência ou não dos indivíduos no mercado de trabalho, enquanto que as variáveis de salário, horas trabalhadas e formalização são definidas somente para as pessoas que permanecem trabalhando.

Consideramos como aposentados os indivíduos que declararam receber rendimento de aposentadoria de instituto de previdência federal (INSS), estadual ou municipal, ou do governo federal, na data de referência da pesquisa, 27 de setembro de 2014. A variável idade foi remodelada para caracterizar-se como uma variável contínua, de forma a facilitar a análise empírica do estudo. Devido às diferenças dos critérios de idade para trabalhadores urbanos e rurais, entre homens e mulheres, nossa análise também é separada para cada um desses quatro grupos. Na PNAD, no entanto, não é possível diferenciar os trabalhadores aposentados pela característica de urbano ou rural relativa ao posto de trabalho que exerciam anteriormente. Consideramos, então, o tipo da área de residência para estabelecer essa separação nos dados. A maior parte da população ocupada das áreas rurais está trabalhando em ocupações agrárias, aproximadamente 64% em 2014. Isso significa que, na nossa estratégia de identificação, boa parte dos trabalhadores dos grupos das áreas rurais está sujeita às regras de aposentadoria para trabalhadores rurais.

Ademais, usamos dados abertos da previdência social e do INSS para obter informações descritivas sobre a maneira como os benefícios de aposentadoria são divididos de acordo com a espécie (tempo de contribuição, invalidez ou por idade), o tipo de trabalho e o valor do benefício.

A estimação dos efeitos da aposentadoria sobre o comportamento dos indivíduos no mercado de trabalho sofre de problemas de endogeneidade, devido especialmente a variáveis omitidas e causalidade inversa. De forma semelhante aos efeitos de aposentadoria sobre indicadores de saúde (Eibich, 2014), é possível que características não observáveis dos indivíduos os induzam à aposentadoria e ao comportamento no mercado de trabalho, tendo como exemplo as condições precárias de saúde que podem levar um indivíduo a procurar a aposentadoria e simultaneamente condicioná-lo a um baixo nível de produtividade e de salário. Além disso, pode haver efeitos de choques no mercado de trabalho sobre a decisão da aposentadoria. Por exemplo, um choque negativo na renda ou uma demissão involuntária pode levar indivíduos ainda produtivos a procurarem a aposentadoria como fonte adicional ou alternativa de renda.

Procuramos atingir uma interpretação causal dos efeitos da aposentadoria sobre o comportamento de mercado de trabalho utilizando a metodologia de

^{5.} A partir da data de nascimento disponível na pesquisa e da data de referência de 27 de setembro de 2014, construímos uma variável de idade contínua em anos, que considera as frações de tempo não contabilizadas na variável de idade calculada pelo IBGE.

^{6.} Definidos pelo grande grupo 6 da Classificação Brasileira de Ocupações (CBO-domiciliar). Disponível em: https://bit.ly/2PGbgmh.

RDD. Entendendo a aposentadoria como o tratamento, exploramos a legislação de aposentadoria por idade que determina a sua atribuição, de modo que descontinuidades nas variáveis de mercado de trabalho em torno dos pontos de descontinuidade podem ser interpretadas como o efeito causal da aposentadoria, sob algumas hipóteses discutidas em seguida.

Nesse contexto, utilizamos a idade como variável de alocação (*enforcing variable*). Conforme descrito na seção 2, em contextos urbanos, os pontos de descontinuidade (*cutoffs*) na variável de idade seriam de 60 anos para as mulheres e de 65 para os homens. Para trabalhadores rurais, os limites seriam de, respectivamente, 55 e 60 anos.

O fato de que os indivíduos não são obrigados a se afastarem do posto de trabalho com o pedido de aposentadoria pelo RGPS implica que a RDD é do tipo *fuzzy*, em que a regra de alocação aumenta a probabilidade de tratamento, porém não necessariamente o atribui para todos (a variação não é de zero para um).⁷ Desse modo, estima-se somente o efeito do tratamento sobre os indivíduos que aderem a ele (*compliers*), chamado efeito de tratamento médio local (*local average tratment effect* – Late).

Há duas hipóteses principais para a viabilização desse tipo de estimação. A primeira delas é de que se faz preciso que a variável de resultado seja uma função suave da variável de alocação. Apesar de não ser uma hipótese testável, no nosso caso é razoável supor que variáveis de mercado de trabalho não sofram saltos ao longo da distribuição de idade dos indivíduos, exceto em casos de aposentadoria compulsória para determinadas categorias.⁸

Um ponto importante sobre a identificação é de que a idade mínima exigida para as aposentadorias por idade dos homens trabalhadores urbanos, 65 anos, coincide com a idade mínima requerida para o benefício de prestação continuada (BPC), prevista na Lei Orgânica da Assistência Social (Loas). Segundo os dados do AEPS de 2015, foram concedidos 61.530 benefícios de amparo assistencial (BPC) para idosos e deficientes com idade entre 65 e 69 anos na data de início do benefício em 2014 (Brasil, 2016b). De acordo com a PNAD 2014, havia 212.767 indivíduos ocupados não aposentados na zona urbana com 64 anos (possíveis beneficiários da aposentadoria por idade). O AEPS não fornece dados de benefícios concedidos para idades exatas, o que seria importante para avaliar se há problema de identificação nas RDDs para os homens da zona urbana. Contudo, pode-se afirmar que os beneficiários do BPC podem enviesar os efeitos da aposentadoria nas variáveis relacionadas ao mercado de trabalho.

^{7.} Há somente três situações em que o trabalhador que se aposenta deve parar de trabalhar: quando a aposentadoria é por invalidez, quando o trabalhador é um servidor público estatutário, ou quando a aposentadoria é especial. Nesse último caso, o aposentado não pode seguir exercendo atividades que colocam em risco sua saúde ou integridade física.

8. A Lei Complementar nº 152/2015 alterou a idade para a aposentadoria compulsória dos segurados dos RPPS de 70 para 75 anos, como explicado em: https://bit.ly/2Z66K3o. Para os dados de 2014, portanto, não há problema de identificação.

A segunda hipótese é de que os agentes não possuem controle preciso sobre a variável de alocação. Como construímos a variável de idade com base nas informações de data de nascimento da PNAD, avaliamos que é razoável supor que ela não está sendo manipulada, ou seja, que os indivíduos não possuem incentivos para declarar incorretamente sua data de nascimento na pesquisa amostral. Adicionalmente, é possível testar essa hipótese de forma gráfica observando a distribuição dos indivíduos por idade.

Estimaremos o efeito da aposentadoria com o desenho fuzzy, por meio de uma regressão em dois estágios, em que instrumentalizamos a variável indicadora de aposentadoria (Apos) pela variável indicadora de o indivíduo ter idade maior do que o limite c, denotada por 1(Idade > c), definido pela legislação de maneira distinta para homens e mulheres, trabalhadores urbanos e rurais. O modelo a ser estimado é o seguinte:

$$Apos = \delta_0 + \delta_1 Idade + \delta_2 Idade \times 1(Idade > c) + \delta_3 1(Idade > c) + X\delta_4 + u. \tag{1}$$

$$Y = \beta_0 + \beta_1 \text{Idade} + \beta_2 \text{Idade} \times 1(\text{Idade} > c) + \beta_3 \widehat{\text{Apos}} + X\beta_4 + \varepsilon. \tag{2}$$

Em que Y é a variável de resultado do mercado de trabalho e o parâmetro de interesse, o efeito da aposentadoria sobre essa variável, é representado por β_3 . Nota-se que permitimos que a inclinação da regressão linear local seja diferenciada abaixo e acima no limite. Como robustez, empregamos adicionalmente polinômios de ordem mais alta nas regressões, até o quarto grau, sempre permitindo que seus parâmetros sejam diferenciados acima e abaixo do ponto de descontinuidade. Em nossas regressões, utilizamos a janela (*bandwidth*) ótima que minimiza o erro quadrático médio para regressões *fuzzy* (Calonico *et al.*, 2017).

4 RESULTADOS

4.1 Resultados descritivos

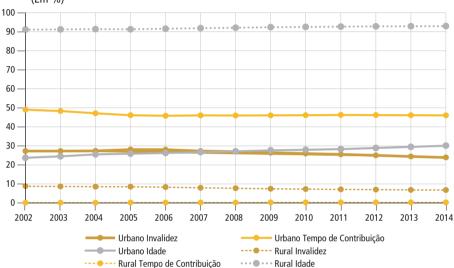
A PNAD não distingue o *status* de recebimento dos benefícios de aposentadoria por tipo (contribuição, idade ou invalidez), mas a previdência social fornece dados abertos, como o AEPS, que são utilizados neste trabalho. Esses dados complementam a análise sobre as aposentadorias precoces na medida em que fornecem estimativas sobre a quantidade e o valor dos benefícios concedidos por cada tipo de aposentadoria no país.

A composição das aposentadorias por tipo é muito diferenciada para trabalhadores rurais e urbanos. O gráfico 1 mostra que, entre os trabalhadores rurais, 90% das aposentadorias são por idade. As aposentadorias por tempo de

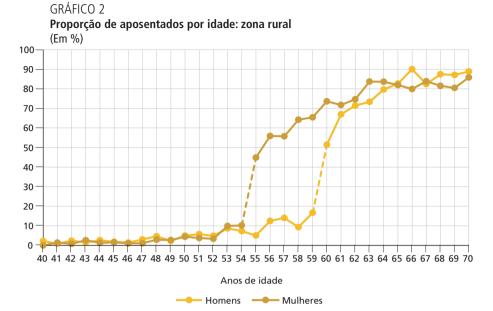
^{9.} Por exemplo, tomando a equação do primeiro estágio, poderíamos representá-la por: Apos $= \delta_0 + p_k(\text{Idade}, c) + \delta_3 1(\text{Idade} > c) + X\delta_4 + u$, em que o termo $p_k(\text{Idade}, c)$ seria o polinômio de ordem k da variável de idade, com a interação para a *dummy* relativa ao ponto de descontinuidade. Para k = 1, temos: $p_7(\text{Idade}, c) = \delta_1 \text{Idade} + \delta_2 \text{Idade} \times 1(\text{Idade} > c)$. Para k = 2, por exemplo, temos: $p_2(\text{Idade}, c) = \gamma_1 \text{Idade} + \gamma_2 \text{Idade} \times 1(\text{Idade} > c) + \gamma_3 \text{Idade}^2 \times \gamma_4 \text{Idade}^2 \times 1(\text{Idade} > c)$.

contribuição entre esses trabalhadores não chegam a somar 1% do total. No que se refere aos trabalhadores urbanos, predominam as aposentadorias por tempo de contribuição, que representam quase 50% do total. As restantes ficam divididas entre as aposentadorias por invalidez e por idade, sendo esse segundo tipo levemente mais presente na composição total a partir de 2008.

GRÁFICO 1 Distribuição de aposentadorias ativas por espécie e localidade (2002-2014) (Em %)



Fonte: Brasil (2016a). Elaboração dos autores.



Fonte: Dados da PNAD 2014. Disponível em: https://bit.ly/2EDRodc. Elaboração dos autores.

Os trabalhadores de áreas rurais, no geral, só irão se aposentar após o limite da idade mínima, de 55 anos para mulheres e de 60 para homens. No gráfico 2, fica clara a descontinuidade do número de aposentadorias nessas áreas. Nota-se que há uma expressiva mudança na proporção de aposentados a partir dos 55 anos para as mulheres e 60 anos para os homens, ponto de idade mínima para elegibilidade da aposentadoria por idade. Para as mulheres, enquanto que aos 54 anos há 10% de aposentadas, a proporção sobe para cerca de 45% aos 55 anos; já para os homens, ela salta de, aproximadamente, 16% para 51% entre os 59 e 60 anos. Aos 60 anos, em torno de 74% das mulheres e 51% dos homens estão aposentados na zona rural; já entre os indivíduos provenientes da zona urbana, cerca de 39% das mulheres e 36% dos homens estão aposentados aos 60 anos.

Os trabalhadores das áreas urbanas, por sua vez, se aposentam mais por tempo de contribuição e, por esse motivo, possuem uma distribuição mais suave de aposentados por idade. O gráfico 3 mostra que há uma quantidade elevada de indivíduos da zona urbana aposentados que ainda possuem idade igual ou inferior a 60 anos. Nessa faixa etária, 36% dos homens e 39% das mulheres já declaram receber benefícios da aposentadoria. O gráfico 3 também apresenta, de forma clara, uma proporção maior de homens que recebe esse tipo de benefício em comparação com as mulheres, para quase todas as idades e especialmente para as idades acima de 60 anos.

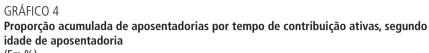


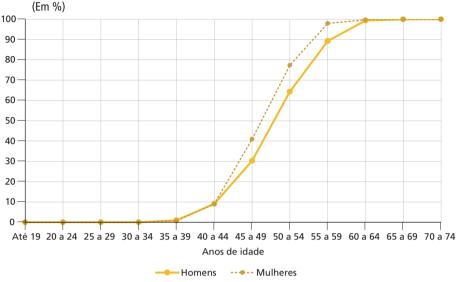
Fonte: Dados da PNAD 2014. Disponível em: https://bit.ly/2EDRodc. Elaboração dos autores.

10

Dessa forma, é possível que, entre os trabalhadores urbanos, haja uma proporção significativa de indivíduos recebendo o benefício da previdência social antes de perder sua capacidade laboral, o que não seria o objetivo do programa. É importante lembrar que os trabalhadores aposentados por invalidez, seja por doença ou acidente no trabalho, de fato não têm capacidade para seguir trabalhando.

As chamadas aposentadorias precoces seriam constituídas por benefícios recebidos por tempo de contribuição. Nesse sentido, o gráfico 4 aponta que 89% dos homens aposentados por tempo de contribuição receberam o benefício com idade igual ou inferior aos 59 anos (valores referentes ao estoque de aposentadorias ativas em 2014). Entre as mulheres, 77% se aposentaram com idade igual ou inferior aos 54 anos.





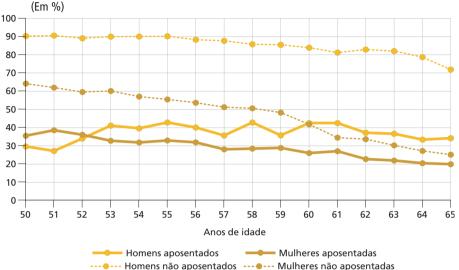
Fonte: Brasil (2016a).

As aposentadorias por tempo de contribuição também possuem, em média, valores maiores. O valor médio dos benefícios da ATC é de R\$ 1.681, valor 42% superior ao valor médio de todos os benefícios, que fica em R\$ 1.181, de acordo com o AEPS de 2014. Tendo em vista que o regime previdenciário já estabelece uma relação direta entre contribuições e valor dos benefícios, é evidente que, dado que esse tipo de aposentadoria possui valores médios relativamente altos, os indivíduos aposentados por tempo de contribuição possuem salários relativamente mais altos que os demais, em média.

Adicionalmente, estudos mostram que as aposentadorias por tempo de contribuição são, no geral, associadas a indivíduos com maior renda (Constanzi e Ansiliero, 2016; Nery, 2016). Adicionalmente, Caetano *et al.* (2016) indicam que 63% daqueles que se aposentaram precocemente (homens com menos de 60 anos e mulheres com menos de 55 anos, segundo os autores) pertencem aos quatro décimos mais altos da renda domiciliar *per capita* brasileira, ou seja, estão entre os 40% mais ricos.

Dessa maneira, acredita-se que a imposição de idade mínima para ATC afetaria, no futuro, os trabalhadores com maior inserção no mercado de trabalho formal, com mais escolaridade e em ocupações mais produtivas em comparação com os demais trabalhadores, potenciais beneficiários do RGPS.





Fonte: Dados da PNAD 2014. Disponível em: https://bit.ly/2EDRodc. Elaboração dos autores.

Uma questão que permanece em aberto, todavia, diz respeito ao impacto do *status* de recebimento do benefício da aposentadoria na situação dos aposentados no mercado de trabalho. O gráfico 5 demonstra que homens e mulheres aposentados têm taxas de ocupação expressivamente inferiores àqueles não aposentados para a mesma idade. Entre os 50 e 65 anos, os homens não aposentados apresentam taxas de ocupação bastante elevadas, com média superior a 85%. Entre os aposentados, a taxa de ocupação cai para menos da metade, 37%. Para as mulheres, a queda na taxa de ocupação é menos acentuada e percebe-se certa convergência entre aposentadas e não aposentadas em relação à taxa de ocupação. Fica clara a tendência declinante da ocupação das quatro categorias analisadas com o passar da idade, o que é intuitivo.

O resultado descritivo corrobora as hipóteses de Schwarzer e Paiva (2003) e Paiva, Rangel e Caetano (2016), indicando que aqueles que recebem proventos previdenciários têm uma probabilidade substancialmente menor de estarem empregados em comparação com os indivíduos da mesma idade não aposentados.

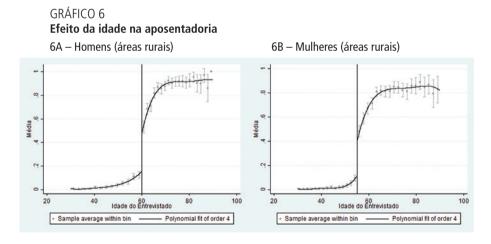
4.2 Resultados econométricos

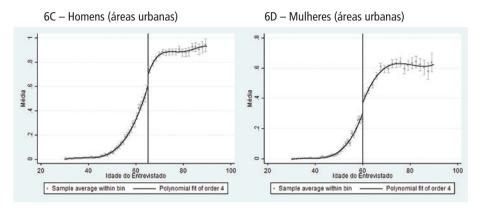
O intuito desta subseção de análise econométrica é avaliar possíveis efeitos da aposentadoria sobre variáveis relacionadas ao mercado de trabalho. Avaliaremos os impactos sobre a ocupação, a atividade, o salário, as horas trabalhadas e as horas

dedicadas a afazeres domésticos. Dado que a variável de aposentadoria é endógena, ou seja, é uma condição determinada pelos próprios indivíduos, exploramos a regra que determina a idade mínima para aposentadorias por idade para identificar os efeitos da aposentadoria sobre as variáveis de resultado. Nesse contexto, a estrutura dessa regressão pode ser representada por uma análise gráfica para fornecer uma compreensão do tamanho da descontinuidade, como fazemos nesta subseção.

Mostraremos na tabela 1 as regressões de primeiro estágio, que são ilustradas pelo gráfico 6. Nesse sentido, é possível observar que a condição de aposentadoria sofre um salto nos pontos representados pelas idades mínimas. Assim, constata-se que a regra desloca os indivíduos da condição de não aposentado para o estado de aposentado nos pontos de corte.

Nos painéis do gráfico 6, foram estimados polinômios de quarto grau, e o intervalo de confiança presente nos gráficos é de 95%. Para homens e mulheres de áreas rurais (respectivamente nos painéis 6A e 6B do gráfico 6), a descontinuidade é evidente. Em ambos os casos, a probabilidade de os indivíduos receberem benefícios da aposentadoria sofre um salto em torno da idade mínima, de mais de cerca de 30 pontos percentuais (p.p.).





Fonte: Dados da PNAD 2014. Disponível em: https://bit.ly/2EDRodc. Elaboração dos autores.

Obs.: 1. Gráficos de proporção de aposentados, condicional à idade, para faixas de seis meses. As linhas cheias representam regressões polinomiais de quarto grau. Os intervalos de confiança das probabilidades são de 95%.

 Figura reproduzida em baixa resolução e cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

Entre indivíduos das áreas urbanas (painéis 6C e 6D do gráfico 6), também se verificam descontinuidades em torno das idades mínimas definidas pelas regras do RGPS, porém com magnitude relativamente menor. A probabilidade de um homem receber proventos de aposentadoria salta de 60% para 73%. Entre as mulheres, o efeito parece ser semelhante, de cerca de 30% para 39%.

TABELA 1 Regressão de primeiro estágio

Variável dependente: aposentadoria					
	Homens (áreas rurais)	Mulheres (áreas rurais)	Homens (áreas urbanas)	Mulheres (áreas urbanas)	
F. 0. 10	0.361***	0.346***	0.130***	0.092***	
Estimativa	(0.034)	(0.027)	(0.019)	(0.014)	
Observações	2,526	3,593	9,827	17,693	
Bandwidth	5.13	6.89	5.05	5.79	
R-quadrado ajustado	0.334	0.383	0.131	0.090	
F	115.9	162.8	48.6	44.2	
p > F	0.000	0.000	0.000	0.000	

Fonte: Dados da PNAD 2014. Disponível em: https://bit.ly/2EDRodc.

Elaboração dos autores.

Obs.: Regressões estimadas sem controles adicionais, com *bandwidth* ótimo (Calonico *et al.*, 2017) e polinômio de primeira ordem. Erros padrão robustos entre parênteses. Significância: ***p<0.01, **p<0.05, *p<0.1.

As regressões de primeiro estágio apontam o expressivo efeito da regra de idade mínima sobre a condição de aposentadoria para os indivíduos provenientes de áreas rurais e urbanas, homens e mulheres. Os resultados indicam que as diferenças de probabilidades de aposentadoria decorrente da descontinuidade em torno dos critérios de idade mínima são de fato estatisticamente significantes. Esses resultados se alteram muito pouco quando repetimos o exercício com controles socioeconômicos e variáveis indicadoras da Unidade Federativa (UF), ou quando utilizamos polinômios até o quarto grau.

A tabela 2 apresenta os resultados das regressões de segundo estágio para os homens de áreas rurais. As variáveis dependentes são ocupação, atividade, salário e número de horas trabalhadas. Verifica-se que, ao nível de significância de 5%, as aposentadorias têm impacto sobre o mercado de trabalho no sentido de diminuir a probabilidade de os indivíduos permanecerem ocupados ou ativos. A proporção de ocupados é reduzida em 19 p.p., o que equivale a 21% da proporção de pessoas ocupadas no grupo de comparação, de pessoas cuja faixa etária se encontra entre a idade de aposentadoria e a diferença entre essa idade e o *bandwidth*. A redução na atividade, de 16 p.p., equivale a 18% do nível de atividade do grupo de controle.

Entre aqueles que continuam trabalhando, por sua vez, não parece haver grandes mudanças nas condições de trabalho. Atividades anteriores encontram associação entre a aposentadoria e as ocupações em trabalhos informais e com salários comparativamente menores (Queiroz e Ramalho, 2009; Moreno e Monte, 2013) e redução da produtividade do trabalhador (Paiva, Rangel e Caetano, 2016).

Há mais de um motivo para que o benefício da aposentadoria altere os salários e as condições de trabalho do indivíduo, em sentidos contraditórios. Por um lado, o rendimento da aposentadoria não é relacionado ao trabalho e aumentaria o salário de reserva do trabalhador, fazendo com que ele reduzisse a quantidade de trabalho ofertada. Por outro lado, os trabalhadores aposentados poderiam optar por ocupações informais que não deduzem a contribuição para a seguridade social, uma vez que já recebem o benefício. Apesar disso, os trabalhadores no setor formal podem fazer retiradas mensais do valor investido no Fundo de Garantia do Tempo de Serviço (FGTS), o que seria um incentivo para permanecer na ocupação formal. O rendimento do trabalho aumentaria em qualquer um desses casos.

A tabela 2 mostra que, entre os homens das zonas rurais, o salário mensal e a formalização não são impactados de maneira estatisticamente significante, enquanto há efeito negativo sobre as horas trabalhadas, embora seja somente marginalmente significante. Os resultados (não apresentados neste artigo) são muito semelhantes quando controlamos por escolaridade, raça, região (UF).

Variável dependente						
	Ocupação	Atividade	Salário	Horas trabalhadas	Formalização	
Estimativa	-0.187**	-0.157**	21.233	-5.286*	0.045	
	(0.076)	(0.077)	(343.884)	(2.971)	(0.070)	
Observações	2,526	2,477	2,976	2,901	2,929	
Randwidth	5 13	5.00	7 70	7 27	7 36	

TABELA 2
Regressões de segundo estágio: homens da zona rural

Fonte: Dados da PNAD 2014. Disponível em: https://bit.ly/2EDRodc.

Elaboração dos autores.

Obs.: Regressões estimadas sem controles adicionais, com bandwidth ótimo (Calonico et al., 2017) e polinômio de primeira ordem. Erros padrão robustos entre parênteses. Significância: ***p<0.01, **p<0.05, *p<0.1.

Com relação às regressões de segundo estágio para as mulheres da zona rural, a tabela 3 mostra que o efeito da aposentadoria não é significativo sobre nenhuma das variáveis analisadas. Isso ocorre apesar de haver forte efeito da regra sobre a condição de aposentado para as mulheres dessa zona, como observa-se no gráfico 1.

Para os indivíduos da zona urbana, as regressões de segundo estágio fornecem coeficientes significativos tanto para ocupação como para atividade. Para os homens, a condição de aposentado parece ter forte impacto no sentido de diminuir a probabilidade de os indivíduos permanecerem ocupados e ativos. Os resultados indicam que a condição de aposentado diminui em 44 p.p. a probabilidade de os homens da zona urbana permanecerem ocupados e ativos, o que equivale a uma redução de 74% em relação ao grupo de controle. Não há evidência, no entanto, de efeitos sobre as condições de trabalho entre aqueles que permanecem trabalhando.

TABELA 3
Regressões de segundo estágio: mulheres da zona rural

Variável dependente						
	Ocupação	Atividade	Salário	Horas trabalhadas	Formalização	
Estimativa	-0.126	-0.103	-133.607	-0.078	0.109	
	(0.094)	(0.083)	(212.493)	(3.660)	(0.091)	
Observações	3,593	4,049	2,232	2,455	1,973	
Bandwidth	6.89	7.79	6.87	7.52	5.98	

Fonte: Dados da PNAD 2014. Disponível em: https://bit.ly/2EDRodc.

Elaboração dos autores.

Obs.: Regressões estimadas sem controles adicionais, com bandwidth ótimo (Calonico et al., 2017) e polinômio de primeira ordem. Erros padrão robustos entre parênteses. Significância: ***p<0.01, **p<0.05, *p<0.1.

No caso dos homens de áreas urbanas, é importante lembrar que o efeito está sendo possivelmente superestimado, visto que o critério de aposentadoria por idade mínima coincide no caso desse grupo com o critério para o recebimento do BPC.

TABELA 4
Regressões de segundo estágio: homens da zona urbana

Variável dependente						
	Ocupação	Atividade	Salário	Horas trabalhadas	Formalização	
Estimativa	-0.440***	-0.438***	1444.605	1.331	-0.063	
	(0.142)	(0.144)	(2437.658)	(4.733)	(0.137)	
Observações	9,827	9,424	5,344	5,162	6,920	
Bandwidth	5.05	4.85	5.44	5.09	6.73	

Fonte: Dados da PNAD 2014. Disponível em: https://bit.ly/2EDRodc. Elaboração dos autores.

Obs.: Regressões estimadas sem controles adicionais, com bandwidth ótimo (Calonico *et al.*, 2017) e polinômio de primeira ordem. Erros padrão robustos entre parênteses. Significância: ***p<0.01, **p<0.05, *p<0.1.

Para as mulheres de áreas urbanas, os resultados revelam que há forte efeito da regra de aposentadoria sobre a ocupação e a atividade, sendo os coeficientes significativos a 1%. A probabilidade de as mulheres de áreas urbanas estarem ocupadas cai em cerca de 59 p.p., enquanto que, em relação à atividade, a probabilidade cai em torno de 63 p.p. Essas reduções são de magnitudes das taxas de ocupação e atividade no grupo de controle. Como no caso dos homens, as demais variáveis não apresentam efeitos significantes. Nota-se, todavia, que a estimativa pontual para o salário é relativamente grande, porém sua dispersão também é elevada. Nesse caso, a dispersão salarial após a idade mínima de aposentadoria aumenta muito e não permite uma inferência clara.

TABELA 5
Regressões de segundo estágio: mulheres da zona urbana

Variável dependente						
	Ocupação	Atividade	Salário	Horas trabalhadas	Formalização	
Estimativa	-0.590***	-0.627***	1537.292	-1.460	-0.281	
	(0.167)	(0.166)	(1331.836)	(8.183)	(0.247)	
Observações	17,693	18,202	8,917	7,067	8,855	
Bandwidth	5.79	5.96	7.36	5.79	7.13	

Fonte: Dados da PNAD 2014. Disponível em: https://bit.ly/2EDRodc.

Elaboração dos autores.

Obs.: Regressões estimadas sem controles adicionais, com bandwidth ótimo (Calonico et al., 2017) e polinômio de primeira ordem. Erros padrão robustos entre parênteses. Significância: ***p<0.01, **p<0.05, *p<0.1.

Os resultados obtidos sugerem que se verificam reduções na taxa de ocupação e de participação dos indivíduos, devido à aposentadoria. Esse resultado é esperado, visto que o rendimento da aposentadoria aumenta o salário de reserva dos indivíduos. No entanto, ao contrário de resultados obtidos por estudos anteriores (Queiroz e Ramalho, 2009), as aposentadorias parecem não exercer efeitos relevantes sobre os salários mensais, as horas trabalhadas e a formalização. As condições de trabalho daqueles que permanecem trabalhando parecem não se modificar de forma significativa, ao menos no curto prazo.

5 CONCLUSÃO

Neste artigo, examinamos os efeitos das aposentadorias sobre o mercado de trabalho. É essencial apontar a importância que o RGPS brasileiro tem como garantidor de renda para aqueles sem capacidade para trabalhar e como um mecanismo de distribuição de renda (Medeiros e Souza, 2013). Nesse contexto, uma possível reforma da previdência deve ser muito bem estudada e planejada antes de entrar em funcionamento. Possíveis mudanças não devem surtir efeito no curto prazo, dado que contribuintes têm direito aos benefícios como direito adquirido e estariam sujeitos a regras de transição, e, portanto, não há necessidade de absoluta urgência na imposição de novas regras para o regime previdenciário.

A princípio, procuramos levantar dados sobre a existência e a importância de aposentadorias precoces, de indivíduos que se aposentam antes de perderem sua capacidade laboral. Essas aposentadorias são possibilitadas pelas aposentadorias por tempo de contribuição. Entre os indivíduos aposentados por essa modalidade, 89% dos homens se aposentam com 59 anos ou menos e 77% das mulheres se aposentam com 54 anos ou menos, sendo que a caracterização desses aposentados precoces, proposta por Caetano *et al.* (2016) e Constanzi e Ansiliero (2016), sugere que eles pertencem aos estratos de renda mais altos e com maior nível de educação. Eles recebem valores de benefício 42% maiores do que a média, começam a receber o benefício relativamente mais jovens e, provavelmente, permanecem recebendo por mais tempo que a média da população.

Ademais, exploramos a descontinuidade colocada na regra da aposentadoria por idade mínima no RGPS e no RPPS para identificar o efeito das aposentadorias voluntárias sobre variáveis de ocupação e participação no mercado de trabalho, além de estudarmos os efeitos sobre variáveis de relativas às condições de trabalho daqueles que permanecem ocupados. Os resultados indicam que a condição de aposentadoria tem efeito significativo sobre a atividade e ocupação dos homens e mulheres, mais acentuados em áreas urbanas.

A probabilidade de os indivíduos dessas áreas estarem ocupados é de, aproximadamente, 50 p.p. para os aposentados. Com relação à atividade, a probabilidade cai em cerca de 46 p.p. para os homens e 56 p.p. para as mulheres. No caso dos

trabalhadores rurais, apesar do grande salto da probabilidade de aposentadoria em torno das idades mínimas, os efeitos são relativamente menores e não significantes para as mulheres. Uma vez que as taxas de ocupação são maiores entre homens e mulheres de áreas rurais em torno das idades mínimas, esse resultado indica que, apesar de se aposentarem, essa população permanece no mercado de trabalho.

Por fim, adicionalmente não encontramos efeitos significantes da aposentadoria sobre variáveis relacionadas ao salário e às condições de trabalho, em contraste com estudos anteriores. Em especial, os efeitos sobre a formalização do trabalho não são significantes em nenhum dos casos estudados.

A imposição de idade mínima para as aposentadorias nos moldes da PEC n° 287/2016 e da PEC n° 6/2019 afetaria mais fortemente os indivíduos provenientes das zonas urbanas (99,6% das ATC), mais escolarizados e com nível de renda relativamente alto. Dessa maneira, essa nova medida poderia manter esses indivíduos por mais tempo no mercado de trabalho. A extinção da ATC sem idade mínima pelo RGPS presente similarmente em ambas as PECs pode ser um fator importante para a sustentabilidade do regime a longo prazo e também para maior igualdade de tratamento entre os trabalhadores de maior rendimento e os mais pobres, que recebem benefícios de menor valor por tempo mais reduzido.

Em 2014, foram concedidos 1.150.880 benefícios de aposentadoria. Desconsiderando aqueles concedidos por invalidez, caso a proposta de idade mínima de 65 anos da PEC nº 287/2016 estivesse valendo em 2014, apenas 17,4% dos benefícios concedidos poderiam de fato ser admitidos, dado que 82,6% desses benefícios foram concedidos para indivíduos com idade inferior aos 65 anos. Isso significa que quase 800 mil indivíduos não teriam se aposentado em 2014 caso as regras propostas pela PEC nº 287/2016 estivessem em vigência. No contexto da PEC nº 6/2019, esses números seriam um pouco menores, uma vez que a idade mínima para as mulheres é de 62 anos.

REFERÊNCIAS

BRASIL. Constituição da República Federativa do Brasil. Brasília: Congresso Nacional, 1988. Disponível em: https://bit.ly/1bIJ9XW. Acesso em: 12 nov. 2019.

______. Ministério do Trabalho e Previdência Social. Instituto Nacional do Seguro Social. Empresa de Tecnologia e Informações da Previdência Social. Anuário Estatístico da Previdência Social — AEPS 2014. Brasília: Dataprev/MTPS, 2016a, v. 23.

_____. Ministério do Trabalho e Previdência Social. Instituto Nacional do Seguro Social. Empresa de Tecnologia e Informações da Previdência Social. Anuário Estatístico da Previdência Social — AEPS 2015. Brasília: Dataprev/MTPS, 2016b. v. 24.

_____. Ministério do Trabalho e Previdência Social. Instituto Nacional do Seguro Social. Empresa de Tecnologia e Informações da Previdência Social. **Anuário Estatístico da Previdência Social – AEPS 2016**. Brasília: Dataprev/MTPS, 2018. v. 23.

CAETANO, M. A.-R. *et al.* **O fim do fator previdenciário e a introdução da idade mínima**: questões para a previdência social no Brasil. Brasília: Ipea, 2016. (Texto para Discussão, n. 2230).

CALONICO, S. *et al.* Rdrobust: software for regression-discontinuity designs. **The Stata Journal**, v. 17, n. 2, p. 372-404, 2017.

CAMARANO, A. A. Perspectivas de crescimento da população brasileira e algumas implicações. *In*: CAMARANO, A. A. (Org.). **Novo regime demográfico**: uma nova relação entre população e desenvolvimento? Rio de Janeiro: Ipea, 2014. p. 177-210.

CAMARANO, A. A.; KANSO, S.; FERNANDES, D. Envelhecimento populacional, perda de capacidade laborativa e políticas públicas brasileiras entre 1992 e 2011. Rio de Janeiro: Ipea, 2013. (Texto para Discussão, n. 1890).

CONSTANZI, R. N.; ANSILIERO, G. As idades médias de aposentadoria urbana por Unidade da Federação e região. Brasília: Ipea, 2016. (Nota Técnica, n. 29).

EIBICH, P. Understanding the effect of retirement on health using regression discontinuity design. Berlim: SOEP, 2014. (SOEPpapers on Multidisciplinary Panel Data Research, n. 669).

FERNANDES, R. *et al.* Reforma da Previdência: sustentabilidade e justiça atuarial. **Estudos Econômicos**, v. 49, n. 3, 2019.

LEME, M. C. S.; MÁLAGA, T. Entrada e saída precoce da força de trabalho: incentivos do regime de previdência brasileiro. **Revista Brasileira de Economia**, Rio de Janeiro, v. 55, n. 2, p. 205-222, 2001.

LIBERATO, V. C. A oferta de trabalho masculina "pós-aposentadoria" Brasil urbano – 1981/2001. 2003. Dissertação (Mestrado) – Faculdade de Ciências Econômicas, Universidade Federal de Minas Gerais, Belo Horizonte, 2003.

MARTELLO, A. Guedes diz que intenção é de economizar R\$ 1 trilhão em dez anos com Reforma da Previdência. **G1**, Economia, 5 de fevereiro de 2019. Disponível em: https://glo.bo/35Gv52i. Acesso em: 18 abr. 2019.

MEDEIROS, M.; SOUZA, P. Gasto público, tributos e desigualdade de renda no Brasil. Brasília: Ipea, 2013. (Texto para Discussão, n. 1844).

MORENO, V. T.; MONTE, P. Avaliando o impacto dos rendimentos de aposentadoria no mercado de trabalho brasileiro. **Revista de Economia**, Curitiba, v. 39, n. 3, p. 59-76, 2013.

NERY, P. F. **Idade mínima**: perguntas e respostas. Brasília: Senado Federal, 2016. (Texto para Discussão, n. 190).

PAIVA, L. H.; RANGEL, L. A.; CAETANO, M. A.-R. O impacto das aposentadorias precoces na produção e na produtividade dos trabalhadores brasileiros. Rio de Janeiro: Ipea, 2016. (Texto para Discussão, n. 2211).

QUEIROZ, V. S.; RAMALHO, H. M. B. A escolha ocupacional dos idosos no mercado de trabalho: evidências para o Brasil. **Revista EconomiA**, Brasília, v. 10, n. 4, p. 817-848, 2009.

RANGEL, L. A. *et al.* Conquistas, desafios e perspectivas da previdência social no Brasil vinte anos após a promulgação da Constituição Federal de 1988. *In*: IPEA – INSTITUTO DE PESQUISA ECONÔMICA APLICADA. **Políticas Sociais**: acompanhamento e análise – vinte anos da constituição federal. Brasília: Ipea, 2009. v. 1.

SCHWARZER, H.; PAIVA, L. H. S. Participação de beneficiários e não beneficiários da previdência social no mercado de trabalho. **Informe de Previdência Social**, v. 15, n. 11, 2003.

TAFNER, P. Desafios e reformas da previdência social brasileira. **Revista USP**, São Paulo, n. 93, p. 137-156, 2012.

TAFNER, P.; BOTELHO, C.; ERBISTI, R. Transição demográfica e o impacto fiscal na previdência brasileira. *In*: CAMARANO, A. A. (Org.). **Novo regime demográfico**: uma nova relação entre população e desenvolvimento. Rio de Janeiro: Ipea, 2014. p. 539-570.

BIBLIOGRAFIA COMPLEMENTAR

CAETANO, M. A. Dinâmica fiscal da previdência social brasileira. *In*: CAMARANO, A. A. (Org.). **Novo regime demográfico**: uma nova relação entre população e desenvolvimento? Rio de Janeiro: Ipea, 2014. p. 574-589.

CAMARGO, J. M.; REIS, M. C. Aposentadoria, pressão salarial e desemprego por nível de qualificação. **Revista Brasileira de Economia**, v. 62, n. 3, p. 299-331, 2008.

CARD, D.; DOBKIN, C.; MAESTAS, N. The impact of nearly universal insurance coverage on health care utilization and health: evidence from Medicare. Cambridge, United States: National Bureau of Economic Research, 2004. (NBER Working Paper, n. 10365).

MARRI, I. G. **Reforma de previdência social**: simulações e impactos sobre os diferenciais de gênero. 2009. Tese (Doutorado) – Centro de Desenvolvimento e Planejamento Regional, Universidade Federal de Minas Gerais, Belo Horizonte, 2009. Disponível em: https://bit.ly/2qoCdRn. Acesso em: 25 out. 2016.

DECOMPOSIÇÃO QUANTÍLICA INCONDICIONAL DOS DIFERENCIAIS DE DESEMPENHO ENTRE ALUNOS DE ESCOLAS PRIVADAS E PÚBLICAS PROFISSIONALIZANTES¹

Ronaldo A. Arraes² Francisca Zilania Mariano³

Este artigo contribui com a literatura, ao trazer novas evidências sobre os diferenciais de desempenho entre os alunos de nível médio de escolas privadas regulares e escolas públicas; estas representadas pelas profissionalizantes que detêm os alunos com melhores resultados escolares. A aplicação empírica diferencia-se pela utilização de regressão quantílica incondicional (RQI) com aplicação da decomposição de Oaxaca-Blinder. Constatou-se haver superioridade dos alunos das escolas privadas em todas as áreas do Exame Nacional do Ensino Médio (Enem), principalmente em redação, cujos diferenciais de *performance* oscilaram entre 12% e 16% ao longo da distribuição. A decomposição detalhada dos efeitos indicou que escolaridade e renda dos pais se destacam para explicar tais diferenciais.

Palavras-chave: desempenho escolar; escolas públicas profissionalizantes; escolas privadas; decomposição quantílica incondicional.

UNCONDITIONAL QUANTILE DECOMPOSITION OF PERFORMANCE DIFFERENTIALS BETWEEN STUDENTS FROM PRIVATE AND VOCATIONAL PUBLIC SCHOOLS

The article contributes to the literature by bringing new evidences on the performance differentials between high school students of vocational public schools and regular private schools. The empirical application is distinguished by applying unconditional quantile regression with the Oaxaca-Blinder decomposition. Private schools' students were found to possess superior performance in all areas of the Enem examination, especially in writing, whose performance differentials ranged from 12% to 16% throughout the distribution. The decomposition of the effects indicated that schooling and parents' earnings are very relevant to explain such differentials.

Keywords: school performance; vocational public schools; private schools; unconditional quantile decomposition.

JEL: C31; I21; I25.

^{1.} Os autores agradecem à Secretaria de Educação do Estado do Ceará e ao Ministério do Trabalho (MTb) por disponibilizarem os dados utilizados na verificação empírica deste trabalho, e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo apoio.

^{2.} Coordenador do Programa de Economia Profissional (PEP) e professor do Programa de Pós-Graduação em Economia e do Departamento de Economia Aplicada da Universidade Federal do Ceará (Caen/UFC). *E-mail*: <ronald@ufc.br>.

^{3.} Professora do Departamento de Economia (campus de Sobral) e do Programa de Economia Profissional da UFC, e pesquisadora do Programa de Pós-Graduação em Economia dessa universidade. E-mail: <zilania@ufc.br>.

1 INTRODUÇÃO

Embora seja consenso que possuir uma boa educação básica seja uma das eficazes vertentes se alcançar o desenvolvimento de forma sustentável, essa ainda se apresenta de baixa qualidade e com alguns entraves, principalmente nos países em desenvolvimento como o Brasil. Muito se tem discutido sobre formas de melhorá-la nas diferentes fases que a compõe, quais sejam, ensino fundamental I e II e ensino médio. Esta última etapa é considerada uma das mais problemáticas, apresentando altos índices de evasão, repetência e baixo desempenho em testes padronizados, principalmente comparada ao ensino privado. Além disso, essa fase é marcada pela transição para o ensino superior e/ou para o mercado de trabalho, o que requer bastante atenção por partes dos gestores e políticos para que se possam desenvolver um ensino médio de qualidade e atrativo aos jovens, e que os capacitem para a escolha que desejarem tomar. Porém, muitos jovens não se sentem preparados a entrarem em alguma universidade e nem de adquirem um emprego de qualidade, levando, muitas vezes, à condição de nem estudar nem trabalhar. Isso eleva os índices de jovens ociosos em situação de vulnerabilidade social, contribuindo, assim, para o ciclo da pobreza e da desigualdade socioeconômica.

De acordo com informações do Ministério da Educação (MEC), o Brasil registrou melhorias no Índice de Desenvolvimento da Educação Básica (Ideb) para o ensino fundamental, ao contrário do índice do nível médio que permaneceu estagnado, considerando todas as redes de ensino em 2015.⁴ Apesar dessa evolução, apenas o ensino fundamental I conseguiu atingir a meta, superando-a em 0,3 pontos.

Estudos que tratam do diferencial de desempenho entre escolas públicas e privadas evidenciam uma superioridade do ensino particular, mesmo após controlar um conjunto de características pessoais e socioeconômicas, bem como a escolaridade dos pais de alunos (França e Gonçalves, 2010; Moraes e Belluzzo, 2014; Costa e Arraes, 2014).

É comum na literatura trabalhos que comparam os diferenciais de notas de testes padronizados entre alunos provenientes das escolas públicas e alunos advindos de escolas privadas. Buscando mensurar as diferenças de desempenho entre esses tipos de escolas na fase inicial do ensino fundamental, França e Gonçalves (2010) estimam o efeito do ensino privado sobre os resultados de alunos da 4ª série, a partir da aplicação do método dos mínimos quadrados ordinários (MQO) ponderados por *propensity score* aos dados do Sistema Nacional de Avaliação da Educação Básica (Saeb) de 2003. De acordo com os resultados encontrados pelos autores, há uma grande diferença de notas entre os dois setores de ensino, sendo que o desempenho acadêmico de alunos da rede privada é superior em 0,9 desviopadrão, aproximadamente, mesmo após controlar um conjunto de características socioeconômicas, escolares e regionais.

^{4.} Informações encontram-se no link disponível em: https://bit.ly/2ErJraR>. Acesso em: 24 set. 2016.

Costa e Arraes (2014) questionam os resultados de estudos que consideram as suposições de identificação da metodologia de pareamento no escore de propensão e regressões lineares na investigação da diferença de resultados entre as escolas públicas e privadas. Para tanto, utilizam o método de identificação parcial de Manski (1989), o qual permite estimar limites com suposições menos restritivas e verificar se as estimações pontuais das metodologias usuais não ultrapassam os limites do intervalo. A partir de informações do Saeb de 2005 para estudantes do 5º e 9º anos do ensino fundamental, os autores comprovaram que essas metodologias usuais sobre-estimam o efeito das escolas privadas, apesar de esse efeito permanecer positivo e significativo.

Moraes e Belluzzo (2014) utilizaram a decomposição por quantis de Melly (2005) para a distribuição condicional de notas, a partir da base de dados do Saeb de 2005 para os alunos do 9º ano do ensino fundamental. Os autores repararam que o diferencial é favorável às escolas privadas, em todos os quantis da distribuição, chegando a 1,5 erro-padrão em torno do percentil 600. Ademais, a análise do contrafactual indica que o desempenho das escolas particulares é relativamente menor na cauda inferior da distribuição.

No Brasil, entre os tipos de políticas públicas que buscam melhorar a qualidade de ensino público e reduzir a quantidade de crianças e adolescentes ociosos, encontra-se o ensino integral, o qual aumenta a carga horária dos alunos nas escolas e fornece outras atividades que sejam atrativas a permanência deles no ambiente escolar. Além desse tipo de ensino, existe o ensino profissionalizante, que busca preparar os jovens para a inserção no mercado de trabalho, por meio de cursos técnicos voltados para diversas áreas.

As escolas de educação de ensino profissional (EEEPs) levam em consideração esses dois tipos de ensino, integral e profissionalizante, e proporcionam ao jovem, durante o ensino médio, cursar as disciplinas básicas dessa etapa e participar de cursos técnicos e profissionalizantes voltados para diversas áreas (saúde, agricultura, computação, finanças etc.). Essa ação possibilita ao aluno concluir o ensino médio e se profissionalizar para o mercado de trabalho, além da alternativa de capacitá-lo para ingressar em universidades.

É, pois, centrado nessas políticas de ensino que este trabalho se desenvolve, com ênfase no ensino profissionalizante. Entretanto, cabe destacar que se optou por conduzir as análises desta pesquisa tendo o estado do Ceará como objeto de estudo, decorrente de sua representatividade e pertinência na geração das amostras para verificações empíricas. Ao considerar esses dois tipos de políticas, esse estado ganhou destaque em âmbito nacional, ao combinar o ensino integral com o profissionalizante e aumentar a participação de jovens no Exame Nacional de Ensino Médio (Enem).

Segundo informações disponibilizadas pelo Ipece (2013), no período 2009-2011, o Ceará apresentou a quarta maior participação de estudantes de escolas públicas estaduais no Enem em 2011. A taxa de participação desses estudantes passou de 2,78%, em 2009, para 8,89%, em 2011, perfazendo a terceira maior variação (219,78%) entre os estados brasileiros. Além disso, em 2011, a média das notas de escolas profissionalizantes cearenses no Enem (480.8) foi superior à média nacional (476.6), enquanto a média das escolas estaduais não profissionalizantes do estado foi de 442.5. Nesse sentido, se considerasse apenas a média dessas escolas para representar o Ceará, o estado ocuparia a oitava posição na classificação geral.

Depreende-se dos trabalhos supracitados que os alunos são influenciados por diversos fatores não controlados, tornando-os bastante heterogêneos, razão da forte explicação de fatores não observáveis. Ao contrário, este estudo busca reduzir a influência desses fatores, ao se construir grupos de alunos mais homogêneos para se aferir o diferencial de rendimento escolar líquido, notadamente entre os alunos provenientes de escolas públicas profissionalizantes e privadas no Ceará, a partir dos resultados destes no Enem de 2014.

Vale ressaltar que esses autores não encontraram trabalhos na literatura que analisem o diferencial de notas entre alunos do ensino médio das escolas públicas e privadas, impedindo que seja feita uma comparação. Cabe também destacar que, embora a análise aqui conduzida lide com apenas um estado da Federação, se observa, por meio do gráfico A.1, no apêndice A, que o desempenho dos alunos das escolas particulares do Ceará no Enem de 2014 possui distribuição semelhante aos de alunos de outros estados de diferentes regiões; por exemplo, Rio de Janeiro, Rio Grande do Norte, Pará, Santa Catarina e Goiás. Assim, há indícios que os resultados aqui encontrados podem ser generalizados para os demais estados.

A pertinência para tal escolha prende-se ao fato de as escolas de ensino profissional apresentarem desempenho superior às demais escolas públicas, uma vez que os estudantes ingressantes passam por um processo seletivo mais rigoroso; consequentemente se acredita que os melhores alunos provenientes do ensino fundamental são selecionados, além de possuírem estruturas mais adequadas e especializadas para proporcionar um ensino de melhor qualidade. Nesse sentido, pretende-se responder aos seguintes questionamentos: esse novo modelo de escola pública provê uma estrutura de ensino equiparável ao da escola privada? Como os alunos dos dois tipos de escola se comparam ao longo da distribuição de rendimento escolar?

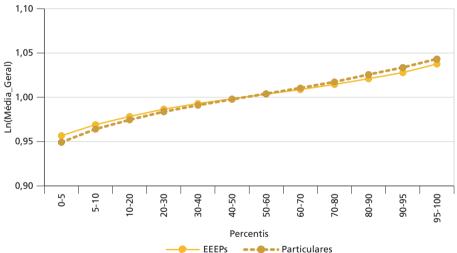
^{5.} Optou-se por comparar o Ceará com um estado de cada região para não expor o gráfico com excesso de informações e ilegível.

^{6.} Para comprovar essa afirmação, optou-se por encontrar o diferencial no desempenho no Enem dos alunos das escolas profissionalizantes e dos estudantes das escolas estaduais regulares (não profissionalizantes), para posteriormente fazer a comparação apenas das escolas particulares com as de ensino profissional.

Quais fatores observáveis mais influenciam os desempenhos dos alunos das duas escolas? Como os fatores não observáveis se comparam quanto ao diferencial de rendimentos entre os alunos?

Os gráficos 1 e 2 destacam o comportamento das notas na média geral (MG) no Enem e os diferenciais segundo decis da distribuição. No gráfico 1, é possível ter uma ideia da evolução das notas ao longo da curva. Uma visualização mais clara dos diferenciais de notas entre as escolas profissionalizantes e particulares é exposta no gráfico 2, em que se verifica uma diferença negativa e positiva nos decis inferiores e superiores, respectivamente. Dessa forma, a aplicação de um modelo quantílico se mostra mais adequada para analisar o diferencial no desempenho entre essas escolas.





Fonte: Microdados do Enem 2014. Disponível em: https://bit.ly/2QGmubm. Elaboração dos autores.

Tendo em vista que a aplicação de distribuição quantílica se torna apropriada, este trabalho buscou inovar ao utilizar o método de regressão quantílica incondicional (RQI) proposto por Firpo, Fortin e Lemieux (2007), o qual é considerada uma generalização do método de Oaxaca-Blinder, com o emprego de regressões baseadas na função de influência recentrada (FIR) combinadas com um procedimento de reponderação inspirado em DiNardo, Fortin e Lemieux (1996). Essa foi escolhida em detrimento do método tradicional de regressão quantílica (RQ) condicional proposto por Koenker e Bassett (1978), pois este não permite interpretar as estimativas como o impacto marginal da variável explicativa sobre a variável dependente (Firpo, Fortin e Lemieux, 2007; 2009).

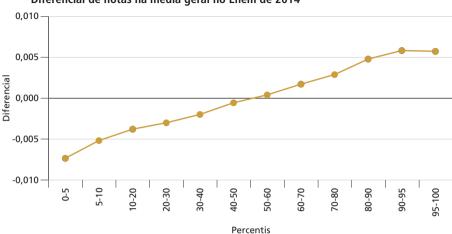


GRÁFICO 2
Diferencial de notas na média geral no Enem de 2014

Fonte: Microdados do Enem 2014. Disponível em: https://bit.ly/2QGmubm. Elaboração dos autores.

Em sequência, o texto está organizado em mais três seções, além desta introdução. A seção 2 aborda o método de decomposição de Firpo, Fortin e Lemieux (2007), já a seção 3 mostra os resultados do modelo. Por fim, na seção 4, apresentam-se as conclusões deste estudo.

2 MÉTODO DE DECOMPOSIÇÃO

Quando se pretende realizar uma decomposição do tipo Oaxaca-Blinder em diferentes quantis ao longo da distribuição, não se pode simplesmente executar RQ nos quantis de interesse e usar os coeficientes estimados no método Oaxaca-Blinder, pois, diferentemente dos obtidos via MQO, que possuem interpretação condicional e incondicional, os coeficientes estimados por regressões quantílicas têm apenas a interpretação condicional.

Sabe-se que o método MQO estima os coeficientes (β) de uma média condicional, $E[Y|X] = X\beta$, da qual se interpreta β como $\partial E[Y|X]/\partial X$. Aplicando a lei das expectativas iteradas, $E_X[E[Y|X]] = E[Y] = E[X]\beta$, β também pode ser explicado como o efeito das alterações na média de X na média incondicional de Y, embora essa lei não seja válida para quantis; ou seja, se aplicar a esperança condicional em quantil τ específico, $Q_{\tau}(X)$, não se encontra o quantil incondicional, $E_X[Q_{\tau}(X)] \neq Q_{\tau}$. Consequentemente, os coeficientes da regressão linear quantílica $Q_{\tau}(X) = X\beta$ têm somente interpretação condicional, de modo que não se pode interpretá-los como o efeito das alterações na média de X no quantil τth marginal (incondicional).

Dessa forma, RQ fornece uma abordagem dos impactos diferenciados de covariáveis ao longo da distribuição de um resultado; assim, esta é utilizada para avaliar o efeito de uma variável em um quantil do resultado condicional a valores específicos de outras variáveis. Entretanto, Borah e Basu (2013) afirmam que, na maioria dos casos, RQ pode gerar resultados que muitas vezes não são generalizáveis ou interpretáveis.

Na média, é suficiente estimar somente o valor médio condicional nos quantis de interesse, embora todo o contrafactual da distribuição acumulada incondicional $F_{Y_B^A}(Y)$ deve ser estimado, e um contrafactual quantil incondicional de interesse pode ser recuperado por inversão, $Q_{\tau,B}^A = F_{Y_B^A}^{-1}(\tau)$, no qual os subscritos A e B representam os dois grupos em comparação.

Vários procedimentos para a obtenção de distribuição contrafactual $F_{Y_B^A}(Y)$ têm sido sugeridos na literatura. Juhn, Murphy e Pierce (1993), Machado e Mata (2005) e Melly (2005) visam substituir cada resultado em B com um contrafactual Y_B^A . Juhn, Murphy e Pierce (1993) utilizam uma abordagem de imputação residual, enquanto Machado e Mata (2005) se baseiam em simulações e regressões de quantil condicional. DiNardo, Fortin e Lemieux (1996) propuseram uma abordagem semiparamétrica reponderada, que usa um fator de reponderação estimado de impor as características dos indivíduos do grupo B sobre os indivíduos do grupo A.

Existem também métodos que primeiro estimam a distribuição condicional $F_{Y_{A|X_A}}(Y|X)\|$ e, em seguida, integram-na sobre a distribuição dos indivíduos do grupo B, $F_{X_B}(X)$, para obter a distribuição contrafactual $F_{Y_B^A}(Y)$. Neles, são incluídos os métodos paramétricos de Donald, Green e Paarsch (2000) e Fortin e Lemieux (1998), bem como uma abordagem menos restritiva de Chernozhukov, Fernandez-Val e Melly (2013).

Se o objetivo for apenas realizar uma decomposição agregada do quantil – ou seja, estimar os efeitos explicáveis e não explicáveis da variável de resultado, sem o interesse de se fazer uma decomposição mais detalhada –, qualquer um desses métodos poderia ser utilizado; caso contrário, esses métodos seriam limitados e inapropriados. Apesar de alguns permitirem, os procedimentos variam em grau de complexidade, e a decomposição deve ser feita sequencialmente; assim, os resultados irão depender da ordem em que as características forem introduzidas. Firpo, Fortin e Lemieux (2007) propuseram uma técnica simples que permite fazer uma decomposição detalhada no quantil de interesse, cujos resultados independem da sequência. O método baseia-se na função de regressão de influência (função de influência – FI) ou RQI de Firpo, Fortin e Lemieux (2009)O método RQI tem sido aplicado em diversos trabalhos na literatura, tais como: Fortin (2008); Edoka (2012); Le e Booth (2014); Borah e Basu (2013); Rubil (2013); Fisher e

Marchand (2014); Hirsch e Winters (2014); Kassenboehmer e Sinning (2014); Morin (2015); e Nguyen (2015). Segundo Firpo, Fortin e Lemieux (2007), existem duas formas de se obter o efeito de uma covariável no quantil incondicional. A primeira delas seria usar as estimativas dos coeficientes da RQ para recuperar o impacto incondicional, que, apesar de intuitivamente simples, muitas vezes se torna intratável. Especificamente, mostra-se que o efeito parcial de uma covariável em um quantil incondicional de *Y* pode ser escrito como uma média ponderada – por meio da distribuição de *X* – do efeito parcial sobre um quantil condicional específico de *Y*, que corresponde ao quantil incondicional de interesse. Portanto, se conseguir mapear todos os quantis incondicionais de *Y* para os quantis condicionais correspondentes sob diferentes argumentos condicionado, então tal abordagem ponderada pode ser facilmente implementada. No entanto, é evidente que isso é uma tarefa árdua, pois requer técnicas não paramétricas, tornando-a intratável.

Uma alternativa para a abordagem mencionada anteriormente foi proposta por Machado e Mata (2005), que assumiram uma mudança na distribuição incondicional ao longo do tempo e a decompuseram em componentes que são atribuíveis a mudanças na distribuição marginal de diferentes Xs. Contudo, esse método captura o efeito total de uma alteração na distribuição marginal de X sobre todos os quantis incondicionais de Y, mas não apenas em um quantil específico.

A segunda forma de se obter o efeito de uma covariável no quantil incondicional é solucionar o problema de intratabilidade e superar a limitação do modelo condicional. Firpo, Fortin e Lemieux (2007) sugerem um modelo RQI baseado nos conceitos de FI e na FIR, tal como utilizado na robusta literatura estatística (Hampel *et al.*, 1986). Uma FI é uma ferramenta analítica que pode ser usada para avaliar o efeito (influência) de uma remoção/adição de uma observação sobre o valor de uma estatística, v(F), sem ter de recalcular a estatística, e é definida por:

$$(y, v(F)) = \lim_{\varepsilon \to 0} \{ \frac{v[(1-\varepsilon).F + \varepsilon. \delta_y] - v(F)}{\varepsilon} \}, 0 \le \varepsilon \le 1,$$
 (1)

em que F representa a função de distribuição acumulada de Ye δ_{y} é uma distribuição que é agregada no valor y.

Uma FIR é obtida adicionando a estatística v(F) a FI:

$$FIR(y,v) = v(F) + FI(y,v).$$
(2)

Uma característica dessa abordagem é que a expectativa é igual a v(F) – ou seja, se a estatística de interesse for a média, então a FI será o residual avaliado em y e a FIR será o próprio valor de y:

$$FI(y,\mu) = \lim_{\varepsilon \to 0} \left\{ \frac{(1-\varepsilon) \cdot \mu + \varepsilon \cdot y - \mu}{\varepsilon} \right\} = y - \mu, e \, FIR(y,\mu) = \mu + y - \mu = y \,. \tag{3}$$

Por conseguinte, a regressão da FIR para a média, em X, produz os mesmos coeficientes da regressão por MQO.

Quando a estatística de interesse é um quantil τ específico da distribuição, tem-se:

$$FI(y, q_{\tau})$$

$$=\frac{\tau - I\{Y \le q_{\tau}\}}{f_{\mathcal{Y}}(q_{\tau})},\tag{4}$$

em que q_{τ} se refere ao quantil τth da distribuição incondicional de Y, $f_{Y}(q_{\tau})$ é a função densidade de probabilidade de Y avaliada nesse quantil (q_{τ}) e I $\{Y \leq q_{\tau}\}$ é uma variável indicadora que delimita a ocorrência de Y até o limite do quantil. Por definição,

$$FIR(y, q_{\tau}) = q_{\tau} + FI(y, q_{\tau}).$$
(5)

Quando a esperança condicional da $FIR(y,q_{\tau})$ é modelada como uma função das variáveis explicativas, $X,E[FIR(y,q_{\tau})|X=x]=m_{\tau}(x)$, uma regressão FIR pode ser vista como uma RQI (Firpo, Fortin e Lemieux, 2009). Por definição da FIR, $E_X E[FIR(y,q_{\tau})|X]=q_{\tau}$; então, $E_X (dm_{\tau}(x)/dX)$ pode ser interpretado como efeito marginal de uma variação na distribuição das covariadas no τ quantil incondicional de Y, mantendo tudo o mais constante.

Firpo, Fortin e Lemieux (2007) afirmam que os procedimentos para estimação da RQI são simples e semelhantes à regressão MQO e se dão da seguinte forma: primeiro, para um quantil τ específico, estima-se a FIR do quantil τ de Y seguindo (3.2.4) e (3.2.5). q_{τ} é estimado usando a estimativa da amostra do quantil τ incondicional, assim como a densidade $f_{\mathcal{Y}}(q_{\tau})$ no ponto q_{τ} é estimada usando o método de Kernel. Segundo, aplica-se MQO na $FIR(y,q_{\tau})$ sobre as covariadas observadas, X. Os autores também destacam as medidas para calcular o efeito quantílico incondicional, que mede o impacto de uma covariável sobre Y em um quantil específico, no qual, na regressão FIR-MQO, se pressupõe que os efeitos parciais quantílicos incondicionais são os coeficientes estimados.⁷

Dessa forma, enquanto a regressão quantílica condicional provê estimativas do retorno de características individuais, *coeteris paribus*, em que este retorno é variável entre os indivíduos de acordo com o quantil incondicional ao qual ele pertence, a regressão quantílica incondicional estima, *coeteris paribus*, o efeito

^{7.} Podem ser implementados pelo software Stata a partir do comando rifreg. Disponível em: https://bit.ly/34DnPU5.

de pequenas mudanças em uma característica dos indivíduos em cada quantil da distribuição, permitindo, assim, avaliar o efeito sobre diversas estatísticas da distribuição (Fournier e Koske, 2012).

Uma característica do método RQI é que os resultados podem ser diretamente aplicados ao modelo de decomposição de Oaxaca-Blinder,8 o qual será utilizado por este trabalho para verificar os fatores que contribuem para os diferenciais de notas no Enem entre os alunos das escolas públicas profissionalizantes e os estudantes das escolas privadas do Ceará ao longo de toda a distribuição, definidos genericamente como A e B, aplicando-se um método de decomposição proposto por Firpo, Fortin e Lemieux (2007). Generalizando-se a decomposição de Oaxaca-Blinder, pode-se representar a diferença nas distribuições de notas entre os dois tipos de escolas por meio de alguma estatística dessas distribuições. Assim, o diferencial total entre as distribuições é escrito como:

$$\Delta^{\nu} = \nu(F_{yA}) - \nu(F_{yB}) \tag{6}$$

em que $v(F_{yr})$ representa uma estatística da distribuição de notas para a escola r=A, B. A decomposição irá dividir Δ^v em um componente associado a diferenças de características observáveis dos indivíduos e um componente associado a diferenças na estrutura de notas entre os dois tipos de escolas. Para tanto, faz-se necessária a construção de um contrafactual (F_{yc}) , para que se possa simular uma distribuição de notas, compondo a estrutura de notas da escola A, e de características – observáveis e não observáveis – dos alunos da escola B. Construindo-se esse contafactual e algumas manipulações algébricas, pode-se obter:

$$\Delta^{v} = [v(F_{yB}) - v(F_{yC}) + [v(F_{yC}) - v(F_{yA})]$$
(7)

$$\Delta^{\nu} = \Delta_{S}^{\nu} + \Delta_{X}^{\nu} \tag{8}$$

O primeiro termo constitui a *parte não explicada* da decomposição (efeito coeficiente (retorno) ou discriminação quando se refere a mercado de trabalho, definido como uma medida da diferença entre os coeficientes estimados das covariadas para o grupo *B* em relação ao *A*). O último termo deste diferencial total é o efeito quantidade, considerado a *parte explicada* da decomposição (efeito dotação ou característico, o qual se refere à diferença nas dotações dos dois grupos, em que a diferença de nota entre os alunos das escolas públicas profissionalizantes perante os das privadas pode ser parcialmente atribuída ao fato de possuírem características mais favoráveis para atingir notas mais elevadas).

^{8.} A regressão FIR quantílica pode ser usada na decomposição de Oaxaca através do comando oaxaca8 no *software* Stata.

Aplicando-se o valor esperado das funções de influência recentrada e assumindose uma especificação linear, têm-se as equações (9) e (10) para as distribuições observadas e para a distribuição contrafactual, respectivamente.

$$v(F_{yB}) = E[FIR(y_r; v_r)|X, R = r] = X_r \beta_r \quad para \quad r = A, B_{e} \quad (9)$$

$$v(F_{yC}) = E[FIR(y_A; v_C)|X, R = B]$$

$$= X_C \beta_C \quad (10)$$

Ao aplicar a decomposição de Oaxaca-Blinder sob a suposição de linearidade, a expectativa da estatística contrafactual em que se adota a distribuição de características da escola B e a estrutura de notas da escola A é estimada de forma que $\bar{X}_C \beta_C = \bar{X}_B \beta_A$. Entretanto, quando a linearidade não se mantém, essa igualdade passa a não ser válida.

A partir de Barsky *et al.* (2002), Firpo, Fortin e Lemieux (2007) propõem uma abordagem que combina o método de regressões FIR com reponderações inspiradas na técnica de DiNardo, Fortin e Lemieux (1996). Esse método estabelece o uso de um fator de reponderação, o qual pode ser definido por:

$$\varphi(X) = \frac{\Pr[R = B|X]}{\Pr[R = A|X]} \cdot \frac{\Pr[R = A]}{\Pr[R = B]},\tag{11}$$

em que $\Pr[R = r | X]$ é a probabilidade de um aluno pertencer à escola r; dadas suas características, X e $\Pr[R = r]$ denotam a proporção de alunos na escola r. Vale ressaltar que esse fator de ponderação pode ser estimado através de um modelo de probabilidade para $\Pr[R = B | X]$, usando as probabilidades preditas para calcular o valor $\widehat{\varphi(X)}$ para cada observação:

$$\widehat{\beta_r} = (\sum_{i \in r} \widehat{w_r}. X_i. X_i')^{-1}. \sum_{i \in r} \widehat{w_r}. FIR(\widehat{y_{r_i}}; v_r). X_i, para \ rA, B,$$
(12)

em que $\widehat{W_r}$ é um fator de ponderação baseado na própria composição da amostra total. Para o contrafactual, estima-se:

$$\widehat{\beta_C} = \left(\sum_{i \in A} \widehat{\varphi(X_i)}.X_i.X_i'\right)^{-1}.\sum_{i \in A} \widehat{\varphi(X_i)}.FIR(\widehat{y_{Ai}}; v_C).X_i$$
(13)

$$\overline{X_c} = \sum_{i \in A} \widehat{\varphi(X_i)} . X_i$$
 (14)

A decomposição é dada por:

$$\hat{\Delta}^{v} = \left[\overline{X_B} \hat{\beta}_B - \overline{X_C} \hat{\beta}_C \right] + \left[\overline{X_C} \hat{\beta}_C - \overline{X_A} \hat{\beta}_A \right]_{e}$$
(15)

$$\hat{\Delta}^{\nu} = \hat{\Delta}_{S}^{\nu} + \hat{\Delta}_{X}^{\nu} \,. \tag{16}$$

A estimativa do efeito composição $\hat{\Delta}_S^v$ pode ser dividida em dois componentes em que um termo padrão $(\bar{X}_C - \bar{X}_A)\hat{\beta}_A$ é somado a um erro de especificação $\bar{X}_C(\hat{\beta}_C - \hat{\beta}_A)$, de forma que:

$$\hat{\Delta}_X^{\nu} = (\bar{X}_C - \bar{X}_A)\hat{\beta}_A + \bar{X}_C(\hat{\beta}_C - \hat{\beta}_A). \tag{17}$$

O erro de especificação tende a ser nulo, se a especificação linear do modelo for correta, uma vez que $plim(\hat{\beta}_C) = plim(\hat{\beta}_A)$. Na decomposição detalhada, em que é possível estimar a contribuição de cada variável explicativa, o efeito composição pode ser escrito como:

$$\widehat{\Delta}_X^{\nu} = \sum_{k=1}^K (\bar{X}_{Ck} - \bar{X}_{Ak}) \widehat{\beta}_A \,. \tag{18}$$

O efeito estrutural das notas passa a ser definido por:

$$\hat{\Delta}_S^{\nu} = \bar{X}_B (\hat{\beta}_B - \hat{\beta}_C) + (\bar{X}_B - \bar{X}_C) \hat{\beta}_C \,. \tag{19}$$

Esse efeito se reduz ao primeiro termo, visto que, em grandes amostras e com a estimação correta do fator de ponderação, o erro de reponderação tende a ser nulo. Destaca-se que o método de regressões quantílicas incondicionais fornece uma estimativa de equilíbrio parcial da variável de interesse, assumindo que mudanças marginais em X não exercem impacto sobre a distribuição conjunta de X e y. Além disso, o método também reconhece a forte independência entre heterogeneidade não observada e as características observadas. De acordo com Fournier e Koske (2012), embora essas suposições não se sustentem na prática, elas argumentam que uma comparação entre estimativas de diferentes quantis permanece válida, desde que o viés potencial seja igual em toda a população da amostra.

2.1 Base de dados

As informações referentes ao desempenho nas cinco áreas de conhecimento e na média geral do Enem de 2014, no tocante às condições socioeconômicas dos alunos das EEEPs da rede estadual de educação e das escolas privadas do Ceará, foram retiradas do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). A identificação das escolas profissionalizantes e de suas dotações de infraestrutura de todas as escolas ocorreu por meio de dados da Secretaria da Educação do Estado do Ceará (Seduc) e do Censo Escolar de 2014, respectivamente.

A amostra utilizada neste trabalho é constituída apenas pelos alunos com as seguintes características: estão cursando o ensino médio das escolas de ensino

^{9.} Último ano com informações disponíveis sobre o desempenho individual no Enem.

profissionalizante e privado do estado; 10 concluíram o ensino médio em 2014; 11 estiveram presentes nos dias de provas do Enem do referido ano; e suas escolas possuem informações disponibilizadas pelo Censo Escolar de 2014. Desse modo, obteve-se um total de 20.947 alunos, sendo 8.955 da rede estadual de ensino profissional e 11.992 das escolas privadas.

2.2 Variáveis do modelo

As variáveis, descritas no quadro 1, foram escolhidas com o intuito de captar os diferenciais nas notas das cinco áreas e na média geral do Enem, compostas por vetores com as características pessoais dos alunos, o *status* educacional dos pais e as condições de infraestrutura das escolas.

QUADRO 1 Descrições das variáveis

Variáveis	Descrição
	notas logaritimizadas do aluno em cada uma das áreas: ciências da natureza (CN); ciências humanas (CH); ódigos (LC); matemática (MT); redação (R); e média geral (MG).
Explicativas	
Alunos	
Priv	1, se o aluno pertence à escola privada; 0, se o aluno pertence à EEEP
Idade	Anos de idade
Sexo	1, se masculino; 0, caso contrário (c.c.)
Raça	1, se branca; 0, c.c.
Estado civil	1, se solteiro; 0, c.c.
Anos_Fund	1, se concluiu o ensino fundamental em oito anos; 0, c.c.
Anos_méd	1, se concluiu o ensino médio em três anos; 0, c.c.
Escolaridade d	lo pai
E_Pai1	1, se o pai não estudou; 0, c.c.
E_Pai2	1, se o pai possui de escolaridade o ensino fundamental incompleto ou completo; 0, c.c.
E_Pai3	1, se o pai possui de escolaridade o ensino médio incompleto ou completo; 0, c.c.
E_Pai4	1, se o pai possui de escolaridade o ensino superior incompleto ou completo; 0, c.c
Escolaridade d	la mãe
E_mae1	1, se a mãe não estudou; 0, c.c.
E_mae2	1, se a mãe possui de escolaridade o ensino fundamental incompleto ou completo; 0, c.c.
E_mae3	1, se a mãe possui de escolaridade o ensino médio incompleto ou completo; 0, c.c.
E_mae4	1, se a mãe possui de escolaridade o ensino superior incompleto ou completo; 0, c.c.

(Continua)

^{10.} Foram excluídos os alunos das escolas estaduais de ensino regular, municipais e federais.

^{11.} Foram excluídos os alunos que já concluíram o ensino médio ou que não irão concluir em 2014.

(Continu	

Variáveis	Descrição								
Condições socioeconômicas									
Renda00	1, se a família não possui renda; 0 ,c.c.								
Renda01	1, se a família possui até 1 salário mínimo (SM); 0, c.c.								
Renda02	1, se a família possui de 1 até 2 SMs; 0, c.c.								
Renda03	1, se a família possui de 2 até 5 SMs; 0, c.c.								
Renda04	1, se a família possui de 5 até 10 SMs; 0, c.c.								
Renda05	1, se a família possui de 10 até 15 SMs; 0, c.c.								
CSE ¹	Índice de condição socioeconômica								
Escolas									
Infra ²	Índice de infraestrutura da escola								

Notas: ¹Índice criado a partir da análise de componente principal, que engloba a existência das seguintes características dos alunos obtidas pelo Enem de 2013: se possui televisão, DVD, rádio, microcomputador, automóvel, máquina de lavar, geladeira, *freezer*, telefone fixo, celular, acesso à internet, televisão por assinatura, aspirador de pó, empregada e banheiro.
²Índice gerado pela análise de componentes principais, que engloba a existência das seguintes características das escolas obtidas no Censo Escolar de 2013: abastecimento de água por rede pública; esgoto sanitário por rede pública; coleta periódica de lixo; sala de diretoria; sala de professores; laboratório de informática e de ciências; quadra de esportes; cozinha; biblioteca; parque infantil; sanitário adequado a alunos portadores de necessidades especiais (PNEs); dependências e vias adequadas a PNEs; televisão; videocassete; DVD; parabólica; copiadora; retroprojetor; impressora; computadores; acesso à internet; e alimentação escolar para os alunos.

A variável que reflete a renda familiar está disponível em intervalos de SMs; nesse contexto, optou-se por fazer um corte de dois salários, tendo em vista que são pais de alunos de escolas públicas, pressupondo-se, assim, não possuírem renda familiar muito elevada. Para facilitar a interpretação da medida dos índices CSE e Infra, as variáveis foram normalizadas, de forma a assumir valores entre 0 a 1; quanto maior for o índice, mais favorável será a disposição de equipamentos.

3 RESULTADOS

3.1 Análise descritiva

Estatísticas descritivas das notas nas cinco áreas e na média geral do Enem e das variáveis de controles relacionadas aos alunos de EEEPs (públicas profissionalizante), regulares (públicas não profissionalizantes) e privadas do Ceará em 2014 estão explicitadas na tabela 1. Observa-se que as médias das notas dos alunos das escolas de ensino privado são superiores às dos alunos de escolas profissionalizantes, e estas são maiores que as médias dos alunos das escolares regulares, em todas as provas. Além disso, as demais características dos alunos das EEEPs se aproximam mais das características dos alunos das escolas particulares. Embora seja apenas a média, esses resultados corroboram a hipótese deste trabalho de que os alunos das escolas profissionalizantes apresentam desempenho superior aos alunos das escolares regulares, e, ao utilizar esse grupo para representar as escolas públicas, isso tornará os grupos mais homogêneos. Fazendo a comparação das escolas de

ensino profissional com as privadas, verifica-se que, na área de ciências da natureza, o diferencial de médias é de 52.24 pontos a favor das escolas privadas, embora estas apresentem uma variação relativa à média 21,7% (0,1572/0,1291) superior à EEEP. As áreas de ciências humanas e linguagens e códigos revelam diferenças semelhantes, 37,39 e 32,67 pontos, respectivamente, com dispersão relativa à média de 4% e 3% superior à EEEP, respectivamente.

Apresentando diferenciais mais elevados e menos homogêneos estão as áreas de matemática e redação, com diferença de 71,92 e 83,97 pontos, respectivamente, e variação relativa semelhante nos dois tipos de escolas, e de 23,5% em matemática, superior à EEEP. Na média geral, os alunos das escolas particulares apresentaram desempenho superior em 55,6 pontos, 11% a mais que os alunos das escolares de ensino profissionalizante, apesar de estas deterem menor coeficiente de variação.

TABELA 1
Estatísticas descritivas das variáveis da amostra (2014)

		Regulares			EEEPs			Privadas	
Variáveis	Média	DP1	CV ²	Média	DP	CV	Média	DP	CV
NOTA_CN	455.52	57.316	0.1258	493.58	63.749	0.1291	545.82	85.8326	0.1572
NOTA_CH	509.365	67.124	0.1318	556.17	63.8829	0.1148	593.56	70.9712	0.1195
NOTA_LC	479.3558	64.26	0.1341	519.56	58.945	0.1134	552.23	64.4928	0.1167
NOTA_MT	434.7047	76.527	0.1760	486.98	95.5583	0.1962	558.9	135.454	0.2423
NOTA_R	436.2812	144.381	0.3309	549.3	139.2176	0.2534	633.27	158.1479	0.2497
NOTA_MG	463.0465	62.199	0.1343	521.12	64.5028	0.1237	576.76	85.6248	0.1484
Alunos									
Idade	18.86	4.43	0.2349	17.467	0.8507	20.532	17.643	2.6064	6.7691
Sexo	0.445	0.4865	1.0933	0.424	0.4942	1.1655	0.4654	0.4988	1.0716
Raça	0.1677	0.3736	2.2278	0.1934	0.3949	2.0422	0.4237	0.4941	1.1661
Estado civil	0.9545	0.2083	0.2182	0.9946	0.073	0.0734	0.9908	0.0953	0.0962
Anos_Fund	0.3663	0.4818	1.3153	0.3794	0.4852	1.2788	0.459	0.4983	1.0855
Anos_Méd	0.3271	0.4691	1.4341	0.357	0.4791	1.3421	0.527	0.4992	0.9473
Escolaridade	do pai								
E_Pai1	0.1206	0.3257	2.7007	0.0692	0.2538	3.6667	0.0149	0.1212	8.124
E_Pai2	0.663	0.4726	0.7128	0.5786	0.4937	0.8533	0.2317	0.4219	1.8208
E_Pai3	0.1844	0.3878	2.1030	0.2944	0.4558	1.5479	0.394	0.4886	1.2399
E_Pai4	0.0277	0.1641	5.9242	0.0501	0.2182	4.3527	0.2655	0.4416	1.6629
Escolaridade	da mãe								
E_mae1	0.0685	0.2527	3.6891	0.0279	0.1647	5.9011	0.0068	0.0824	12.0522
E_mae2	0.6507	0.4767	0.7326	0.4943	0.4999	1.0113	0.1571	0.3639	2.3156
E_mae3	0.2204	0.4145	1.8807	0.3563	0.4789	1.344	0.3922	0.4882	1.2447
E_mae4	0.0466	0.2108	4.5236	0.0904	0.2868	3.1712	0.2952	0.4561	1.5449

(Continua)

		~ \
// on	tinua	ranl
(CUII	unua	Lau,

		Regulares			EEEPs		Privadas			
Variáveis	Média	DP1	CV ²	Média	DP	CV	Média	DP	CV	
Condições so	ocioeconômica	S								
Renda00	0.0306	0.1723	5.6307	0.0155	0.1236	7.9643	0.0049	0.0699	14.2223	
Renda01	0.6899	0.4625	0.6704	0.6391	0.4802	0.7513	0.1323	0.3388	2.5606	
Renda02	0.2279	0.4195	1.8407	0.2797	0.4488	1.6047	0.2635	0.4405	1.6718	
Renda03	0.0469	0.2115	4.5096	0.0599	0.2374	3.9595	0.299	0.4578	1.5311	
Renda04	0.0039	0.0624	16.000	0.0052	0.0722	13.767	0.1752	0.3802	2.1692	
Renda05	0.0003	0.0178	59.333	0.0002	0.0149	66.922	0.0617	0.2407	3.8967	
CSE	0.3777	0.1513	0.4006	0.4806	0.1461	0.304	0.674	0.1847	0.274	
Escolas										
Infra	0.2247	0.0637	0.2835	0.3547	0.0903	0.2546	0.4144	0.149	0.3597	

Notas: ¹ DP – desvio-padrão.

² CV – coeficiente de variação.

Obs.: 1. A proporção de alunos com pai que possui pós-graduação (E_pai5) foi 0.007 para as profissionalizantes e 0.093 para as privadas, perfazendo 100% guando somada às demais.

- 2. A proporção de alunos com mãe que possui pós-graduação (E mae5) foi 0.030 para as profissionalizantes e 0.148 para as privadas, perfazendo 100% quando somada às demais.
- 3. A proporção de alunos com renda familiar maior que 15 SMs (renda_6) foi 0.0001 para as profissionalizantes e 0.0631 para as privadas, perfazendo 100% quando somada às demais.

Os gráficos A.2 a A.7, no apêndice A, mostram a distribuição das proficiências dos alunos nas cinco áreas e na média geral para as escolas particulares, profissionalizantes e regulares. Por meio do gráfico A.7, pode-se notar que a curva para os alunos do ensino privado se encontra mais concentrada para a direita, indicando que a frequência de alunos com notas maiores é superior aos alunos pertencentes ao ensino profissionalizante, e estes estão melhor distribuídos que os alunos das escolas regulares.

Referente às demais variáveis presentes na tabela 1, observa-se que, apesar de apresentar idade média semelhante, 17 anos, o grupo das escolas privadas é menos homogêneo ao das escolas profissionalizantes, pois apresenta variação absoluta de, aproximadamente, dois anos, enquanto, no outro grupo, o desvio-padrão é de meses. Além disso, a proporção de homens e alunos solteiros é semelhante: nas EEEPs e nas escolas privadas, 42% e 46% dos alunos são do sexo masculino, respectivamente, e, em ambos os grupos, 99% são solteiros. Quando se refere à raça, a diferença entre os percentuais aumenta: 19,3% dos estudantes são considerados brancos nas escolas profissionalizantes e 42,3%, nas particulares.

As variáveis *Anos_Fund* e *Anos_Médio* apresentam a proporção de alunos que terminaram o ensino fundamental em oito anos e o médio em três anos, os quais, para as escolas profissionalizantes, foram de 38% e 36%, e, para as particulares, 46% e 53%, respectivamente. Isso mostra indícios que os alunos que fizeram o

3º ano do ensino médio nas escolas profissionalizantes apresentam problemas de repetência maiores do que as escolas particulares.

Ao se analisar as escolaridades do pai e da mãe, percebe-se que os alunos das escolas privadas apresentam pais com níveis mais elevados de estudo. Para as EEEPs, a maior proporção reside em pai e mãe com fundamental incompleto ou completo, 57,86% e 49,43%, respectivamente, enquanto, para as escolas regulares não profissionalizantes, pais com ensino médio incompleto ou completo são maioria, aproximadamente 39%. Ademais, a proporção de pais com ensino superior nas escolas de ensino privado é maior do que no grupo das profissionalizantes, aproximadamente 20 pontos percentuais (p.p.) a mais.

Através das variáveis de condições socioeconômicas, observa-se que a renda familiar dos alunos das escolas profissionalizantes se concentra em até 2 SMs, ao passo que, nas escolas privadas, a concentração se dá a partir de 3 SMs. Além disso, os alunos destas escolas apresentam um índice CSE médio mais elevado do que o observado para as EEEPs, 0.6740 e 0.4806, respectivamente. Apesar de as EEEPs apresentarem boa infraestrutura (Ceará, 2014), as escolas particulares detêm uma melhor oferta de equipamentos aos alunos, conforme atestam os respectivos índices, com valores de 0.3547 e 0.4144.

3.2 Regressão FIR nos quantis incondicionais

Para decompor os diferenciais de notas entre os alunos das escolas regulares e profissionalizantes e entre os alunos destas e das privadas por meio da metodologia de Firpo, Fortin e Lemieux (2007), faz-se necessário estimar as regressões FIR as quais têm o objetivo de estimar efeitos das covariáveis sobre os quantis incondicionais do *log* das notas das provas do Enem, cujos resultados para os quantis 0.1, 0.5 e 0.9 se encontram nas tabelas do apêndice B. Além disso, essas tabelas apresentam as estimativas por MQO. Pode-se observar que os efeitos não são constantes ao longo da distribuição, justificando, assim, a análise por regressões quantílicas, ¹² além de diferirem entre os tipos de escolas (regulares, profissionalizantes e particulares).

A análise desses resultados se concentra na comparação das escolas profissionalizantes e particulares. Os coeficientes estimados para a variável *idade* foram negativos e significativos em todos os quantis analisados e áreas de conhecimentos para ambos os tipos de ensino. Isso indica que quanto maior a idade do aluno, menor será seu desempenho no Enem, independentemente da parte da distribuição na qual o aluno se encontre. Tal efeito decresce no sentido da parte superior da distribuição.

^{12.} O teste de Koenker e Bassett (1982) mostrou que os coeficientes são estatisticamente diferentes, apresentando uma estatística F igual a 25.58.

Os resultados mostram que o efeito gênero difere entre os tipos de escolas, ao longo da distribuição e entre as áreas de conhecimento. Na média geral, nas áreas de ciências humanas, nas ciências da natureza e em matemática, considerando-se a mediana e o quantil 0.9, pode-se observar que o efeito é positivo e significativo em ambas as escolas, embora com efeito maior nas profissionalizantes, enquanto, no quantil 0.1, o efeito difere entre as áreas e entre as escolas. Na média geral, este foi insignificante para as EEEPs e negativo significativo para as escolas particulares. Na área de ciências humanas, o efeito foi positivo e significativo para as profissionalizantes e insignificante para as privadas. Já nas áreas de ciências da natureza e matemática, ambos foram positivos e significativos. Apresentando resultados negativos e significativos em todos os quantis analisados e nos dois tipos de escolas, encontra-se a área de redação; por sua vez, em linguagens e códigos, esse tipo de efeito se sobressai apenas para as escolas particulares, na mediana e no quantil 0.9.

Referente à raça, verificou-se que, no primeiro quantil, o efeito é predominantemente insignificante, exceto em redação, o qual foi positivo significativo para os tipos de escolas. Na mediana, podem-se observar, na média geral e nas áreas de linguagens e códigos, matemática e redação, efeitos insignificantes para as EEEPs e positivos significativos para as privadas, enquanto nas áreas de ciências humanas e ciências da natureza os efeitos foram positivos e significantes nas EEEPs; nas escolas particulares, foram insignificantes e positivos significativos, respectivamente. No quantil 0.9, esse efeito se mostrou positivo e significativo em todas as áreas para as escolas de ensino privado e nas áreas de ciências humanas e linguagens e códigos para as EEEPs.

Além desses resultados, nota-se que alunos cujo estado civil é solteiro não apresentam efeito significativo sobre as notas do Enem no primeiro quantil da distribuição, na mediana, exceto em linguagens e códigos para as EEEPs e em matemática para as escolas particulares. Esse também é o caso referente ao quantil 0.9, nas áreas de ciências humanas e linguagens e códigos para ensino privado, bem como em todas as áreas nas EEEPs, exceto em ciências humanas. Diferentemente da variável que indica se o aluno concluiu o ensino fundamental em oito anos, cujo efeito foi positivo e significativo em todos os quantis analisados, áreas de conhecimentos e escolas, exceto no primeiro quantil de matemática para o ensino profissionalizante, cujo efeito foi insignificante. Já para a variável *Anos_medio*, o efeito foi insignificante para as EEEPs, em todas as áreas e quantis, exceto no quantil 0.1, em redação.

Referente às *dummies* de escolaridade dos pais dos alunos de ensino profissionalizante, observa-se que todos os efeitos foram insignificantes, nas partes inferiores e superiores da distribuição de notas, em todas as áreas, e na mediana, exceto na variável que indica se o pai tem ensino médio incompleto

ou completo para as áreas de conhecimento de média geral e de ciências da natureza, e nas variáveis E_pai2 e E_pai4 em ciências da natureza, cujos efeitos foram positivos. Porém, nas escolas privadas, verificam-se, predominantemente, efeitos negativos e significativos, exceto para as variáveis E_pai1 , nos quantis 0.9 da média geral e 0.1 das ciências humanas, das ciências da natureza e de matemática, e E_pai_3 , na mediana em linguagens e códigos e no quantil 0.1 de linguagens e códigos, enquanto, para E_pai4 , o efeito se mostrou, na maioria dos casos, insignificante.

Resultado semelhante pôde ser constatado para as escolaridades das mães dos alunos das escolas profissionalizantes, com exceção da média geral para E_mae2 e E_mae3 , no quantil 0.1, e da mediana, para Emae3 na média geral e E_mae2 nas ciências da natureza. Já no ensino privado, os efeitos variaram entre insignificantes e negativos significativos, tanto por quantil quanto por área de conhecimento.

Quanto ao efeito das faixas de renda sobre as variações nas notas do Enem, verifica-se que, a partir da média geral, para as escolas profissionalizantes, todos os efeitos foram positivos e significativos, exceto na mediana para renda de 10 a 15 SMs e no quantil 0.9, para famílias que não possuem renda e para aquelas que possuem de 5 a 15 SMs. Nas escolas particulares, todavia, o efeito foi predominantemente negativo e significativo, com exceção de algumas faixas de renda no primeiro quantil e na mediana.

Ademais, efeitos insignificantes puderam ser observados ao longo da distribuição para o índice de condição socioeconômica e de infraestrutura das escolas, considerando a média geral para os alunos das escolas profissionalizantes, enquanto, nas escolas privadas, esse efeito foi positivo e significativo em todos os quantis analisados.

3.3 Decomposição da regressão FIR: efeito composição e estrutural

A decomposição tradicional de Oaxaca-Blinder e quantílica irá identificar quais fatores e efeitos composição (característicos) ou estrutural (coeficientes) determinam os diferenciais de notas por tipo de escola (regulares e profissionalizantes; privadas e EEEPs), na média e nos quantis da distribuição, utilizando as estimativas de MQO e FIR, respectivamente. Para tanto, esse procedimento foi estimado para os dezoito quantis estabelecidos, 0.05 a 0.95, cujos diferenciais de notas regulares e profissionalizantes se encontram nos gráficos A.8 a A.13, no apêndice A; os diferenciais entre as escolas privadas e as EEEPs, juntamente com os efeitos característicos e coeficientes nos quantis, são ilustrados através dos gráficos A.2 a A.7, embora somente alguns quantis foram selecionados e expostos na tabela 2, somados aos resultados da decomposição na média. Em todas as notas, as estimativas dos

diferenciais mostraram-se significativas a 1%; assim, é possível verificar vantagens dos alunos das escolares particulares em relação aos alunos do ensino profissional, bem como destas sobre os alunos das escolas regulares.¹³

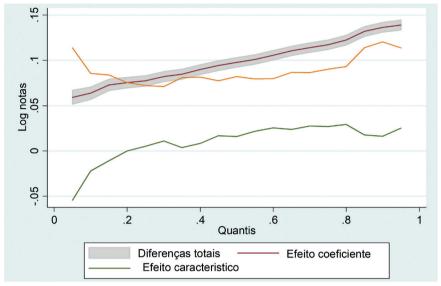
A partir da tabela 2, em matemática, nas ciências da natureza e na média geral, o aumento da diferença de notas na parte superior da distribuição indica o efeito teto de vidro (*glassceiling*). Verifica-se que o diferencial total entre as notas dos alunos das escolas profissionalizantes e privadas na média geral é monotonamente crescente nos quantis. No quantil 0.1, a diferença é de 0.0637, atingindo 0.0978 na mediana e 0.1365 no quantil 0.9 — ou seja, 6,57%, 10,27% e 14,62% a mais para os alunos das escolas privadas, respectivamente. Dessa forma, pode-se dizer que, em média, a diferença está aumentando à medida que se move a quantis mais elevados da distribuição — isto é, o diferencial de notas a favor dos alunos das escolas particulares é maior para aqueles que apresentam desempenhos superiores.

Interpretação semelhante pode ser utilizada para as áreas de ciências da natureza e matemática, cujo diferencial total também é positivo e crescente. Porém, observa-se que, em matemática, o hiato cresce com o quantil da distribuição. Nessa área, o quantil 0.10 apresenta uma diferença de 5,51% e atinge 21,55% no quantil 0.9, enquanto, em ciências da natureza, essas diferenças são de 6,2% e 13,63%, respectivamente.

Em ciências humanas e em linguagens e códigos, apesar de apresentarem diferenciais positivos, estes praticamente permanecem constantes ao longo da distribuição, com leves oscilações. No 10º, 50º e 90º quantis, as diferenças são 0.0698, 0.0602 e 0.0767, respectivamente, para ciências humanas, e 0.0606, 0.0637 e 0.0636, respectivamente, para linguagens e códigos. Apresentando diferenciais mais elevados ao longo de toda a distribuição, encontra-se a área de redação, na qual há a oscilação de 0.1504, equivalente a 16,22%, no quantil 0.1, para 0.1135 (12,01%) na mediana e 0.1503 (16,21%) no quantil 0.9.

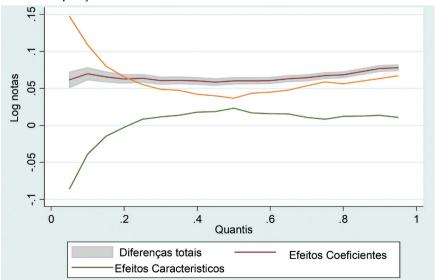
^{13.} Ver apêndice A.

GRÁFICO 3 **Decomposição dos efeitos: MG**



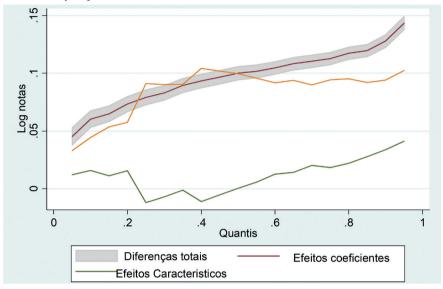
Obs.: Gráfico cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

GRÁFICO 4 **Decomposição dos efeitos: CH**



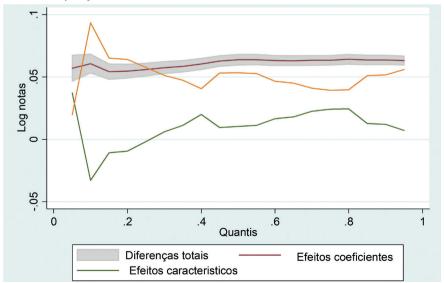
Elaboração dos autores.

GRÁFICO 5 Decomposição dos efeitos: CN



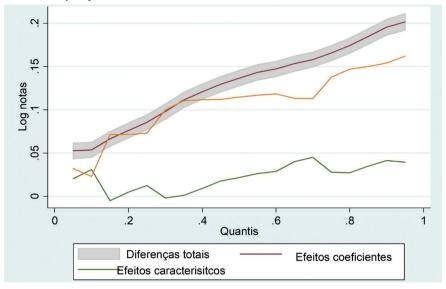
Obs.: Gráfico cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

GRÁFICO 6 Decomposição dos efeitos: LC



Elaboração dos autores.

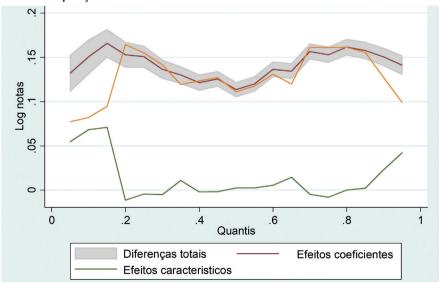
GRÁFICO 7 **Decomposição dos efeitos: MT**



Obs.: Gráfico cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

GRÁFICO 8

Decomposição dos efeitos: R



Elaboração dos autores.

Os resultados dos efeitos característicos para a média geral e a área das ciências humanas apresentam sinais negativos e significantes nos primeiros quantis (0.05, 0.10 e 0.15) da distribuição, nos quantis 0.10 e 0.15 de linguagens e códigos e no percentil 0.25 das ciências da natureza, indicando que estes agem no sentido de diminuir as diferenças entre as notas. Isso implica que as diferenças de notas entre as escolas em características observáveis preveem vantagens aos estudantes das EEEPs nessas áreas e na parte inferior da distribuição, com destaque para o quantil 0.1 de ciências humanas e de linguagens e códigos, que agem reduzindo o diferencial em 56% e 54%, respectivamente, como pode ser observado na tabela 2 e nos gráficos 3 a 8.

Dessa forma, os alunos das escolas profissionalizantes apresentam características produtivas melhores que os alunos das escolas particulares, e o efeito positivo para o efeito estrutural sugere que os retornos das características foram maiores para os alunos das escolas particulares. Com isso, nesses quantis, o que explica os diferenciais de notas são os fatores não observáveis, pois os alunos das EEEPs apresentaram atributos observáveis melhor do que os alunos das escolas privadas. Portanto, o sinal positivo da ação dos efeitos estruturais foi suficiente para anular as características (efeito composição).

Os efeitos composição nos quantis 0.2, 0.25, 0.35 e 0.5 da média geral, 0.2 e 0.25 das ciências humanas, 0.2, 0.25, 0.3, 0.45, 0.5 e 0.55 de linguagens e códigos, 0.15 a 0.6 de matemática, 0.15, 0.3 a 0.65 e 0.75 das ciências da natureza e 0.2 a 0.85 da redação não foram estatisticamente significantes. Para os demais, os efeitos composição foram positivos e significantes, atuando de forma a contribuir com o diferencial de notas — ou seja, os alunos das escolas privadas apresentam melhores características que ajudam a obter um bom desempenho, principalmente nos quantis mais elevados da distribuição, pois, em todas as áreas de conhecimento e na média geral, esses efeitos foram positivos e significativos nos percentis 0.9 e 0.95. Já nos menores quantis de matemática e de redação, observam-se os maiores efeitos, 57% e 45%, respectivamente, com destaque para o primeiro, o qual superou o efeito estrutural em contribuir para o aumento do diferencial de notas.

Contudo, os efeitos não observáveis (coeficientes) são todos positivos e significativos em todos os quantis da distribuição e em todas as áreas do Enem, sendo responsáveis por explicar a maior parte do diferencial das notas. Em matemática, o efeito da variável é crescente ao longo da curva; diferentemente das demais áreas, cujos efeitos oscilam entre os percentis. Esses resultados indicam que os alunos das escolas privadas são mais capazes de converter insumos educacionais em maiores pontuações em testes padronizados e que existem outros fatores não observáveis pertencentes a estes alunos que contribuem para melhores desempenhos.

TABELA 2

Decomposição agregada na distribuição das notas nas áreas do Enem

Quantil	MQO	%	0.1	%	0.25	%	0.5	%	0.75	%	0.9	%
	MG											
Diferencial	0.0980*		0.0637*	100	0.0774*	100	0.0978*	100	0.1173*	100	0.1365*	100
Características	0.0833*		-0.022*	-35	0.0053	7	0.0158	16	0.0270**	23	0.0161**	12
Coeficientes	0.0146*		0.0857*	135	0.0720*	93	0.0820*	84	0.0903*	77	0.1204*	88
							СН					
Diferencial	0.0644*		0.0698*	100	0.0635*	100	0.0602*	100	0.0672*	100	0.0767*	100
Características	0.0574*		-0.039**	-56	0.0084	-	0.0232*	39	0.0083**	12	0.0135*	18
Coeficientes	0.0070*		0.1088*	156	0.0551*	87	0.0369*	61	0.0588*	88	0.0632*	82
							CN					
Diferencial	0.0965*		0.0602*	100	0.0791*	100	0.1000*	100	0.1127*	100	0.1278*	100
Características	0.0790*		0.0158*	26	-0.0120*	-15	0.0004	-	0.0184	-	0.0337***	26
Coeficientes	0.0174*		0.0443*	74	0.0912*	115	0.0996*	100	0.0943*	84	0.0940*	74
							MT					
Diferencial	0.1273*		0.0537*	100	0.0855*	100	0.1365*	100	0.1655*	100	0.1952*	100
Características	0.1245*		0.0308**	57	0.0126	-	0.0216	-	0.0278*	17	0.0414*	21
Coeficientes	0.0028*		0.0228*	43	0.0728*	85	0.1148*	84	0.1377*	83	0.1538*	79
							LC					
Diferencial	0.0605*		0.0606*	100	0.0558*	100	0.0637*	100	0.0633*	100	0.0636*	100
Características	0.0496*		-0.0328*	-54	-0.0018	-	0.0104	-	0.0240**	38	0.0120*	19
Coeficientes	0.0108*		0.0934*	154	0.0577*	103	0.0533*	84	0.0392*	62	0.0516*	81
							R					
Diferencial	0.1429*		0.1504*	100	0.1506*	100	0.1135*	100	0.1528*	100	0.1503*	100
Características	0.1090*		0.0683*	45	-0.0044	-	0.0024	-	-0.0079	-	0.0230*	15
Coeficientes	0.0338*		0.0821*	55	0.1550*	103	0.1111*	98	0.1608*	105	0.1273*	85

Obs.: Significância (*) = 1%; (**) = 5%; e (***) = 10%.

Os dois efeitos, característico e estrutural, podem ser decompostos de forma detalhada, como mostra a tabela 3 e os gráficos 9 e 10 para a média geral.

Variávai-		C	aracterístico			Estrutural				
Variáveis	0.1	0.25	0.5	0.75	0.9	0.1	0.25	0.5	0.75	0.9
Total	-0.0222*	0.0053	0.0158	0.0270**	0.0161**	0.0857*	0.0720*	0.0820*	0.0903*	0.1204*
Idade	-0.0061*	-0.0053*	-0.0052*	-0.0037*	-0.0022*	0.2099*	0.2913*	0.3843*	0.3092	0.1951*
Sexo	0.0002	0.0007*	0.0012*	0.0012*	0.0011*	-0.0077*	-0.0093*	-0.0094*	-0.0075*	-0.007*
Raça	0.0012	0.0021**	0.0012	0.0027*	0.0003	-0.0002	-0.0018	0.0014	0.0002*	0.0052**
Estado civil	-0.001	-0.0002	-0.0001	-0.0000	-0.0000	-0.0855	-0.0452	-0.0325	-0.0473**	-0.0641*
Anos_Fund	0.0014*	0.0018*	0.0003*	0.0027*	0.0011*	0.0043	0.0048**	0.0055*	0.0088*	0.0042**
Anos_ Médio	0.0007	-0.0001	0.0001	0.0015	-0.0009	0.0026	0.0080*	0.0029	0.0012	0.0107*
E_Pai1	0.0012	0.0001	-0.0006	0.0005	-0.0002	-0.0006	-0.0007	-0.0011*	-0.0008**	-0.0003
E_Pai2	0.0026	-0.0032	-0.0078	-0.0022	-0.0038	-0.0084	-0.0135**	-0.0163*	-0.0150*	-0.0115***
E_Pai3	0.0006	0.0028	0.0042**	0.0016	0.0016	-0.0138	-0.0209**	-0.0290*	-0.0256*	-0.0202**
E_Pai4	0.0009	0.0040	0.0075	0.0034	0.0029	-0.0042	-0.0073	-0.0118**	-0.0088	-0.0045
E_mae1	-0.0002	0.0002	-0.0001	0.0002	0.0004	-0.0004	-0.0002	-0.0000	-0.0000	-0.0000
E_mae2	-0.0078***	-0.0001	-0.0019	0.0066	0.0071	-0.0097*	-0.0056*	-0.0061*	-0.0021	-0.0013
E_mae3	0.0009***	0.0001	0.0007***	0.0000	-0.0000	-0.0130**	-0.0065	-0.0172*	-0.0109**	-0.0098
E_mae4	0.0037	0.0005	0.0031	-0.0009	-0.0002	-0.0068	0.0002	-0.0057	0.0002	-0.0002
Renda00	-0.0066*	-0.0030*	-0.0011*	-0.0005*	-0.0001	-0.0035*	-0.0017*	-0.0010	-0.0008*	-0.0005*
Renda01	-0.3294*	-0.01503*	-0.0666*	-0.0382*	-0.0176*	-0.0965*	-0.0516*	-0.0320*	-0.0263*	-0.0175*
Renda02	-0.0110*	-0.0054*	-0.0027*	-0.0017*	-0.0011*	-0.1909*	-0.1041*	-0.0685*	-0.0609*	-0.0447*
Renda03	0.1634*	0.0810*	0.0435*	0.0318*	0.0205*	-0.2054*	-0.1092*	-0.0693*	-0.0655*	-0.0499*
Renda04	0.1166*	0.0523*	0.0296*	0.0127*	0.0086	-0.1217*	-0.0554*	-0.0334*	-0.0216*	-0.0154*
Renda05	0.0475*	0.0277*	0.0105	0.0123	-0.0003	-0.0479*	-0.0280*	-0.0110	-0.0144	-0.0015
CSE	-0.0037	-0.0022	-0.0010	-0.0008	-0.0008	0.0382**	0.0326**	0.0546*	0.0555*	0.0563*
Infra	0.0016	0.0014	-0.0006	0.0005	-0.0000	0.0419*	0.0567*	0.0802*	0.0445*	0.0137
Constante	-	-	=	-	-	0.6056*	0.1398***	-0.1019***	-0.0212	0.0842

TABELA 3

Decomposição dos efeitos característico e retorno: MG

Obs.: Significância (*) =1%; (**) = 5%; (***) = 10%.

Na parte inferior da distribuição, no quantil 0.1, o hiato de notas foi reduzido devido a características referentes a idade, escolaridade da mãe e faixas de rendas para os que não possuem renda e os que possuem até 2 SMs, com destaque para a variável que indica se o aluno possui renda mensal domiciliar até 1 SM, cujo efeito foi responsável por tornar o efeito composição negativo. Resultado semelhante pode ser observado para os demais quantis, embora a característica do aluno ter mãe com o ensino fundamental completo ou incompleto não afete mais, no sentido de reduzir os diferenciais de notas e na parte superior da distribuição, quantil 0.9. Efeito semelhante é constatado ao se considerar família destituída de renda.

Por sua vez, a variável *gênero* age de forma a contribuir para o aumento do diferencial, positiva e significativa, exceto no quantil 0.1, cuja maior contribuição se encontra no quantil 0.25, aproximadamente 13. Em partes da distribuição, a raça também possui esse efeito, aumentando os diferenciais em 39,62 e 10, nos quantis 0.25 e 0.75, assim como as escolaridades da mãe e do pai, nas quais o efeito sobre o hiato no quantil 0.5 foi de 4,43 e 26,58, respectivamente.

Ao longo de toda a distribuição, a variável que indica se o aluno terminou o ensino fundamental em oito anos apresentou sinal positivo, contribuindo para o aumento das disparidades de notas entre os alunos das profissionalizantes e privadas, cujo efeito foi maior no quantil 0.25, aproximadamente 40. Desse modo, ter uma boa trajetória no ensino fundamental afeta o desempenho no Enem, favorecendo os alunos das escolas particulares. Além disso, possuir renda a partir de 2 SMs, para alguns percentis, também age de forma a aumentar as diferenças de notas, cujo maior efeito pode ser observado no quantil 0.25, nas três faixas de renda. Com isso, entre as variáveis que contribuem para elevar o hiato de notas na média geral do Enem, pertencer a elevadas faixas de rendas é o principal fator.

Como se pôde observar na tabela 2, a parte do diferencial atribuído a fatores não observáveis foi bastante elevada em todos os percentis: 135 no quantil 0.1; 93 no quantil 0.25; 84 na mediana; e 77 e 88 nos quantis 0.75 e 0.9, respectivamente. Isso pode ser comprovado por meio da decomposição detalhada desse efeito, apresentado na tabela 3, o qual mostra a influência significativa de fatores não observáveis na maioria das variáveis em boa parte da distribuição, principalmente em percentis mais elevados.

Na parte inferior da distribuição, o efeito não observável seria menor e contribuiria para a redução do hiato entre as notas na média geral, caso os alunos das escolas profissionalizantes tivessem um aumento no desempenho para os alunos do sexo masculino, com mães que possuem ensino fundamental completo ou incompleto, e um maior retorno em todas as faixas de rendas, principalmente para aqueles com renda familiar entre 2 e 5 SMs. Esse efeito não observável age de forma a reduzir o efeito estrutural em aproximadamente 240.

Considerando-se a mediana da distribuição, percebe-se que o efeito estrutural de notas poderia ser reduzido e, assim, contribuir para a queda da desigualdade de notas entre as escolas. Isso ocorreria se os alunos do sexo masculino das escolas profissionalizantes – entre aqueles com pais em todos os níveis de escolaridade, exceto os que possuem nível superior completo e incompleto – melhorassem o desempenho. O destaque seria para aqueles alunos cujos pais possuem ensino médio completo ou incompleto; nesse segmento, a redução sobre os efeitos não observáveis seria 35, 36. Além destes, o efeito estrutural reduziria o aumento do efeito dos alunos com mães que possuem ensino fundamental e médio completo

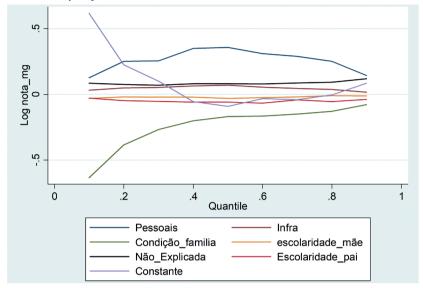
ou incompleto em, aproximadamente, 7 e 21, respectivamente; com um maior retorno em notas na média geral dos alunos em todas as faixas de rendas, exceto na *Renda05*, cujo efeito mais expressivo foi para aqueles com renda familiar entre 2 e 5 SMs, reduzindo o efeito não observado em aproximadamente 85.

Resultado similar pode ser verificado na parte superior da distribuição. Porém, o efeito estrutural das notas seria reduzido com o aumento no desempenho para os alunos das EEEPs do sexo masculino, solteiros, com pais que possuem ensino fundamental e médio completo ou incompleto. Além disso, esse efeito também se reduziria com um maior retorno em notas na média geral dos alunos em todas as faixas de rendas, exceto para os que possuem renda familiar entre 10 e 15 SMs, cujo efeito mais expressivo foi para aqueles com renda familiar entre 2 e 5 SMs, diminuindo o efeito não observado em, aproximadamente, 41.

Assim, é importante verificar a distribuição de notas por quantis para que se possa identificar o efeito heterogêneo das características dos alunos e das escolas, bem como do retorno dessas características no hiato de notas entre as escolas particulares e profissionalizantes do Ceará. As contribuições das covariáveis na redução do hiato podem servir como forma de identificar os fatores em que as políticas públicas poderiam agir quando se pretende equiparar o ensino público ao privado. Porém, não se pôde analisar características relacionadas à habilidade inata do aluno e ao esforço exercido pelos estudantes de ambas as redes de ensino, uma vez que estes podem ser fatores relevantes para a determinação da nota do aluno e do diferencial, mas não são observáveis.

Para melhor visualização da contribuição desses efeitos, foram construídos os gráficos 9 e 10, nos quais se optou por agregar algumas variáveis em grupos, em que *idade*, *sexo*, *raça*, *estado_civil*, *Anos_Fund* e *Anos_Médio* pertencem às características pessoais; *E_Pai1*, *E_Pai2*, *E_Pai3* e *E_pai4* inserem-se no grupo que representa a escolaridade do *p*; e a escolaridade da mãe e as variáveis *renda* e *CSE* pertencem ao grupo da condição da família.

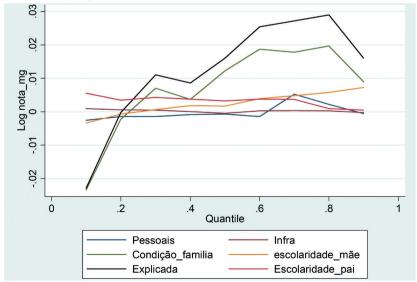
GRÁFICO 9 **Decomposição detalhada do efeito característico**



Elaboração dos autores utilizando os dados da tabela 3.

Obs.: Gráfico cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

GRÁFICO 10 Decomposição detalhada do efeito estrutural



Elaboração dos autores utilizando os dados da tabela 3.

Observa-se que as condições socioeconômicas da família e as escolaridades dos pais são fatores determinantes para a redução dessa desigualdade. Nesse contexto, optou-se por fazer simulações para entender como uma melhoria de renda poderia compensar deficiências históricas no perfil educacional no desempenho dos alunos. Desse modo, serão simuladas distribuições com mudanças na escolaridade dos pais e na renda; fatores que individualmente aparecem com grande importância, de forma que os alunos das escolas profissionalizantes não possuam diferenças significativas nessas variáveis quando comparados aos alunos das escolas particulares.

Para tanto, aplicou-se o balanceamento por entropia. ¹⁴ Conforme Hainmueller (2012), esse método permite ponderar um agrupamento de dados em que as distribuições das variáveis e as observações reponderadas satisfaçam um conjunto de condições especiais de momentos, de forma que exista equilíbrio exato sobre o primeiro (média), o segundo (variância) e o terceiro (assimetria) momentos das distribuições de variáveis independentes nos grupos de tratamento e controle. Assim, é possível se especificar um nível de equilíbrio desejável para as covariadas, usando um conjunto de condições associadas aos momentos da distribuição.

No cenário I, os alunos das escolas profissionalizantes apresentariam a mesma distribuição dos estudantes das escolas particulares nas variáveis referentes à renda da família. No cenário II, além dessas variáveis, foram balanceadas as que são relacionadas à educação da mãe. Por fim, no cenário III, acrescentaram-se ao cenário anterior as variáveis concernentes à educação do pai. Após a aplicação do balanceamento por entropia, estimou-se o diferencial total a partir das estimativas da FIR, ponderada pelos pesos da entropia, e comparou-se este aos resultados da tabela 2, cujos resultados se encontram na tabela 4.

TABELA 4
Diferencial total ponderado pela entropia: MG

	Cenários											
I					ı			III				
Antes		Após		Antes		Após		Antes		Após		
EEEPs	Particulares	EEEPs	Particulares	EEEPs	Particulares	EEEPs	Particulares	EEEPs	Particulares	EEEPs	Particulares	
0.0155	0.0049	0.0049	0.0049	0.0155	0.0049	0.0049	0.0049	0.0155	0.0049	0.0049	0.0049	
0.6392	0.1323	0.1323	0.1323	0.6392	0.1323	0.1323	0.1323	0.6392	0.1323	0.1323	0.1323	
0.2797	0.2635	0.2635	0.2635	0.2797	0.2635	0.2635	0.2635	0.2797	0.2635	0.2635	0.2635	
-	=	-	=	0.0279	0.0068	0.0068	0.0068	0.0279	0.0068	0.0068	0.0068	
-	=	-	-	0.4944	0.1572	0.1572	0.1572	0.4944	0.1572	0.1572	0.1572	
-	-	-	-	-	-	-	-	0.0692	0.0149	0.0149	0.0149	
-	-	-	-	-	-	-	-	0.5787	0.2317	0.2317	0.2317	
	0.0155 0.6392	EEEPs Particulares 0.0155 0.0049 0.6392 0.1323	EEEPs Particulares EEEPs 0.0155 0.0049 0.0049 0.6392 0.1323 0.1323	EEEPs Particulares EEEPs Particulares 0.0155 0.0049 0.0049 0.0049 0.6392 0.1323 0.1323 0.1323 0.2797 0.2635 0.2635 0.2635	EEEPs Particulares EEEPs Particulares EEEPs 0.0155 0.0049 0.0049 0.0049 0.0155 0.6392 0.1323 0.1323 0.1323 0.6392 0.2797 0.2635 0.2635 0.2635 0.2797 0.0049 0.0049 0.0049 0.0049 0.0049 0.0279 0.2635 0.2635 0.2635 0.2637	Note Note	Note Note	Note Note	No. No	Note Note	The state of the law of the la	

(Continua)

^{14.} Pareamento desenvolvido por Hainmueller (2012), cuja especificação do método se encontra no apêndice C.

(Continuação	0)					
Desigualdade	es .					
Quantil 0.1	0.0637	0.0406	0.0637	0.0384	0.0637	0.038
Quantil 0.5	0.0978	0.0655	0.0978	0.0617	0.0978	0.0609
Quantil 0.9	0.1365	0.1072	0.1365	0.1044	0.1365	0.1028

A partir da tabela 4, verifica-se que, antes do pareamento, os alunos das escolas profissionalizantes apresentavam desvantagens referentes às variáveis *renda* e *educação dos pais*. No cenário I, após o balanceamento, esses alunos passaram a evidenciar a renda da família equivalente aos alunos das escolas particulares, reduzindo, assim, a desigualdade na média geral em, aproximadamente, 36% no primeiro quantil das distribuições de notas, 33% na mediana e 21,4% no quantil 0.9. No cenário II, as variáveis da escolaridade da mãe também não apresentaram diferenças significativas entre os dois grupos, após o pareamento. Dessa forma, a desigualdade reduziria para 39,7%, 36,9% e 23,5%, respectivamente, nos quantis, 0.1, 0.5 e 0.9, enquanto, no cenário III, após balancear as variáveis *renda*, *escolaridade da mãe* e *escolaridade do pai*, o diferencial total diminuiria 40,3%, 37,7% e 24,6% nos quantis analisados.

Dessa forma, percebe-se a importância da decomposição detalhada dos fatores sobre o diferencial total; em destaque, a renda da família e a escolaridade dos pais, uma vez que estes fatores foram capazes de reduzir a desigualdade de notas dos alunos na média geral, em todos os quantis. Além disso, cabe destacar que esse efeito é maior nos quantis iniciais da distribuição — ou seja, entre os alunos com menores desempenho nos dois grupos.

4 CONCLUSÕES

Este trabalho procurou ampliar o debate existente na literatura da economia da educação sobre o ensino público, com ênfase no ensino profissionalizante desenvolvido no Ceará. Para tanto, buscou-se trazer novas evidências sobre os diferenciais de desempenho entre o ensino público e privado, ao analisar as diferenças nas distribuições de notas no Enem, para as cinco áreas de conhecimento e na média geral, entre alunos das EEEPs e das escolas privadas.

Esta análise foi feita visando reduzir a influência de fatores não observados, ao se construir grupos de alunos mais homogêneos para se aferir o diferencial de rendimento escolar líquido. A pertinência para tal escolha prende-se ao fato de as escolas de ensino profissional apresentarem desempenho superior às demais escolas públicas. Dessa maneira, busca-se verificar se esse novo modelo de escola pública provê uma estrutura de ensino equiparável ao da escola privada, ao analisar

como os alunos dos dois tipos de escola se comparam ao longo da distribuição de rendimento escolar e quais fatores observáveis mais influenciam os desempenhos dos alunos das duas instituições.

Além de propor uma nova construção para o grupo das escolas públicas, este trabalho buscou inovar ao utilizar o método RQI proposto por Firpo, Fortin e Lemieux (2007), o qual emprega regressões com base na FIR combinadas com um procedimento de reponderação inspirado em DiNardo, Fortin e Lemieux (1996) e permite aplicar o método de decomposição de Oaxaca-Blinder.

Como não se possuem informações identificadas dos alunos do ensino privado para uma análise longitudinal, utilizaram-se apenas as informações referentes ao desempenho dos alunos no Enem para 2014, as quais foram extraídas do Inep. Já a identificação das escolas profissionalizantes e de suas dotações de infraestrutura de todas as escolas ocorreu por meio de dados da Seduc e do Censo Escolar de 2014, respectivamente.

Os resultados encontrados mostram que, em todas as áreas, as estimativas dos diferenciais foram significantes, apresentando vantagens aos alunos das escolas particulares em relação aos alunos do ensino profissional. Nas áreas de matemática, ciências da natureza e na média geral, o aumento da diferença de notas na parte superior da distribuição indica o efeito *glassceiling*, pois observa-se que o diferencial total entre as notas dos alunos das escolas profissionalizantes e privadas é crescente nos quantis. Dessa forma, a diferenciação de notas, a favor dos alunos das escolas particulares, é maior para aqueles que apresentam desempenho superior. Nas áreas de ciências humanas e linguagens e códigos, apesar de apresentarem diferenciais positivos, estes praticamente permanecem constantes ao longo da distribuição, com leves oscilações; enquanto, em redação, os diferenciais foram os mais elevados ao longo de toda a distribuição.

A decomposição quantílica de Firpo, Fortin e Lemieux (2007) permite identificar os fatores que contribuem para esse diferencial de notas, os quais podem ser observados através dos efeitos composição (caraterísticos) e estruturais (coeficientes). Dos resultados, percebeu-se que os efeitos não observáveis (coeficientes) são todos positivos e significativos em todos os quantis da distribuição e em todas as áreas do Enem, sendo responsáveis por explicar a maior parte da diferença das notas, com destaque para matemática, cujo efeito é crescente ao longo da curva. Esses resultados indicam que os alunos das escolas particulares são mais capazes de converter insumos educacionais em maiores pontuações em testes padronizados e que existem outros fatores não observáveis pertencentes a estes alunos que contribuem para melhores desempenhos.

Além disso, verificou-se que o efeito característico age no sentido de diminuir as diferenças entre as notas apenas em alguns quantis da parte inferior da distribuição destas, principalmente para o quantil 0.1 de ciências humanas e de linguagens e códigos, cujo efeito reduz o diferencial em 56 e 54, respectivamente. Porém,

na maioria dos casos, esse efeito foi positivo e significante, atuando de maneira a contribuir com o diferencial de notas. Ou seja, os alunos das escolas privadas apresentam melhores características que contribuem para um bom desempenho, principalmente nos quantis mais elevados da distribuição, pois, em todas as áreas de conhecimento e na média geral, esses efeitos foram positivos e significativos nos percentis 0.9 e 0.95.

A partir da decomposição detalhada dos fatores que afetam o diferencial entre os grupos, identificou-se que melhorias na renda da família e na escolaridade dos pais poderiam minimizar as deficiências no perfil educacional no desempenho dos alunos. Dessa forma, foram construídos três cenários com mudanças nessas variáveis, e utilizou-se o balanceamento por entropia para equiparar o grupo dos alunos das escolas profissionalizantes aos alunos das escolas particulares. Após esse procedimento, estimou-se o diferencial total a partir da FIR ponderada pelos pesos da entropia. Dos resultados, observou-se que, ao se melhorar as faixas de renda e a educação dos pais, a desigualdade entre esses grupos poderia reduzir 40,3%, 37,7% e 24,6%, respectivamente, no primeiro quantil, na mediana e no extremo da distribuição das notas na média geral.

Embora este estudo tenha buscado inovar ao comparar grupos teoricamente mais homogêneos, alunos das escolas profissionalizantes com alunos das escolas particulares, o diferencial de notas no desempenho do Enem ainda é significativo, ao longo de toda a distribuição. Com isso, políticas públicas educacionais, como o ensino profissionalizante, devem ser complementadas com políticas sociais que não só visem à melhoria no tipo de ensino que é ofertado, mas também foquem no desenvolvimento das famílias, por meio de melhorias na renda e na escolaridade dos pais.

REFERÊNCIAS

BARSKY, R. *et al.* Accounting for the black-white wealth gap: a nonparametric approach. **Journal of the American Statistical Association**, v. 97, n. 459, p. 663-673, 2002.

BORAH, B. J.; BASU, A. Highlighting differences between conditional and unconditional quantile regression approaches through an application to assess medication adherence. **Health Economics**, v. 22, n. 9, p. 1052-1070, 2013.

CEARÁ. Secretaria da Educação do Estado do Ceará. **O pensar e o fazer da educação profissional no Ceará**: 2008-2014. Relatório de gestão. Fortaleza: Seduc, 2014. (Coordenadoria de Educação Profissional, 1. ed.).

CHERNOZHUKOV, V.; FERNANDEZ-VAL, I.; MELLY, B. Inference on counterfactual distributions. **Econometrica**, v. 81, n. 6, p. 2205-2268, 2013.

COSTA, L.; ARRAES, R. Identificação parcial do efeito das escolas privadas brasileiras. **Pesquisa e Planejamento Econômico**, v. 44, n. 2, p. 337-371, ago. 2014.

DINARDO, J; FORTIN, N. M.; LEMIEUX, T. Labor market institutions and the distribution of wages, 1973-1992: a semiparametric approach. **Econometrica**, v. 64, n. 5, p. 1001-1044, Sept. 1996.

DONALD, S. G.; GREEN, D. A.; PAARSCH, H. P. Differences in wage distributions between Canada and the United States: an application of a flexible estimator of distribution functions in the presence of covariates source. **Review of Economic Studies**, v. 67, n. 4, p. 609-633, 2000.

EDOKA, I. P. Decomposing differences in cotinine distribution between children and adolescents from different socioeconomic backgrounds. Heslington: HEDG/University of York, 2012. (Working Paper, n. 12/29).

FRANÇA, M. T. A.; GONÇALVES, F. O. Provisão pública e privada de educação fundamental: diferenças de qualidade medidas por meio de propensity score. **Economia Aplicada**, v. 14, n. 4, p. 373-390, 2010.

FISHER, J.; MARCHAND, J. Does the retirement consumption puzzle differ across the distribution? **The Journal of Economic Inequality**, v. 12, p. 279-296, 2014.

FIRPO, S.; FORTIN, N.; LEMIEUX, T. **Decomposing wage distributions using recentered influence function regressions**. Vancouver: University of British Columbia, 2007. Mimeographed.

_____. Unconditional quantile regressions. **Econometrica,** v. 77, n. 3, p. 953-973, 2009.

FORTIN, N. M. The gender wage gap among young adults in the United States: the importance of money vs. people. **The Journal of Human Resources**, v. 43, n. 4, p. 886-920, 2008.

FORTIN, N. M.; LEMIEUX, T. Rank regressions, wage distributions, and the gender gap. **The Journal of Human Resources**, v. 33, n. 3, p. 610-643, 1998.

FOURNIER, J.; KOSKE, I. Less income inequality and more growth: are they compatible? Part 7. The drivers of labour earnings inequality: an analysis based on conditional and unconditional quantile regressions. Paris: OECD Publishing, 2012. (Working Papers, n. 930).

HAINMUELLER, J. Entropy balancing for causal effects: a multivariate reweighting method to produce balanced samples in observational studies. **Political Analysis**, v. 20, n. 1, p. 25-46, 2012.

HAMPEL, F. R. *et al.* **Robust statistics**: the approach based on influence functions. New York: Wiley, 1986.

HIRSCH, B. T.; WINTERS, J. V. An anatomy of racial and ethnic trends in male earnings in the U.S. **Review of Income and Wealth**, v. 60, n. 4, p. 930-947, 2014.

IPECE – INSTITUTO DE PESQUISA E ESTRATÉGIA ECONÔMICA DO CEARÁ. Análise da participação das escolas públicas estaduais cearenses no Exame Nacional do Ensino Médio (Enem): 2009-2011. Fortaleza: Ipece, fev. 2013 (Ipece Informe, n. 54). Disponível em: https://bit.ly/2PQuLYa.

JUHN, C.; MURPHY, K. M.; PIERCE, B. Wage inequality and the rise in returns to skill. **Journal of Political Economy**, v. 101, n. 3, p. 410-442, June 1993.

KASSENBOEHMER, S. C.; SINNING, M. G. Distributional changes in the gender wage gap. **Industrial & Labor Relations Review**, v. 67, n. 2, p. 335-361, Apr. 2014.

KOENKER, R.; BASSETT, J. R. Regression quantiles. **Econometrica**, v. 46, n. 1, p. 33-50, 1978.

_____. Robust tests for heteroscedasticity based on regression quantiles. **Econometrica**, v. 50, n. 1, p. 43-61, 1982.

LE, H. T.; BOOTH, A. L. Inequality in vietnamese urban-rural living standards, 1993-2006. **Review of Income and Wealth**, v. 60, n. 4, p. 862-886, Dec. 2014.

MACHADO, J. F.; MATA, J. Counterfactual decomposition of changes in wage distributions using quantile regression. **Journal of Applied Econometrics**, v. 20, n. 4, p. 445-465, May/June 2005.

MANSKI, C. Anatomy of the selection problem. **The Journal of Human Resource**, v. 24, n. 3, p. 343-360, 1989.

MELLY, B. Decomposition of differences in distribution using quantile regression. **Labour Economics**, v. 12, n. 4, p. 577-590, Aug. 2005.

MORAES, A. G. E.; BELLUZZO, W. O diferencial de desempenho escolar entre escolas públicas e privadas no Brasil. **Nova Economia**, v. 24, n. 2, p. 409-430, maio/ago. 2014.

MORIN, L. P. Do men and women respond differently to competition? Evidence from a major education reform. **Journal of Labor Economics**, v. 33, n. 2, p. 443-491, 2015.

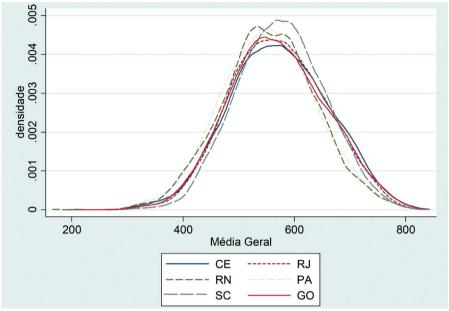
NGUYEN, H. T. **The evolution of the gender test score gap through seventh grade**: new insights from Australia using quantile regression and decomposition. Bentley: BCEC, 2015. (Working Paper, n. 15/7).

RUBIL, I. **The great recession and the public-private wage gap**: distributional decomposition evidence from Croatia 2008-2011. MPRA Paper, 2013. (Working Paper, n. 46798).

APÊNDICE A

GRÁFICO A.1

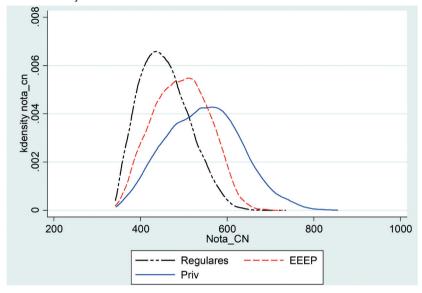
Desempenho no Enem dos alunos das escolas particulares: CE, RJ, RN, PA e GO



Elaboração dos autores.

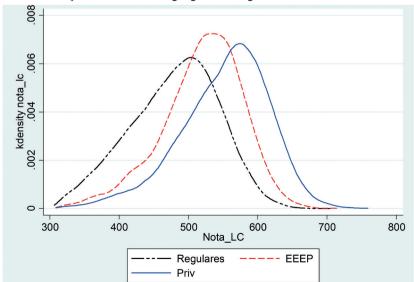
- Obs.: 1. Enem Exame Nacional do Ensino Médio; CE Ceará; RJ Rio de Janeiro; RN Rio Grande do Norte; PA Pará; e GO Goiás.
 - Gráfico cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

GRÁFICO A.2 Distribuição das notas em ciências naturais



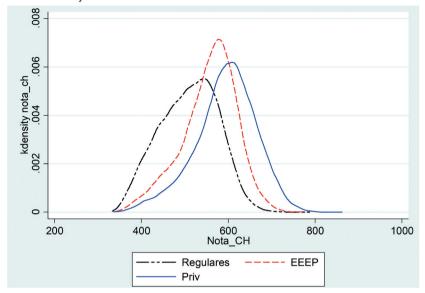
Obs.: Gráfico cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

GRÁFICO A.3 Distribuição das notas em linguagens e códigos



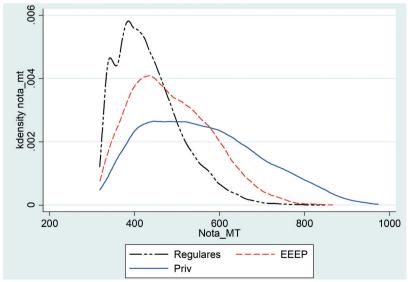
Elaboração dos autores.

GRÁFICO A.4 Distribuição das notas em ciências humanas



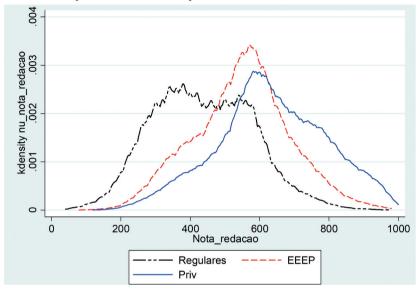
Obs.: Gráfico cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

GRÁFICO A.5 Distribuição das notas em matemática



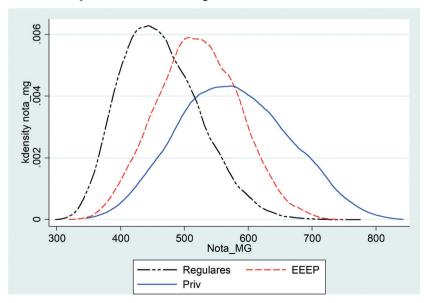
Elaboração dos autores.

GRÁFICO A.6 Distribuição das notas em redação



Obs.: Gráfico cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

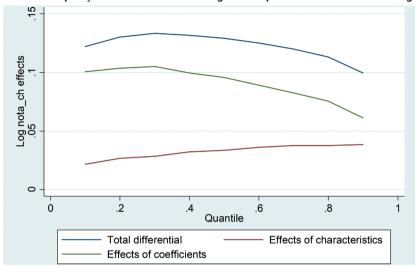
GRÁFICO A.7 Distribuição das notas em média geral



Elaboração dos autores.

GRÁFICO A.8

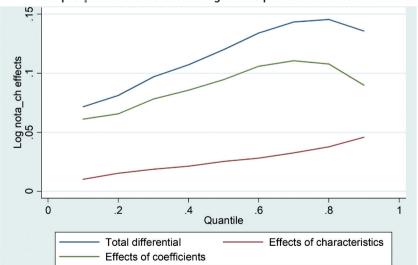
Decomposição dos efeitos: escolas regulares e profissionalizantes – média geral



Obs.: Gráfico cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

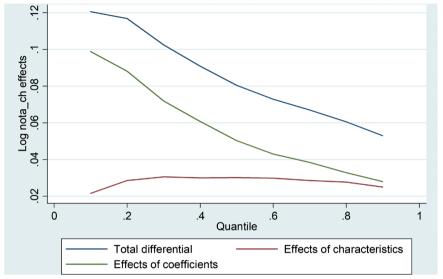
GRÁFICO A.9

Decomposição dos efeitos: escolas regulares e profissionalizantes – matemática



Elaboração dos autores.

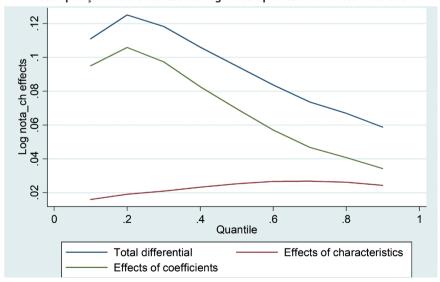
GRÁFICO A.10 Decomposição dos efeitos: escolas regulares e profissionalizantes – linguagens e códigos



Obs.: Gráfico cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

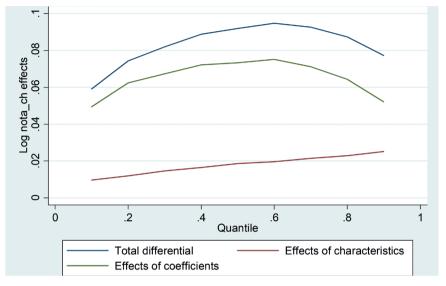
GRÁFICO A.11

Decomposição dos efeitos: escolas regulares e profissionalizantes – ciências humanas



Elaboração dos autores.

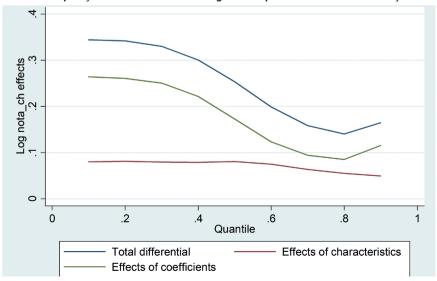
GRÁFICO A.12 Decomposição dos efeitos: escolas regulares e profissionalizantes – ciências da natureza



Elaboração dos autores.

Obs.: Gráfico cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

GRÁFICO A.13 Decomposição dos efeitos: escolas regulares e profissionalizantes – redação



Elaboração dos autores.

Obs.: Gráfico cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

APÊNDICE B

 $\label{eq:thm-problem} \text{TABELA B.1}$ Estimativas da regressão quantílica incondicional por escola: média geral

		MQ0 ²			0.1			0.5			6:0	
ValidVels	EEEPs1	Particulares	Regulares	EEEPs	Particulares	Regulares	EEEPs	Particulares	Regulares	EEEPs	Particulares	Regulares
Idade	-0.024*	*7600.0-	-0.0037*	-0.0348*	-0.0229*	-0.0029*	-0.0296*	-0.0078*	-0.0042*	-0.0125*	-0.0015*	-0.0036*
Sexo	0.020*	0.0033	-0.0058*	0.0064	-0.0102**	-0.0209*	0.0291*	0.0088**	-0.0076*	0.0276*	0.0123*	0.0088*
Raça	0.005***	*0600.0	*9600.0	0.0054	0.0049	0.0056**	0.0054	0.0088**	0.0110*	0.0016	0.0140*	0.0136*
Estado civil	0.0283	-0.0128	-0.0172*	0.0426	-0.0436	0.0269*	0.0271	-0.0057	-0.0226*	0.0251**	-0.0396**	-0.0028
Anos_Fund	0.0192*	0.0303*	0.0156*	0.0187*	0.0282*	*6800.0	0.0253*	0.0374*	0.0170*	0.0149*	0.0242*	-0.0230*
Anos_Médio	-0.0007	0.0093*	-0.0007	0.0043	0.0093**	-0.0062*	0.0011	0.0067***	-0.0013	-0.0055	0.0148*	0.0031
E_Pai1	0.0005	-0.0494*	-0.0365*	-0.0237	-0.0644**	-0.0209**	0.0115	-0.0665*	-0.0428*	0.0045	-0.0180	-0.0509**
E_Pai2	0.0098	-0.0458*	-0.0291*	-0.0077	-0.0441*	-0.0077	0.0227	-0.0479*	-0.0344*	0.0111	-0.0386*	-0.0491**
E_Pai3	0.0233	-0.0325*	-0.0039	0.0060	-0.0291*	0.0081	0.0425**	-0.0311*	-0.0057	0.0170	-0.0343*	-0.0204
E_Pai4	0.0208	-0.0100**	0.0055	0.0041	-0.0118**	0.0068	0.0351	9600:0-	-0.0004	0.0136	-0.0033	0.0049
E_mae1	-0.0045	-0.0272***	-0.0462*	0.0117	-0.0581	-0.0187*	0.0093	0.0001	-0.0439*	-0.0197	-0.0243***	-0.0644*
E_mae2	-0.0035	-0.0330*	-0.0395*	0.0232***	-0.0387*	-0.0151*	0.0057	-0.0336*	-0.0330*	-0.0212	-0.0298*	-0.0590*
E_mae3	0.0086	-0.0190*	-0.013**	0.0274**	-0.0058	0.0040	0.0198***	-0.0241*	-0.0014	-0.0025	-0.0275*	-0.0328*
E_mae4	0.0042	-0.0026	-0.0071	0.0183	-0.0047	0.0001	0.0153	-0.0042	0.0044	-0.0010	-0.0020	-0.0166
Renda00	0.1937*	-0.0931*	0.0353	0.6233*	-0.0928**	0.0022	0.1094*	-0.0983*	0.0103	0.0168	-0.1034*	0.0016
Renda01	0.2117*	-0.0992*	0.0448	0.6499*	-0.0796*	0.0130	0.1314*	-0.1107*	0.0179	0.0348*	-0.0978*	0.0133
Renda02	0.2456*	-0.0780*	0.0685	0.6789*	-0.0454*	0.0319	0.1720*	-0.0881*	0.0449	0.0706*	-0.0993*	0.0384
Renda03	0.2561*	-0.0458*	0.0797	0.6838*	-0.0032	0.0231	0.1821*	-0.0499*	0.0498	0.0858*	-0.0810*	0.0836
Renda04	0.2327*	-0.0220*	0.0777	0.6861*	-0.0086	0.0261	0.1742*	-0.0165**	0.0460	0.0507	-0.0375*	0.0987
Renda05	0.2837*	-0.0144**	0.098***	0.7715*	-0.0048	0.0371	0.1707	-0.0072	0.0208	-0.0056	-0.0307***	0.1424
CSE	-0.0054	0.0575*	0.0810*	-0.0196	0.0371**	0.0789*	-0.0052	0.0758*	0.0973*	-0.0042	0.0792*	0.0780*
Infra	0.0103	0.1304*	0.1165*	0.0269	0.1282*	0.0724*	-0.0107	0.1828*	0.1080*	-0.0014	0.0317**	0.1785*
Constante	6.3922	6.5196*	6.1587*	5.9553*	6.5609*	5.9992*	6.5421*	6.4402*	6.1906*	6.5380*	6.6223*	6.3618*

Elaboração dos autores.

Notas: 1 EEEPs = Escolas de educação de ensino profissional. 2 MQO = Mínimos quadrados ordinários. Obs.: Significância (*) = 1%; (**) = 5%; e (***) = 10%.

 ${\sf TABELA~B.2} \\ {\sf Estimativas~da~regress\~ao~quant\'ilica~incondicional~por~escola:~ci\^encias~humanas} \\$

tability EEEP Particulares Regulates REPA Particulares Regulates REPA Particulares Regulates Regulates Regulates Regulates Particulares Regulates Particulares Regulates Particulares Regulates Co.0034 A.00483 A.00403 A.00403 A.00404 A.00404<			MQO			0.1			0.5			6.0	
10,0154 0,00483 0,00079 0,0029° 0,00079° 0,	Vallavels	EEEPs	Particulares	Regulares	EEEPs	Particulares	Regulares	EEEPs	Particulares	Regulares	EEEPs	Particulares	Regulares
10,0256 0,00730 0,00791 0,02099 0,0028 0,0024 0,0004 0,0004 0,0004 0,0004 0,0004 0,0004 0,0004 0,000730 0,00169 0,00083 0,00028 0,00028 0,00028 0,00029 0,	Idade	-0.0154	-0.00483	-0.0007	-0.0219*	-0.0087*	-0.0010*	-0.0198*	-0.0047*	*8000:0-	-0.0074*	-0.0010*	-0.0010*
1	Sexo	0.0253	0.01350	0.00791	0.0209*	-0.0022	-0.004***	0.0276*	0.0175*	*0600.0	0.0306*	0.0184*	0.0192*
60058 0.0051 0.0054 0.0058 0.0081* 0.0024* 0.0101* 0.0103* 0.0058* 0.0011* 0.0103* 0.0058* 0.0014* 0.0024 0.01351 0.01351 0.01351 0.0132 0.0034* 0.0145* 0.0161* 0.01151 0.01151* 0.01151 0.0024 0.0022 0.0034 0.0022 0.0034 0.0024 0.0034	Estado civil	0.0260	-0.00730	-0.0169	0.0885	0.0028	-0.0294*	-0.0004	-0.0045	-0.0187*	0.0166***	-0.0042	***900'0-
dig 0.0152 0.02244 0.01351 0.0197* 0.0289* 0.01017* 0.0178* 0.0164* 0.01615 0.0152 0.0032 0.0038* 0.01017* 0.0178* 0.0164* 0.0161 0.0011 0.0012 0.0032 0.0036 0.0034 0.0034* 0.00404 0.00404 0.00404 0.00404 0.00404 0.00404 0.00404 0.00404 0.0024 0.0021* 0.00404 0.00404 0.0024 0.0021* 0.0024 0.00249 0.0021* 0.0024 0.00249 0.0021* 0.0024 0.00249 0.0021* 0.0024 0.00249 0.0021* 0.0024 0.00249 0.0021* 0.0024 0.00249 0.0021* 0.0024 0.00249 0.0021* 0.0024 0.00249 0.0021* 0.0024 0.0034	Raça	0.0058	0.00511	0.00797	0.0045	0.0078	0.0028	0.0081**	0.0028	0.0101*	0.0103*	0.0050***	0.0082*
dio 0.0012 0.00396 -0.0020 -0.0036 -0.0020 -0.0036 -0.0036 -0.0036 -0.0036 -0.0036 -0.00396 -0.00396 -0.00396 -0.00396 -0.00349 -0.00396 -0.00349 -0.00340 -0.00349 -0.00340 -0.00340 -0.00340 -0.00340 -0.0034 -0.00349 -0.00340 -0.00340 -0.00349 -0.00340 -0.00344 -0.00340 -0.00344 -0.00344 -0.00344 -0.00344 -0.00344 -0.00344 -0.00344 -0.00344 -0.00344 -0.00344 -0.00344 -0.00344 -0.00344 -0.00344 -0.00344 -0.00344 <td>Anos_Fund</td> <td>0.0152</td> <td>0.02224</td> <td>0.01351</td> <td>0.0197*</td> <td>0.0289*</td> <td>0.0121*</td> <td>0.0178*</td> <td>0.0245*</td> <td>0.0161*</td> <td>0.0101*</td> <td>0.0151*</td> <td>0.0137*</td>	Anos_Fund	0.0152	0.02224	0.01351	0.0197*	0.0289*	0.0121*	0.0178*	0.0245*	0.0161*	0.0101*	0.0151*	0.0137*
0.0061 -0.02949 -0.0296 -0.0479 -0.0234** -0.0214 -0.02949 -0.0298* -0.0214 -0.0214 -0.02405 -0.02404 -0.0214	Anos_Médio	0.0012	0.00396	-0.0022	-0.0036	-0.0001	-0.0058**	0.0049	0.0050***	-0.0020	0.0011	0.0074*	0.0019
0.0085 -0.03102 -0.0247 -0.02018 -0.0214 -0.04312 -0.04102 -0.04312 -0.04102 -0.04104 -0.04014 -0.04014 -0.04014 -0.04014 -0.04014 -0.04014 -0.04014 -0.0414 -0.0214 -0.0414 -0.0214 -0.0121 -0.0114 -0.0014 -0.0124 -0.0114 -0.0014 -0.0014 -0.0114 -0.0014 -0.0124 -0.0114 -0.0014 -0.0124 -0.0114 -0.0014 -0.0014 -0.0124 -0.0114 -0.0014 -0.0014 -0.0124 -0.0121* -0.0114* -0.0014 -0.0124 -0.0214* -0.0214 -0.0214* -0.0214* -0.0214* -0.0214* -0.0214* -0.0214* -0.0214* -0.0214* -0.0214* -0.0214* -0.0214* -0.0214* -0.0214* -0.0214* -0.0214* -0.0214* -0.0114* -0.0114* -0.0114* -0.0114* -0.0114* -0.0114* -0.0114* -0.0114* -0.0114* -0.0114* -0.0114* -0.0114* -0.0114* -0.0114* -0.0114* -	E_Pai1	0.0061	-0.02949	-0.0290	0.0087	-0.0479	-0.023***	0.0022	-0.0405*	-0.0472*	-0.0045	-0.0295*	-0.0327**
0.00154 -0.00121 -0.00249 -0.0018*** -0.0016 -0.0029* -0.0015 -0.0015 -0.0015 -0.0015 -0.0015 -0.0015 -0.0015 -0.0015 -0.0015 -0.0014 -0.0016 -0.0016 -0.0016 -0.0016 -0.0016 -0.0017 -0.0016 -0.0017 -0.0018 -0.0018 -0.0014 -0.0018 -0.0014	E_Pai2	0.0085	-0.03102	-0.0247	-0.0002	-0.0298*	-0.0214	0.0053	-0.0332*	-0.0401*	-0.0042	-0.0378*	-0.030***
0.0175 0.00558 0.0034 0.0112 0.0007 0.0176 0.0059 0.00558 0.0024 0.0112 0.0007 0.0176 0.0059 0.0024 0.0112 0.0005 0.0024 0.0124 0.0121* 0.0112*** 0.0005 0.00176 0.00248 0.00176 0.00248 0.00176 0.00248 0.00248 0.00176 0.0034 0.00177	E_Pai3	0.0234	-0.0212	-0.0025	0.0249	-0.0218**	-0.0016	0.0210	-0.0229*	-0.0125	0.0056	-0.0314*	-0.0114
0.0003 0.00059 0.0029 0.0191 0.0183 0.00053 0.00148 0.00348 0.00148 0.00248 0.00248 0.00154 0.00154 0.00154 0.00154 0.00154 0.00154 0.00154 0.00154 0.00154 0.00154 0.00154 0.00154 0.00154 0.00284 0.00284 0.0044 0.00171 0.0041 0.0044 0.00174 0.0044<	E_Pai4	0.0175	-0.00558	0.0034	0.0122	-0.0112	0.0007	0.0170	-0.0050	-0.0034	-0.0075	-0.0125***	0.0073
0.0003 -0.0176 -0.0244 -0.0251** -0.016*** -0.0058 -0.0195* -0.0195* -0.0044 -0.0036 0.0084 -0.0129 -0.0019 0.0284 -0.0058 0.0041 0.0070 -0.0171* 0.0041 0.0044 0.0044 0.0016 -0.0026 0.0031 0.0048 -0.0052 0.0045 0.0113* -0.0619* 0.0049 0.0044 0.31398 -0.0755 0.0075 0.0171* -0.0653* 0.113* -0.0619* 0.0049 0.0048 0.32080 -0.0715 0.0179 0.0178* -0.0678* 0.0162* 0.1111* 0.0052* 0.118* 0.0052* 0.0054* 0.0064	E_mae1	0.0043	0.00059	-0.0279	0.0191	0.0183	-0.0201**	-0.0053	-0.0148	-0.0300*	-0.0033	0.0161	-0.0316*
0.0084 -0.0129 -0.0019 0.0284 -0.0058 0.0041 0.0070 -0.0171* 0.0041 0.0044 -0.0044 0.0016 -0.00063 0.0019 0.00108 -0.0002 0.0031 0.0016 -0.0619* 0.0071 0.0049 0.31398 -0.07255 0.0075 0.8756* -0.1121** -0.0663* 0.1132* -0.0619* 0.0055** 0.0059* -0.0598* 0.32080 -0.07152 0.0179 0.8716* -0.0678* -0.0675* 0.1181* -0.0619* 0.0059* -0.0598* 0.32080 -0.0756 0.0176* -0.0468* -0.0518* 0.1181* -0.0507* 0.1359* -0.0598* 0.32028 -0.02764 0.0466 0.9392* -0.0031 -0.0518* 0.1404* -0.0139* 0.0458* -0.0518* 0.334520 -0.01205 0.04624 0.9109* -0.0048 -0.044* 0.1404** 0.0108* 0.0159* -0.0598* 0.0166 0.0279 0.0239* -0.0248 0.029	E_mae2	0.0003	-0.02176	-0.0224	0.0242	-0.0251**	-0.016***	-0.0058	-0.0248*	-0.0195*	0.0084	-0.0036	-0.0287*
0.0016 -0.00063 0.0019 -0.0002 0.0031 0.0016 -0.0065** 0.0019 -0.0058* 0.0002 -0.0052* 0.0172* 0.0056*** 0.0058** 0.0058* -0.0588* 0.0058*	E_mae3	0.0084	-0.0129	-0.0019	0.0284	-0.0058	0.0041	0.0070	-0.0171*	0.0041	0.0027	-0.0044	-0.0088
0.31398 -0.07255 0.0075 0.8756* -0.1121** -0.0762* 0.1132* -0.0619* 0.0965*** 0.01121* 0.0762* 0.0132* 0.00563* 0.0132* 0.00563* 0.0132* 0.0054*	E_mae4	0.0016	-0.00063	0.0019	0.0108	-0.0002	0.0031	0.0016	-4.65e-06	0.0071	0.0040	0.0048	-0.0023
0.32080 -0.07152 0.01794 0.8716* -0.0878* -0.0663* 0.1181* -0.0704* 0.1110*** 0.0294* -0.0694* 0.35243 -0.05325 0.03856 0.9167* -0.0468* -0.0515* 0.1529* -0.0527* 0.1359* 0.0451* -0.0648* 0.36238 -0.02764 0.0466 0.9392* -0.0031 -0.0524* 0.1604* -0.0139* 0.0459* -0.0648* -0.0548* 0.1414* -0.0149** 0.1408** 0.0559* -0.0576* -0.0576* 0.34520 -0.0126 0.04624 0.9109* -0.0048 -0.0518* 0.1414* -0.0103* 0.1408** 0.0359* -0.0563* -0.0044* -0.0144** 0.0144** 0.0144** -0.0156* 0.0564* -0.0549* -0.0540* 0.0540* -0.0044* 0.0044* 0.0164* 0.0164* 0.0164* 0.0164* 0.0164* 0.0164* 0.0166* 0.0164* 0.0164* 0.0164* 0.0164* 0.0164* 0.0164* 0.0166* 0.0164* 0.0164* 0.0166*	Renda00	0.31398	-0.07255	0.0075	0.8756*	-0.1121**	-0.0762*	0.1132*	-0.0619*	0.0965***	0.0029	-0.0598*	0.0139
0.35243 -0.05325 0.03856 0.9167* -0.0468* -0.0515* 0.1529* -0.0527* 0.1359** 0.0451* -0.0648* 0.36238 -0.02764 0.0466 0.9392* -0.0031 -0.0524* 0.1604* -0.0149** 0.1429** 0.0559* -0.0576* 0.34520 -0.01205 0.04624 0.9109* -0.0048 -0.0518* 0.1414* -0.0104** 0.1408** 0.0341 -0.0563* 0.3416 -0.0089 0.06258 1.0233* -0.0045 -0.0144* -0.0104** 0.1408** 0.0150* -0.0150* 0.0005 0.02750 0.06278 1.0233* 0.0254* -0.0144* 0.0455* 0.0165* 0.0150* 0.0164* 0.0005 0.0807 0.09219 -0.0540* 0.1207* 0.0540* -0.0006 0.0914* 0.1002* 0.0150 0.0259* 0.0150 0.0259* 0.0150 0.0259* 0.0150 0.0259* 0.0150 0.0259* 0.0150 0.0250 0.0150 0.0254* 0.0160	Renda01	0.32080	-0.07152	0.0179	0.8716*	-0.0878*	-0.0663*	0.1181*	-0.0704*	0.1110***	0.0294*	-0.0694*	0.0157
0.36238	Renda02	0.35243	-0.05325	0.03856	0.9167*	-0.0468*	-0.0515*	0.1529*	-0.0527*	0.1359**	0.0451*	-0.0648*	0.0365
0.34520 -0.01205 0.04624 0.9109* -0.00518* 0.1414* -0.01044*** 0.1408** 0.0341 -0.0263* 0.39146 -0.00889 0.06258 1.0233* -0.0063 0.0745 0.2959* -0.0035 0.1707** 0.0150 0.0299* 0.0005 0.02750 0.06376 -0.0540* 0.1207* 0.0540* 0.00904* 0.1002* 0.0100* 0.0914* 0.1002* 0.0150 0.0299* 0.0005 0.0807 0.09219 -0.0546* 0.3368* 0.1360* 0.5554* 0.4567* 0.1399* 0.4987* 0.5563*	Renda03	0.36238	-0.02764	0.0466	0.9392*	-0.0031	-0.0524*	0.1604*	-0.0303*	0.1429**	0.0559*	-0.0576*	0.0581
0.39146 -0.00889 0.06258 1.0233* -0.0063 -0.0745 0.2959* -0.0035 0.1707** -0.0150 -0.0299*	Renda04	0.34520	-0.01205	0.04624	0.9109*	-0.0048	-0.0518*	0.1414*	-0.0104***	0.1408**	0.0341	-0.0263*	0.0720
-0.0156 0.02750 0.06376 -0.0530** 0.0202 0.0545* -0.0104 0.0455* 0.0837* 0.0116 0.0464* 0.0005 0.0807 0.09219 -0.0540 0.1207* 0.0540* -0.0006 0.0914* 0.1002* 0.0150 0.0299* 0.1976 6.4667 6.2148 5.5466* 6.3368* 6.1360* 6.5554* 6.4567* 6.1399* 6.4987* 6.5563*	Renda05	0.39146	-0.00889	0.06258	1.0233*	-0.0063	-0.0745	0.2959*	-0.0035	0.1707**	-0.0150	-0.0299*	0.0502
-0.0005 0.0807 0.09219 -0.0540 0.1207* 0.0540* -0.0006 0.0914* 0.1002* 0.0150 0.0299* 0.1976 6.4667 6.2148 5.5466* 6.3368* 6.1360* 6.5554* 6.4567* 6.1399* 6.4987* 6.5563*	CSE	-0.0156	0.02750	0.06376	-0.0530**	0.0202	0.0545*	-0.0104	0.0455*	0.0837*	0.0116	0.0464*	0.0464*
6.1976 6.4667 6.2148 5.5466* 6.3368* 6.1360* 6.5254* 6.4567* 6.1399* 6.4987* 6.5563*	Infra	-0.0005	0.0807	0.09219	-0.0540	0.1207*	0.0540*	-0.0006	0.0914*	0.1002*	0.0150	0.0299*	0.1339*
	Constante	6.1976	6.4667	6.2148	5.5466*	6.3368*	6.1360*	6.5254*	6.4567*	6.1399*	6.4987*	6.5563*	6.3764*

Elaboração dos autores. Obs.: Significância (*) = 1%; (***) = 5%; (***) = 10%.

TABELA B.3 Estimativas da regressão quantílica incondicional por escola: ciências da natureza

	,											
		MQO			0.1			0.5			6.0	
Validvels	EEEPs	Particulares	Regulares	EEEPs	Particulares	Regulares	EEEPs	Particulares	Regulares	EEEPs	Particulares	Regulares
Idade	-0.01619	-0.00665	-0.00141	-0.014*	-0.0111*	*8000.0-	-0.0208*	-0.0074*	-0.0017*	-0.008*	-0.0011*	-0.0020*
Sexo	0.04458	0.030123	0.01743	0.0300*	0.0091***	0.0035**	0.0564*	0.0409*	0.0174*	0.0453*	0.0363*	0.0325*
Estado civil	0.02499	-0.01561	-0.00872	0.0509	-0.0451	-0.0116**	0.0251	-0.0119	-0.0131*	0.0093	-0.0210***	-0.0032
Raça	0.00415	0.008923	0.00697	-0.0042	0.0019	0.0030	0.0089***	0.0106*	0.0075*	0.0022	0.0183*	*6600.0
Anos_Fund	0.01332	0.029400	0.01038	0.0149*	0.0218*	0.0062*	0.0138*	0.0389*	*8600.0	0.0190*	0.0232*	0.0180*
Anos_Médio	0.00289	0.008523	0.00087	0.0049	0.0066	0.0004	0.0025	0.0061	-0.0002	-0.0000	0.0145*	0.0015
E_Pai1	0.02350	-0.045682	-0.01941	0.0436	-0.0296	-0.0137	0.0323	-0.0582*	-0.0275**	-0.0184	-0.0375*	-0.034***
E_Pai2	0.02603	-0.042944	-0.01779	0.0449	-0.0325*	-0.016***	0.0405***	-0.0457*	-0.022***	-0.0223	-0.0396*	-0.035***
E_Pai3	0.03070	-0.036050	-0.00682	0.0517	-0.0286*	-0.0093	0.0474**	-0.0370*	-0.0116	-0.0198	-0.0381*	-0.0140
E_Pai4	0.032032	-0.009626	0.00791	0.0464	-0.0120	-0.00007	0.0517**	-0.0042	0.0073	-0.0075	-0.0053	0.0074
E_mae1	-0.00394	-0.036917	-0.03318	0.0182	-0.0971**	-0.0246*	-0.0156	-0.0113	-0.0346*	-0.0140	-0.0117	-0.0428*
E_mae2	-0.0054635	-0.029006	-0.02840	0.0230	-0.0222**	-0.011***	-0.021***	-0.0367*	-0.0323*	-0.0116	-0.0267*	-0.0400*
E_mae3	-0.0016787	-0.0186841	-0.01423	0.0218	-0.0076	-0.0042	-0.0171	-0.0226*	-0.0171**	-0.0025	-0.0167**	-0.0215**
E_mae4	0.0018264	-0.0012844	-0.00943	0.0268	0.0030	-0.0054	-0.0122	-0.0005	-0.0108	-0.0034	0.0025	-0.0119
Renda00	0.0653425	-0.090357	0.01734	-0.114*	-0.0559	0.0657	0.1368*	-0.1103*	0.0022	0.0216	-0.1152*	0.0776*
Renda01	0.0797549	-0.09416	0.02013	-0.113*	*6590.0-	0.0644	0.1526*	-0.1034*	0.0113	0.0237*	-0.1081*	0.0807*
Renda02	0.1014531	-0.07831	0.03431	-0.107*	-0.0297*	0.0729	0.1816*	-0.0892*	0.0269	0.0498*	-0.1114*	0.0993*
Renda03	0.1083583	-0.04581	0.04464	-0.092*	-0.0046	0.0744	0.1864*	-0.0481*	0.0354	*6790.0	-0.0938*	0.1272*
Renda04	0.0788905	-0.02117	0.05204	-0.081*	-0.0027	0.0816	0.1653*	-0.0193**	0.0292	0.0000	-0.0526*	0.1455*
Renda05	0.2043453	-0.01205	0.03445	-0.0117	-0.0120	0.0707	0.1891	-0.0120	0.0642	0.2761	-0.0336**	0.0924
CSE	-0.012838	0.05229	0.0318113	-0.0133	0.0566*	0.0108***	-0.0416*	0.0552*	0.0386*	0.0026	0.0681*	0.0363*
Infra	0.00586	0.13008	0.086118	-0.0084	0.1288*	0.0555*	-0.0109	0.1705*	0.0868*	0.0111	0.0368**	0.1292*
Constante	6.31781	6.401588	6.11947	6.2495*	6.2488*	5.9143*	6.3473*	*6068.9	6.1419*	6.4610*	6.5434*	6.2345*

Elaboração dos autores. Obs.: Significância (*) = 1%; (**) = 5%; (***) = 10%.

 ${\sf TABELA~B.4} \\ {\sf Estimativas~da~regress\~ao~quant\'ilica~incondicional~por~escola: linguagens~e~c\'odigos} \\$

vallavels		MQO			0.1			0.5			6.0	
	EEEPs	Particulares	Regulares	EEEPs	Particulares	Regulares	EEEPs	Particulares	Regulares	EEEPs	Particulares	Regulares
Idade	-0.0183	-0.00688	-0.00312	-0.0291*	-0.0134*	-0.0034*	-0.0210*	-0.0057*	-0.0035*	-0.0101*	-0.0010*	-0.0024*
Sexo	-0.0091	-0.01930	-0.02130	-0.0265*	-0.0347*	-0.0361*	-0.0035	-0.0167*	-0.0232*	0.0018	-0.0017	-0.0070*
Estado civil	0.03945	0.00691	-0.01398	0.0715	0.0148	-0.0361*	0.0347***	-0.0004	-0.0091**	-0.0036	-0.0097	-0.0041
Raça	0.00595	0.00489	0.00851	0.0044	0.0057	0.0100*	0.0051	0.0060**	0.0084*	**9600.0	*6900.0	*0800.0
Anos_Fund	0.01377	0.01925	0.01065	0.0116**	0.0241*	0.0017	0.0156*	0.0198*	0.0119*	0.0123*	0.0114*	0.0132*
Anos_Médio	-0.0012	0.00358	-0.00225	-0.0069	0.0118**	-0.0020	0.0016	0.0015	-0.003***	-0.0014	0.0009	-0.0015
E_Pai1	-0.0107	-0.03899	-0.02176	-0.0454	-0.0593**	-0.0436*	-0.0093	-0.0438*	-0.0141	-0.0020	-0.0393*	-0.028***
E_Pai2	-0.0038	-0.03556	-0.01487	-0.0204	-0.0273*	-0.0324**	-0.0116	-0.0411*	-0.0095	0.0011	-0.0374*	-0.027***
E_Pai3	0.0078	-0.0255	0.00750	-0.0111	-0.0220*	-0.0098	0.0044	-0.0274*	0.0174	0.0087	-0.0302*	-0.0061
E_Pai4	-0.0018	-0.00943	0.01352	-0.0354	-0.0100	-0.0135	-0.0027	-0.0027	0.0248***	0.0033	-0.0140**	0.0025
E_mae1	-0.0041	-0.00460	-0.03291	-0.0166	-0.0335	-0.0060	0.0055	0.0055	-0.0450*	-0.0059	0.0028	-0.0352*
E_mae2	-0.0064	-0.01651	-0.02888	-0.0022	-0.0234**	-0.0104	-0.0014	-0.0014	-0.0355*	-0.0117	-0.0093**	-0.0310*
E_mae3	0.00486	-0.0100	-0.00791	0.0146	-0.0101	0.0101	0.0109	0.0109	-0.0113	-0.0052	-0.0077***	-0.0135
E_mae4	-0.0033	-0.00050	-0.00458	0.0048	-0.0020	0.0068	0.0012	0.0012	-0.0100	-0.0085	0.0037	-0.0037
Renda00	0.17614	-0.06865	0.03453	0.8101*	-0.0216	0.1040	0.1198*	0.1198*	-0.0060	0.0121***	*0690.0-	0.0466*
Renda01	0.19784	-0.06354	0.04222	0.8470*	-0.0484*	0.1124	0.1353*	0.1353*	0.0056	0.0378*	-0.0514*	0.0582*
Renda02	0.22196	-0.043608	0.06379	0.8742*	-0.0125	0.1340	0.1634*	0.1634*	0.0305	0.0652*	-0.0512*	0.0771*
Renda03	0.23301	-0.02629	0.06658	0.8804*	0.0009	0.1269	0.1753*	0.1753*	0.0321	0.0728*	-0.0399*	0.1011*
Renda04	0.22363	-0.013051	0.07079	*0968.0	0.0028	0.1492	0.1512*	0.1512*	0.0275	0.0449**	-0.0191**	0.1208*
Renda05	0.27203	-0.011407	0.07223	0.9740*	0.0001	0.1046	0.1610	0.1610	0.0211	0.0073	-0.0258*	0.1315**
CSE	-0.0086	0.032172	0.06686	-0.0246	0.0473*	0.0795*	-0.0065	-0.0065	0.0765*	-0.0002	0.0389*	0.0401*
Infra	0.04038	0.04917	0.1042	0.0588***	0.0636*	0.0553*	0.0423**	0.0423**	0.1152*	0.0368**	0.0086	0.1219*
Constante	6.3137	6.4396	6.1766	5.6827*	6.3425*	5.9520*	6.4370*	6.4370*	6.2322*	6.4977*	6.4973*	6.3135*

Elaboração dos autores. Obs.: Significância (*) = 1%; (**) = 5%; (***) = 10%.

 $\label{eq:total_total} \text{TABELA B.5}$ Estimativas da regressão quantilica incondicional por escola: matemática

EEEPs Particulares Regulares CORDA -0.00427 -0.00147 -0.00428 -0.01047 -0.00428 -0.00407 -0.00324 -0.00324 -0.00407 -0.00324 -0.00324 -0.00407 -0.00408 -0.00407 -0.00324 -0.00407 -0.00407 -0.00324 -0.00407 -0.00407 -0.00324 -0.00407			MQO			0.1			0.5			6.0	
- 0.03228 - 0.01032 - 0.02040 - 0.0140* - 0.0140* - 0.0040* - 0.0020* - 0.0032* - 0.0032* - 0.0032* - 0.0032* - 0.0180* - 0.0160* - 0.0032* - 0.0032* - 0.0032* - 0.0180* - 0.0180* - 0.0108* - 0.0183* - 0.0033* - 0.0034* - 0.0184* - 0.0034* <t< th=""><th>Validveis</th><th>EEEPs</th><th>Particulares</th><th>Regulares</th><th>EEEPs</th><th>Particulares</th><th>Regulares</th><th>EEEPs</th><th>Particulares</th><th>Regulares</th><th>EEEPs</th><th>Particulares</th><th>Regulares</th></t<>	Validveis	EEEPs	Particulares	Regulares									
10,000 1	Idade	-0.03228	-0.01073	-0.00342	-0.0270*	-0.0155*	-0.0014*	-0.0428*	-0.0140*	-0.0040*	-0.0204*	-0.0023*	-0.0052*
1	Sexo	0.09407	0.07791	0.05372	0.0537*	0.0534*	0.0282*	0.1180*	*9960.0	0.0556*	0.1058*	0.0833*	*0870.0
dio 0.00266 0.01344 0.00697 0.00294 0.00026 0.01344 0.00697 0.00126* 0.0134* 0.00039 0.0039* 0.00034 0.00034 0.00148* 0.00649* 0.00369 0.00394* 0.00036 0.00144* 0.0026 0.00148* 0.00364 0.00149* 0.00148* 0.00149* 0.00149* 0.00036 0.00149*	Estado civil	-0.00199	-0.04258	-0.01265	0.0270	-0.0806	-0.0069	-0.0197	-0.0647**	-0.0187*	-0.0159	-0.0423**	-0.0133**
dig 0.02065 0.04452 0.01712 0.0089 0.0199* 0.0037*** 0.0238* 0.01644* 0.0264* 0.0264* 0.0264* 0.0264* 0.0264* 0.0284* 0.0238* 0.00144* 0.0036 0.00194 0.00199 0.00196 0.0018 0.0018* 0	Raça	0.00296	0.01344	0.00697	0.0021	0.0034	0.0004	0.0001	0.0126**	0.0087*	-0.0020	0.0239*	0.0136*
dio 0.00136 0.00036 0.00036 0.00046 0.00048 0.00038 0.00138 0.00038 0.00138 0.00138 0.00138 0.00138 0.00138 0.00138 0.00138 0.00138 0.00138 0.00144 0.	Anos_Fund	0.02065	0.04652	0.01712	0.0089	0.0199*	0.0037***	0.0278*	0.0605*	0.0164*	0.0264*	0.0458*	0.0332*
0.00360 -0.01340 -0.01340 -0.00136 -0.01340 -0.00136 -0.01348 -0.00226 -0.0328 -0.01341 -0.00341 -0.00144*** -0.00340 -0.00340 -0.01513 -0.01513 -0.01513 -0.01513 -0.01513 -0.01513 -0.01514 -0.00340 -0.01513 -0.01514 -0.00340 -0.00340 -0.00340 -0.00340 -0.00340 -0.00340 -0.00340 -0.01579 -0.00380 -0.01579 -0.00380 -0.01579 -0.00380 -0.01579 -0.00380 -0.01579 -0.00380 -0.01579 -0.00380 -0.01579 -0.00380 -0.01579 -0.00380 -0.01579 -0.00380 -0.01634 -0.00446 -0.00381 -0.00380 -0.01634 -0.0184 -0.00381 -0.01840 -0.00381 -0.01840 -0.00381 -0.01840 -0.00381 -0.01840 -0.00381 -0.01840 -0.00381 -0.01840 -0.00381 -0.01840 -0.00381 -0.01840 -0.00381 -0.01840 -0.00381 -0.01840 -0.01840 -0.00381 -0.01840	Anos_Médio	-0.00138	0.01945	-0.00036	0.0029	0.0090	-0.0016	-0.0084	0.0238*	0.0006	0.0012	0.0135**	-0.0028
0.00785 0.0538 0.0151 0.0144*** 0.0040 0.026 -0.0688* -0.0210 0.0034 -0.0034 -0.0034 -0.0034 -0.0034 -0.0034 -0.0034 0.0048 -0.0034 -0.0034 -0.0034 -0.0034 -0.0034 -0.0034 -0.0034 -0.0034 -0.0034 -0.0034 -0.0034 -0.0034 -0.0034 -0.0244 -0.0034 -0.0284 -0.0349 -0.0377 -0.0034 -0.0374 -0.0354 -0.0354 -0.0354 -0.0354 -0.0354 -0.0354 -0.0354 -0.0354 -0.0354 -0.0364 -0.0364 -0.0344	E_Pai1	-0.00360	-0.05130	-0.01742	0.0085	-0.0139	0.0002	-0.0202	-0.0738*	-0.0226	0.0128	-0.0431***	-0.0014
0.01911 -0.04084 -0.00312 0.0214 -0.0034 0.0075 0.0163 -0.0548 -0.0034 0.0033 -0.00840 -0.0031 -0.00840 -0.00444 -0.0038 -0.0255 -0.0157 -0.0279 -0.0279 -0.0254 -0.0057 -0.0284 -0.0284 -0.0284 -0.0284 -0.0284 -0.0284 -0.0284 -0.0284 -0.0284 -0.0284 -0.0284 -0.0274 -0.0567 -0.0577 -0.0277 -0.0377 <	E_Pai2	0.00785	-0.05358	-0.01513	0.0125	-0.0214***	0.0047	-0.0026	-0.0688*	-0.0210	0.0238	-0.0736*	-0.0004
0.0332 -0.00840 0.00484 0.0454 0.0068 0.0055 -0.0167 -0.0165 -0.0167 -0.0058 -0.0478 -0.00478 -0.0284 -0.0281 -0.0344 -0.0324 -0.0354 -0.0579 -0.05779 -0.05779 -0.01775 -0.0347 -0.0446* -0.0281 -0.0281 -0.0458 -0.0446* -0.0281 -0.0260 -0.0629* -0.04417* -0.0577	E_Pai3	0.01911	-0.04084	-0.00312	0.0214	-0.0034	0.0075	0.0163	-0.0548*	-0.0094	0.0178	-0.0603*	0.0204
-0.01755 -0.0497 -0.05779 -0.0287 -0.0287 -0.05779 -0.02779 -0.02879 -0.0497 -0.0497 -0.0477 -0.02877 -0.0287 -0.0287 -0.0560 -0.0287 -0.0412* -0.0560 -0.0560 -0.0580 -0.0448 -0.0231 -0.0240 -0.0469 -0.0412* -0.0469 -0.0412* -0.0469 -0.0412* -0.0469 -0.0412* -0.0469 -0.0412* -0.01634 -0.01834 -0.01834 -0.01834 -0.01834 -0.01834 -0.01834 -0.01634 -0.01862* -0.0089 -0.04055 -0.04054 -0.01634 -0.01862* -0.01864 -0.01864 -0.01864 -0.01864 -0.01874	E_Pai4	0.0332	-0.00840	0.00484	0.0454	0.0068	0.0057	0.0255	-0.0167	-0.0058	0.0544	-0.0089	0.0497
-0.01201 -0.05807 -0.04708 0.0061 -0.0446* -0.0231* -0.02609 -0.0412* -0.0166 -0.0590* -0.0405 -0.0412* -0.0166 -0.0590* -0.0590* -0.0405 -0.0405 -0.0405 -0.0405 -0.0405 -0.0405 -0.0405 -0.0405 -0.0405 -0.0405 -0.0153 -0.0153 -0.0153 -0.0153 -0.0153 -0.0154 -0.0154 -0.0156 -0.0160 -0.0153 -0.0154	E_mae1	-0.01755	-0.0497	-0.05779	-0.0127	-0.0289	-0.0281*	-0.0374	-0.0562	-0.0573*	-0.0277	-0.0577*	-0.1057*
0.00469 0.0331 0.1957 0.0198*** 0.0084 0.0016 0.0140 0.0045 0.0141 0.00527* 0.0527* 0.00002 0.00030 0.1634 0.0330 -0.0130 0.0160 0.016 0.017 0.015 0.0155 0.0155 0.0155 0.0153 0.0153 0.0153 0.0153 0.01573 0.	E_mae2	-0.01201	-0.05807	-0.04708	0.0061	-0.0446*	-0.0231*	-0.0260	-0.0629*	-0.0412*	-0.0166	-0.0590*	-0.0963*
0.00002 0.01634 0.0036 -0.0130 -0.0076 -0.0166 0.0017 -0.0152 0.0165 -0.0056 -0.0056 -0.0154 0.0015 -0.0153* -0.0153* -0.0153* -0.0153* -0.01518* -0.01528* -0.01528* -0.01528* -0.01529* -0.01	E_mae3	0.00469	-0.03321	-0.01957	0.0198	-0.0178***	-0.0084	0.0018	-0.0405*	-0.0141	0.0054	-0.0527*	-0.0531*
0.13414 -0.11765 0.04348 -0.1262* -0.0261 0.0553 0.1573* 0.0536 0.0330 -0.1518* 0.1555 -0.16037 0.08341 -0.1390* -0.0810* 0.0955 0.0202* 0.0647 0.0647 0.0559** -0.1607* 0.2026 -0.1364 0.10514 -0.1334 0.1134* -0.0134 0.0954* 0.0957 0.0877 0.0826* -0.1539* -0.1529* 0.21470 -0.09055 0.11875 -0.1134* -0.0144 0.0892* -0.0444* 0.0947* 0.0977 0.0826* -0.1529* 0.17182 -0.04954 0.11875 -0.0144 0.0893 0.2182* -0.0484* 0.1047 0.0755 -0.0848** 0.19145 -0.0257 0.0267 0.0011 0.1116 0.0136 0.0256 0.0433** 0.0239 0.0169 0.0162* 0.0169 0.1162** 0.0458 0.0175 0.0448** 0.0175 0.0448** 0.0176 0.0448** 0.0176 0.0448** 0.0176 0.0448*	E_mae4	-0.00002	0.00030	-0.01634	0.0030	-0.0130	-0.0076	-0.0160	0.0017	-0.0152	0.0165	-0.0059	-0.037***
0.1550	Renda00	0.13414	-0.11765	0.07438	-0.1262*	-0.0261	0.0753	0.1580*	-0.1573*	0.0536	0.0300	-0.1518*	0.0073
0.20208	Renda01	0.15550	-0.16037	0.08341	-0.1390*	-0.0810*	0.0797	0.2055*	-0.2002*	0.0647	0.0259**	-0.1607*	0.0206
0.21470	Renda02	0.20208	-0.1394	0.105161	-0.1123*	-0.0537*	0.0926	0.2546*	-0.1646*	0.0877	0.0826*	-0.1751*	0.0542
0.17182 0.04954 0.13176 0.1326* 0.0087 0.1076 0.2182* 0.0484* 0.1047 0.0775 0.0840* 0.19145 0.02332 0.16781 0.0267 0.0011 0.1316 0.2556 0.0169 0.1612*** 0.0458 0.0488*** 0.0947 0.07704 0.06268 0.0261 0.0433*** 0.0343* 0.0245** 0.0093 0.2219* 0.0769* 0.0763* 0.0365* 0.01488 0.18425 0.09563 0.0968 0.1327* 0.0245** 0.0696 0.5513* 0.0769* 0.0343 0.0665* 0.05249 0.09563 0.0356 0.0356 0.0357 0.0357 0.0353 0.0655*	Renda03	0.21470	-0.09055	0.11875	-0.1134*	-0.0144	0.0893	0.2892*	-0.0947*	0.0921	0.0993*	-0.1529*	0.0818
0.19145 -0.0233	Renda04	0.17182	-0.04954	0.13176	-0.1426*	-0.0087	0.1076	0.2182*	-0.0484*	0.1047	0.0775	-0.0840*	0.1072
0.00407 0.07704 0.06268 0.0261 0.0433*** 0.0343* -0.0239 0.0828* 0.0748* 0.0157 0.1490* 0.01488 0.18425 0.09563 -0.0087 0.1327* 0.0245*** 0.0093 0.2219* 0.0769* 0.0343 0.0665* 6.50249 6.4935 6.0235 6.4286* 6.5253* 5.7783* 6.6696* 6.5513* 6.0425* 6.6641* 6.7093*	Renda05	0.19145	-0.02332	0.16781	-0.0267	0.0011	0.1316	0.2556	-0.0169	0.1612***	-0.0450	-0.0488***	0.1753
0.01488 0.18425 0.09563 -0.0087 0.1327* 0.0245*** 0.0093 0.2219* 0.0769* 0.0343 0.0665* 6.52249 6.4935 6.0235 6.4286* 6.2523* 5.7783* 6.6696* 6.5513* 6.0425* 6.6641* 6.7093*	CSE	0.00407	0.07704	0.06268	0.0261	0.0433***	0.0343*	-0.0239	0.0828*	0.0748*	0.0157	0.1490*	0.0955*
6.50249 6.4935 6.0235 6.4286* 6.2523* 5.7783* 6.6696* 6.5513* 6.0425* 6.6641* 6.7093*	Infra	0.01488	0.18425	0.09563	-0.0087	0.1327*	0.0245***	0.0093	0.2219*	*69/0.0	0.0343	0.0665*	0.2022*
	Constante	6.50249	6.4935	6.0235	6.4286*	6.2523*	5.7783*	*9699.9	6.5513*	6.0425*	6.6641*	6.7093*	6.3245*

Elaboração dos autores. Obs.: Significância (*) = 1%; (**) = 5%; (***) = 10%.

TABELA B.6 Estimativas da regressão quantílica incondicional por escola: redação

)	•		•		n						
		MQO			0.1			0.5			6.0	
Validvels	EEEPs	Particulares	Regulares	EEEPs	Particulares	Regulares	EEEPs	Particulares	Regulares	EEEPs	Particulares	Regulares
Idade	-0.04414	-00.01073	-0.01197	-0.0749*	-0.0605*	-0.0156*	-0.0356*	-0.0103*	-0.0130*	-0.0227*	-0.0034*	-0.0052*
Sexo	-0.04618	00.07791	-0.10008	-0.0602*	*8660.0-	-0.1309*	-0.0471*	-0.0835*	-0.1050*	-0.0402*	-0.0589*	-0.0597*
Estado civil	0.04923	-00.04258	-0.04391	0.0853	0.0314	-0.1038*	0.0605	0.0078	-0.030**	0.0606	-0.0434***	-0.0015
Raça	0.00923	00.01344	0.01889	0.0485*	0.0263**	0.0130***	0.0037	0.0121**	0.0193*	-0.0041	0.0132**	0.0207*
Anos_Fund	0.03442	00.04652	0.02935	0.0505*	0.0429*	0.0383*	0.0264*	*0650.0	0.0334*	0.0229*	0.0269*	0.0248*
Anos_Médio	-0.00039	00.01945	-0.00145	0.0263***	0.0230***	-0.0198*	0.0000	0.0136**	0.0026	-0.0120	0.0131**	0.0068***
E_Pai1	-0.00842	-00.05130	-0.10362	-0.0352	-0.1461***	-0.0728**	-0.0238	-0.0775*	-0.1067*	-0.0584	-0.0513**	-0.0704**
E_Pai2	0.01465	-00.05358	-0.07663	0.0117	-0.1112*	-0.0408	-0.0113	-0.0677*	-0.083**	-0.0233	-0.0408*	-0.062***
E_Pai3	0.04047	-00.04084	-0.01195	0.0477	*69/0.0-	0.0244	0.0078	-0.0352*	-0.0075	0.0008	-0.0288**	-0.0185
E_Pai4	0.02934	-00.00840	-0.00007	0.0489	-0.0537*	0.0217	-0.0034	-0.0152	-0.0022	-0.0132	-0.0143	0.0090
E_mae1	-0.00935	-00.0497	-0.08757	-0.0880	-0.3109**	-0.0734*	-0.0125	-0.0202	-0.1065*	0.0266	-0.0099	-0.1103*
E_mae2	0.00237	-00.05807	-0.07848	-0.0065	-0.0762*	-0.0516*	-0.0042	-0.0380*	*8680.0-	0.0146	-0.0470*	-0.1048*
E_mae3	0.02456	-00.03321	-0.02257	0.0204	-0.0073	0.0032	0.0205	-0.0264*	-0.0272	0.0302	-0.0373*	-0.0668*
E_mae4	0.01951	00.00030	-0.00823	0.0223	-0.0088	0.0024	0.0030	-0.0162***	-0.0014	0.0340	-0.0166	-0.042***
Renda00	0.23637	-00.11765	0.06942	-0.2363*	-0.1798	0.2184	0.2622*	-0.0813***	-0.1041	0.0601**	-0.1281*	0.0045
Renda01	0.26106	-00.16037	0.09074	-0.2192*	-0.1468*	0.2403	0.2844*	-0.0867*	-0.0905	0.1211*	-0.1045*	0.0195
Renda02	0.30796	-00.1394	0.13654	-0.1581*	-0.0932*	0.2843	0.3227*	-0.0583*	-0.0375	0.1689*	-0.0966*	0.0477
Renda03	0.31917	-00.09055	0.16031	-0.1710*	-0.0148	0.2761	0.3301*	-0.0273**	-0.0046	0.2225*	-0.0682*	0.0897
Renda04	0.29146	-00.04954	0.11805	-0.3329*	-0.0224	0.1828	0.3209*	-0.0025	-0.0305	0.2209*	-0.0353***	0.1064
Renda05	0.34757	-00.02332	0.18990	0.0329	-0.0071	0.1755	0.3123***	-0.0114	0.0663	0.0324**	-0.0643*	0.2382
CSE	0.01120	00.07704	0.20595	0.0946***	0.1780*	0.2451*	0.0073	0.1106*	0.2328*	-0.0335	0.0959*	0.1102*
Infra	-0.00549	00.18425	0.22785	0.0535	0.2588*	0.2727*	-0.0392	0.2269*	0.2204*	-0.0403	0.1097*	0.1832*
Constante	6.68635	6.49352	6.20832	7.2347*	7.0221*	5.6730*	6.6280*	6.5642*	6.4721*	6.8421*	6.8747*	6.5846*

Elaboração dos autores. Obs.: Significância (*) = 1%; (**) = 5%; (***) = 10%.

APÊNDICE C

BALANCEAMENTO POR ENTROPIA

Considere amostras aleatórias extraídas de unidades de tratados e de controle, nas quais se inserem as definições a seguir.

- 1) D_i é uma variável binária que assume o valor 1, se a unidade pertence ao tratamento, e 0, se pertence ao controle.
- 2) X é uma matriz composta pelos elementos x_{ij} , referentes aos valores da variável exógena predeterminada j na unidade i.
- 3) A densidade das covariadas nas amostras de tratamento e controle são dadas por $f_{X|D_i=1}(x)$ e $f_{X|D_i=0}(x)$, respectivamente.
- 4) A variável resultado observado é expressa por $Y_i = Y_i(1)D_i + (1 D_i)Y_i(0)$.

O efeito médio do tratamento sobre os tratados (EMTT) é dado por $\tau = E[Y_i(1)|D_i=1] - E[Y_i(0)|D_i=1]$, em que a primeira média pode ser diretamente identificada do grupo de tratados, mas a segunda corresponde ao contrafactual, o qual não é observado. Rosenbaum e Rubin (1983) apontam que, assumindo seleção nos observáveis, $Y(0) \perp D|X$, e sobreposição, $Pr(D_i=1|X_i=x) < 1$ para todo x no suporte de $f_{X|D_i=1}$, o EMTT é identificado como:

$$\tau = E[Y_i|D_i = 1] - \int E[Y_i|X_i = x, D_i = 0] f_{X|D_i = 1}(x) dx.$$
 (1)

Para estimar o contrafactual, a distribuição da covariável no grupo de controle necessita ser ajustada para torná-la semelhante à distribuição no grupo de tratamento, tal que o indicador de tratamento D se torne mais perto de ser ortogonal em relação às covariáveis. Uma variedade de métodos de pré-processamento de dados – tais como pareamento pelo vizinho mais próximo, cem e escore de propensão – tem sido proposta para reduzir o desequilíbrio na distribuição de variáveis de controle. Uma vez ajustadas, as distribuições de variáveis independentes e os métodos de análise padrão, tais como a regressão, podem ser posteriormente utilizados para estimar o tratamento com menor erro e o modelo de dependência (Imbens, 2004; Rubin, 2006; Ho *et al.*, 2007; Sekhon, 2009).

Considere o caso mais simples em que o efeito do tratamento nos dados préprocessados é estimado usando a diferença nos resultados médios entre os grupos de tratados e o controle ajustado, cujo método muito utilizado na literatura é o escore de propensão ponderado (Hirano, Imbens e Ridder, 2003). A média contrafactual é estimada como segue:

$$E[Y_i(0)|D_i = 1] = \frac{\sum_{\{i|D = 0\}} Y_i d_i}{\sum_{\{i|D = 0\}} d_i}$$
 (2)

As unidades de controle recebem um peso dado por $d_i = \frac{\hat{p}(x_i)}{1-\hat{p}(x_i)}$, $\hat{p}(x_i)$ na equação (2), que é o escore de propensão, o qual é comumente estimado através de uma regressão *probit* ou *logit*. Se este modelo estiver corretamente especificado, então o peso estimado d_i assegura que a distribuição da covariável das unidades de controle reponderadas corresponda à distribuição no grupo de tratamento. No entanto, na prática, essa abordagem muitas vezes não consegue equilibrar conjuntamente todas as covariáveis.

O balanceamento por entropia generaliza a abordagem de ponderação do escore de propensão, ao estimar os pesos diretamente de um conjunto de restrições de equilíbrio que exploram o conhecimento do pesquisador sobre os momentos amostrais. Considere W_i o peso do balanceamento por entropia escolhido para cada unidade de controle, o qual foi encontrado pelo seguinte esquema de reponderação que minimiza a distância métrica de entropia:

$$\min_{w_i} H(w) = \sum_{\{i|D=0\}} w_i \log(w_i/q_i)$$
(3)

Sujeito às restrições de equilíbrio e normalização:

$$\sum_{\{i|D=0\}} w_i c_{ri}(X_i) = m_r \qquad com \, r \in 1, ..., R$$
(4)

$$\sum_{\{i|D=0\}} w_i = 1 \tag{5}$$

$$w_i \ge 0$$
 para todo i, tal que $D = 0$, (6)

em que $q_i=1/n$ é um peso base, sendo n o tamanho da amostra das unidades de controle, e $c_{ri}(X_i)=m_r$ descreve um conjunto de R restrições referentes aos momentos das covariadas no grupo de controle reponderados. Inicialmente, escolhe-se a covariada que será incluída na reponderação. Para cada covariada, especifica-se um conjunto de restrições de balanceamento (equação 4) para equiparar os momentos das distribuições das covariadas entre os grupos de tratamento e controles reponderados. As restrições de momentos podem ser a média (primeiro momento), a variância (segundo momento) e a assimetria (terceiro momento).

Dessa forma, o balanceamento por entropia procura, para um conjunto de unidades, pesos $W = [w_i, ..., w_{n_0}]'$, nos quais minimiza a equação (3), distanciase de entropia entre W e o vetor-base de pesos $Q = [q_i, ..., q_{n_0}]'$ e sujeita as

restrições de balanceamento (equação 4), de normalização (equação 5) e de não negatividade (equação 6). Devido à sua propriedade de minimizar os desequilíbrios entre os grupos, esse método foi aqui aplicado.

REFERÊNCIAS

HIRANO, K.; IMBENS, G.; RIDDER, G. Efficient estimation of average treatment effects using the estimated propensity score. **Econometrica**, v. 71, n. 4, p. 1161-1189, 2003.

HO, D. *et al.* Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. **Political Analysis**, v. 15, n. 3, p. 199-236, 2007.

IMBENS, G. Nonparametric estimation of average treatment effects under exogeneity: a review. **Review of Economics and Statistics**, v. 86, n. 1, p. 4-29, 2004.

ROSENBAUM, P.; RUBIN, D. Assessing sensitivity to an unobserved binary covariate in an observational study with binary outcome. **Journal of the Royal Statistical Society**, v. 45, n. 2, p. 212-218, 1983.

RUBIN, D. B. **Matched sampling for causal effects**. Cambridge: Cambridge University Press, 2006.

SEKHON, J. S. Opiates for the matches: matching methods for causal inference. **Annual Review of Political Science**, v. 12, p. 487-508, 2009.

ANÁLISE DA EFICIÊNCIA TÉCNICA E DA PRODUTIVIDADE DOS SERVIÇOS DE ÁGUA E ESGOTOS NO BRASIL DE 2006 A 2013

Felipe Ponciano da Cruz¹ Ronaldo Seroa da Motta² Alexandre Marinho³

O marco regulatório da Lei Nacional do Saneamento Básico introduz mudanças fundamentais na prestação dos serviços de abastecimento de água, coleta e tratamento de esgotos. O objetivo deste trabalho é estudar a eficiência do setor após a implantação da referida lei, utilizando o método análise envoltória de dados (data envelopment analysis — DEA) em uma amostra contendo as 27 prestadoras representativas de cada ente da Federação entre 2006 e 2013. Os resultados na análise estática indicam escores de eficiência baixos e com grande disparidade entre as regiões do país. Já a análise dinâmica demonstra que o período de maior avanço de produtividade ocorreu entre 2010 e 2013, liderado pela expansão da fronteira.

Palavras-chave: análise de envoltória de dados; eficiência; água e esgotos; saneamento no Brasil.

ANALYSIS OF TECHNICAL EFFICIENCY AND PRODUCTIVITY OF WATER AND SEWAGE SERVICES IN BRAZIL, 2006 TO 2013

The regulatory framework of the National Basic Sanitation Law introduces fundamental changes in the provision of water supply, sewage collection and treatment services. The objective of this text is to study the efficiency and productivity of the sector after this law applying the Data Envelopment Analysis (DEA) method in a sample with 27 representative providers of each entity of the federation between the years of 2006 and 2013. The results for the static analysis indicate low efficiency scores with great disparity between the regions. The dynamic analysis showed that the period of greatest productivity improvement occurred between the years of 2010 and 2013 led by the expansion of the frontier.

Keywords: data envelopment analysis; efficiency; water and sewerage; sanitation in Brazil.

JEL: L90; L95; L50.

1 INTRODUÇÃO

Os serviços de distribuição de água potável e coleta e tratamento de esgoto influenciam tanto a saúde da população quanto a qualidade do meio ambiente e, portanto, são fundamentais para a redução da pobreza e para o desenvolvimento sustentável. Desde a captação da água bruta até o descarte dos esgotos, os serviços de água e

^{1.} Mestre em economia pelo Programa de Pós-Graduação em Ciências Econômicas (PPGCE) da Universidade do Estado do Rio de Janeiro (UERJ). *E-mail*: <f1ponciano@qmail.com>.

^{2.} Professor associado do PPGCE da UERJ. E-mail: <seroadamotta.ronaldo@gmail.com>.

^{3.} Técnico de planejamento e pesquisa do Ipea e professor associado da Faculdade de Ciências Econômicas da UERJ. *E-mail*: <a href="mailto: calexandre.marinho@ipea.gov.br >.

esgotos dependem e interferem nas condições ambientais. Consequentemente, diversas doenças são influenciadas pelas condições inadequadas desses sistemas, tais como diarreias, febres entéricas, hepatite A, esquistossomose, leptospirose, teníases, micoses e conjuntivites (Mendonça e Seroa da Motta, 2007; Brasil, 2002). Dessa forma, o desempenho do setor de saneamento tem sido considerado crucial para minimizar esses efeitos.

O referido setor, no Brasil, alcançou níveis crescentes de cobertura de serviços desde a criação das companhias estaduais de saneamento básico (Cesbs), quando a maioria dos serviços municipais foram absorvidos por essas companhias. Essa mudança na estrutura produtiva foi motivada pelo Plano Nacional de Saneamento Básico (Planasa) (Salles, 2008; Seroa da Motta, 2005). Hoje, mais de duas décadas após o fim do programa, ainda existe a predominância das Cesbs.

Com o fim do Planasa e do Sistema Financeiro do Saneamento (SFS), houve um vazio institucional e também de investimentos. Desse modo, as décadas de 1990 e 2000 foram marcadas por intensos debates sobre o novo marco regulatório para o saneamento. Finalmente, em 2010, com a regulamentação da Lei Nacional do Saneamento Básico (LNSB), Lei nº 11.445, de 5 de janeiro de 2007 (Brasil, 2007; 2010), o país passou a contar com um novo marco regulatório.

A LNSB define que os serviços de saneamento serão prestados com eficiência e sustentabilidade econômica e os contratos de concessão devem prever metas progressivas e graduais de expansão de serviços, de qualidade, de eficiência e de uso racional da água e da energia. E, mais ainda, que essa prestação deve conter mecanismos e procedimentos para a avaliação sistemática da eficiência e da eficácia das ações programadas. A política tarifária deve, portanto, incluir incentivos à eficiência dos prestadores de serviços. Além da LNSB, a Constituição Federal Brasileira de 1988 (CF/1988), em seu art. 37, determina que a busca da eficiência na prestação de serviços públicos é um imperativo. No *caput* do referido artigo, consta que: "A administração pública direta e indireta de qualquer dos Poderes da União, dos estados, do Distrito Federal e dos municípios obedecerá aos princípios de legalidade, impessoalidade, moralidade, publicidade e *eficiência* (...)" (Brasil, 1988, grifo nosso).

Há também na LNSB uma complexa atribuição de competências entre os entes federados.⁴ Os municípios são os titulares dos serviços de saneamento e podem executá-los diretamente ou por meio de concessão. É de competência municipal a organização, regulação e fiscalização da prestação de tais serviços, bem como a elaboração do Plano Municipal de Saneamento Básico (PMSB). Assim, os municípios podem instituir suas próprias agências reguladoras ou delegar a

^{4.} Inclusive, a definição dessas atribuições foi um dos principais obstáculos na elaboração do marco regulatório e ainda é fonte de litígio entre estados e municípios. Ver Estache, Garsous e Seroa da Motta (2016).

responsabilidade às entidades reguladoras de seus estados. No caso de a operação dos serviços do município ser realizada por uma operadora estadual (Cesb), a regulação é de responsabilidade dessa agência estadual.

Em suma, a LNSB define o município como titular das quatro funções essenciais para a coordenação dos serviços de saneamento: prestação dos serviços, regulação, fiscalização e planejamento. A única função não delegável é a de planejamento, representada pelo PMSB. Entretanto, em regiões metropolitanas (RMs) e microrregiões, os estados compartilham essas funções regulatórias e de competência com os municípios.

A gestão dos recursos hídricos que concede outorga para a captação de água e para o despejo de esgotos, e determina a qualidade das águas, é regulada pela Política Nacional dos Recursos Hídricos (PNRH) com competência federal, para os cursos d'águas que banham mais de um Estado ou país, ou dos estados, nos casos que se limitam ao seu território.

Em que pese à evolução da cobertura nessas décadas, ressalta-se que ainda 18,3% das crianças de 0 a 14 anos de idade residentes em domicílios particulares permanentes não possuem acesso à rede geral de abastecimento de água; apenas 55,5% dos domicílios têm esgotamento sanitário de rede geral, ou fossa séptica, e somente 39% dos esgotos do país são tratados (IBGE, 2014; Brasil, 2014).

Além das restrições regulatórias e de financiamento, a evolução do setor também depende da eficiência das empresas provedoras. Inúmeros estudos nas últimas duas décadas tentaram explicar a evolução da eficiência das operadoras dos serviços de saneamento no Brasil. Esses estudos analisaram o desempenho dessas empresas; e, em muitos casos, medindo o efeito da natureza da propriedade (privada *versus* pública) e da jurisdição da operação (regional *versus* local).

Carmo e Távora Junior (2003), utilizando dados de 2000, calcularam a eficiência técnica e de escala de 26 companhias estaduais de saneamento básico com um modelo DEA de *input* orientado. Tupper e Resende (2004), com um modelo com retornos variáveis de escala e DEA *output* orientado, com dados de 1996 a 2000, mensuraram a eficiência de vinte Cesbs. Seroa da Motta e Moreira (2006) utilizaram o índice de Malmquist para analisar 104 companhias brasileiras de saneamento, abrangendo tanto as regionais quanto as locais, entre 1998 e 2002. Sampaio e Sampaio (2007) analisaram 36 companhias de saneamento brasileiras, entre 1998 e 2003, por meio de um modelo com retornos variáveis de escala e DEA *output* orientado. Utilizando os dados de 57 companhias de saneamento paulistas, entre 1995 e 2004, Grigolin (2007) usou modelos DEA *output* orientados com retornos constantes e variáveis de escala. Sato (2011) analisou a eficiência de 26 companhias de saneamento das capitais dos estados brasileiros com dados referentes aos anos de 2005 e 2008. Barbosa (2012) procurou explicar a eficiência por meio

da estrutura regulatória com uma amostra de 41 prestadoras de serviço de água e esgoto. Carvalho (2014) utilizou o índice de Malmquist, com modelo DEA *input* orientado, para mensurar a eficiência de 29 prestadoras entre 2006 e 2011.⁵

Nesses estudos, se observa que há dominância do método DEA, e os resultados, na sua maioria, confirmam a evolução da produtividade e variam quanto à dominância e ao efeito de emparelhamento e de mudança de fronteira e em como essa evolução se deu regionalmente.

Literatura similar internacional existe também com a aplicação de modelos DEA. Lin e Berg (2008) utilizaram as seguintes técnicas para avaliar a eficiência de 38 prestadoras de serviço de distribuição de água peruanas entre 1998 e 2001. Byrnes et al. (2010) analisaram os efeitos de diversas ações políticas na eficiência de 52 prestadoras de serviços de água nas regiões australianas de New South Wales e Victoria. Portela *et al.* (2011) utilizaram o índice de meta-Malmquist circular para analisar a mudança na produtividade de 22 prestadoras de serviços de distribuição de água da Inglaterra e do Reino Unido entre 1993 e 2007. Também na Austrália, Abbott, Cohen e Wang (2012) analisaram o desempenho dos setores de água e esgotos das cidades de Sydney, Melbourne, Brisbane, Adelaide, Perth e Canberra, utilizando o índice de Malmquist, entre 1995 e 2008. Em geral, os trabalhos citados encontram avanços modestos de produtividade nos períodos analisados. Houve uma alternância quanto à primazia entre o efeito de emparelhamento e o de mudança de fronteira, inclusive com comportamentos erráticos, em que os efeitos ora contribuíram positivamente ora negativamente para a produtividade total dos fatores (PTF) ao longo do intervalo analisado (Lin e Berg, 2008; Portela et al., 2011).

Este trabalho tem como objetivo analisar a eficiência técnica das prestadoras estaduais de saneamento no Brasil para o período 2006-2013, que inclui a vigência da LNSB, a qual alterou o marco regulatório do setor.

Ademais, o estudo aplica originalmente na avaliação de desempenho do setor de saneamento no Brasil o método DEA e o índice de Malmquist, em conjunto com a técnica *bootstrap*, para acrescentar intervalos de confiança a métodos anteriormente determinísticos.

No geral, observamos escores de eficiência baixos e dispersos, mas crescentes ao longo do período analisado. A única região que apresentou diminuição da média dos escores foi a Centro-Oeste. As empresas da região Sudeste do país apresentam, em média, os maiores escores de eficiência, enquanto as da região Norte

^{5.} Ademais, há também estudos com modelos DEA na avaliação de eficiência de empresas distribuidoras de energia elétrica no Brasil (Banker, 2011; Souza e Legey, 2010), nos serviços hospitalares (Marinho e Façanha, 2002; Marinho, 2003) e na eficiência da gestão (Seroa da Motta e Moreira, 2009; Mendes e Sampaio de Sousa, 2006; Sampaio de Sousa e Stosic, 2005).

apresentam os menores. A PTF foi decomposta em dois componentes: mudança tecnológica e mudança de eficiência (emparelhamento ou *catch up*). A contribuição da eficiência das empresas é deletéria para a PTF no início do período e positiva ao fim. A tecnologia reduz a PTF no início e no fim do período analisado e aumenta a PTF nos anos intermediários. Estimamos que, nos quatro primeiros anos após a promulgação, no ano de 2007, da Lei Nacional do Saneamento Básico (LNSB) não ocorreram avanços robustos na produtividade. Para o período 2010-2013, os resultados apresentados indicam um aumento significativo de produtividade com mediana de 6,7%, que o modelo atribuiu ao componente tecnológico, apesar de um ligeiro decréscimo da eficiência (*catch up*). Em geral, as estatísticas-resumo apontam um avanço na produtividade total dos fatores no período entre 2010 e 2013, caracterizado principalmente pela mudança tecnológica.

Este trabalho se divide em mais quatro seções além desta introdução: metodologia, modelo e dados, resultados e conclusão.

2 METODOLOGIA

A eficiência econômica é entendida como o máximo de *outputs* que a empresa pode produzir dada quantidade de *inputs* utilizada, ou, da mesma maneira, como o mínimo de *inputs* que podem produzir uma dada quantidade de *outputs*, ou seja, os valores ótimos de *outputs* e *inputs* utilizados.

Marinho e Façanha (2001), ao estudarem os problemas de avaliação em programas sociais, relacionaram efetividade "à implementação e ao aprimoramento de objetivos, independentemente das insuficiências de orientação e das falhas de especificação rigorosa dos objetivos iniciais declarados do programa" (Marinho e Façanha, 2001, p. 6). Dessa maneira, o ideal seria avaliar a efetividade e não a eficiência.

Para essa avaliação, as análises da fronteira estocástica (SFA) e da envoltória dos dados (DEA) são amplamente utilizadas na literatura. A primeira mensura uma fronteira estocástica, seguindo formas funcionais, enquanto a última adota técnicas não paramétricas. Segundo Coelli (1995), utilizar a DEA pode ser mais atrativo do que estimar a fronteira estocástica no caso de multiprodutos, sobretudo quando há dificuldades para se obter os preços.

Os dados do Sistema Nacional de Informações sobre Saneamento (SNIS), por exemplo, apresentariam essa limitação porque se restringem a valores de custos e receitas sem distinguirem quantidades ou preços. Ademais, a característica de monopólio público das Cesbs permite que muitas vezes seus preços sejam decididos politicamente (Estache, Garsous e Seroa da Motta, 2016).

Assim, a escolha da DEA em detrimento da SFA se deve às suas características de não necessitar de especificações sobre a função de produção com maior flexibilidade no tratamento de multiprodutos e de dispensar a utilização de preços para calcular a eficiência técnica.

A descrição dessa aplicação para as análises estática e dinâmica é apresentada a seguir.

2.1 Análise estática

Conforme Bogetoft e Otto (2011), em um modelo racional ideal, temos uma entidade econômica em busca dos melhores meios para atingir seus objetivos, e, dessa forma, podemos assumir que uma firma transforma m *inputs* $x^* \in R^m$ em *n outputs* $y^* \in R^n$; e que a função objetivo ou função de utilidade é dada por: $U: R^{m+n} \to R$, em que U(x,y) é a utilidade ligada ao plano de produção (x,y). Devemos assumir também que o conjunto de vetores (x,y) possíveis é $T \subseteq R^{m+n}$.

Sendo assim, o vetor (x^*, y^*) é ótimo se resolve o problema de decisão básico a seguir se, e somente se, a firma realizar o melhor e mais efetivo uso do seu potencial:

$$máx U(x,y) (1)$$

sujeito a: $(x, y) \in T$.

Entretanto, é difícil encontrar na prática as condições ideais para realizar avaliações de efetividade, pois o plano de produção factível, T, e/ou o critério de avaliação agregado, U(.), são desconhecidos. Outro problema apontado é que os múltiplos *outputs* são de difícil agregação.

A proposta da moderna teoria da análise de produtividade é focar em eficiência no lugar da efetividade, de modo a construir uma tecnologia usando a análise da eficiência relativa. Esse raciocínio pode ser sumarizado da seguinte maneira: $(x',y') \in T \subseteq R^{m+n}$ é eficiente relativamente à tecnologia T se, e somente se, para todo $(x,y) \in T$: $x \le x'$; $y \ge y' \to x = x'$ e y = y', e, portanto, não é possível encontrar melhor vetor do que (x',y').

Assim, a eficiência é uma condição necessária, embora não suficiente, para a efetividade, e utiliza-se eficiência ao invés de efetividade devido ao desconhecimento de U(·) e da real tecnologia *T*, de modo a encontrar o conjunto de planos eficientes, T^E, que é o "menor conjunto suficiente de alternativas a considerar" na busca da efetividade (Bogetoft e Otto, 2011, p. 53).

O método utilizado é a DEA, metodologia não paramétrica (não propõe a estimação de uma função com parâmetros) e não estocástica (não é associada

a qualquer distribuição de probabilidade). A fronteira de produção é gerada implicitamente via programação linear.

Para atenuar a desvantagem dessa abordagem não paramétrica, é utilizada a técnica *bootstrap*, a partir de Simar e Wilson (1998), para corrigir o viés dos escores e estimar intervalos de confiança.

As decison making units (DMUs) são os elementos básicos das análises de eficiência. Elas devem ser homogêneas, utilizarem os mesmos inputs e outputs, bem como apresentarem autonomia na tomada de decisão.

A eficiência de cada DMU é medida pela otimização da soma ponderada das saídas (*outputs*) dividida pela soma ponderada das entradas (*inputs*), ou seja, pela

combinação entre *output* virtual e *input* virtual de modo a maximizar <u>input virtual</u> , <u>input virtual</u>

ou, de maneira equivalente, minimizar output virtual.

Por definição, não é possível apresentar o *input* virtual maior do que o *output* virtual, e, portanto, o *output* virtual deve ser menor ou igual ao *input* virtual. Isto posto, uma DMU plenamente eficiente é aquela que consegue transformar todo

o *input* virtual em *output* virtual, e assim temos
$$\frac{input\ virtual}{output\ virtual} = 1$$
.

Os modelos de DEA admitem diferentes naturezas de retornos de escala. Os modelos mais conhecidos e utilizados são: *i*) o modelo CCR⁶ (Charnes, Cooper e Rhodes, 1978) impõe retornos constantes de escala; e *ii*) o modelo BCC⁷ (Banker, Charnes e Cooper, 1984) admite retornos variáveis de escala. Existem também modelos com retornos decrescentes ou não crescentes de escala.

Existe ainda a opção de escolher a orientação do modelo de DEA, entre *input* ou *output* orientado. Um modelo *output* orientado maximiza os *outputs* mantendo os *inputs* inalterados, enquanto um modelo *input* orientado minimiza os *inputs* mantendo os *outputs* inalterados.

O modelo (2) apresenta o modelo da envoltória, *output* orientado, e com retornos variáveis de escala. Onde: $h \ge 1$ é a medida de expansão radial, quando $h_0 = I$ a DMU é considerada eficiente, λ 's são os pesos das combinações convexas das DMUs e \mathbf{x}_{ik} e \mathbf{y}_{jk} são os *inputs* i e *outputs* j da DMU \mathbf{k} , $\mathbf{k} = 1, \ldots, \mathbf{n}$; \mathbf{x}_{io} e \mathbf{y}_{jo} são os *inputs* i e *outputs* j da DMU 0. A DMU 0 é a DMU sob análise, e o problema será resolvido n vezes, uma vez para cada DMU.

^{6.} Ver Charnes, Cooper, Rhodes (1978).

^{7.} Ver Banker, Charnes, Cooper (1984).

A eficiência é representada pela função objetivo na qual se estima o valor que deve multiplicar todos os *outputs* para colocar as DMUs sobre a fronteira eficiente. As duas restrições garantem que cada aumento dos *outputs* não ultrapasse a fronteira definida pelas DMUs eficientes e também que não seja modificado o atual nível dos *inputs* da DMU.

 $\text{Máx}\,h_0$

sujeito a:

$$-h_0 y_{j0} + \sum_{k=1}^n y_{jk} \lambda_k \ge 0, \forall j$$

$$x_{i0} - \sum_{k=1}^n x_{ik} \lambda_k \ge 0, \forall i$$

$$\sum_{k=1}^n \lambda_k = 1$$

$$\lambda_k \ge 0 \ \forall k$$

$$(2)$$

Os estimadores da DEA, apesar de serem consistentes, possuem taxas muito lentas de convergência e, dessa forma, o *bootstrap* poderia auxiliar a encontrar a variância dos escores de eficiência que possuem distribuição desconhecida. Entretanto, conforme Simar e Wilson (1998), o *bootstrap* usual não se aplica ao método DEA. Não é recomendado reamostrar diretamente o conjunto de escores da DEA, pois dessa maneira seria admitido que os escores são independentes e identicamente distribuídos, o que iria contra a suposição de que eles dependem da relação entre *inputs* e *outputs*. Tampouco é aconselhável reamostrar simplesmente os *inputs* (ou *outputs*) para estimar a tecnologia e a eficiência de determinada DMU, pois existe a possibilidade de encontrar escores maiores do que a unidade, ou seja, vetores que não pertencem à tecnologia. Tal ocorrência é mais provável em DMUs mais próximas à fronteira original.

Os escores do DEA também apresentam o problema de serem bastante concentrados em torno do valor unitário. Para enfrentar esta dificuldade, Simar e Wilson (1998) propõem os métodos de alisamento da distribuição e de reflexão, bem como a correção do viés.

O alisamento consiste em corrigir a amostra gerada pelo *bootstrap* com a utilização do fator $h\epsilon$, onde h é a faixa de amplitude (*bandwidth*) $e \in \acute{e}$ gerado a partir de uma distribuição normal. O alisamento visa evitar o surgimento de muitos valores repetidos, que gerariam picos na distribuição do estimador. Por

meio do método da reflexão, o problema do surgimento de muitos escores de eficiência próximos da unidade é evitado, o que introduziria viés e inconsistência no estimador. A cada escore de eficiência E_k é calculado um reflexo $2 - E_k$, e utilizase sempre o valor menor ou igual à unidade. Assim, a média será $[(\widetilde{E}_k + 2 - \widetilde{E}_k)]/2 = 1$, e simétrica em torno da unidade.

O escore de eficiência do DEA é benevolente, ou seja, superestima a eficiência real. Isso ocorre devido à tecnologia T̂ gerada pelo modelo ser sempre o menor conjunto suficiente, pois algumas observações podem não constar da amostra. O estimador com o viés corrigido é dado por (Bogetoft e Otto, 2011, p. 173):

$$\tilde{\theta}^{k} = 2\hat{\theta}^{k} - \bar{\theta}^{k*}. \tag{3}$$

Em que:

- θ̂^k é o estimador de eficiência obtido a partir da tecnologia T̂, estimada pela DEA;
- $\overline{\theta}^{k*}$ é a média de θ^{k*} , que é a eficiência estimada por *bootstrap* de θ^k .

O *bootstrap* que será aplicado é o composto das seguintes etapas (Bogetoft e Otto, 2011, p. 173).

- 1) Computar $\hat{\theta}^k$ como soluções para $min\{\theta | \theta(x^k, y^k) \in \hat{T}\}$ para k=1, ..., n;
- 2) Usar o *bootstrap* utilizando a amostra alisada de $\hat{\theta}^1, ..., \hat{\theta}^k$ para obter uma réplica $\theta^{1*}, ..., \theta^{k*}$ do modo a seguir:
 - usar o *bootstrap* na amostra com reposição a partir de $\hat{\theta}^1, ..., \hat{\theta}^k$, e chamar os resultados de $\beta^1, ..., \beta^k$;
 - simular as variáveis independentes aleatórias ε¹, ..., ε^k com distribuição normal-padrão;
 - calcular (utilizando o alisamento e a reflexão):

$$\tilde{\theta}^k = \begin{cases} \beta^k + h\varepsilon^k \text{ se } \beta^k + h\varepsilon^k \leq 1\\ 2 - \beta^k - h\varepsilon^k \text{ caso contrário} \end{cases} \tag{4}$$

Por construção $\tilde{\theta}^k \leq 1$.

• ajustar $\tilde{\theta}^k$ para obter parâmetros com variância assintoticamente correta, e, então, estimar a variância $\hat{\sigma}^2 = \frac{1}{n} \sum_{k=1}^K (\hat{\theta}^k - \overline{\hat{\theta}})^2$ e calcular $\bar{\beta} = \frac{1}{n} \sum_{k=1}^K \beta^k$:

$$\theta^{k*} = \overline{\beta} + \frac{1}{\sqrt{1 + h^2/\widehat{\sigma}^2}} (\widehat{\theta}^k - \overline{\beta}). \tag{5}$$

- 3) Calcular o *bootstrapped input* $x^{kb} = \frac{\hat{\theta}^k}{\theta^{k*}} x^k$ baseado na eficiência calculada no *bootstrap*;
- 4) Solucionar o problema de programação da DEA para estimar $heta^{kb}$ como:

$$\begin{split} \theta^{kb} &= \min \big\{ \theta \geq 0 \, \big| y^k \leq \sum_{j=1}^K \lambda_j y_j, \theta x^k \geq \sum_{j=1}^K \lambda_j x_j^{\, kb} \,, \lambda_j \geq 0, \sum_{j=1}^K \lambda_j = 1 \big\} \\ (k &= 1, \dots, n); \end{split}$$

- 5) Repetir os passos a partir de (2.a) para obter as estimativas *bootstrap* $(\theta^{1b}, ..., \theta^{Kb})$ (b = 1, ..., B);
- 6) Calcular a média e variância de $(\theta^{1b}, ..., \theta^{Kb})$ para conseguir a estimativa *bootstrap* θ^{k*} , e o estimador com o viés corrigido $\tilde{\theta}^{k*}$, e a variância:

$$\widehat{\sigma 2} = \frac{1}{B} \sum_{b=1}^{B} (\theta^{kb} - \overline{\theta}^{k*})^2. \tag{7}$$

Ao final do processo, a distribuição de $(\hat{\theta}^k - \theta)$ é desconhecida, mas pode ser aproximada pela distribuição de $(\theta^{k*} - \hat{\theta}^k)$. θ é o escore de eficiência real, mas desconhecido, $\hat{\theta}^k$ é o escore de eficiência tradicional obtido na DEA, e θ^{k*} é a estimativa do escore de eficiência da DEA obtida no *bootstrap*.

2.2 Análise dinâmica

Para entender a evolução dinâmica da eficiência, é utilizado o índice de Malmquist, que representa a variação na produtividade total dos fatores entre os períodos s (inicial) e t (final) dessa forma (Bogetoft e Otto, 2011):

$$M(s,t) = \sqrt{\frac{E(t,s)E(t,t)}{E(s,s)E(s,t)}}$$
(8)

Em que:

- E(t,s) é a eficiência da DMU medida no período t com tecnologia do período s;
- E(t,t) é a eficiência da DMU medida no período t com tecnologia do período t;
- E(s,t) é a eficiência da DMU medida no período s com tecnologia do período t;

 E(s,t) é a eficiência da DMU medida no período s com tecnologia do período s.

Dessa forma, podemos tomar as seguintes conclusões sobre os valores do índice:

- M(s,t) > 1: houve uma mudança positiva na PTF entre os períodos s
 e t;
- M(s,t) = 1: n\u00e1o houve qualquer mudan\u00e7a na PTF entre os per\u00edodos s
 e t;
- M(s,t) < 1: houve mudança negativa na PTF entre os períodos $s \in t$.

Uma característica importante do índice de Malmquist é que ele pode ser decomposto em dois fatores:

$$M(s,t) = \sqrt{\frac{E(t,s)E(s,s)}{E(t,t)E(s,t)}} \frac{E(t,t)}{E(s,s)} = TC(s,t)EC(s,t)$$
(9)

Onde:

$$TC(s,t) = \sqrt{\frac{E(t,s)E(s,s)}{E(t,t)E(s,t)}}$$
(10)

$$EC(s,t) = \frac{E(t,t)}{E(s,s)} \tag{11}$$

Se E(t,s) > E(t,t) é possível afirmar que a tecnologia progrediu, pois a mudança de tecnologia de t para s foi suficiente para aumentar a eficiência. A mesma lógica serve para E(s,s) > E(s,t).

O índice de mudança tecnológica, TC(s,t), representa a média geométrica de dois índices, E(t,s)/E(t,t) e E(s,s)/E(s,t). Portanto, TC mede a mudança tecnológica, e, se TC(s,t) > 1, houve um progresso tecnológico, ou seja, uma mudança (expansão) da fronteira.

O índice de mudança de eficiência, EC(s,t) (catch up), mede a aproximação da firma da fronteira, efeito emparelhamento, caso E(t,t) > E(s,s) consequentemente EC(s,t) > 1 e a DMU se aproximou da fronteira.

O *bootstrap* é utilizado para encontrar intervalos de confiança do índice de Malmquist (Simar; Wilson, 1999). Dessa maneira, devido ao desconhecimento da distribuição verdadeira do índice de Malmquist, o *bootstrap* é empregado para estimar a distribuição e definir o intervalo de confiança seguindo os seguintes passos.

1) Estimar o índice de Malmquist $\widehat{M}_{t,s}$ (x^s, y^s, x^t, y^t) das k DMUs da maneira descrita acima.

- 2) Obter as pseudoamostras (x^{t*} , y^{t*}) de cada DMU para construir a tecnologia via *bootstrap* com alisamento e reflexão.
- 3) Calcular, utilizando a amostra obtida no passo 2, a estimativa *bootstrap* do índice de Malmquist $\widehat{M}_{t,s}^{kb*}(\mathbf{x}^s, \mathbf{y}^s, \mathbf{x}^t, \mathbf{y}^t)$ de cada uma das k DMUs.
- 4) Repetir B vezes os passos 2 e 3 para obter um conjunto de estimativas $\widehat{M}_{t,s}^{kb*}(\mathbf{x}^s,\mathbf{y}^s,\mathbf{x}^t,\mathbf{y}^t)$.
- 5) Obtém-se o intervalo de confiança a partir deste conjunto de estimativas.

Com o intervalo de confiança é possível verificar a significância do índice de Malmquist. Caso o intervalo de confiança encontrado tenha os dois extremos menores do que a unidade, podemos afirmar que a produtividade decaiu; caso o intervalo de confiança tenha os dois extremos maiores do que a unidade, podemos afirmar que a produtividade aumentou. Enfim, se o intervalo de confiança tiver um extremo menor do que a unidade e outro maior, não podemos afirmar coisa alguma sobre o comportamento da produtividade entre os períodos em questão.

3 MODELO E DADOS

A decisão sobre o modelo a ser utilizado se baseou, principalmente, na revisão de literatura dos trabalhos brasileiros e no entendimento do funcionamento do setor e seus objetivos.

No modelo, as DMUs são as prestadoras de serviços de distribuição de água e coleta e tratamento de esgotos. Essa escolha é justificada pela decisão sobre a alocação de recursos a ser tomada pelas empresas. Pressupomos que as DMUs se defrontam com a mesma tecnologia (T), que é desconhecida, e estimamos a tecnologia (\hat{T}) com o uso da DEA e do índice de Malmquist.

O modelo será *output* orientado, pois as empresas são majoritariamente públicas, nas quais o corte de gastos é dificultado e a demanda pelos serviços de fornecimento de água e de esgotamento sanitário não está plenamente atendida.

Quanto à escolha entre o modelo CCR (que impõe retornos constantes de escala) e o modelo BCC (que admite retornos variáveis de escala), optou-se pelos retornos variáveis de escala devido ao reconhecimento da existência de dificuldades para as prestadoras mudarem de tamanho no curto prazo, bem como pela grande variação de dimensão entre as prestadoras analisadas, conforme o apontado por Sampaio e Sampaio (2007).

Os dados que serão utilizados neste trabalho serão da série histórica do SNIS⁸ para 27 prestadoras estaduais no período 2006-2013.⁹

Os *output*s y_k =(y_{1k} ,..., y_{5k}) utilizados são listados a seguir, conforme a nomenclatura do SNIS:

- item AG003: quantidade de economias¹⁰ ativas de água unidade: economia:
- item AG010: volume de água consumido unidade: 1.000 m³/ano;
- item ES003: quantidade de economias ativas de esgoto unidade: economia;
- item ES005: volume de esgoto coletado unidade: 1.000 m³/ano; e
- item ES006: volume de esgoto tratado unidade: 1.000 m³/ano.

Apesar da cobrança pela coleta e pelo tratamento de esgoto ser feita de maneira conjunta, decidiu-se por tratá-los como produtos separados, a exemplo de Seroa da Motta e Moreira (2006).

Os *output*s com os números de economia de água e esgoto (AG003 e ES003) são os mais utilizados na literatura e captam o objetivo da universalização que preconiza a LNSB.

Como *input*, x_{Ik} , utilizaremos o item FN015 – despesas com exploração – que se aproxima do conceito de custo operacional com unidade: R\$/ano. Os valores apresentados no FN015 foram atualizados para 2013 utilizando o índice Geral de Preços – disponibilidade interna (IGP-DI) da Fundação Getulio Vargas (FGV).

A aplicação do método DEA será realizada por meio do pacote FEAR (frontier efficiency analysis with *R*), do programa *R*. O *R* é um *software* livre criado por Ross Ihaka e por Robert Gentleman, da Universidade de Auckland (Nova Zelândia), que permite a operação de métodos estatísticos e a formatação de gráficos, podendo ser instalado e distribuído livremente.

O pacote FEAR nos permitirá a aplicação *bootstrap* no DEA e no Malmquist, que foi desenvolvido por Wilson (2008) e é disponibilizado em sua página pessoal.¹¹

^{8.} Mais dados em: http://www.snis.gov.br/>. Acesso em: 11 abr. 2019.

^{9.} Essa amostra representa o atendimento de 79,3% dos municípios que responderam ao SNIS em 2013. Não se estendeu a cobertura temporal para anos mais anteriores porque a base de dados evoluiu em termos de conferência da consistência das informações que são declaratórias.

^{10.} Economia — imóvel de uma única ocupação, ou subdivisão de imóvel com ocupação independente das demais, perfeitamente identificável ou comprovável em função da finalidade de sua ocupação legal, dotado de instalação privativa ou comum para o uso dos serviços de abastecimento de água ou de coleta de esgoto. Exemplo: um prédio com dez apartamentos possui uma ligação e dez economias. Disponível em: https://bit.ly/2oKaCcH. Acesso em: 11 abr. 2019.

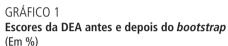
^{11.} Disponível em: https://bit.ly/36tMvQN. Acesso em: 11 abr. 2019.

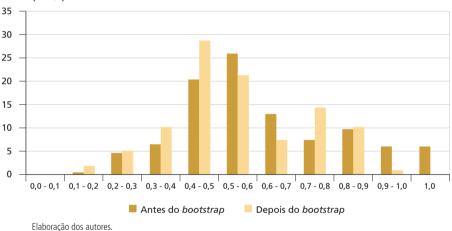
4 RESULTADOS

A seguir são apresentados os resultados da análise estática, seguidos pelos da análise dinâmica.

4.1 Análise estática

O gráfico 1 retrata os resultados da análise estática antes e depois do *bootstrap*, que foi realizado com mil repetições. Conforme o esperado, após a utilização do método de reamostragem, houve a correção do viés que superestimava as *performances*, o que pode ser identificado pela diminuição considerável das DMUs com escores próximos ou iguais à unidade.





A tabela 1 apresenta os escores de eficiência calculados conjuntamente para as 216 DMUs, no período 2006-2013 (27 distribuidoras em oito anos), após a utilização do *bootstrap*. Observa-se que o nível de eficiência, em torno de 0,40 e 0,56, é muito baixo, e embora se note uma melhora ao longo do período, essa não é contínua. Ademais, se observa também um alto desvio-padrão, demonstrando grande heterogeneidade entre o desempenho das prestadoras.

			•	•			
Ano	Mínimo	Máximo	1º Quartil	Mediana	Média	3º Quartil	Desvio-padrão
2006	0,168	0,817	0,416	0,504	0,519	0,645	0,162
2007	0,179	0,820	0,436	0,537	0,541	0,647	0,155
2008	0,184	0,850	0,450	0,485	0,549	0,663	0,162
2009	0,243	0,821	0,433	0,532	0,531	0,633	0,150
2010	0,195	0,904	0,410	0,493	0,542	0,677	0,182
2011	0,212	0,921	0,452	0,503	0,545	0,667	0,182
2012	0,232	0,892	0,483	0,558	0,584	0,737	0,183
2013	0,240	0,865	0,486	0,536	0,574	0,748	0,172

TABELA 1 Estatísticas-resumo dos escores (2006-2013)

Elaboração dos autores.

O gráfico 2 apresenta como esses escores estão correlacionados com a quantidade de economias de água que serve de aproximação para o tamanho da prestadora. Percebe-se, com isso, que há a tendência de a eficiência aumentar quanto maior for a quantidade de economias de água.

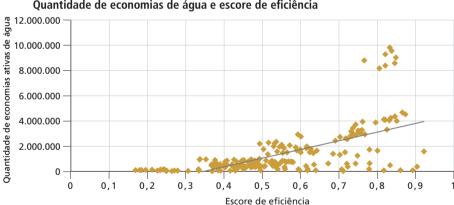
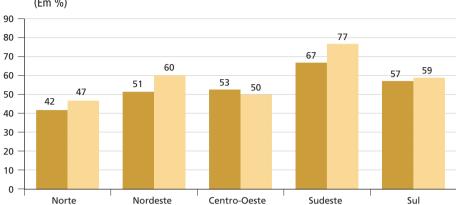


GRÁFICO 2 Quantidade de economias de água e escore de eficiência

Elaboração dos autores.

O gráfico 3 revela que a região Sudeste possui a maior média dos escores de eficiência, tanto em 2006 quanto em 2013. Já a região Norte apresenta a menor média dos escores de eficiência, tanto no início quanto no fim do período analisado. É relevante observar que a única região que apresentou diminuição da média dos escores foi a Centro-Oeste.



2006

2013

GRÁFICO 3 **Média dos escores de eficiência por região** (Em %)

Elaboração dos autores.

4.2 Análise dinâmica

A tabela 2 apresenta os valores do índice de Malmquist, em conjunto ao *bootstrap*, com mil repetições. Observa-se que a produtividade total dos fatores, em geral, cresceu durante o período analisado, exceto para o período 2008-2009. Já entre 2012 e 2013, somente o índice medido pela média é também menor que a unidade.

TABELA 2	
Estatísticas da dinâmica da produtividade total dos fatores (2	(006-2013)

Ano inicial-final	Mínimo	Máximo	1º Quartil	Mediana	Média	3º Quartil	Desvio-padrão
2006-2007	0,829	1,488	0,980	1,033	1,042	1,095	0,122
2007-2008	0,832	1,482	0,980	1,026	1,041	1,073	0,130
2008-2009	0,766	1,765	0,858	0,913	0,941	0,964	0,178
2009-2010	0,648	1,292	0,949	1,039	1,035	1,130	0,162
2010-2011	0,703	2,256	0,931	1,001	1,064	1,118	0,272
2011-2012	0,808	2,379	0,987	1,048	1,153	1,259	0,294
2012-2013	0,437	1,436	0,946	1,019	0,970	1,041	0,173

Elaboração dos autores.

Como mostra a tabela 3, é possível verificar que o componente de emparelhamento teve uma contribuição positiva entre 2006 e 2009, e destacadamente, no período final de 2012 a 2013. Já entre 2009 e 2012 a contribuição do *catch up* para a produtividade total dos fatores foi negativa.

			3	,		-,	
Ano inicial-final	Mínimo	Máximo	1º Quartil	Mediana	Média	3º Quartil	Desvio-padrão
2006-2007	0,890	1,442	0,987	1,052	1,055	1,096	0,112
2007-2008	0,792	1,448	1,000	1,041	1,069	1,123	0,124
2008-2009	0,830	2,093	0,940	1,000	1,033	1,024	0,222
2009-2010	0,539	1,117	0,842	0,926	0,909	1,000	0,139
2010-2011	0,612	1,659	0,725	0,801	0,848	0,922	0,200
2011-2012	0,501	1,235	0,621	0,779	0,789	0,977	0,192
2012-2013	0,819	2,049	1,312	1,471	1,474	1,821	0,363

TABELA 3
Estatísticas da dinâmica da mudança de eficiência (2006-2013)

Elaboração dos autores.

Já a dinâmica da mudança de tecnologia apresenta uma evolução oposta à dinâmica da mudança de eficiência, como apresenta a tabela 4. A contribuição da tecnologia é negativa no início e no fim do período analisado e positiva nos anos intermediários.

TABELA 4
Estatísticas-resumo da dinâmica da mudança de tecnologia (2006-2013)

Ano inicial-final	Mínimo	Máximo	1º Quartil	Mediana	Média	3º Quartil	Desvio-padrão
2006-2007	0,888	1,101	0,949	0,993	0,987	1,012	0,051
2007-2008	0,881	1,156	0,909	0,958	0,976	1,033	0,076
2008-2009	0,802	1,005	0,904	0,923	0,914	0,926	0,042
2009-2010	1,068	1,292	1,122	1,132	1,139	1,133	0,055
2010-2011	0,991	1,457	1,167	1,346	1,260	1,349	0,142
2011-2012	1,001	1,957	1,330	1,362	1,486	1,623	0,237
2012-2013	0,512	1,112	0,565	0,668	0,680	0,714	0,141

Elaboração dos autores.

Para explorar melhor as estimativas, a análise foi dividida em dois períodos, a saber: 2006-2010 e 2010-2013, na tentativa de reduzir o impacto do período 2008-2009, que teve resultados muito baixos, na dinâmica da produtividade total dos fatores, e permitir a análise a partir de 2010 que foi o ano de introdução da LNSB.

TABELA 5
Estatísticas-resumo do índice Malmquist (2006-2010)

Índice de Malmquist	Mínimo	Máximo	1º Quartil	Mediana	Média	3º Quartil	Desvio-padrão
Eficiência	0,753	1,595	0,938	1,000	1,030	1,082	0,166
Tecnológica	0,919	1,251	0,934	0,946	1,006	1,042	0,108
Total	0,705	1,639	0,899	1,006	1,037	1,151	0,203

Elaboração dos autores.

Conforme a tabela 5, entre 2006 e 2010, o resultado das estatísticas de tendência central da produtividade total dos fatores foi em torno da unidade, indicando pouco avanço de eficiência no período.

Os resultados da tabela 6, a seguir, estimam para o período 2006-2010 a quantidade percentual de prestadoras que se enquadram nos três casos: *i*) o intervalo de confiança encontrado tenha os dois extremos menores do que a unidade, o que significa que a produtividade piorou de maneira estatisticamente significativa; *ii*) o intervalo de confiança tenha os dois extremos maiores do que a unidade, o que implica que a produtividade melhorou de maneira significativa; *e iii*) o intervalo de confiança possui um extremo menor do que a unidade e outro maior, de modo que não houve mudança estatisticamente significativa de produtividade.

Assim, verificamos que 40,7% das prestadoras apresentaram melhora estatisticamente significativa, enquanto 48,1% apresentaram piora estatisticamente significativa entre 2006 e 2010. Já a mesma análise realizada nos índices decompostos entre os efeitos mudança de eficiência e mudança tecnológica revela que a maioria das prestadoras, 70,3% e 62,9%, respectivamente, não apresentou mudança estatisticamente significativa.

Em resumo, no período 2006-2010, apesar das estatísticas-resumo indicarem pequeno avanço na produtividade, já que a mediana indica crescimento de 0,6%, houve mais prestadoras com piora significativa de produtividade do que com melhora. Assim, os quatro primeiros anos após a LNSB não representaram avanços robustos na produtividade.

TABELA 6 Mudança significativa de eficiência (2006-2010) (Em %)

Índice de Malmquist	Melhora significativa	Piora significativa	Sem mudança significativa	
Eficiência	22,2	7,4	70,3	
Tecnologia	11,1	25,9	62,9	
Total	40,7	48,1	11,1	

Elaboração dos autores.

Já os resultados para o período 2010-2013 apresentados na tabela 7 indicam um aumento significativo de produtividade; a mediana indica um aumento de 6,7%, somente atribuído ao componente tecnológico.

Na tabela 8, as estatísticas resumidas mostram que em torno de 55% das unidades analisadas apresentaram melhoria tecnológica estatisticamente significativa de produtividade contra e que 29,6% com piora estatisticamente significativa.

Para os efeitos de emparelhamento, ao contrário, não há melhora estatisticamente significativa, enquanto 48% estatisticamente pioraram.

TABELA 7
Estatísticas-resumo do índice Malmquist (2010-2013)

Índice de Malmquist	Mínimo	Máximo	1º Quartil	Mediana	Média	3º Quartil	Desvio-padrão
Eficiência	0,538	1,659	0,820	0,905	0,925	0,994	0,198
Tecnológica	0,886	1,804	0,997	1,209	1,239	1,350	0,258
Total	0,702	2,994	0,968	1,067	1,146	1,163	0,412

Elaboração dos autores.

Os resultados positivos para a mudança tecnológica no período 2010-2013 indicam que houve expansão da fronteira produtiva do setor. Entretanto, os resultados negativos do efeito emparelhamento apontam que as DMUs não ficaram mais próximas da fronteira expandida.

TABELA 8 Mudança significativa de eficiência (2010-2013) (Em %)

Malmquist	Melhora significativa	Piora significativa	Sem mudança significativa
Eficiência	0	48,1	51,8
Tecnológica	59,2	0	40,7
Total	55,5	29,6	14,8

Elaboração dos autores.

Agora, finalmente, é apresentada a análise de todo o período da amostra, ou seja, entre 2006 e 2013. Observa-se, na tabela 9, que as estatísticas-resumo apontam um avanço na produtividade total dos fatores no período. Esse aumento, como entre 2010 e 2013, é explicado pela mudança tecnológica, já que a mudança de eficiência apresentou média e mediana menores do que a unidade.

TABELA 9
Estatísticas-resumo do índice Malmquist (2006-2013)

Índice de Malmquist	Mínimo	Máximo	1º Quartil	Mediana	Média	3º Quartil	Desvio-padrão
Eficiência	0,487	1,698	0,818	0,894	0,952	1,008	0,245
Tecnológica	0,939	1,969	1,106	1,220	1,248	1,369	0,221
Total	0,802	2,841	0,979	1,067	1,194	1,150	0,451

Elaboração dos autores.

A análise da mudança significativa de cada prestadora, tabela 10, também indica avanço entre 2006 e 2013, principalmente na mudança tecnológica, onde mais de 51% das prestadoras apresentaram melhora significativa e nenhuma piorou significativamente.

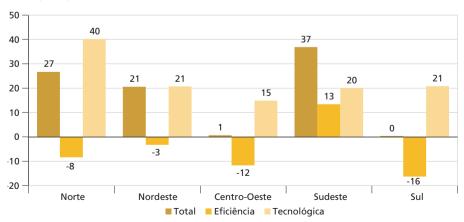
TABELA 10 Mudança significativa de eficiência (2006-2013) (Em %)

Malmquist	Melhora significativa	Piora significativa	Sem mudança significativa	
Eficiência	7,4	33,3	59,2	
Tecnológica	51,8	0	48,1	
Total	40,7	22,2	37,0	

Elaboração dos autores.

Resumindo, os sete anos analisados apresentaram avanço na produtividade total dos fatores; a mediana do índice de Malmquist aponta para um crescimento de 6,7%. É relevante que quase 41% das prestadoras tenham aumentado sua produtividade no período. Os resultados também indicam que houve uma expansão na fronteira; entretanto, mais de 59% das DMUs não apresentaram mudança significativa no que se refere à mudança de eficiência. Portanto, os anos após o novo marco regulatório foram caracterizados pelo crescimento tecnológico, mas não pelo aumento do escore de eficiência das prestadoras, no sentido de que a maioria manteve distância em relação à fronteira.

GRÁFICO 4 **Média do índice de Malmquist por região (2006-2013)** (Em %)



Elaboração dos autores.

As informações do gráfico 4 indicam que a região que mais avançou foi a Sudeste, com um crescimento médio de 37% na produtividade total dos fatores, seguida pela região Norte, com incremento de 27%. Essas duas regiões apresentam, respectivamente, maior e menor média dos escores de eficiência, conforme o observado no gráfico 3. Essa situação indica que, apesar da região Norte ter avançado bastante em termos de produtividade, a sua distância para a região Sudeste aumentou em termos de eficiência.

É relevante observar que somente a região Sudeste obteve a média da mudança de eficiência positiva, 13%, ou seja, esta foi a única região que as prestadoras de fato se aproximaram da fronteira. Assim, a região mais eficiente foi a que mais progrediu após a LNSB, o que pode ter aumentado ainda mais a heterogeneidade das *performances* entre as regiões.

Cabe ressaltar também o fraco desempenho das regiões Centro-Oeste e Sul, onde, inclusive, os resultados da mudança de eficiência indicam que, em média, as prestadoras dessas regiões se afastaram consideravelmente da fronteira produtiva.

Em suma, os resultados aqui apresentados indicam avanço na produtividade, embora ainda com baixos níveis de escores. E a evolução é dominada pelo efeito mudança de fronteira com menor importância dos efeitos emparelhamento. Essa evolução é mais acentuada no período recente, quando já estava em vigor o marco regulatório do setor.

Na comparação com estudos anteriores, que utilizaram índice de Malmquist, há convergência na dominância dos resultados de baixos escores de eficiência.

Por sua vez, os resultados aqui estimados, de pouco avanço da produtividade total para o período 2006-2010, confirmam os apresentados por Carvalho (2014) e Barbosa (2012) para período equivalente, embora esses autores não encontrem a mesma dinâmica da produtividade regional.

Em outros estudos para períodos anteriores, como Sato (2011) e Seroa da Motta e Moreira (2006), há, contudo, alinhamento de resultados quanto à dinâmica da produtividade no tempo que aqui se observou para o período 2010-2013.

5 CONCLUSÃO

Com o objetivo de analisar a eficiência na prestação dos serviços de água e esgoto depois da Lei Nacional do Saneamento Básico, foram selecionadas as maiores prestadoras de serviço de cada estado entre 2006 e 2013. Foi aplicada a DEA, em conjunto com o *bootstrap*, para a análise estática, e o índice de Malmquist, também em conjunto com o *bootstrap*, para uma análise dinâmica.

Na análise estática, verificou-se que os escores no último ano permaneceram baixos, heterogêneos, e a análise da trajetória da eficiência, por meio das medidas de tendência central, foi inconclusiva.

A análise da dinâmica, realizada ano a ano com a utilização do índice de Malmquist, informa que, em cinco dos sete anos, os valores da média e da mediana apontaram para o avanço da produtividade total dos fatores.

Adicionalmente, o padrão de eficiência estimado apresenta grande disparidade entre as regiões do país, pois as provedoras das regiões menos desenvolvidas são geralmente menos eficientes. Tal indicação sugere diferenças locacionais de gestão técnica e corporativa e, portanto, os esforços de expansão do setor devem também considerar ajustes regionais, e mesmo locais, diferenciados, no desenho e na adoção de instrumentos de gestão.

Os primeiros três anos (2010-2013) da LNSB foram caracterizados pela expansão da fronteira produtiva, mas não pelo aumento do escore de eficiência das prestadoras, ou seja, a maioria delas não se aproximou da fronteira ampliada. Dessa forma, o período após a LNSB apresentou um aumento estatisticamente significativo de produtividade, no qual se concentrou também a expansão da fronteira do setor. Contudo, observamos um nível de eficiência relativamente baixo e heterogêneo, e, portanto, o avanço no período não se mostrou suficiente para mudar a realidade dos serviços de água e esgotos no país. Então, não seria possível descartar a hipótese de que o novo marco regulatório introduzido pela LNSB tenha expandido o potencial do setor. Por seu turno, também não é possível descartar a hipótese de que as empresas operadoras não foram capazes de se tornar mais eficientes no sentido de Pareto-Kopmans. Dito de outro modo, teria sido possível produzir mais resultados com os mesmos recursos empregados ou, alternativamente, produzir os mesmos resultados com menores dispêndios de recursos.

Da leitura do parágrafo anterior, deduzimos que existiria a necessidade de introdução de incentivos econômicos no processo regulatório para reversão dessa estagnação. A LNSB define explicitamente que os serviços de saneamento serão prestados com eficiência e sustentabilidade econômica. Além disso, o referido diploma legal preconiza que os contratos de concessão devem prever metas progressivas e graduais de expansão de serviços, de qualidade, de eficiência e de uso racional da água e da energia. E, mais ainda, que essa prestação deve conter mecanismos e procedimentos para a avaliação sistemática da eficiência e eficácia das ações programadas. Este texto descreve, e utiliza, metodologias de avaliação de eficiência e eficácia, no estado-da-arte, que vêm sendo utilizados no Brasil e no exterior, em diversos setores regulados, incluindo o fornecimento e tratamento de água e esgoto. Portanto, já existem a lei e os métodos. Cabe, então, tornar efetiva a melhoria de eficiência setorial.

REFERÊNCIAS

ABBOTT, M.; COHEN, B.; WANG, W. C. The performance of the urban water and wastewater sectors in Australia. **Utilities Policy**, v. 20, n. 1, p. 52-63, 2012.

BANKER, R. D.; CHARNES, A.; COOPER, W. W. Some models for estimating technical and scale inefficiencies in data envelopment analysis. **Management Science**, v. 30, p. 1078-1092, 1984.

BANKER, R. D. Report on Aneel's proposal for electricity distribution tariff regulation. [s.l.]: Aneel, 2011. (Nota Técnica, n. 101). Disponível em: https://bit.ly/2Ni6OsO. Acesso em: 11 abr. 2019.

BARBOSA, A. Pode a regulação econômica melhorar o desempenho econômicofinanceiro e a universalização dos serviços de águas e esgotos no Brasil? Brasília: Seae, 2012. 67 p.

BOGETOFT, P.; OTTO, L. Benchmarking with DEA, SFA and R. New York: Springer, 2011.

BRASIL. **Constituição da República Federativa do Brasil**. Brasília: Senado, 1988. 140 p.

_____. Ministério da Saúde. **Programa de saneamento básico**. Brasília: MS, 2002. Disponível em: https://bit.ly/32kPW95.
_____. Lei nº 11.445, de 5 de janeiro de 2007. Estabelece diretrizes nacionais para

o saneamento básico; altera as Leis nº 6.766, de 19 de dezembro de 1979, 8.036, de 11 de maio de 1990, 8.666, de 21 de junho de 1993, 8.987, de 13 de fevereiro de 1995; revoga a Lei nº 6.528, de 11 de maio de 1978; e dá outras providências. **Planalto.gov**, Brasília, 5 jan. 2007. Disponível em: https://bit.ly/2NkLtyO. Acesso em: 11 abr. 2019.

_____. Decreto nº 7.217, de 21 de junho de 2010. Regulamenta a Lei nº 11.445, de 5 de janeiro de 2007, que estabelece diretrizes nacionais para o saneamento básico, e dá outras providências. **Diário Oficial da União**, Brasília, 22 jun. 2010. Seção 1. Disponível em: https://bit.ly/2r8ZpDl. Acesso em: 11 abr. 2019.

_____. Ministério das Cidades. **Sistema nacional de informações sobre saneamento**: diagnóstico dos serviços de água e esgotos — 2013. Brasília: SNIS, 2014. 58 p. Disponível em: https://bit.ly/34yNrRW>.

BYRNES, J. *et al.* The relative economic efficiency of urban water utilities in regional New South Wales and Victoria. **Resource and Energy Economics**, v. 32, n. 3, p. 439-455, 2010.

CARMO, C. M.; TÁVORA JUNIOR, J. L. Avaliação da eficiência técnica das empresas de saneamento brasileiras utilizando a metodologia DEA. *In*: ENCONTRO NACIONAL DE ECONOMIA DA ANPEC, 31., 2003, Porto Seguro, Bahia. **Anais...** Porto Seguro: Anpec, 2003. Disponível em: https://bit.ly/2rgq2Xa. Acesso em: 11 abr. 2019.

CARVALHO, A. E. C. Caminhos para a universalização dos serviços de água e esgotos no Brasil: a atuação das entidades reguladoras para indução da eficiência dos prestadores de serviços. Dissertação (Mestrado) – Programa de Pós-Graduação em Gestão Pública, Universidade Federal do Rio Grande do Norte, Natal, 2014, 140 p. Disponível em: https://bit.ly/2NGIDDn. Acesso em: 11 abr. 2019.

CHARNES, A.; COOPER, W. W.; RHODES, E. Measuring the efficiency of decision making units. **European Journal of Operational Research**, v. 2, n. 4, p. 429-444, 1978.

COELLI, T. J. Recent developments in frontier modelling and efficiency measurement. **Australian Journal of Agricultural and Resource Economics**, v. 39, n. 3, p. 219-245, 1995.

ESTACHE, A.; GARSOUS, G.; SEROA DA MOTTA, R. Shared mandates, moral hazard, and political (mis)alignment in a decentralized economy. **World Development**, v. 83, p. 98-110, 2016.

GRIGOLIN, R. **Setor de água e saneamento no Brasil: regulamentação e eficiência**. São Paulo: FGV, 2007. Dissertação (Mestrado) – Escola de Economia de São Paulo, Fundação Getulio Vargas, São Paulo, 2007, 64 p.

IBGE – INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Síntese de indicadores sociais**: uma análise das condições de vida da população brasileira 2014. Rio de Janeiro: IBGE, 2014. Disponível em: https://bit.ly/2JTpPQ0. Acesso em: 11 abr. 2019.

LIN, C.; BERG, S. V. Incorporating service quality into yardstick regulation: an application to the Peru water sector. **Review of Industrial Organization**, v. 32, n. 1, p. 53-75, 2008.

MARINHO, A. Avaliação da eficiência técnica nos serviços de saúde dos municípios do estado do Rio de Janeiro. **Revista Brasileira de Economia**, Rio de Janeiro, v. 57, n. 3, p. 515-534, 2003.

MARINHO, A.; FAÇANHA, L. O. **Programas sociais**: efetividade, eficiência e eficácia como dimensões operacionais da avaliação. Rio de Janeiro: Ipea, 2001. (Texto para Discussão, n. 787).

_____. Hospitais universitários: indicadores de utilização e análise de eficiência. **Revista de Economia Aplicada**, São Paulo, v. 6, n. 3, p. 607-638, 2002.

MENDES, C. C.; SAMPAIO DE SOUSA, M. C. Estimando a demanda por serviços públicos nos municípios brasileiros. **Revista Brasileira de Economia**, v. 60, n. 3, p. 281-296, 2006.

MENDONÇA, M. J.; SEROA DA MOTTA, R. Saúde e saneamento no Brasil. **Planejamento e Políticas Públicas**, v. 30, p. 15-30, 2007.

PORTELA, M. *et al.* Productivity change in the water industry in England and Wales: application of the meta-Malmquist index. **Journal of the Operational Research Society**, v. 62, n. 12, p. 2173-2188, 2011.

SALLES, M. J. **Política nacional de saneamento**: percorrendo caminhos em busca da universalização. 2008. Tese (Doutorado) — Escola Nacional de Saúde Pública, Rio de Janeiro, 2008. Disponível em: https://bit.ly/2NkKKh8. Acesso em: 11 abr. 2019.

SAMPAIO, B.; SAMPAIO, Y. Influências políticas na eficiência de empresas de saneamento brasileiras. **Economia Aplicada**, v. 11, n. 3, p. 369-386, jul.-set., 2007.

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, n. 2, p. 157-181, 2005.

SATO, J. M. **Utilização da análise envoltória de dados (DEA) no estudo de eficiência do setor de saneamento**. 2011. Dissertação (Mestrado) — Programa de pós-graduação em Economia Regional, Universidade Católica de Brasília, Brasília, 2011.

SEROA DA MOTTA, R. A ausência de regulação econômica em saneamento no Brasil. *In*: SALGADO, L. H.; SEROA DA MOTTA, R. (Eds.). **Marcos regulatórios no Brasil**: o que foi feito e o que falta fazer. Rio de Janeiro: Ipea, 2005.

SEROA DA MOTTA, R.; MOREIRA, A. Efficiency and regulation in the sanitation sector in Brazil. **Utilities Policy**. v. 14, p. 185-195, 2006.

_____. Political factor an the efficiency of municipal expenditures in Brasil. **Economia Aplicada**, v. 13, n. 4, p. 353-371, 2009.

SIMAR, L.; WILSON, P. W. Sensitivity analysis of efficiency scores: how to *bootstrap* in nonparametric frontier models. **Management science**, v. 44, n. 1, p. 49-61, 1998.

_____. Theory and methodology: estimating and bootstrapping Malmquist index. **European Journal of operational research**, v. 115, p. 459-471, 1999.

SOUZA, F. C.; LEGEY, L. F. L. Dynamics of risk management tools and auctions in the second phase of the Brazilian Electricity Market reform. **Energy Policy**, v. 38, p. 1715-1733, 2010.

TUPPER, H. C.; RESENDE, M. Efficiency and regulatory issues in the Brazilian water and sewage sector: an empirical study. **Utilities Policy**, n. 12, p. 29-40, 2004.

WILSON, P. W. FEAR 1.0: a software package for frontier efficiency analysis with R. **Socio-economic planning sciences**, v. 42, n. 4, p. 247-254, 2008.

BIBLIOGRAFIA COMPLEMENTAR

MARINHO, A.; RESENDE, M. Service quality in electricity distribution in Brazil: a malmquist approach. **CESIFO Working Paper**, n. 6276, 2016.

A DURAÇÃO DAS EXPORTAÇÕES BRASILEIRAS: EVIDÊNCIA E DETERMINANTES

Hélio de Sousa Ramos Filho¹ Rennan Kertlly de Medeiros²

Este artigo analisa a duração das exportações brasileiras e seus determinantes utilizando dados desagregados em nível de produtos a Harmonized System (HS) 6 dígitos. Durante o período 1992-2013, as relações comerciais brasileiras exibiram baixa taxa de sobrevivência com mediana e média de um e três anos, respectivamente. Sob uma amostra reduzida, foram aplicados dois modelos para identificar os determinantes da duração das exportações: modelo Cox com covariáveis tempo-variantes e um modelo paramétrico com heterogeneidade não observada. Os resultados encontrados sugerem que o tamanho econômico (PIB *per capita*) e a proximidade geográfica do país-parceiro comercial do Brasil são importantes para maior duração das relações. Ademais, novos acordos preferenciais de comércio, política cambial favorável às exportações e melhoria na qualidade dos produtos elevam a duração das exportações brasileiras.

Palavras-chave: duração; exportações; análise de sobrevivência.

THE DURATION OF BRAZILIAN EXPORTS: EVIDENCE AND DETERMINANTS

This paper analyzes the duration of Brazilian exports and their determinants using disaggregated data for 6-digit HS (Harmonized System). During the period of 1992-2013, the Brazilian trade relations exhibit low rate of survival with median and average of one and three years, respectively. Under a reduced sample, two models were applied to identify determinants of the duration of exports: Cox models with covariates time-variants and a parametric model with unobserved heterogeneity. The results suggest that the economic size (GDP per capita) and geographical proximity with the Brazil are important for longer duration of relations. In addition that new trade agreements, exchange-rate policy favourable to exports and improvement in the quality of the products increase the duration of Brazilian exports.

Keywords: duration; exports; survival analysis.

JEL: F14; F19; C41.

1 INTRODUÇÃO

As novas teorias do comércio concentram-se em compreender os determinantes e as implicações microeconômicas e macroeconômicas da presença de firmas heterogêneas em atividades de exportação. Essa literatura concentra-se nas diferenças entre firmas exportadores e não exportadores (Bernard e Jensen, 1999; Eaton e Kortum, 2002; Melitz, 2003). As firmas exportadoras tendem a apresentar maior produtividade quando comparadas às firmas não exportadoras; porém, há divergência

^{1.} Professor adjunto do Departamento de Economia da Universidade Federal da Paraíba (UFPB). *E-mail*: <helioramos@gmail.com>.

^{2.} Doutorando em economia aplicada pelo Programa de Pós-Graduação em Economia da UFPB. *E-mail*: <rennanmedeiros22@hotmail.com>.

quanto à explicação deste fenômeno, se seria causa ou consequência do comércio internacional.³ Entretanto, quando as firmas entram no mercado, podem, dentro de um curto período de tempo, cessar a atividade de exportação, logo a análise da duração dos fluxos comerciais merece maior atenção.

Ao longo da década de 2000, inúmeros estudos discutem a importância de análise da duração das relações de exportação com o intuito de encontrar causas e consequências do sucesso ou fracasso das relações comerciais dos países (Besedes e Prusa, 2006a; Caron e Anson, 2008; Brenton, Pierola e Uexkull, 2009; Fugazza e Molina, 2011). Besedes e Prusa (2006a) foram precursores na discussão de análise de sobrevivência do comércio.⁴ Segundo os autores, a duração média das relações comerciais dos Estados Unidos com os países integrantes da Organização para a Cooperação e Desenvolvimento Econômico (OCDE) é muito curta, situandose entre dois a quatro anos. Os estudos também apontaram que a sobrevivência relacionada ao nível do produto deve ser mais curta do que ao nível da indústria.

Brenton, Pierola e Uexkull (2009) chegaram à confirmação de que os laços culturais e geográficos entre os parceiros comerciais, juntamente com o tamanho do mercado e a experiência de exportação desempenham papel importante na sobrevivência de exportação de países. Por sua vez, em um estudo pioneiro sobre a duração das exportações brasileiras com valores de até US\$ 20 mil, Caron e Anson (2008) encontraram evidências de que serviços de isenção e de reduções dos custos de transporte são fundamentais para a inserção e sobrevivência das exportações brasileiras em países em desenvolvimento. Essas políticas mencionadas contribuíram significativamente para a duração da exportação de produtos de 8 mil empresas brasileiras, ou seja, 40% das firmas do Brasil puderam manter-se no mercado após a implementação dessas ações.

Diante do exposto, este artigo pretende analisar a duração das exportações brasileiras e seus determinantes em nível de produtos no período 1992-2013. Espera-se que este estudo contribua com a literatura incipiente, inclusive apontando sugestões de políticas comerciais e/ou industriais.

Além desta introdução, este trabalho está dividido em quatro seções. A seção 2 dedica-se à revisão da literatura empírica, tendo como finalidade apresentar as evidências nacionais e internacionais sobre a duração do comércio. A seção 3 apresenta os principais instrumentos metodológicos utilizados para analisar a duração dos produtos comercializados pelo Brasil. A seção 4 é reservada à apresentação dos resultados e das discussões da pesquisa. Por fim, a seção 5 destina-se às considerações finais.

^{3.} Embora a evidência empírica se mostre mais favorável ao efeito autosseleção.

^{4.} Para esses autores, quanto maior a desagregação dos dados e o período de tempo analisado, menor tenderá a ser o tempo de sobrevivência de um produto comercializado por um país.

2 REVISÃO DE LITERATURA: EVIDÊNCIAS EMPÍRICAS SOBRE A DURAÇÃO DO COMÉRCIO

As teorias tradicionais do comércio internacional têm pouco a dizer sobre a duração dos fluxos comerciais. Só a partir da década de 2000, alguns trabalhos surgiram buscando descrever a duração dos fluxos comerciais e seus principais determinantes. Esta seção tenta realizar uma breve revisão da literatura a respeito do tema.

Besedes e Prusa (2006a; 2006b) foram pioneiros na aplicação das técnicas de análise de sobrevivência para estudar a duração das importações dos Estados Unidos, e, assim, observaram o período de tempo até que um país deixe de exportar um produto para os Estados Unidos, um evento referido como "fracasso". Os autores utilizaram os dados de importação dos Estados Unidos e aplicaram o estimador não paramétrico Kaplan-Meier aos períodos 1978-1988 e 1989-2001. Os resultados da análise sugerem que, aproximadamente, metade de todas as relações comerciais duram apenas um ano. No entanto, uma vez que se considera a censura na amostra, a duração média das relações com os Estados Unidos foi de, aproximadamente, dois anos.

Brenton, Pierola e Uexkull (2009) procuraram encontrar explicações para a baixa taxa de sobrevivência das relações comerciais. Para a análise, utilizou-se um conjunto de dados que se estendem por vinte anos, de 1985 a 2005, e o modelo de estimação Prentice-Gloeckler.⁵ Os autores chegaram à conclusão de que a experiência com a exportação do mesmo produto para outros mercados ou produtos diferentes contribuem para aumentar fortemente a possibilidade de duração da sobrevivência de exportação.

Por sua vez, Besedes e Blyde (2010) analisaram a duração das exportações na América Latina. Os dados utilizados pelos autores são desagregados em nível de 4 dígitos da Standard International Trade Classification (SITC), abrangendo o período 1975-2005, concentrados ao setor manufatureiro. Os resultados encontrados mostram que a taxa de crescimento das exportações na América Latina tem sido relativamente baixa, sobretudo quando comparada com outras regiões, como a União Europeia ou a Ásia Oriental. Entretanto, quando comparada com os Estados Unidos, o Brasil ainda apresenta desvantagem nos resultados. Por fim, comprovouse que todos os outros países da América Latina têm funções de sobrevivência que são expressivamente menores que as funções de sobrevivência das três regiões de referência. No entanto, os autores não apresentaram explicações subjacentes ao fenômeno observado no Brasil.

^{5.} Por se tratar de dados organizados anualmente e com grandes possibilidades de falhas, utilizou-se esse modelo, pois ele trata de duração discreta. Ver Lancaster (1990).

Com o intuito de analisar o sucesso ou a falha das exportações africanas, Cadot *et al.* (2013) analisam dados de quatro países africanos (Malawi, Mali, Senegal e Tanzânia). É possível destacar que há um elevado grau de consonância entre os baixos índices de sobrevivência das relações comerciais e a alta e média renda dos países desenvolvidos. Sendo assim, a diferenciação do produto por parte das empresas é uma das maneiras favoráveis à propensão de maior duração das transações comerciais.

Já Fugazza e Molina (2011) estudaram os determinantes da sobrevivência das exportações, dando ênfase ao custo fixo e à dificuldade de exportar. A análise é baseada em dados de comércio bilaterais desagregados para uma amostra de 96 países desenvolvidos e países em desenvolvimento, durante o período 1995-2004. A aplicação da análise de sobrevivência, utilizando a estratégia de estimação na versão canônica do modelo de Cox, encontrou os seguintes resultados: *i*) a duração das relações comerciais de economias mais ricas se caracteriza por taxas de risco menores; *ii*) a inserção de produtos diferenciados atenua a taxa de risco no comércio; *iii*) maiores custos na exportação aumentam sistematicamente a probabilidade de falha de exportação; e, por fim, *iv*) quanto maior o volume de transação comercial, maior tenderá a ser a probabilidade de sobrevivência.

Falentina e Ichihashi (2013) examinam o comércio da Indonésia buscando investigar a duração e os fatores que afetam sua taxa de risco. A partir de uma análise descritiva acerca dos dados das exportações e importações da Indonésia, no período 1990-2011, comprovou-se que a maioria dos fluxos comerciais indonésios são de curta duração e, além disso, em muitos casos não há evidência de comércio sendo frequentemente interrompido (parou e reiniciou). Para identificar os fatores que afetam a taxa de risco nos fluxos comerciais, utilizou-se um modelo de sobrevivência de tempo discreto.⁶

Stirbat, Record e Nghardsaysone (2013) buscaram investigar os fatores que motivam a sobrevivência das exportações do Laos. Os dados coletados correspondem ao curto período 2005-2010, e o principal instrumento para análise é um modelo de tempo discreto que representa heterogeneidade não observada. Assim, o estudo centrou-se sobre o impacto da experiência e os efeitos de redes sobre as possibilidades de que um contrato comercial (operacionalizado como um trio empresa-produtodestino) sobrevive nos meses posteriores. Em linha com os resultados encontrados em literaturas citadas anteriormente, também foi confirmado que a riqueza do parceiro comercial e o crescimento de seu produto interno bruto (PIB) são favoráveis para spells de exportação mais longos, enquanto o oposto é verdadeiro para a distância.

^{6.} Uma característica convencional desses modelos de sobrevivência é que eles se tornam modelos para respostas dicotômicas, quando os dados foram ampliados para chamadas de dados do produto do período. Modelos *logit* e *probit* padrão podem, então, ser utilizados, bem como modelos de *log-log* complementares (Beck, Katz e Tucker, 1998).

Tratando-se dos estudos sobre a sobrevivência das exportações do Brasil, Caron e Anson (2008) procuraram entender o efeito de políticas comerciais que viessem a promover o aumento das exportações de baixo valor agregado. O banco de informações utilizadas corresponde ao período 1999-2007, tendo como instrumento básico de análise técnicas econométricas. Constatou-se que, através de um serviço facilitação de comércio (isenção de impostos e exportação em quantidades limitadas), é possível aumentar o tempo de sobrevivência dos países em desenvolvimento, já que esses enfrentam dificuldades para entrar no mercado externo, devido aos elevados custos de transporte.

Medeiros e Ramos (2016) analisaram a relação entre a diferenciação de mercados (localização geográfica e nível de renda), a diferenciação de produtos (diferenciado ou homogêneo) e a duração das exportações brasileiras. Os resultados encontrados sugerem que as relações de exportação são mais duradoras se forem estabelecidas com países de renda alta integrantes da OCDE ou se envolverem países da América do Norte. Ademais, constatou-se que a duração média das exportações brasileiras envolvendo produtos diferenciados foram superiores aos outros dois tipos de produtos (homogêneo e diferenciado organizado em mercado homogêneo).

Em suma, a duração das exportações parece diferir entre países desenvolvidos e em desenvolvimento, mas não de maneira expressiva. Ademais, a revisão sugere ainda que, alguns fatores, a depender do país, afetam a duração: PIB *per capita*, distância, qualidade e experiência.

3 PROCEDIMENTOS METODOLÓGICOS DA PESQUISA

3.1 Análise de duração (sobrevivência)

A análise de duração é geralmente definida como um conjunto complexo de métodos de análise de dados cuja variável principal é o tempo até a ocorrência de um evento de interesse. No contexto de comércio, o evento de interesse (ou falha) se refere à interrupção de relações comerciais do Brasil com o resto do mundo, isto é, a exportação de um produto específico para determinado parceiro deixa de existir, em outras palavras, o evento corresponde à saída de um produto específico de determinado mercado.

Em contrapartida, o número de períodos em que um produto específico foi exportado para um dado mercado, ou cujos valores das exportações positivas foram positivos, é denominado de *export spell*. Portanto, corresponde ao número de vezes em que o Brasil comercializou determinado produto. Quando a exportação de um produto específico é interrompida – saída do mercado –, ocorre a *morte* do *export spell*. Portanto, a duração das exportações (ou tempo de sobrevivência/vida das exportações) corresponde ao número de anos consecutivos cujas exportações foram positivas.

Ademais, dado que a falha ou o evento de interesse (saída de um produto de determinado mercado) pode não ocorrer necessariamente para todas as relações comerciais do Brasil durante todo o período em análise, então os dados de exportação podem sofrer censura. A censura ocorre quando o tempo de vida/duração de um produto não é observado. Os dados usados correspondem a produtos específicos, logo as taxas de sobrevivência são *taxas de sobrevivência de produtos*. A censura presente nos dados foi tratada.

3.2 Estratégia empírica

Esta subseção dedica-se a apresentar a estratégia empírica para analisar a duração das exportações brasileiras. Dado que o objetivo é investigar o tempo de vida dos produtos exportados pelo Brasil para o resto do mundo, a estratégia usada consiste em: *i*) calcular a taxa de sobrevivência dos produtos exportados ou igualmente a duração das relações de exportação; e *ii*) identificar os principais determinantes da duração das exportações brasileiras (ou a duração das relações de exportação).

No contexto da duração das exportações, a taxa de sobrevivência ou função de sobrevivência pode ser definida como a probabilidade de uma relação de exportação entre Brasil e resto do mundo sobreviver até o tempo t, ou seja, é a probabilidade de que não ocorram falhas na relação comercial entre os dois antes da data t. Define-se T como tempo até a ocorrência de uma falha (saída do produto de determinado mercado). Admite-se que T seja uma variável aleatória discreta assumindo valores t_i , i=1,2,...,n em que: $t_1 < t_2 < \cdots < t_n$. A função de sobrevivência S da variável aleatória T é dado por:

$$S(t) = \Pr(T > t) = \sum_{t_i > t} p(t_i), \tag{1}$$

em que $p(t_i)$ é a função densidade de probabilidade. Ademais, a função risco ou taxa de risco é a razão entre a probabilidade de risco/falha (saída do produto do mercado) ocorra em T condicionado à probabilidade de sobrevivência (permanência do produto no mercado) até o tempo t_i , sendo definido por:

$$h(t_i) = \Pr(T = t_i | T > t_i) = \frac{p(t_i)}{s(t_i - 1)}, i = 1, 2, \dots, n,$$
 (2)

em que $s(t_0) = 1$. Assim, a função de sobrevivência e a função de risco estão relacionadas por meio da seguinte expressão:

$$S(t) = \prod_{t_i < t} [1 - h(t_i)] \tag{3}$$

Em outras palavras, a análise de sobrevivência e seus parâmetros (taxa de risco e taxa de sobrevivência) dependem de uma relação comercial bilateral duradoura. A função de sobrevivência é geralmente obtida utilizando o estimador limite de Kaplan-Meier não paramétrico. Frequentemente, admite-se que a amostra

contenha n observações independentes (t_i, c_i) , em que $i=1,2,\ldots,n,$ t_i é o tempo de sobrevivência e c_i é a variável indicadora de censura da observação c_i assumindo valor de 1 se ocorrer a falha (saída do produto de determinado mercado exportador) e 0 caso contrário. Ademais, admite-se que são observados c_i tempo de falhas e o tempo de sobrevivência ordenado do menor para o maior seja denotado por c_i con c_i o número de indivíduos (produtos exportados) em risco de falhar em c_i e representado por c_i e seja c_i o número de falhas observadas, então o estimador Kaplan-Meier da função de sobrevivência é expresso por:

$$\widehat{S(t)} = \prod_{t_i < t} \left[\frac{n_i - d_i}{n_i} \right]. \tag{4}$$

Por convenção, $\widehat{s(t)} = 1$ se $t < t_{(1)}$. Conforme Besedes e Prusa (2006a; 2006b), o estimador Kaplan-Meier é robusto à censura e usa informações de observações censuradas e não censuradas. Por sua vez, a função de risco é estimada a partir da razão entre o número de indivíduos que falham (produto deixa de ser exportado) em relação ao número de indivíduos em risco (com potencial de saída) em um dado ano:

$$\widehat{h(t)} = \frac{d_i}{n_i} \,. \tag{5}$$

Todavia, o estimador não paramétrico de Kaplan-Meier não assume qualquer relação funcional entre s(t) e covariáveis, isto é, ele permite apenas calcular a duração das relações de exportação. Besedes e Prusa (2006b) propõem o uso de modelos de duração semiparamétricos de Cox (1972) para analisar os determinantes da duração das relações comerciais. Este modelo tem a vantagem de não requerer uma especificação para a distribuição da duração; porém, ele apresenta a desvantagem de admitir que as variáveis afetam proporcionalmente a função de risco. A hipótese de proporcionalidade pode ser relaxada ao incluir termos de iteração entre a(s) variável(is) independente(s) que variam no tempo e a variável de tempo de sobrevivência.

Portanto, o modelo Cox semiparamétrico permite identificar quais fatores contribuem para maior sobrevivência de um produto específico no mercado exportador. Em outras palavras, o que poderia reduzir a taxa de risco dos produtos brasileiros? O modelo Cox de riscos proporcionais é dado por:

$$h(t|x_i) = h_0(t)\exp(x(t_i), \beta_x), \tag{6}$$

em que t é o tempo de sobrevivência, x(t) é um vetor de covariáveis explicativas que variam em t e representam as características do produto i, $h_0(t)$ é denominada de função base de risco⁷ e β_x é um vetor de coeficientes que mede o efeito das características do produto i sobre a sua taxa de risco de sobrevivência.

^{7.} O termo $h_0(t)$ representa o risco no tempo t quando x(t) = 0.

Contudo, a presença de efeitos específicos não observáveis em cada grupo produto-mercado provoca heterogeneidade não observada e viesa as estimativas do modelo Cox. Desse modo, a avaliação da robustez dos resultados requer um modelo de duração com heterogeneidade não observada específica ao grupo produto-mercado. Tais modelos são conhecidos por modelos *frailty shared*⁸ e são similares aos modelos de dados em painel com efeitos aleatórios. Especificamente, seguindo a literatura, assume-se que a distribuição seja do tipo Weibull e que a heterogeneidade não observada segue uma distribuição Gama.

Besedes e Prusa (2006a; 2006b), Breton, Pierola e Uexkull (2009) e Fugazza e Molina (2011) analisaram vários determinantes da duração das exportações de países desenvolvidos e em desenvolvimento. Seguindo de perto esses autores, foram utilizadas as variáveis, a seguir, ajustadas aos dados brasileiros.

- 1) Variáveis gravitacionais: PIB per capita do Brasil e dos países parceiros proxies para nível de desenvolvimento, bem como para o potencial do mercado do comprador, da distância, da contiguidade (fronteira) e da língua comum. Espera-se um efeito positivo das variáveis língua comum e fronteira (ambas contribuem para reduzir os custos de comércio e facilitam a estabilidade das exportações) e tamanho da economia (quanto maior o mercado exportador, mais fácil é de encontrar compradores em potencial), e um impacto negativo da variável distância. Conforme Rauch (1999), estes fatores contribuem para o estabelecimento e/ou fortalecimento de uma relação comercial.
- 2) Variáveis de qualidade do produto exportado: o valor unitário é geralmente uma *proxy* para qualidade do produto. De acordo com essa abordagem, altos valores unitários refletem produtos com qualidade elevada que poderiam garantir maior estabilidade às exportações. De acordo com essa abordagem, altos valores unitários refletem produtos com qualidade elevada que poderiam garantir maior estabilidade às exportações.
- 3) Variáveis de experiência na atividade exportadora: o valor total exportado de um dado produto tem sido utilizado para capturar a experiência na exportação. Quando um dado produto já é exportado para outros países, os custos atrelados aos novos mercados (similares) seriam menores (efeito aprendizado), elevando as chances de relações comercias mais duradouras (Fugazza e Molina, 2011). Assim, espera-se que o valor total exportado e a duração de exportação apresentem relação positiva.

^{8.} A heterogeneidade não observada também é referida como *frailty* (tradução: fragilidade), especialmente na bioestatística (indivíduos mais *frágeis* apresentam maior risco de mortalidade e menor possibilidade de sobrevivência). Para mais detalhes sobre essa classe de modelos, consultar Cleves, Gould e Gutierrez (2008).

^{9.} Ver, por exemplo, Hallak (2006) e Bastos e Silva (2010).

^{10.} Valores unitários são calculados ao dividir o valor das exportações pela quantidade exportada do produto. Hallak e Schott (2011), Khandelwal (2010), Verhoogen (2008) e Feenstra (1994) desenvolveram metodologias alternativas ao uso do valor unitário para mensurar a qualidade dos produtos. Contudo, o emprego destas metodologias foge ao escopo deste artigo.

- 4) Variável de confiança: os valores iniciais de exportação, isto é, o valor das exportações no início de uma relação comercial, são considerados proxy para o nível de confiança entre os parceiros comerciais. Assume-se que as relações comerciais tendem a ser mais duradoras se dois países iniciam o comércio envolvendo grandes valores.
- 5) Variáveis de política: acordo preferencial comercial (RTA) e taxa de câmbio real. O RTA é uma variável *dummy* que indica a participação do Brasil em acordo comercial comum. Espera-se que ser membro de RTA afete positivamente as exportações, permitindo maior duração das relações comerciais. A duração das exportações pode ainda depender positivamente da variação da taxa de câmbio real. Isto é, quanto maior a depreciação cambial, maior a duração das relações de exportação (menor o risco de saída de um produto de determinado mercado). Desse modo, as oscilações (primeira diferença) da referida variável foram incluídas na análise como medida de volatilidade do câmbio.

Assim, seguindo a literatura internacional, essas variáveis foram utilizadas na identificação dos determinantes da duração das exportações brasileiras.

3.3 Dados, fontes e variáveis utilizadas

Seguindo a literatura, a base de dados utilizada é oriunda do sítio das Nações Unidas para o Comércio (United Nations Comtrade Database – UN Comtrade). Os dados são desagregados a 6 dígitos e seguem o sistema de classificação de produtos (*Harmonized System* – HS 6). A base de dados compreende fluxos de relações comerciais entre o Brasil e 193 países parceiros. A base completa é formada por 1.099.973 observações. O quadro 1 reporta a descrição de variáveis, fontes e definições.

QUADRO 1 **Descrição das variáveis**

Variáveis	Sigla	Descrição	Fonte
PIB per capita	PIBpcapita	Log do PIB do país de destino e do Brasil (preços constantes de 2000)	WDI
Valor inicial	Exp_inicio	Log das exportações do produto no início do export spell	UN Comtrade
Valor unitário	V_unitario	Log do valor unitário do produto exportado	UN Comtrade
Experiência	TotExport	Log do valor total exportado pelo Brasil do produto	UN Comtrade
Distância	Dist	Log da distância entre os países	CEPII
Contiguidade	Contig	Variável <i>dummy</i> : 1 para país que faz fronteira com o Brasil e 0 caso contrário	CEPII
Língua comum	Comlang	Variável <i>dummy</i> : 1 para país que possui a língua portuguesa como oficial e 0 caso contrário	CEPII
RTA	RTA	Variável <i>dummy</i> : 1 se país possui acordo comercial com Brasil e 0 caso contrário	CEPII
Taxa de câmbio	∆tx_cambio	Log da variação da taxa de câmbio real (R\$/US\$)	WDI

Fontes: Centro de Estudos Prospectivos e Informações Internacionais (CEPII), disponível em: https://comtrade, disponível em: https://comtrade.un.org. Indicadores de desenvolvimento mundial do Banco Mundial (WDI), disponível em: https://data.worldbank.org.

Para análise da duração das relações de exportação, é fundamental, ainda, a definição do período amostral. Neste artigo, a disponibilidade de dados limitou-se ao período 1992-2013, no qual Brasil comercializou até 4.868 diferentes produtos. Conforme mencionado, as observações estão sujeitas a censura à esquerda ou à direita, dependendo do período de análise escolhido.

Se o coeficiente da regressão é menor (maior) que 1, isto implica que a variável independente tem um efeito negativo (positivo) sobre a taxa de risco (ou falha) e um efeito positivo (negativo) sobre a taxa de sobrevivência; outras coisas iguais. Valores menores significam menor risco (ou falha) e, portanto, maior sobrevida dos produtos exportados.

4 ANÁLISE DOS RESULTADOS E DISCUSSÕES

4.1 Duração das exportações brasileiras: análise descritiva

A duração das exportações corresponde ao número de anos em que um produto é continuamente exportado do Brasil para um país de destino em particular. Dado que o período de análise compreende de 1992 a 2013, a duração máxima de uma relação comercial seria 21 anos. É importante destacar que algumas dessas relações comerciais poderiam parar e se reestabelecer, ou seja, exibiriam múltiplos intervalos de tempo de comércio (denominado de múltiplos *spells*). Contudo, 73% das relações de exportação brasileiras – portanto, a maioria delas – ocorrem em um período de tempo único (*spell* único). Por essa razão, este estudo utilizou todos os tipos de *spells*.

O primeiro resultado em relação à duração das relações comerciais advém da estatística descritiva. A tabela 1 apresenta a duração das exportações brasileiras a 2, 4 e 6 dígitos sob classificação HS. Conforme a tabela 1, os dados consistem de 311.176 spells observados a 6 dígitos, 188.090 spells observados a 4 dígitos e 28.856 spells observados a 2 dígitos. Como esperado, quanto maior a agregação dos dados, menor é a possibilidade de observar as entradas e saídas de produtos de determinados mercados, logo a duração tende a se elevar.

De acordo com a tabela 1, a duração média das exportações brasileiras a 6 dígitos é relativamente curta, com mediana e média de 1 e 3,15 anos, respectivamente. A probabilidade de uma relação de exportação durar – taxa de sobrevivência – por um ano é de, aproximadamente, 55%, enquanto a probabilidade de sobreviver por seis anos é de apenas 23%.

Ao analisar as taxas de sobrevivência por nível de renda e distância entre Brasil e países parceiros, os resultados da tabela 1 indicam como são as relações de exportações brasileiras, conforme comentado a seguir.

- 1) Mais duradoras quando o Brasil comercializa com países de renda alta duração média de 3,35 anos, com probabilidade de sobrevivência por um ano de 56%. Outro fato interessante é que relações comerciais com países de renda média alta garantem maior probabilidade de sobrevivência dos produtos brasileiros, embora sua duração seja inferior às envolvendo países de renda alta. A existência do BRICS e dos acordos preferenciais poderiam explicar tal resultado.
- 2) Mais duradora quando o Brasil comercializa com países próximos (até 5 mil km) duração média de 4,3 anos e probabilidade de sobrevivência de 62% (um ano), 49% (dois anos) e 32% (seis anos). Desse modo, o resultado sugere que reduções nos custos de transporte sejam importantes para maior sobrevida dos produtos exportados (relações comerciais mais duradouras).

TABELA 1

Duração das exportações brasileiras: análise descritiva (1992-2013)

Nível de agregação	Tamanho (anos) observado de <i>spell</i>		Taxa de sobrevivência (km) estimada			Número total de spells	Número total de códigos de
	Média	Mediana	1 ano	2 anos	6 anos	зрепз	produtos
HS6	3,15	1	0,55	0,40	0,23	311.176	4.868
HS4	4,03	2	0,57	0,43	0,28	188.090	1.239
HS2	5,73	2	0,65	0,53	0,39	28.856	97
HS6							
Nível de renda¹							
Renda alta	3,35	1	0,56	0,42	0,24	155.982	4.800
Renda média alta	3,3	1	0,56	0,43	0,25	89.980	2.123
Renda média baixa	2,82	1	0,52	0,37	0,21	52.566	4.149
Renda baixa	1,94	1	0,40	0,25	0,10	12.738	4.595
Distância (km)							
Até 5 mil	4,3	1	0,62	0,49	0,32	81.158	4.640
Entre 5 mil e 10 mil	3,19	1	0,55	0,41	0,22	124.303	4.656
Mais de 10 mil	2,85	1	0,51	0,36	0,18	61.436	3.768

Fonte: UN Comtrade. Disponível em: https://comtrade.un.org/>.

Elaboração dos autores.

Nota: 1 Å classificação dos países por nível de renda per capita seguiu a definição do Banco Mundial.11

Obs.: Os dados utilizados sofreram censura tipo I, na qual não é possível observar o ano de início do spel/ de exportação.

^{11.} Em 2013, a classificação do Banco Mundial foi atualizada e definiu: i) países de renda baixa são aqueles com renda per capita inferior ou igual a US\$ 1,035; ii) países com renda média baixa apresentam renda per capita entre US\$ 1,036 e US\$ 4,085; iii) países de renda média alta com renda per capita de US\$ 4,086 a US\$ 12,615; e, iv) países de renda alta com renda per capita superior ou igual US\$ 12,616.

As análises agregadas e desagregadas demostram que as taxas de sobrevivência dos produtos estão se reduzindo à medida que o tempo transcorre. A probabilidade de uma relação de exportação durar até um ano é relativamente alta (55%), porém decresce em ritmo acelerado até o 14º ano. A partir desse ponto, a probabilidade se estabiliza em, aproximadamente, 16%. Dessa maneira, uma vez estabelecida uma relação comercial que dure até o 14º ano, é provável que ela sobreviva por um período maior. Há um tipo de efeito limiar no 14º ano, que Pakes e Erickson (1998) denominam de efeito limiar de comércio (the trade threshold effect).

O resultado sugere, ainda, a presença do princípio da dependência negativa em relação à duração – a probabilidade de falha diminui com a duração das exportações brasileiras. Isto é, quanto maior a permanência de um produto em determinado mercado, menor é a probabilidade de que essa relação de exportação seja interrompida nos anos subsequentes. O gráfico 1 exibe as curvas de sobrevivência considerando diferentes níveis de desagregação dos dados de exportação brasileira.

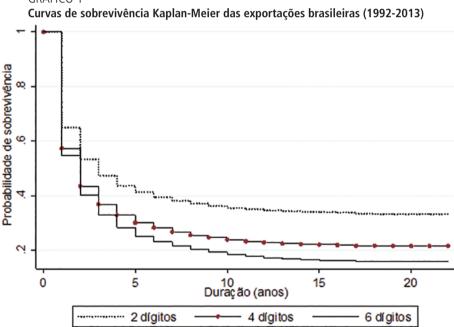


GRÁFICO 1

Fonte: UN Comtrade. Disponível em: https://comtrade.un.org/>.

Elaboração dos autores.

Obs.: Figura reproduzida em baixa resolução e cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

Portanto, a maioria das relações comerciais envolvendo as exportações brasileiras parece exibir curta duração, dependência negativa e efeito limiar no 14º ano. De acordo com Besedes e Prusa (2011), o Brasil apresenta relações comerciais mais duradoras quando comparado aos demais países latino-americanos e próximas da duração das exportações da União Europeia e da Ásia, embora ainda inferior aos Estados Unidos.

4.2 Determinantes da duração das exportações brasileiras

Na tentativa de identificar os determinantes da duração das relações de exportação brasileira, foi aplicado o modelo semiparamétrico Cox com covariáveis tempovariante. No entanto, a exportação do mesmo produto para diferentes destinos pode gerar taxas de sobrevivência diferentes, ou seja, a duração da relação de exportação envolve efeitos específicos não observáveis da relação produto-mercado. Desse modo, estimou-se, ainda, um modelo paramétrico com heterogeneidade não observada. Os resultados estão apresentados na tabela 2.

Foram estimados quatro modelos Cox (modelos 1 a 4) e um modelo paramétrico com heterogeneidade não observada (modelo 5). A segunda e terceira coluna (modelos 1 e 2) utilizam a versão padrão e ampliada da equação gravitacional, respectivamente. Na quarta coluna, são incluídas variáveis *proxies* para qualidade do produto, experiência e confiança. Na quinta coluna, o modelo completo com a presença de variáveis de política comercial e de custos de entrada. Na última coluna, um modelo paramétrico com distribuição Weibull para riscos e distribuição gama para heterogeneidade não observada ou *frailty*. Vale destacar que, ao incluir variáveis gravitacionais (PIB *per capita*, distância etc.) aos dados de exportações brasileiras, a amostra ficou restrita ao período 2000-2013.

Os coeficientes dos modelos estimados representam taxas de risco (ou falha de uma relação comercial/saída do produto). A interpretação é a seguinte: um coeficiente maior (menor) que 1 implica que uma dada variável aumente (reduz) a taxa de risco, enquanto um coeficiente igual a 1 deveria ser interpretado como não afetando a taxa de risco. Assim, um coeficiente menor (maior) que 1 implica que a variável contribua para aumentar (reduzir) a duração das exportações brasileiras.

Os modelos 1 e 2 avaliam apenas as variáveis gravitacionais tradicionais como determinantes da duração. Os coeficientes apresentam os sinais esperados pela literatura gravitacional, exceto o coeficiente do PIB brasileiro.

Os coeficientes do PIB *per capita* dos importadores apresentam-se inferiores a 1 e significativos. Assim, caso haja uma elevação em 100% no PIB *per capita* do importador, a taxa de risco se reduziria em, aproximadamente, 2% (isto é, $0.97^{\ln(2)} - 1 = -0.02$). Logo, afetaria negativamente o risco e positivamente a duração das exportações, ou seja, as chances de sobrevivência da relação comercial crescem atreladas ao tamanho do mercado comprador (parceiro comercial). Tal resultado corrobora a análise descritiva da duração das exportações brasileiras na subseção 4.1.

TABELA 2	
Determinantes da duração das exportações brasileiras: modelos Cox e de dist	tribuição
Weibull	

Variáveis	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
PIBpcapita (log)	0,95	0,95	0,98	0,97	0,93
	(0,002)	(0.002)	(0,001)	(0,002)	(0,005)
PIBpcapitaBR (log)	3,05	2,99	1,01	0,34	0,45
	(0,034)	(0,038)	(0,012)	(0,004)	(0,009)
Distância (log)	1,31	1,08	1,09	1,09	1,16
	(0,005)	(0,007)	(0,008)	(0,008)	(0,016)
Contig		0,82	0,89	0,91	0,76
		(0,008)	(0,009)	(0,009)	(0,014)
Comlang		0,94	0,91	0,89	0,83
		(0.01)	(0,009)	(0,009)	(0,017)
RTA		0,71	0,75	0,78	0,62
		(0,009)	(0,009)	(0,009)	(0,013)
Exp_inicio (log)			2,036	2,124	4,97
			(0,007)	(0,007)	(0,029)
TotExport(log)			0,41	0,42	0,15
			(0,001)	(0,001)	(0,001)
V_unitario (log)			0,95	0,98	0,97
			(0,001)	(0,001)	(0,002)
Δtx_cambio Real/dólar (<i>log</i>)				1,02	1,01
				(0,027)	(0,016)
Número de indivíduos (spells)	184.627	184.627	184.627	157.106	157.106
Número de observações (N)	752.978	752.978	746.251	746.191	746.191

Elaboração dos autores.

Obs.: 1. Desvio-padrão em parênteses.

- 2. Níveis de significância: p<0,01; p<0,05; e p<0,1.
- 3. As variáveis de cada modelo possuem nível de significância p<0,01.
- 4. A indisponibilidade de dados para alguns países resultou em diferentes números de observações nos modelos.

No que se refere à distância, seu efeito é positivo e significativo sobre as taxas de risco e negativo em relação à sobrevivência (duração) das exportações. Isto é, a duração das exportações é menor para relações comerciais envolvendo países distantes do Brasil. Mais uma vez, esse resultado confirma a análise descritiva na subseção 4.1.

Por sua vez, o coeficiente de contiguidade (contig.) menor que 1 e significativo sugere que o efeito fronteira reduz o risco e eleva a duração das exportações. Relações comerciais com países não limítrofes ao Brasil apresentam 10% mais risco de falha do que relações com países vizinhos. Isto é, a probabilidade de uma relação comercial durar é maior entre países que se avizinham ao Brasil. Em relação à influência dos

acordos preferenciais de comércio sobre a duração, o coeficiente da variável RTA inferior a 1 e significativo sugere que a existência de acordos reduz o risco de uma relação comercial falhar em, aproximadamente, 12%. Isto é, a participação do Brasil em acordos preferenciais tende a elevar a duração das exportações.

Nos modelos 3 e 4 (tabela 2), além dos tradicionais fatores gravitacionais, foram incluídas variáveis que buscam captar a influência da qualidade do produto exportado, da confiança e experiência na atividade exportadora e das políticas comerciais para maior estabilidade às exportações.

Com a experiência, representada pela variável *proxy Totexports*, a duração das exportações apresenta uma relação positiva com o volume total exportado do bem. Um aumento de 100% no volume total exportado (confiança) dentro da mesma categoria de produtos a 6 dígitos provoca uma redução em 45,2% (isto é, $0.42^{\ln(2)} - 1 = -0.452$) na taxa de risco. Esse resultado sugere que a experiência na exportação de produtos é relevante para duração das relações comerciais.

Fugazza e Molina (2011) assumem que os valores iniciais de exportação representam os custos fixos de entrada. Segundo os autores, altos valores iniciais são resultados de baixo custo de acesso ao mercado. Sob essa interpretação, o coeficiente dos valores iniciais de exportação superiores a 1 e significativos indica que a taxa de risco (ou falha) é maior para fluxos de exportação com valores iniciais mais elevados, ou seja, as exportações brasileiras são operações arriscadas apesar dos custos fixos de entrada menores.

De acordo com Di Giovanni e Levchenko (2010), países com vantagens comparativas em setores seguros ou que apresentem elevada vantagem em setores arriscados tenderão a especializar-se, enquanto países com vantagens comparativas em setores de alto risco diversificam suas estruturas de exportação para assegurar as receitas de exportação. Portanto, esse resultado sugere a necessidade de intensificação do ciclo do produto – descoberta, rápido crescimento, maturação e declínio – e de diversificação da pauta das exportações brasileiras.

Por sua vez, o coeficiente variável valor unitário, *proxy* para qualidade do produto, inferior a 1 e significativo, sugere que a qualidade afeta inversamente a taxa de risco e diretamente a duração das exportações, uma vez que uma elevação em 100% no valor unitário implica uma redução de 2% (isto é, $0.98^{\ln(2)} - 1 = -0.02$) na taxa de risco. Diante desse resultado, a baixa probabilidade de sobrevivência das relações comerciais brasileiras estaria associada à qualidade dos produtos comercializados.

A despeito da taxa de câmbio, o coeficiente estimado indica que a depreciação cambial reduz a taxa de risco (ou falha da relação comercial), pois uma depreciação em 100% provocaria uma redução do risco em, aproximadamente, 2% (isto é, 1,02-1=-0,02). Uma possível explicação para esse fenômeno estaria relacionada

à melhora no poder de compra do importador, causada pela depreciação cambial, na medida em que a depreciação torna os produtos exportados pelo Brasil mais atrativos e os fluxos comerciais mais estáveis e duradouros.

O modelo 5 (tabela 2) considerou na estimação os efeitos heterogêneos específicos à combinação produto-mercado. Os resultados deste diferem do modelo 4 apenas na magnitude dos parâmetros. Isso sugere que os achados são robustos à heterogeneidade não observada.

5 CONSIDERAÇÕES FINAIS

O objetivo deste artigo consistiu em analisar a duração das exportações brasileiras e seus determinantes em nível de produto. Ao considerar o período 1992-2013, os resultados encontrados indicam que a duração das exportações brasileiras é relativamente curta: a maioria das relações comerciais sobrevivem em média três anos (com mediana de um ano). Contudo, sob certas condições, a duração pode ser maior. É o caso de relações comerciais envolvendo países geograficamente próximos ou países importadores de maior poder aquisitivo (PIB *per capita*).

Foram avaliados, ainda, os efeitos do tamanho do mercado (PIB *per capita*), distância, contiguidade (fronteira), língua comum, acordos preferenciais de comércio, qualidade do produto exportado, confiança e experiência na atividade exportadora, valor inicial da exportação e taxa de câmbio sobre a duração das exportações brasileiras. Dada a limitação de dados, a análise ficou restrita ao período 2000-2013.

Os principais resultados encontrados foram: *i*) comercializar com países de maior poder aquisitivo eleva a duração das exportações; *ii*) quanto maior a distância do Brasil em relação ao parceiro comercial, maior o risco de falha da atividade exportadora; *iii*) as relações comerciais com países vizinhos ao Brasil ou que utilizam a mesma língua, o português, contribuem para estabilidade das exportações; *iv*) a presença de acordos preferenciais de comércios proporciona maior duração das exportações; *v*) a baixa probabilidade de sobrevivência das relações comerciais brasileiras estaria associada à qualidade dos produtos comercializados; *vi*) se as exportações começam com valores elevados, as relações tendem a ser menos duradouras; *vii*) a experiência na atividade exportadora parece importante para aumentar a duração das exportações; e *viii*) a depreciação cambial estimula o crescimento das exportações e poderia garantir maior estabilidade às relações comerciais brasileiras.

Com base na análise da duração das exportações, é possível sugerir algumas implicações de política que podem ser relevantes para reduzir a mortalidade das relações comerciais brasileiras. A primeira se refere às exportações para países desenvolvidos. Embora o Brasil potencialmente comercialize com os países emergentes, políticas voltadas ao crescimento do comércio com países desenvolvidos, tanto da margem

extensiva (novos produtos ou novos parceiros comerciais) quanto intensiva (volume exportado), poderiam proporcionar maior estabilidade às exportações brasileiras. A segunda, diretamente relacionada à primeira, consiste em políticas de inovação que agreguem maior conteúdo tecnológico e qualidade aos produtos exportados pelo Brasil. Desse modo, a política de inovação em produtos contribuiria para aumentar a competitividade, tornando as relações de exportação mais duradouras e rentáveis.

A terceira implicação está relacionada à política de inserção das firmas, isto é, às medidas atuais de incentivos fiscais aos exportadores. Ao que parece, seriam necessárias medidas complementares para reduzir o risco de exportar altos valores. A quarta sugestão de política concentra-se na necessidade de incentivar e promover novos acordos comerciais. Um acordo contribui não somente para reduzir tarifas e aumentar o volume de comércio, mas também para diminuir a incerteza da transação e favorecer a estabilidade das exportações, ou seja, garantir relações comerciais de longo prazo. A quinta e última sugestão diz respeito à política cambial. Oscilações acentuadas da taxa de câmbio podem ser prejudiciais à duração da exportação, assim, a estabilidade dessa taxa, no curto prazo, contribuiria para reduzir os riscos de exportação.

As conclusões deste estudo têm implicações importantes para empresários brasileiros e formuladores de políticas. Em primeiro lugar, os resultados encontrados aqui e as novas pesquisas podem ajudar os empresários a avaliar com maior precisão suas chances de sucesso nos mercados internacionais antes que esses decidam se inserir no mercado exterior. Em segundo lugar, os formuladores de políticas não devem apenas direcionar incentivos à criação de novos exportadores, mas também precisam se preocupar com a sobrevivência dos novos exportadores, especialmente no período inicial, para sustentar o crescimento das exportações brasileiras.

A saída efêmera de novos exportadores implica custos econômicos e sociais elevados. As políticas citadas anteriormente poderiam ser direcionadas para melhorar o produto e o acesso aos mercados estrangeiros e fornecer infraestrutura de exportação, a fim de reduzir o custo de persistência das empresas brasileiras nos mercados estrangeiros.

Por fim, é importante destacar uma das principais limitações do estudo: esta pesquisa é baseada em dados secundários. Isso limitou nossas medidas da duração exata da exportação. Pesquisas futuras devem ser realizadas utilizando dados sobre o tempo de entrada e saída de empresas nos mercados de exportação no Brasil.

REFERÊNCIAS

BASTOS, P.; SILVA, J. The quality of a firm's exports: where you export to matters. **Journal of International Economics**, v. 82, n. 2, p. 99-111, Nov. 2010.

BECK, N.; KATZ, J. N.; TUCKER, R. Taking time seriously: time-series-cross-section analysis with a binary dependent variable. **American Journal of Political Science**, v. 42, n. 4, p. 1260-288, Oct. 1998.

BERNARD, A. B.; JENSEN, J. B. Exceptional exporter performance: cause, effect, or both? **Journal of International Economics**, v. 47, n. 1, p. 1-25, Feb. 1999.

BESEDES, T.; BLYDE, J. What drives export survival? An analysis of export duration in Latin America. Washington, D.C.: World Bank, 2010. p. 1-43.

BESEDES, T.; PRUSA, T. J. Ins, outs, and the duration of trade. **Canadian Journal of Economics**, v. 39, n. 1, p. 266-295, 2006a.

_____. Product differentiation and duration of US import trade. **Journal of International Economics**, v. 70, n. 2, p. 339-358, Dec. 2006b.

_____.The role of extensive and intensive margins and export growth. **Journal of Development Economics**, v. 96, n. 2, p. 371-379, 2011.

BRENTON, P.; PIEROLA, M. D.; UEXKULL, E. **The life and death of trade flows**: undestanding the survival rates of developing-country exporters. Washington, D.C.: The World Bank, 2009. p. 127-144.

CADOT, O. *et al.* Success and failure of African exporters. **Journal of Development Economics**, v. 101, p. 284-296, Mar. 2013.

CARON, J.; ANSON, J. Trade facilitation for low-valued exports in Brazil: lessons to be learned from simplified export declarations and the use of postal networks through "Exporta Fácil". **International Postal Union**, 2008. Mimeografado. Disponível em: https://bit.ly/2NN14X2.

CLEVES, M. A.; GOULD, W. W.; GUTIERREZ, R. G. An introduction to survival analysis using stata. College Station: Stata Press, 2008.

COX, D. R. Regression models and life-tables. **Journal of the Royal Statistical Society: Series B (Methodological)**, v. 34, n. 2, p. 187-202, 1972.

DI GIOVANNI, J.; LEVCHENKO, A. A. **The risk content of exports**: a portfolio view of international trade. Cambridge, United States: NBER, May. 2010. (NBER Working Paper Series, n. 16005). Disponível em: https://bit.ly/32jGGBX.

EATON, J.; KORTUM, S. Technology, geography, and trade. **Econometrica**, v. 70, n. 5, p. 1741-779, 2002.

FALENTINA, A.; ICHIHASHI, M. **Indonesian trade**: understanding the duration and the determinants of its hazard rate. Higashihiroshima: IDEC Hiroshima University, Sept. 2013. (Discussion Paper). Disponível em: https://bit.ly/2qwHG8t.

FEENSTRA, R. C. New product varieties and the measurement of international prices. **The American Economic Review**, v. 84, n. 1, p. 157-177, Mar. 1994.

FUGAZZA, M.; MOLINA, A. C. **On the determinants of exports survival**. Geneva: UNCTAD, 2011. (Policy Issues in International Trade and Commodities Study Series, n. 46). Disponível em: https://bit.ly/2NmRdrN>.

HALLAK, J. C. Product quality and the direction of trade. **Journal of International Economics**, v. 68, n. 1, p. 238-265, 2006.

HALLAK, J. C.; SCHOTT, P. Estimating cross-country differences in product quality. **Quarterly Journal of Economics**, v. 126, n. 1, 417-474, 2011.

KHANDELWAL, A. The long and short (of) quality ladders. **The Review of Economic Studies**, v. 77, n. 4, p. 1450-476, 2010.

LANCASTER, T. **The econometric analysis of transition data**. Cambridge, United Kingdom: Cambridge University Press, 1990.

MEDEIROS, R. K.; RAMOS, H. S. Duração das exportações brasileiras e a diferenciação de produtos e mercados. *In*: ENCONTRO REGIONAL DE ECONOMIA, 9., 2016, Fortaleza. **Anais**... Fortaleza: Anpec, 2016.

MELITZ, M. J. The impact of trade on intra-industry reallocations and aggregate industry productivity. **Econometrica**, v. 71, n. 6, p. 1695-1725, 2003.

PAKES, A.; ERICSON, R. Empirical implications of alternative models of firm dynamics. **Journal of Economic Theory**, v. 79, n. 1, p. 1-45, Mar. 1998.

RAUCH, J. E. Networks versus market in international trade. **Journal of International Economics**, v. 48, n. 1, p. 7-35, June 1999.

STIRBAT, L.; RECORD, R.; NGHARDSAYSONE, K. Determinants of export survival in the Lao PDR. **Review of Policy Research**, n. 6301, 2013.

VERHOOGEN, E. A. Trade, quality upgrading and wage inequality in the Mexican manufacturing sector. **The Quarterly Journal of Economics**, v. 123, n. 2, p. 489-530, May 2008.

BIBLIOGRAFIA COMPLEMENTAR

HELPMAN, E. **Understanding global trade**. Cambridge, United States: Harvard University Press, 2011.

INTRA-INDUSTRY TRADE: ANALYSIS OF THE SMOOTH ADJUSTMENT HYPOTHESIS WITH RESPECT TO BRAZIL¹

Priscilla Belle Oliveira Pinto² Cláudio R. Foffano Vasconcelos³ Ricardo da Silva Freguglia⁴

The aim of this study is to analyze the smooth adjustment hypothesis (SAH) for the Brazilian economy during the period from 1997 to 2008. This study incorporates two approaches to the construction of the variable for labor market adjustment. Theoretically, we expect an inverse relationship between marginal intra-industry trade and the movement of workers in the industries that engage in foreign trade. Therefore, we use the econometric approach and dynamic panel data. The empirical results show that the hypothesis does not hold true for Brazil in the period examined.

Keywords: intra-industry trade; trade flows; displacement of workers; dynamic panel data.

COMÉRCIO INTRA-INDÚSTRIA: ANÁLISE DA HIPÓTESE DE AJUSTE SUAVIZADO PARA O BRASIL

O trabalho tem como objetivo analisar a hipótese de ajustamento suavizado (SAH) para a economia brasileira no período entre 1997 a 2008. Este estudo incorpora duas abordagens para a construção da variável de ajustamento do fator trabalho. Teoricamente, espera-se uma relação inversa entre o comércio intraindústria marginal e o movimento de trabalhadores nas indústrias engajadas no comércio externo. Para tanto, utilizou-se a abordagem econométrica de dado em painel dinâmico. Os resultados empíricos mostram que a hipótese de ajustamento suavizado não se confirma para a economia brasileira no período analisado.

Palavras-chave: comércio intraindústria; fluxo de comércio; deslocamento de trabalhadores; dados em painel dinâmico.

JEL: F12; F16; J62.

1 INTRODUCTION

The impact of greater trade openness or an adverse shock in an economy means that there is a reallocation of employment in industries in the short-term. Balassa (1966) was one of the pioneers to discuss this issue by showing that when there is greater trade openness, industries with intra-industrial characteristics have an advantage due to the lower cost of adjustment. Considering that intra-industry trade flow is characterized by the low cost of adjustment in the pre-existing productive structures (Brülhart and Elliott, 2002), the intra-industry trade expansion implies

^{1.} The authors are grateful for the financial support of the National Council for Scientific and Technological Development (CNPg).

^{3.} Professor at the PPGE/UFJF. E-mail: <claudio.foffano@ufjf.edu.br>.

^{4.} Professor at the PPGE/UFJF. E-mail: <ricardo.freguglia@ufjf.edu.br>.

a soft reallocation of resources and, therefore, low cost of transition. According to these authors, considering two small and open economies that suffer a demand shock due to a reduction in trade protection, their relative prices would be impacted and would act as a signal for resources to move from one sector to another. If this shock provokes an increase in imports in the competition of a certain industry, this will cause a reduction in the demand for factors of production of this sector. Thus, assuming work as the factor that most reacts in the short term, it will tend to feel the first effects of the cost adjustment pressure. The exact impact will depend on the structure of the labor market, but overall it will be a combination of a change in wages and employment. Since this observation, several authors such as Brülhart (1999, 2002), Brülhart and Thorpe (1999), Cabral and Silva (2006), Erlat and Erlat (2006), Faustino and Leitão (2009), and Thorpe and Leitão (2011), among others, have investigated the smooth adjustment hypothesis (SAH).

A question that seeks a better understanding of the hypothesis is the establishment of a measure for evaluating adjustment costs, which has led to disagreements about the implications of intra-industry trade (IIT). In particular, Cabral and Silva (2006) proposed a new variable of adjustment cost in order to ascertain a less biased analysis.

The analysis is guided by IIT, particularly in manufacturing industries, and the importance of this industry in the development of Brazil has been growing over time. In this context, considering that empirical evidence for the SAH with respect to the Brazilian economy was not found in the reviewed literature, the objective of this study is to assess the impact of trade on the intra-industry adjustment of workers in sectors from the manufacturing industry in Brazil from 1997 to 2008. Specifically, the purpose is also to discuss adjustment costs by focusing on a new variable to measure the effects of job change in the Brazilian industry, as proposed by Cabral and Silva (2006). To achieve the proposed objectives, we adopt a dynamic panel methodology to conduct a longitudinal analysis over a period of 12 years, following the movement of workers within the same sectors of the manufacturing industry throughout the analysis period.

This study is expected to expand the understanding of the subject in the context of the national and international literature *i*) by bridging the gap in the literature on the SAH in Brazil, which is a country with great potential for growth and ranks among the 20 largest exporters and global importers, and *ii*) by broadening its scope and comparing the two measures of the evaluation of adjustment cost by using the variable proposed by Cabral and Silva (2006) and Brülhart and Elliott (1998).

^{5.} The SAH assumes that the higher the proportion of IIT type, the shorter is the distance of the movement of workers and the lower is the cost of adjustment resulting from this stream of commerce.

^{6.} The period considers the maturing of the process of trade liberalization from the productivity growth in the beginning of the 1990s to until the beginning of the international financial crisis in 2008.

The remainder of the paper is organized as follows: in Section 2, the theoretical framework and a review of the empirical literature on the hypothesis-related adjustment are presented. Section 3 describes the methodology and the database used, while Section 4 presents the main results. Lastly, the concluding remarks are provided in Section 5.

2 THEORETICAL AND EMPIRICAL FRAMEWORK

In the model of monopolistic competition with production specialization in different varieties of products, each country will export different kinds of products to each other. Therefore, this type of flow of commerce trade is referred to as intraindustry (Dixit and Stiglitz, 1977; Krugman, 1981). Complete specialization and IIT do not occur in the Heckscher–Ohlin model with two sectors: Countries can produce the same sectors, but export or import occurs in each industry or sector and not in both sectors/industries.⁷

According to Brülhart (2002), the flow characteristic of IIT is the low cost of the adjustment of pre-existing productive structures. Hence, instead of having sectoral specialization across countries, as provided by the Heckscher–Ohlin model, national economies seem to preserve their broad industrial structures and predominantly specialize at an intra-sectoral level.

According to Brülhart and Elliott (2002), the expansion of IIT implies a smooth reallocation of resources, and thus low cost adjustment factors resulting in the movement of this stream of commerce. This proposition is known in the literature as the SAH. According to these authors, the understanding of the SAH can be obtained when it assumes a demand shock (e.g., a reduction of tariff barriers) in two small open economies. This shock can affect relative prices, acting as a flag to the movement of resources from one firm to another within the same industry. With reduced rates, there is an increase in imports by some firms and exports by others. The first factor that will experience the effects will work, because it is the factor that responds to changes in the short-term. The actual impact will depend on the structure of the existing market, but, in general, this will be a combination of the change in earnings and employment.

^{7.} Moreover, from the analysis of the Heckscher–Ohlin model of continuous goods, in the case of real numbers greater than factors, we have complete specialty products at different prices when the factors are not equal (Feenstra, 2004, chap. 3).

^{8.} Balassa (1966) studied the trade of the European Economic Community and realized that the adjustment was more understated where IIT prevailed, because the capital/labor ratio is likely to be more similar to inter-industry than intraindustrial structures.

^{9.} For this work, we consider the industry classification CNAE with a maximum of three digits; however, in the compatibility of databases, there exist some industries with four digits. These firms are part of an industry.

The central idea of the SAH is the assumption that the greater the proportion of new commerce, intra-industry type, the smaller is the relative cost of adjustment. Brülhart (2002) explained that adjustment costs arise from temporary inefficiencies when there are changes in demand or supply. More specifically, adjustment costs that are studied in the context of expanding trade are those for which the welfare loss is temporary unemployment due to labor market rigidities in the factor prices or costs incurred by job search relocation or the retraining of workers.

The analysis of the SAH lists three possible reasons for IIT implying a lower cost of adjustment: *i*) labor mobility between firms and occupations may be greater within industries than between industries; *ii*) relative wages are more flexible within industries than between industries; and *iii*) other factors of production can be more mobile within industries than between them.

With respect to the empirical verification of the SAH, Brülhart and Thorpe (1999) analyzed the expansion of trade in the high-growth period 1970–1994 in Malaysia and its implications in labor market adjustment. An econometric methodology was adopted to analyze panel data with fixed effects. The main conclusion was that the data for Malaysia did not support the notion that high marginal intra-industry trade (MIIT) is negatively related to job changes.

Brülhart (2000) investigated the dynamic aspects of the SAH associated with MIIT in Ireland during the period 1977–1991, aiming to clarify three important points on the subject. First, to review the importance of using MIIT rather than the traditional IIT index to analyze the adjustments to conclude that the signs of the MIIT estimates agree with those expected in the literature and are significant, and that the IIT index is not statistically significant. Second, to discuss the appropriate choice of the time horizon to calculate MIIT both conceptually and on the basis of the empirical results, reaching the conclusion that the data are best with the least amount of time – year after year intervals. Third, to investigate the relative time of trade and changes in the labor market, assuming different lag structures, with the results suggesting that the labor market effects follow changes in trade structure with a one-year lag.

Greenaway, Haynes, and Milner (2002) studied the relationship between the flow of foreign trade and the labor market of the United Kingdom between 1982 and 1998 and found the occurrence of relative changes when considering employment at the industry, firm, occupation, and regional level. They then related this incidence of job adjustment "within industries/sectors" on any of these dimensions to the characteristics of the labor market indicators and trade exhibition. To do this, they used a panel approach with a fixed effects model. The main findings of their work are that the overall result is not consistent with the notion that there is a systematic relationship between the type of trade expansion

(inter-or intra-industry) and type of adjustment (adjustment between or within industries). Further, there is less adjustment in the labor market associated with intra-industry than the associated inter-industry.

Brülhart and Elliott (2002) studied the link between trade and labor market changes in the manufacturing industry in the United Kingdom during the period 1979-1991. For this purpose, they used measures to evaluate the variability of unemployment at the industry level and variability of wages, relating these measures to IIT. Their evidence supported the SAH.

Brülhart, Murphy, and Strobl (2004) investigated the SAH with respect to Ireland and used the turnover of workers in IIT as the dependent variable. Their study covered the period from 1979 to 1990 with 64 Irish manufacturing industries. To estimate the model, they used panel data with fixed effects and GMM to lag behind the dependent variable and include in the model. The results showed a positive and statistically significant MIIT with the dependent variable, giving empirical support to the SAH.

Brülhart, Elliott, and Lindley (2006) tested whether the expanding intraindustry in the United Kingdom was related to the displacement of intra-sectoral workers during the period from 1986 to 2000. They used both IIT and MIIT as indexes, and in both cases, the sign was found to be negative, confirming the SAH.

Erlat and Erlat (2006) investigated the SAH for manufacturing industries in Turkey during the years 1974-1999. The study periods were divided into 1974-1975 and 1998-1999. For the adjustment of the cost proxy, the same variable proposed by Brülhart and Thorpe (2000) was used; however, the study result contradicted the SAH, obtaining a positive MIIT with a statistically significant cost adjustment. To perform the estimation methodology, the authors used panel data with fixed effects.

Cabral and Silva (2006), while looking for evidence of the SAH for manufacturing industries in Portugal, found that various studies of the subject have been inconclusive due to the improper use of a measure of cost adjustment. Thus, the authors proposed an alternative measurement of the variable cost of employment adjustment. The motion variable is most appropriate for capturing the possibility that adjustment costs are lower with IIT. According to the authors, the expansion of trade may entail a variation not only in total demand for labor, but also in the composition of labor demand in each industry. Thus, the new variable considers the proposed relocation on the job due to the change in the composition of the workforce in each industry and changes in the level of total employment in each industry, eliminating the labor movement that is not associated with the changes induced by trade.

Faustino and Leitao (2009) tested the SAH for manufacturing industries in Portugal from 1996 to 2003. They employed the MIIT index used by Brülhart (1994) as the dependent variable and the change in the total employment of manufacturing industries and developed a dynamic panel model to perform the data analysis. The results obtained in some analyses confirmed the SAH and not others, considering the contemporaneous effects, and the findings did not support the SAH. However, when considering one- and two-year lag effects, the results were different and sensitive to the size of the gap. The comparison of the results of empirical studies by these authors suggested that the validity of the SAH depends on the choice of variable cost index adjustment work, the structure of the time lag, and the set of control variables.

Rasekhi and Ghaderi (2012) investigated the SAH by using a panel for manufacturing industries in Iran for the period 2001-2006. They made the distinction between total vertical and horizontal IIT and used the variable proposed by Cabral and Silva (2006) as a proxy for cost adjustment. Their results did not show statistical significance for the country, as confirmed by the SAH study; however, when there was differentiation between vertical and horizontal IIT, the hypothesis was confirmed.

Through this literature review, it was observed that in developed economies such as Ireland, Portugal, and United Kingdom, the SAH is verified, while it is not supported in developing economies such as Iran, Malaysia, and Turkey. One way to understand this likely dichotomy is that developing economies benefit from trade through a greater structural change of exports and that advanced economies may be closer to their technological frontier. The analysis of another developing economy (i.e., Brazil) strengthens the motivation to broaden the basis of studies contributing to an increased empirical understanding of the topic.

3 METHODOLOGY

3.1 Empirical model

Initially, to test the SAH due to the flow of trade to the Brazilian economy, we analyze the following relationship:

$$ET_{j,t} = f(MIIT_{j,t}, GA_{j,t}, GA_MIIT_{j,t}, AD_{j,t}, S_{j,t})$$

$$(1)$$

where the subscript j represents the industry, t denotes time analysis, ET_j is the total effect of the variation of workers, MIIT_{jt} is MIIT , GA_{jt} is the degree of the exposure of the economy, $\mathrm{GA_MIIT}_{jt}$ is the interaction between degree of exposure and MIIT , AD_{jt} is the variation in apparent demand, and S_{jt} is the share of skilled workers relative to the total workforce. As a proxy qualification, this work uses the schooling of the worker with a college degree.

Cabral and Silva (2006) proposed measuring the magnitude of the cost adjustment, which is the total effect (ET_i) of the reallocation of workers. That is,

$$ET_{jt} = DE_{jt} + CE_{jt}$$
(2)

where DE_j denotes the effect size and represents net workers in the industry and CE_j indicates the effect of composition, that is, the net of workers in the industry as an absolute value. Thus, in Eq. (3) the cost adjustment will be equal to the sum of the net changes in workers with qualifications (in absolute value) weighted by the average total number of workers employed in the industry in the period:

$$ET_{jt} = \frac{\left|L_{j}^{1} - L_{j}^{0}\right|}{\left(L_{j}^{0} + L_{j}^{1}\right) \times 0.5} + \frac{\left(\sum_{k} \left|L_{jk}^{1} - L_{jk}^{0}\right|\right) - \left|L_{j}^{1} - L_{j}^{0}\right|}{\left(L_{j}^{0} + L_{j}^{1}\right) \times 0.5} = \frac{\sum_{k} \left|L_{jk}^{1} - L_{jk}^{0}\right|}{\left(L_{j}^{0} + L_{j}^{1}\right) \times 0.5}$$
(3)

where L corresponds to workers and the subscripts j and k represent the industry workers (j = 1,...,81) and qualifications (k = 1 full graduation, incomplete graduate, and high school), respectively. The 1 and 0 superscripts correspond to year 1 and year 0, respectively. L^1_{jk} and L^0_{jk} are the number of skilled workers (k) of industry j who were present in the initial and final years, respectively. This variable will be zero if there is no change in skilled workers from one year to another.

Regarding the independent variables of the specified empirical model, Eq. (1), must comprise two elements: a variable representing the MIIT index and a set of variables to control for other exogenous influences on $\mathrm{ET_{jt}}$. According to Brülhart (1994; 2000) MIIT is more appropriate for the analysis of the SAH than the traditional static intra-industry trade index IIT developed by Grubel and Lloyd (1975). Thus, considering the literature on the subject in this work, we used only MIIT.

The construction of this variable is based on that described by Brülhart (1994; 2000). That is,

$$MIIT_{j} = A_{j} = \sum_{i=1}^{j} (w_{i} \times A_{i})$$

$$\tag{4}$$

and

$$w_{i} = \frac{\left|\Delta_{I}X_{t}\right|_{i} + \left|\Delta_{I}M_{t}\right|_{i}}{\sum_{i=1}^{j} \left(\left|\Delta_{I}X_{t}\right|_{i} + \left|\Delta_{I}M_{t}\right|_{i}\right)}; \ A_{i} = 1 - \frac{\left|\Delta_{I}X_{ti} + \Delta_{I}M_{ti}\right|}{\left|\Delta_{I}X_{t}\right|_{i} + \left|\Delta_{I}M_{t}\right|_{i}}$$
(5)

where Δ denotes changes in constant prices, t indicates the base year, I indicates the length of the time interval between the base and end years; i denotes the level of industry disaggregation, ¹⁰ and X and M represent exports and imports, respectively.

MIIT has the same interpretation as IIT, that is, when the MIIT index is zero, marginal trade is exclusively inter-industry, and when the index is one, it is entirely intra-industry. According to Brülhart (2000), one methodological issue is the choice of the time interval for the calculation of MIIT measures. The author stated that "there is no systematic relationship between the A index calculated over a certain time interval and A indices calculated over constituent subintervals" (Brülhart, 2000, p. 243). A way to mitigate this issue is to work with the smallest possible interval. Therefore, we calculate Eq. (5) over one-year intervals.

According to theory, it is expected that its sign is negative, that is, the higher the proportion of MIIT, the lower is the cost of labor market adjustment. To measure this, the type of trade is assumed to be an intra-industry commerce variable, which is marginally more recommended for this type of study, as proposed by Brülhart (1994).

For the set of variables to control for other exogenous influences on ET_{jb} the following is a description of these variables as well as their expected signs.¹¹

This variable, suggested by Brülhart and Elliot (2002), represents the degree of the exposure of the economy and measures the intensity of total trade (exports plus imports) to GDP. That is,

$$GA_{j} = \frac{\left[\left(X_{j} + M_{j} \right)^{1} + \left(X_{j} + M_{j} \right)^{0} \right]}{ITV_{j}^{1} + ITV_{j}^{0}}$$
(6)

where ITV denotes the industrial transformation value, which is a proxy for the production of sector j.

Moreover, the strategy of using this variable allows inferences about other additional assumptions such as higher intensity in trade openness, which affects the cost adjustment in the same way as that presupposed by the SAH (Cabral and Silva, 2006). For this variable, we expect a positive sign, because a higher exposure of economic openness means more competition, which will lead to an increase

^{10.} The industry i was disaggregated in the four-digit level – or denomination "class" – of the CNAE 1.0.

^{11.} No theoretical model is defined in the literature as the best specification. Brülhart and Thorpe (2000) worked with variables in log; however, in the present study, we adopted the variables in levels because this way fit the data better. This sensibility analysis was not included in the text to save space.

in the pressure adjustment of industries (Brülhart, 2000). The studies by Cabral and Silva (2006) and Thorpe and Leitão (2011) showed a positive sign for this variable, corroborating the theory.

This variable was suggested by Brülhart and Elliot (2002), and it represents the interaction of MIIT with trade exposure. The variable allows the analysis of other additional hypotheses, because a major change in the structure of the stream of commerce exchange affects the cost adjustment in the same manner as that presupposed by the SAH (Cabral and Silva, 2006). This variable is expected to have a negative value, because trade expansion can strengthen MIIT and hence result in a lower cost adjustment.

3.1.3 AD_{it}

For this variable, which represents apparent industry demand, a negative relationship with the domestic expansion of the movement of workers is expected (Brülhart, Elliot and Lindley, 2006; Cabral and Silva, 2006). Eq. (7) shows the measurement of the variation of AD:

$$\Delta AD_{j} = \frac{\left[\left(ITV - X_{j} + M_{j} \right)^{1} - \left(ITV - X_{j} + M_{j} \right)^{0} \right]}{\left[\left(ITV - X_{j} + M_{j} \right)^{0} + \left(ITV - X_{j} + M_{j} \right)^{1} \right] \times 0.5}$$
(7)

$$3.1.4 S_{jt}$$

This variable represents the share of unskilled workers in the total workforce. That is,

$$S_{j} = \frac{H_{j}}{L_{j}} \tag{8}$$

where *H* and *L* denote the unskilled and total workforce, respectively.

Cabral and Silva (2006) and Brülhart *et al.* (2004) used this variable as a reverse proxy for technological intensity; thus, we expect a positive relationship with labor movement.

3.2 Econometric model

The empirical strategy used was the panel data method often employed in this type of study. This methodology also allows us to analyze manufacturing over time and provides an increase in estimation accuracy because it gives a greater amount of information over time (Cameron and Trivedi, 2005).

This method also allows us to obtain estimators of the variables *MIIT*, *GA*, *AD*, *S*, and *MIIT*_*GA*, which are consistent even in the presence of unobserved variables that are related to the characteristics of the manufacturing sectors. The no-observed-effect model can be described as shown in Eq. (9):

$$ET_{it} = \beta_0 + \beta_1 MIIT_{it} + \beta_2 GA_{it} + \beta_3 GA MIIT_{it} + \beta_4 AD_{it} + \beta_5 S_{it} + c_i + u_{it}$$
(9)

where ET_{jt} represents the dependent variable in sector j during year t; $MIIT_{jt}$, GA_{jt} , $GA_{m}IIT_{jt}$, AD_{jt} , and S_{jt} are the observed explanatory variables that do not change over time, but assume heterogeneous characteristics across industries; c_j is the observed explanatory variable that does not change over time, but assumes heterogeneous characteristics between industries (it is observed that the term is not related to the characteristics of the manufacturing sectors that do not change over time (also known as unobserved heterogeneity)); and u_j corresponds to the idiosyncratic error term.

The econometric strategy for this study consists of an analysis with a static model (which does not allow us to consider the time lag of the dependent variable) and panel data with fixed and dynamic effects compared with a dynamic panel model. This is because the analysis of economic behavior often involves the specification of dynamic econometric models, that is, models with lagged dependent variables. Such time series in which the model is dynamic and estimation methods are based on the ordinary least squares (OLS) method using lags rarely allow estimators to maintain good properties (Harris, Matyas and Sevestre, 2008). Therefore, the dynamic model seems to have an advantage over the other models discussed earlier, and it allows us to analyze the total displacement of workers in an earlier period, affecting the displacement current in the context of the analysis of the SAH.

The dynamic panel used for the analysis is shown in Eq. (10):

$$ET_{jt} = \delta ET_{j,t-1} + \beta_0 + \beta_1 MIIT_{jt} + \beta_2 GA_{jt} + \beta_3 GA_MIIT_{jt} + \beta_4 AD_{jt} + \beta_6 S_{jt} + c_j + u_{jt}$$
(10)

where $j = \{1, ..., 81\}$ and $t = \{1,...,6\}$. The construction $ET_{j, t-1}$ is correlated with the no-observed effect of each industry (c_i) .

Regardless of the pattern of c_j (being a fixed or a random effect), the OLS estimation of δ and β is inconsistent because the regressor $ET_{j,t-1}$ is correlated with c_j and thus with the composition of the error term $(c_j + u_{jt})$ (Cameron and Trivedi, 2005).

The term c_j could be a random effect, as stated earlier, but the choice of the fixed effects model is generally more appropriate. Judson and Owen (1996) proposed two reasons for this choice: (1) the effects of individual industries represent omitted variables; as a result, specific features of industries may be correlated with other

independent variables; and (2) a macro panel is likely to contain more industries of interest, that is, if all the industries of interest in this study were present, no a random sample of the universe of manufacturing industries could be made.

Thus, the first-difference model (10) can be represented by Eq. (11):

$$ET_{j,t} - ET_{j,t-1} = \delta \left(ET_{j,t-1} - ET_{j,t-2} \right) + \left(Z_{j,t} - Z_{j,t-1} \right) \beta + \left(u_{j,t} - u_{j,t-1} \right)$$
(11)

For simplification, Z denotes the regressors $MIIT_{jt}$, GA_{jt} , $GA_{MIIT_{jt}}$, AD_{jt} , and S_{jt} , and t = 1,...,6. This transformation removes the unobserved effect, keeping only the correlation effect of the lagged variable with the error term in an earlier period. However, without the no-observed effect, the lagged variable becomes a strong instrumental variable estimator (Baum, 2006).

To circumvent the problem of bias in the estimations for OLS (i.e., inconsistent and biased), fixed effects (which is biased), and random effects (also biased), it is necessary to use instrumental variables and instruments for equations in the first-difference predetermined lagged explanatory variables, as proposed by Arellano and Bond (1991).

However, Arellano and Bover (1995), similar to Blundell and Bond (1998), showed that models estimated in first differences by using instruments in levels may be weak instruments. To reduce potential bias generated by the difference estimator GMM, Arellano and Bover (1995) and Blundell and Bond (1998) suggested an estimator that is more sophisticated than that proposed by Arellano and Bond (1991), namely the system GMM. They showed that there is bias in finite samples when using the first-difference estimator and that the system GMM presents stronger results.

The proposed changes include lag level as well as lag difference. The estimation is carried out by combining system-level regressions and instruments are the explanatory variable lags of endogenous differences, that is, $(y_{j,t-1}-y_{j,t-2})$ is the instrument with the regression level.

The second step in the GMM estimation of the standard errors of the difference GMM has shown bias. Hence, to fix this problem, a correction to the Windmeijer finite sample and the standard errors was made.

Importantly, Roodman (2009a) noted that there are two important issues in the estimation. The first is with respect to using time dummies to prevent a recurrence of more correlation across individuals – contemporary effects. The second is that the test depends on the assumption that the sample size (N) is large. The meaning of "large" is not very accurate, but estimates with a panel of N = 20 is a concern.

3.3 Database

The database contains information on 81 Brazilian manufacturing industries. ¹² To calculate the number of workers in the sectors used, the entire population of the Annual Report of Social Information ¹³ (RAIS) for industries, classified as manufacturing according to the criteria of the National Classification of Economic Activities ¹⁴ (CNAE) 1.0 and 2.0, totaling 918,310 workers a year, was considered and the data were balanced so that all workers could be observed at any time.

Exports and imports were obtained from the Aliceweb system of the Ministry of Development, Industry, and Foreign Trade (MDIC, 2011). The analysis was restricted to the Brazilian manufacturing industry and the period from 1997 to 2008 due to data availability in RAISMIGRA after trade liberalization. Data for production (industrial transformation value) were obtained from the website of the Brazilian Institute of Geography and Statistics (IBGE, 2011) from the Annual Survey of Industries.

To estimate the econometric model, first it was necessary to ensure that the Aliceweb CNAE 1.0 and 2.0 databases for the period 1997–2008 were compatible. Six two-year periods were generated (1997-1998, 1999-2000, 2001-2002, 2003-2004, 2005-2006, and 2007-2008). To make the Aliceweb system, CNAE 1.0 and CNAE 2.0 were made compatible, and the changes were made to the CNAE 1.0 as the base. The IBGE provided a table of correspondence for CNAE 1.0 with CNAE 2.0. The MDIC had a translator to transform the Mercosur Common Nomenclature for CNAE 1.0.

To build the database, first we withdrew the Aliceweb system's values of exports and imports from Brazil to the rest of the world, and then used the MDIC (2011) translator to select the products that were considered to be manufacturing industry, according to the criteria of CNAE 1.0.

After the separation of the products of the Harmonized System that corresponded to the manufacturing sector, it was found that some products were exported but not imported and vice versa. Thus, we proceeded to ensure the compatibility of imports and exports in order not to lose information from any database.

With the data organized, exports and imports departed from the compatibility of the Harmonized System with CNAE 1.0, with reference to the translator of the MDIC. Compatibility was made with 3903 Brazilian products in international

^{12.} According to CNAE (2003), the manufactoring industry corresponds to the activities that involve the significant mechanical, physical or chemical transformations of materials, substances or components in order to obtain new products.

^{13.} The base RAISMIGRA includes the Ministry of Labor and Employment.

^{14.} The CNAE was structured with reference to the International Standard Industrial Classification of All Economic Activities of the United Nations.

trade. The four-digit CNAE 1.0 had 245 sectors for manufacturing. Thus, the products of the Harmonized System were aggregated within these 245 sectors. For 2007 and 2008, it was necessary to match CNAE 1.0 with CNAE 2.0, and to this end, the board's conversion MTE (2011) was used.

The final compatibility was made from the four digits of CNAE 1.0 to reach 81 three-digit industries.¹⁵ Exports and imports were used in current values, and therefore, to make the conversion from VTI, the average exchange rate for each year of sales, obtained from Ipeadata (Ipea, 2011), was used.

4 RESULTS

4.1 Characterization of Brazilian trade flows

To start the trade analysis, we used the Grubel-Lloyd index (1975). This index allows us to calculate IIT when considering the trade flow for a specific time period. Figure 1 presents an overview of the flow behavior of commercial industries. Throughout the series, the percentage of industries that showed higher IIT than inter-industry trade occurred only in 2006. Therefore, there is a predominance of inter-industry trade flows for the 81 industries analyzed in the Brazilian economy for the period examined.

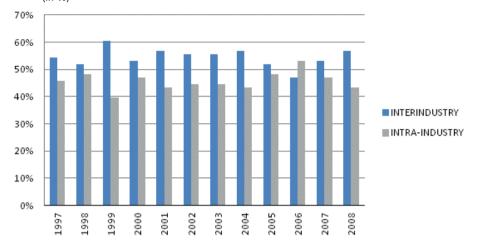
As the static index does not explain the dynamics of the periods and as the focus of this study was to analyze the dynamics of employment in different industries, a descriptive analysis of MIIT was necessary. The same static index showing predominant inter-industry behavior, in general, did not show large differences in the amount of industries whose flows were characterized by IIT and in those that had a characterization of inter-industry. Unlike the static index, the index showed that trade industries in MIIT, compared with the dynamism of time, are much greater than those in inter-industry trade. Figure 2 clearly shows that Brazil is a country with an inter-industry characteristic.

Importantly, from figures 1 and 2, the general behavior of industries can be observed, and even those exhibiting intra-industry-type behavior in a given period may not present the same characteristics in the following period. Fifty-four industries, or 66.67% of the total, had, at some point, intra-industrial characteristics, but none presented continuous behavior in the period 1997-2008.

^{15.} The tables of compatibility and list of 81 sectors are available from the authors.

^{16.} The behavior change of the Grubel-Lloyd index in 2006 is due to the excellent performance of exports of manufactured and semi-manufactured products in that year. Consequently, given the volume of imports of manufactured and semi-manufactured products, the growth of exports of this type of products increases the intra-industry trade.

FIGURE 1
Representation of the types of trade in the Brazilian manufacturing industry with the rest of the world – Grubel-Lloyd Index (1997-2008) (In %)

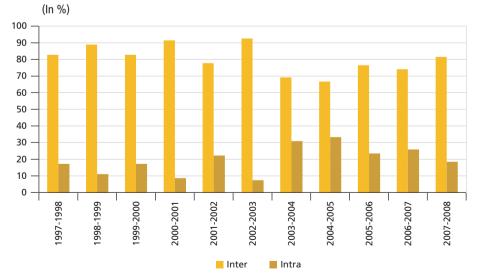


Source: MDIC/Aliceweb - raw data.

Authors' elaboration.

Publisher's note: Figure displayed in low resolution and whose layout and texts could not be formatted and proofread due to the technical characteristics of the original files.

FIGURE 2 Amount of manufacturing industries that were classified as inter- or intra-industry by MIIT



Source: MDIC/Aliceweb – raw data. Authors' elaboration.

4.2 Results of the economics models

The test of the SAH was carried out by using the dynamic model with the estimators proposed by Arellano and Bover (1995) and Blundell and Bond (1998). Table 1 shows a descriptive analysis of the correlation of the variables. Although no inference can be made based on it, one can get a hint of how the variables tend to behave. The Spearman coefficient¹⁷ was chosen because it considers the order of the data and not the intrinsic value, and this is an advantage when there are outliers among the variables (if openness denotes discrepant values of the other variables). The variable MIIT showed a statistically significant correlation with the variable ET; this variation was positively correlated and it did not demonstrate convergence for the SAH. An important issue to note is that the positive magnitude is very low. The correlation between the variables openness (GA) and skilled labor (S) indicates that they are positively correlated with the variable ET. This finding indicates that initially the variables are as expected in the literature review.

TABLE 1 Spearman's correlation

	ET	ETB	ET _{t-1}	ETB _{t-1}	MIIT	GA	AD	S	MIIT_GA
ET	1								
ETB	0.3643*	1							
ET_{t-1}	0.4415*	0.2090*	1						
ETB_{t-1}	0.2722*	0.1605*	0.3949*	1					
MIIT	0.0751*	-0.0513	-0.0145	-0.1076*	1				
GA	0.1967*	0.1212*	0.2095*	0.1664*	0.0175	1			
AD	0.1086*	-0.0101	0.027	-0.007	0.0227	0.0699	1		
S	0.3342*	0.0044	0.1171*	0.0714	0.1086*	0.0147	0.0348	1	
MIIT_GA	0.1616*	0.0252	0.1074*	-0.0127	0.8015*	0.5067*	0.0641	0.0738	1

Authors' elaboration.

Obs.: 1. ETB is the effect of the variation of works measured as in Brülhart (2000).

Table 2 shows the behavior of the variables. On average, the variable that represents the total effect of the displacement of skilled manufacturing workers varies by 5.2%. MIIT and its interaction with degree of openness have a greater variation static compared with the variable ET. The variable with the highest average of variation is degree of openness, explaining the greater trade openness that occurred over the study period.

^{2. *} represents 10% significance.

^{17.} Brülhart (2002) and Greenaway, Haynes and Milner (2002) also used such a correlation.

Variable	Obs.	Average (%)	Standard deviation	Min	Max
ET	486	5.2	0.119	0.00012	1.346
MIIT	486	26.9	0.267	0.00000	0.992
GA	486	86.4	0.835	0.00005	8.955
GA_MIIT	486	23.3	0.397	0.00000	3.566
DA	486	0.20	2.258	-25.227	23.703
S	486	2.10	0.090	-0.418	1.834

TABLE 2
Characteristics of the variables

Authors' elaboration.

Table 3 presents the main estimated results. It shows all the main steps to reach the result shown in Eqs. (9), (10), and (11). First, we estimated the pooled OLS model with the same variables proposed by Cabral and Silva (2006), and the difference in the model specification of this work occurred at the level of the variables. As the literature does not have a standard analysis method, we chose to work with all level variables.

Subsequently, the test was reset to verify that the linear regression model was well specified. The flaw in the specification may be due to omitted variables, incorrect functional form, the problem of simultaneous equations, and heteroskedasticity. The results in table 3 show that one should reject the null hypothesis that the template has no fault in the specification, giving an indication that there is a need to improve the model specification.

Subsequently, two tests were carried out: the Breusch-Pagan test indicated that the best econometric model would be the unobserved effects and the Hausman test indicated that we cannot reject the null hypothesis, suggesting using the random effects model (table 3).

It is noteworthy that in the fixed effects model, as in the random effects model, the hypothesis that MIIT implies a movement of workers was not statistically significant. It was found that there was heteroskedasticity in the fixed effects model for such controls, turning the robust model.

Considering that the model needs further information due to the RESET test, as explained earlier, the test was carried out again for a pooled OLS model with the inclusion of the lagged dependent variable. The results did not allow us to reject the null hypothesis that the model has no omitted variable at the 5% significance level, indicating that the phase variable movement of industrial workers in an earlier period can affect the model.¹⁸

^{18.} Table 3 reports the main results of the estimation and the variations of the estimates, including the one that includes the lag. These results are available from the authors.

Another fact that indicates that the model should have the lag as an explanatory variable on the dependent variable is the main objective of the work, which is a dynamic issue by nature. Arellano and Bond (1991), when deriving employment equations estimated by using unbalanced data for a panel of 140 UK companies for the period 1974–1984, used two lags to explain employment in companies in each year of the analysis. In the present study, for the calculation of the dependent variable (which represents variation in employment), we chose to use only one lag.

The dynamic model also did not confirm the SAH (table 3). However, before we explore these results, a brief explanation of the statistics is necessary. The Arellano-Bond test is appropriate to verify the existence of serial autocorrelation (table 3 – Arellano-Bond AR), and this is especially important when lags are used as instruments. This test rejects the null hypothesis that there is no first-order autocorrelation; however, there is no evidence of second-order autocorrelation, and hence, it is not possible to reject the null hypothesis.

The analysis was carried out by testing various instruments. The Hansen test and Sargan test, which examine the validity of the joint GMM instruments, did not reject the hypothesis that the instruments are valid and they had high p-values. Roodman (2009b) stated that there is a debate among researchers regarding the notion that a high p-value validates the test results of Hansen. As the "difference in Hansen" test, which analyzes subsets of instruments, also does not reject the hypothesis of exogeneity instruments, one can conclude that the estimation is efficient.¹⁹ It is known that the reliability of the tests cited earlier may be affected if the number of instruments is greater than that of the units of cross-sections. However, the number of instruments in this study is 16 and the number of cross-sections is 81.

The main model was the dynamic model proposed by Arellano and Bover and Blundell and Bond, with the two-stage estimator, correcting the variance and covariance matrix to treat heteroskedasticity and presenting estimates of the standard errors corrected for finite samples (Arellano–Bover). Table 5 shows that MIIT is not statistically significant for explaining the variation in intra-industry workers, while the other estimators (fixed effects, random effects, and OLS) were used to verify the robustness of the result, which was maintained. This finding shows that the SAH is not valid for Brazil. The result finds support in the empirical literature, because Brazil has little intra-industrial characterization. The result of the study by Brülhart and Thorpe (2000) for Malaysia, a country that has little intra-industrial characterization, to determine the validity of the SAH was also not statistically significant.

^{19.} These results are available from the authors.

Degree of openness had a positive coefficient and was statistically significant for the dynamic model. It was noted that there was a difference in the coefficient found for the Arellano-Bover (1995) and Blundell-Bond (1998) estimators compared with the other estimators (fixed effects, random effects).

The variable that represents apparent demand (AD) and the participation of skilled workers in relation to the total workforce of the manufacturing industry (S) was not statistically significant, similar to the results of the fixed effects and random effects models. The interaction coefficient (MIIT_GA) was not statistically significant in any model (table 3). Similar to the variable MIIT at low characterization, the Brazilian intra-industry may have affected the statistics of this index.

TABLE 3
Results of OLS, FE, AE, and dynamic model

	OL C	FF (DOD)	4.E. (DOD)	Dynamic
Variables/dependent variable	OLS	FE (ROB)	AE (ROB)	A-Bover
	ET	ET	ET	ET
Instruments (endogenous/exogenous)				ET _{t-1} /year
ET _{t-1}	-	-	-	0.445*
				(0.239)
MIIT	-0.00721	0.00995	-0.00145	-0.702
	(0.0274)	(0.0225)	(0.0302)	(0.452)
GA	0.0425***	0.0291	0.0387	0.035*
	(0.0085)	(0.0313)	(0.0281)	(0.020)
AD	0.00432*	0.000511	0.00257	-0.062
	(0.00234)	(0.00288)	(0.00275)	(0.047)
S	-0.0255	-0.0225	-0.0239	-0.368
	(0.0588)	(0.0777)	(0.0719)	(2.371)
MIIT_GA	-0.0358	-0.025	-0.0293	0.133
	(0.022)	(0.0308)	(0.036)	(0.179)
Constant	0.00995	0.0126	0.00969	0.138
	(0.0151)	(0.0259)	(0.0216)	(0.101)
Observations	486	486	486	405
R-squared	0.1	0.056		
Sector No.	-	81	81	81
Test Reset (Prob>F)	0.0000	-	-	-
Breusch–Pagan (Prob>Chibar2)	-	-	0.0000	-

(Continues)

(Continued)

	OLS	EE (DOD)	ΛΓ /DOD\	Dynamic
Variables/dependent variable	ULS	FE (ROB)	AE (ROB) -	A-Bover
	ET	ET	ET	ET
Hausman (Prob>chi2)	-	0.078	-	-
Test Heteroskedasticity (Prob>Chi2)	-	0.0000	-	-
Arellano–Bond AR(1) : z	-	-	-	-1.71*
Arellano–Bond AR(2) : z	-	-	-	-0.05
Hansen Test	-	-	-	2.97 0.685 ⁺
Sargan Test	-	-	-	3.10 0.705 ⁺
No. of instruments	-	-	-	16
Hansen Test – groups exclusion	-	-	-	0.13
Difference ($h_0 = exogenous$)	-	-	-	2.97
Hansen Test – groups exclusion	-	-	-	2.43
Difference (h ₀ = exogenous)	-	-	-	0.67

Authors' elaboration.

Obs.: 1. *, **, and *** represent 10%, 5%, and 1% significance, respectively.

- 2. ET: dependent variable proposed by Cabral and Silva (2006).
- 3. Values in parentheses correspond to the standard error corrected for each variable.
- 4. In all models, year dummies were estimated.
- 5. The + sign indicates the p-value of the test.

To check the robustness of the sensitivity analysis, analyses were performed by removing the control variables, and the results are shown in table 4. The results of the sensitivity analysis for the fixed effects and dynamic models suggested that there is robustness in the results found in column 4, dynamic model, of table 4. The result of the random effects model was significant, and this finding is in agreement with that presented in the literature. This result is important because it does not contradict the hypothesis, although this model cannot be accepted for the explanation presented this study, as mentioned earlier when using the Hausman test.

In addition to removing the variables, other sensitivity analyses were also performed. Roodman (2009b) suggested the withdrawal of the instruments to verify that the analysis would be sensitive, because a large number of instruments can make the analysis invalid; thus, the "gold standard" dynamic model would be fewer instruments than cross-section units (sector number). This study used fewer instruments (16) than the sector number (81), as can be seen in table 4. Thus, instead of removing the instruments, lags of all the explanatory variables were added as instruments to calculate the sensitivity of the dynamic model, generating 86 instruments. The result that there is no significant relationship between labor

movement and IIT remained marginal, but in the model with instruments greater than the number of sectors, degree of openness (GA) was not significant, which may have occurred as a result of excessive instruments.

To check for changes in the results compared with the variable proposed by Cabral and Silva (2006), the SAH was also tested with a variable that represents the gross variation of workers in the manufacturing industry. This variable was used by Brülhart and Elliott (1998) and it represents a part of the composition of the index proposed by Cabral and Silva (2006), the "size effect." It can be seen from table 4 that the coefficient of MIIT was not statistically significant, reinforcing, once again, the idea that there is evidence for the hypothesis that Brazil's adjustment costs have softened. In particular, according Cabral and Silva (2006), because of the "global" variable, ETB would be biased toward accepting the SAH.

Degree of openness, GA, was not statistically significant for the variable "gross"²⁰ (column 2 of table 5). This finding can be understood by the fact that trade liberalization affects more skilled workers, implying a statistically non-significant variable that considers total workers. The variables S, AD, and MIIT_GA were not significant, corroborating the estimate made earlier with the variable ET (table 5).

It can be recalled that the results to verify the model specification were good, as shown in table 5. The test to check for the presence of autocorrelation did not reject the null hypothesis that there is no first-order serial autocorrelation in residual variation and did not reject the hypothesis for second-order serial autocorrelation, implying the validity of the instruments.

^{20.} The variable "gross" referred to in this work is the variable that does not compute the level of skill/occupation.

 $\ensuremath{\mathsf{TABLE}}\ 4$ Sensitivity analysis of the model estimated by using OLS, EF, EA, and dynamic model

		ä			V				imcani	Johnsonic model (A Douge)
Variables		ב			Ä				Dynaiii	IC IIIOdel (A-Dovel)
variables	ET	ET	ET	ET	ET	ET	ET	ET	ET	ET
Instruments (endogenous/ exogenous)		ı	ı	ı	1	,	ET _{t-1} /ano	ET _{t-1} /ano	ET _{t-1} /ano	$ET_{e1},MIIT_{e1},S_{e1},AD_{e1},GA_{e1},MIIT_GA_{e1}$
ET _{t-1}							0.445*	0.120	0.108	-0.030
MIIT	0.010	-0.016	-0.003	-0.001	-0.0267*	-0.044**	-0.702	-0.315	-0.05	-0.123
GA	0.029	ı	ı	0.038	1	1	0.035*	1	1	0.023
AD	0.001			0.002			-0.062			0.010*
S	-0.023	,		-0.023		1	-0.368		1	0.901*
MIIT_GA	-0.025	ı	-0.015	-0.029	ı	0.019***	0.133		-0.09	0.01
Constant	0.013	0.034***	0.034***	0.009	0.037***	0.037***	0.138	0.114	0.072	0.053
No. of Sectors	81	81	81	81	81	81	81	81	81	81
Arellano—Bond AR(1): z				1			-1.71*	-2.33**	-2.13**	-2.75***
Arellano—Bond AR(2): z		ı	ı	1	ı	1	-0.05	-0.7	-0.29	-0.090
Hansen Test/(p-value)					1		2.97/	9.99/ 0.351	4.480	0.496
Sargan Test							3.1	20.39**	22.58**	0.394
Number of instruments				1			16	16	16	98
Hansen test – groups exclusion		ı	ı	1	ı	1	0.13	2.98	1.64	53.78
Difference ($h_0 = exogenous$)							2.97	7.01	2.84	20.67
Hansen test – groups exclusion						,	2.43	7.53	4.45	70.87
Difference ($h_0 = exogenous$)		ı	ı	1	1		0.67	2.46	0.03	3.57
-										

Authors' elaboration. Obs.: 1. *, **, and *** indicate the values with 10%, 5%, and 1% significance, respectively.

TABLE 5
Results of estimating the dynamic model with Windjmeijer's correction

	mic models		
- Verichler	A-E	Bover	
Variables -	ET	ETB	
_	(1)	(2)	
Instruments (endogenous/exogenous)	ET _{t-1} /year	ET _{t-1} /year	
ET _{t-1}	0.445*	0.096	
	(0.239)	(0.078)	
MIIT	-0.702	-0.705	
	(0.452)	(0.786)	
GA	0.035*	-0.150	
	(0.020)	(0.188)	
AD	-0.062	-0.005	
	(0.047)	(0.006)	
S	-0.368	2.397	
	(2.371)	(2.139)	
MIIT_GA	0.133	0.61	
	(0.179)	(0.759)	
Constant	0.138	0.174	
	(0.101)	(0.170)	
Observation	405	405	
Sector No.	81	81	
Arellano–Bond AR(1)	-1.71*	-1.41	
Arellano–Bond AR(2)	-0.05	-0.750	
Hansen Test	2.97	2.840	
Sargan Test	3.1	2.490	
Number of instruments	16	16	
Hansen test – groups exclusion	0.13	0.60	
Difference ($h_0 = exogenous$)	2.97	2.24	
Hansen test - groups exclusion	2.43	2.61	
Difference ($h_0 = exogenous$)	0.67	0.23	

Authors' elaboration

Obs.: 1. *, **, and *** indicate the values with 10%, 5%, and 1% significance, respectively. 2. Values in parentheses correspond to the corrected standard error of each variable.

The Hansen test did not reject the null hypothesis of the joint validity of the GMM instruments. The second test (difference) to check if the instruments were exogenous also did not reject the null hypothesis, confirming the exogeneity of the instruments. Furthermore, there was no problem with the "golden rule" of the model, ensuring that the instrument number was lower than that of the sectors.

After the tests were performed, it was not possible to demonstrate the SAH with respect to Brazil. It should be noted that the models used to test the hypothesis were robust to purge the heterogeneity present in the model; that is, the results from these models were constrained to accept the hypothesis.

5 CONCLUSION

The aim of this paper was to analyze the effects of the expansion of MIIT on the variation in skilled labor in the Brazilian manufacturing industry in the period 1997-2008. The most important results showed that the hypothesis is not fit for Brazil in the examined period.

There were several specifications to assess the robustness of the model and, largely, the result of the specifications was not significant. Only three analyses, namely two OLS and one random effects results, confirmed the SAH, but these results may be biased. The important point to note is that in all tests, the result was significant and positive, which is contrary to the hypothesis, in this case.

The study found answers with the sign expected in the literature for the control variables, such as degree of openness and interaction between degree of openness and MIIT. As noted in the descriptive analysis of the Brazilian trade flow in various industries, greater openness largely implies inter-industry characteristics, which is consistent with the results found in the literature. Degree of openness was found to be associated with lower rates of MIIT, which suggested a positive association with the displacement of workers.

The results, in fact, were expected due to reduced Brazilian IIT (from the marginal perspective) and also considering that the analysis of trade flows was with the rest of the world, which makes IIT lose its characteristics. The perspectives of future studies can be developed from this initial study of the SAH with respect to Brazil. Analysis of economic partners with characteristics more similar to Brazil is important to state whether the SAH is confirmed for Brazil.

This study aimed to evaluate the variation in shift workers with MIIT, expanding the scope by incorporating two approaches to the construction of the dependent variable. It is argued that the cost ratio smoothed adjustment with the structure of trade flows does not support the hypothesis that an expansion of IIT implies a lower cost adjustment, suggesting that the hypothesis is not valid for Brazil in the period analyzed, confirming the work of Brülhart and Thorpe (2000).

REFERENCES

ARELLANO, M.; BOND, S. Some tests of specification for panel data: monte carlo evidence and an application to employment equations. **The Review of Economic Studies,** v. 58, n. 2, p. 277-297, 1991. Retrieved from: https://bit.ly/36Izjrm.

ARELLANO, M.; BOVER, O. Another look at the instrumental variable estimation of error-components models. **Journal of Econometrics**, v. 68, n. 1, p. 29-51,1991. Retrieved from: https://bit.ly/2NM8Hgi.

BALASSA, B. Tariff reductions and trade in manufacturers among the industrial countries. **The American Economic Review**, v. 56, n. 3, pp. 466-473, 1966.

BAUM, C. F. An introduction to modern econometrics using Stata. Stata Press Books, StataCorp LP, 2006.

BLUNDELL, R.; BOND, S. Initial conditions and moment restrictions in dynamic panel data models. **Journal of Econometrics**, v. 87, n. 1, p. 115-143, 1998. Retrieved from: https://bit.ly/2WYAnCP>.

BRÜLHART, M. Marginal intra-industry trade: measurement and relevance for the pattern of industrial adjustment. **Review of World Economics**, v. 130, n. 3, p. 600-613, 1994. Retrieved from: https://bit.ly/2WRnvOI.

BRÜLHART, M.; ELLIOTT, R. Adjustment to the European single market: inferences from intra-industry trade patterns. **Journal of Economic Studies**, v. 25, n. 3, pp. 225-247, 1998.

BRÜLHART, M.; ELLIOTT, R. Labour-market effects of intra-industry trade: evidence for the United Kingdom. **Review of World Economics**, v. 138, n. 2, p. 207-228, 2002. Retrieved from: https://bit.ly/36M2r0D>.

BRÜLHART, M. Marginal intra-industry trade and trade-induced adjustment: a survey. *In*: BRÜLHART, M.; HINE, R. C. (Eds.). **Intra-industry trade and adjustment**: the European experience. Basingstoke: Macmillan, 1999.

BRÜLHART, M. Dynamics of intraindustry trade and labor-market adjustment, **Review of International Economics**, v. 8, n. 3, p. 420-435, 2000. Retrieved from: https://bit.ly/2PTkJHr.

BRÜLHART, M. Marginal intra-industry trade: towards a measure of non-disruptive trade expansion. *In*: LLOYD, P. J.; LEE, H. H. (Eds.). **Frontiers of research on intraindustry trade**. Basingstoke: Palgrave-Macmillan, 2002.

BRÜLHART, M.; MURPHY, A.; STROBL, E. **Intra-industry trade and job turnover**. Unpublished Manuscript: University of Lausanne; University College Dublin/CORE; University of Louvain, 2004.

BRÜLHART, M.; ELLIOTT, R. J. R.; LINDLEY, J. Intra-industry trade and labour-market adjustment: a reassessment using data on individual workers. **Review of World Economics**, v. 142, n. 3, p. 521-545, 2006. Retrieved from: https://bit.ly/2JZ9RnD.

BRÜLHART, M.; THORPE, M. East-Asian growth, intra-industry trade and adjustment. Asian Pacific Journal of Economics and Business, v. 3, n. 2, 1999.

BRÜLHART, M.; THORPE, M. Intra-industry trade and adjustment in Malaysia: puzzling evidence. **Applied Economics Letters**, v. 7, n. 11, p. 729-733, 2000.

CABRAL, M.; SILVA, J. Intra-industry trade expansion and employment reallocation between sectors and occupations, **Review of World Economics**, v. 142, n. 3, p. 496-520, 2006. Retrieved from: https://bit.ly/2NqKWeD>.

CAMERON, A. C.; TRIVEDI, P. K. **Microeconometrics**, Cambridge Books, Cambridge University Press, 2005.

CNAE. Classificação Nacional de Atividade Econômica versão 1.0. IBGE – Concla. Rio de Janeiro, 2003. Retrieved from: https://bit.ly/32lUfAU.

DIXIT, A. K.; STIGLITZ, J. E. Monopolistic competition and optimum product diversity. **The American Economic Review**, v. 67, n. 3, p. 297-308, 1977. Retrieved from: https://bit.ly/2PVWVmj.

ERLAT, G.; ERLAT, H. Intra-industry trade and labor market adjustment in turkey: another piece of puzzling evidence? **Emerging Markets Finance & Trade**, v. 42, n. 5, p. 5-27, 2006. Retrieved from: https://bit.ly/2WRzZpE>.

FAUSTINO, H, C.; LEITÁO, N. C. **Intra-industry trade and labor costs**: the smooth adjustment hypothesis. Technical University of Lisbon, Department of Economics. Working Paper: WP17/2009/DE/SOCIUS, 2009.

FEENSTRA, R. C. **Advanced internatinal trade**: theory and evidence. Princeton: Princeton University Press, 2004.

GRUBEL, H. G.; LLOYD, P. J. Intra-industry trade. London: Macmillan, 1975.

GREENAWAY, D.; HAYNES, M.; MILNER, C. Adjustment, employment characteristics and intra-industry trade, **Review of World Economics**, v. 138, n. 2, p. 254-276, 2002. Retrieved from: https://bit.ly/36F4rIh.

HARRIS, M. N.; MÁTYÁS, L.; SEVESTRE, P. Dynamic models for short panels the econometrics of panel data. *In*: MÁTYÁS, L.; SEVESTRE, P. (Eds.). **Springer Berlin Heidelberg**, v. 46, p. 249-278, 2008. (Advanced Studies in Theoretical and Applied Econometrics).

IBGE – INSTITUTO BRASILEIRO DE BANCO DE DADOS AGREGADOS. **Sistema IBGE de Recuperação Automática** - *SIDRA*, 2011. Retrieved from: http://www.ibge.gov.br.

IPEA – INSTITUTO DE PESQUISA ECONÔMICA APLICADA. 2011. Retrieved from: http://www.ipeadata.gov.br/.

JUDSON, R. A.; OWEN, A. L. Estimating dynamic panel models: a practical guide for macroeconomists. **Federal Reserve Board of Governors mimeo**, 1996.

KRUGMAN, P. R. Intra-industry specialization and the gains from trade. **Journal of Political Economy**, University of Chicago Press, v. 89, Oct., p. 959-973, 1981.

MDIC – MINISTÉRIO DO DESENVOLVIMENTO, DA INDÚSTRIA E DO COMÉRCIO EXTERIOR. 2011. Retrieved from: https://bit.ly/2ns5eum.

MTE – MINISTÉRIO DO TRABALHO E EMPREGO. **Relação Anual de Informações Sociais**, 2011. Retrieved from: http://www.mte.gov.br/RAIS/>.

RASEKHI, S.; GHADERI, S. Marginal intra-industry trade and employment reallocation: the case study of Iran's manufacturing industries. **Journal of Industry, Competition and Trade**, p. 1-13, 2012.

ROODMAN, D., How to do xtabond2: An introduction to difference and system GMM in Stata, **Stata Journal**, College Station, TX, v. 9, n. 1, p. 86-136, 2009a. Retrieved from: https://bit.ly/2PVWGrf.

ROODMAN, D. A note on the theme of too many instruments. **Oxford Bulletin of Economics and Statistics**, Department of Economics, University of Oxford, v. 71, p. 135-158, 2009b.

THORPE, M.; LEITÃO, N. C. Marginal intra-industry and adjustment costs: the Australian experience, CLM Discussion Paper Series 2011/02, 2011.

MEDINDO A EFICIÊNCIA PRODUTIVA DO TRANSPORTE POR FERROVIAS BRASILEIRAS: UMA APLICAÇÃO DOS MODELOS DEA E TOBIT

Francisco Gildemir Ferreira da Silva¹ José Kleber Macambira² Carlos Henrigue Rocha³

Este estudo mede a eficiência produtiva das ferrovias de carga brasileiras por meio da aplicação de modelos DEA (data envelopment analysis), considerando retornos constantes de escala e retornos variáveis de escala. Dados da Agência Nacional de Transportes Terrestres (ANTT) foram utilizados por possuírem atualização mensal. O período de estudo desta pesquisa é de janeiro de 2006 a dezembro de 2011. Doze concessionárias de transporte ferroviário foram consideradas nesta análise. Primeiro, foram selecionadas variáveis proxies que representam os fatores de produção utilizados pelas concessionárias ferroviárias na produção de transporte. Essas variáveis foram utilizadas como insumos para a produção e o transporte, e produção de modelos DEA. Após a aplicação dos modelos DEA, foram obtidos os ganhos de eficiência produtiva. Com o modelo Tobit, modelos de regressão foram preparados para avaliar o nível de correlação entre os inputs dos modelos DEA e das eficiências produtivas. Os resultados dos modelos DEA mostram que as estradas de ferro que se especializam no transporte de minerais e commodities agrícolas têm níveis mais elevados de eficiência produtiva do que as ferrovias que operam no transporte de um portfólio mais diversificado de produtos. Por sua vez, os modelos de regressão Tobit mostram que todas as variáveis usadas como insumos na modelagem DEA têm correlação significativa com a eficiência produtiva. Desse modo, analisar os significados de correlações pode-se inferir que o pessoal e a gestão operacional são questões significativas no desempenho do frete brasileiro nas ferrovias.

Palavras-chave: modelos truncados e censurados; modelos de regressão de limiar; produção; produtividade multifatorial; ferrovias.

MEASURING THE PRODUCTIVE EFFICIENCY OF BRAZILIAN RAILWAYS TRANSPORT: AN APPLICATION OF DEA AND TOBIT MODELS

This study measures the productive efficiency of Brazilian freight railways by the application of DEA (Data Envelopment Analysis) models considering constant returns to scale and variable returns to scale returns. Data from the National Transportation Agency was used because it has a monthly actualization. The study period is among the months from January 2006 to December 2011 and the 12 dealerships of railway transportation were considered for this study. First, proxy variables representing the factors of production used by railway concessionaires in the production of transport were selected. These variables were used as inputs to production and transport as output of DEA models. After application of DEA models and obtaining the productive efficiencies, Tobit censored regression models were prepared to assess the level of correlation between inputs

^{1.} Coordenação do Programa de Pós-Graduação em economia da Universidade Federal do Ceará (Caen/UFC). *E-mail*: <qildemir@qmail.com>.

^{2.} Mestre em transportes pela Universidade de Brasília (UnB). *E-mail*: <josekleber.pjl@gmail.com>.

^{3.} Programa de Pós-Graduação em Transportes (PPGT) da UnB. E-mail: <chrocha@unb.br>.

of the DEA models and productive efficiencies. The results of DEA models show that the railroads that specialize in the transportation of minerals and agricultural commodities have higher levels of productive efficiency than the railroads that operate in the transport of a more diversified portfolio of products. On the other hand, the Tobit regression models show that all variables used as inputs in modeling DEA, have significant correlation with productive efficiencies, and analyzing the meanings of correlations can be inferred that the personnel and operational management are significant issues in performance of Brazilian freight railways.

Keywords: truncated and censored models; threshold regression models; production; multifactor productivity; railroads transportation.

JEL: C24; D24; L92.

1 INTRODUÇÃO

Entre 1996 e 1998 foram concedidos 28.614 km de ferrovias à iniciativa privada. Atualmente, a gestão e operação dessa malha são realizadas por doze concessões que atuam em diferentes mercados e regiões. A expectativa do Estado com a privatização do setor ferroviário era a de aumento da participação desse setor na matriz de transporte de carga brasileira, com o aumento da eficiência produtiva ocasionado por investimentos que seriam realizados pelas concessionárias. Porém, o que se observou foi uma especialização por parte das concessões no atendimento a dois segmentos específicos do mercado, minério de ferro e grãos agrícolas (soja e milho). Esse arranjo fez com que o Estado iniciasse uma série de reformas no setor ferroviário com o objetivo de aumentar sua eficiência com a desmonopolização do setor e a criação de condições para um sistema ferroviário mais aberto e mais competitivo. Os estudos de eficiência do setor ferroviário indicam um significativo nível de ineficiência do sistema e um efeito positivo é observado com as reformas (Pavlyuk, 2008; Coelli e Perelman, 2000; Friebel, Ivaldi e Vibes, 2005; Asmild et al., 2008; Wetzel, 2008; Silva et al. 2010). Apesar do número significativo de estudos internacionais nessa área, não há qualquer uniformidade na compreensão dos resultados de funcionamento do sistema de transporte ferroviário, considerando parâmetros relacionados à eficiência, como o emprego do material rodante, consumo de combustível e emprego de mão de obra.

Objetivando analisar a eficiência produtiva do sistema ferroviário brasileiro de carga, que é composto por doze concessionárias, durante o período de 2006 a 2011, este trabalho determina parâmetros que influenciam a eficiência produtiva do sistema ferroviário brasileiro de carga, utilizando modelos não paramétricos de análise de eficiência, aferindo o nível de correlação entre *inputs* dos modelos não paramétricos e a eficiência produtiva do sistema ferroviário brasileiro de carga parametricamente.

Para consecução do proposto, este artigo está estruturado em cinco tópicos, incluindo essa introdução. O tópico dois apresenta um histórico sobre as ferrovias brasileiras desde o período do Império até o processo de elaboração do novo marco

regulatório para implantação do modelo ferroviário *open access* no Brasil. O terceiro tópico apresenta os conceitos de produtividade e eficiência, e descreve as técnicas não paramétricas e paramétricas de análise de eficiência utilizadas na mensuração de eficiência produtiva das concessões ferroviárias brasileiras. No quarto tópico, faz-se uma descrição dos indicadores do setor, desenvolve-se a metodologia e os modelos para estudar a eficiência das concessões ferroviárias brasileiras, expondo os resultados encontrados. O último tópico conclui o estudo.

2 ASPECTOS HISTÓRICOS DAS FERROVIAS BRASILEIRAS

No Brasil, o modal ferroviário ocupou lugar de destaque no transporte de cargas, tendo o seu desenvolvimento favorecido pela hegemonia da tecnologia ferroviária de transportes no mundo, a expansão da produção de café e os incentivos governamentais para a construção de ferrovias (Buri *et al.*, 2006). Tal quadro se reverteu com a deflagração da Primeira Guerra Mundial, implicando uma desaceleração das obras ferroviárias por falta de recursos financeiros. Com isso, em 1958, a rede ferroviária brasileira atingia a marca recorde de 37.967 km em tráfego; enquanto a malha americana, de mesmo período de construção, representava quase dez vezes esse valor. Além disso, a partir daquele ano, a malha de ferrovias começou a encolher, com a erradicação de linhas e ramais antieconômicos, sendo catalisada na década de 1950. Com os grandes volumes de investimentos direcionados à construção e pavimentação de rodovias, as ferrovias começaram a perder importância no transporte brasileiro (Buri *et al.*, 2006).

Tavares (2003) indica que, durante a década de 1980 e início da década de 1990, o governo federal realizou vultosos aportes na Rede Ferroviária Federal S/A (RFFSA), que contribuíram significativamente para o enorme deficit público, obrigando o governo a repensar a gestão do setor ferroviário brasileiro. Em 1992, a RFFSA foi incluída no Programa Nacional de Desestatização (PND), criando a expectativa de superar um dos principais gargalos ao crescimento econômico do país, a infraestrutura de transportes, pois objetivava transferir a gestão do setor ferroviário à iniciativa privada, almejando que esta realizasse investimentos para reverter o quadro de deterioração do setor (Souza, 1997). A partir de 1994, as concessões foram consideradas prioritárias e, em fevereiro e julho de 1995, foram aprovadas as Leis nºs 8.987 e 9.074, que dispunham sobre o regime de concessão e permissão dos serviços públicos, favorecendo o início das desestatizações cujo modelo consistia em: licitação, no mesmo leilão, da concessão do serviço público de transporte ferroviário de cargas prestado pela RFFSA; arrendamento dos ativos operacionais e venda de bens de pequeno valor de propriedade da RFFSA; manutenção da empresa estatal encarregada de administrar seus ativos não operacionais, de modo a auferir receita sobre eles e amortizar o seu endividamento, que não seria repassado à iniciativa privada; e criação de uma entidade encarregada de resolver,

em caráter administrativo, eventuais conflitos entre concessionária, usuários e o poder concedente (Souza, 2003).

Segundo Oliveira (2005) *apud* Lang (2007), o modelo proposto de divisão da malha da RFFSA, para realização das concessões, baseou-se nas superintendências regionais (SRs) da RFFSA, como apresentado na tabela 1. O objetivo era obter melhor configuração operacional para se constituírem em unidades de negócio independentes e autossuficientes, segundo os principais fluxos de transporte atuais e potenciais.

TABELA 1
Composição da RFFSA antes do processo de concessão iniciado em 1995

Malha	Superintendências regionais	Sede das superintendências regionais	Ferrovias após a concessão	Extensão (km)	
Tereza Cristina	SR 9	Tuburão/SC	Ferrovia Tereza Cristina	169	
	SR 2	Belo Horizonte/MG			
Malha Centro-Leste	SR 7	Salvador/BA	Ferrovia Centro- Atlântica	7.207	
	SR 8	Campos/RJ			
	SR 1	Recife/PE			
Malha Nordeste	SR 11	Fortaleza/CE	Companhia Ferroviária do Nordeste¹	4.654	
	SR 12	São Luiz/MA			
Malha Oeste	SR 10	Bauru/SP	Ferrovia Novoeste ²	1.621	
Malha Sudeste	SR 3	Juiz de Fora/MG	MDC Logistics	1.622	
Maina Sudeste	SR 4	São Paulo/SP	MRS Logística	1.633	
Malha Cul	SR 5	Curitiba/PR	Farrania Cul Atlântica3	6.705	
Malha Sul	SR 6	Porto Alegre/RS	Ferrovia Sul-Atlântico ³	6.785	
Total	Administração geral (AG)	Rio de Janeiro/RJ	Todas administradas pela RFFSA	22.069	

Fonte: Silveira (2003) apud Silveira (2007).

Elaboração dos autores.

Notas: 1 Atual Ferrovia Transnordestina S.A. (FTL).

Nos editais de concessão, foram estabelecidas as obrigações em atender metas de desempenho que norteavam os investimentos, pois não existiam indicações claras dos investimentos a serem realizados. Em março de 1996, foi estabelecido o Regulamento dos Transportes Ferroviários (RTF) (Brasil, 1996). Este regulamento foi publicado antes dos leilões dos ativos da RFFSA. Segundo Castro (2000) *apud* Lang (2007), o Decreto nº 1.832, de 4 de março de 1996, pouco avançava no sentido de detalhar os critérios de regulamentação dos pontos críticos do setor,

² Atual América Latina Logística Malha Paulista (ALLMP-RMP).

³ Atual América Latina Logística Malha Sul (ALLMS-RMS).

como: tarifas, desativação de ramais antieconômicos, interpenetração,4 tráfego mútuo e direito de passagem. Como parte do processo de privatização e frente à necessidade de regulamentação e fiscalização da prestação do serviço público de transporte ferroviário, em 2001, foi criada a Agência Nacional de Transportes Terrestres (ANTT). A criação desta agência permitiu a regulamentação de penalidades, bem como dos direitos e deveres previstos nos contratos de concessão. Até então, houve um lapso de mais ou menos dez anos para ocorrer mudanças significativas no setor, notoriamente no marco regulatório e no plano de investimentos público privados, iniciados somente em julho de 2011 com as novas resoluções para o setor, resumidas no quadro 1.

QUADRO 1 Resoluções nº 3.694/2011, 3.695/2011 e 3.696/2011

Tema	Antes	Depois
Legislação	Decreto nº 1.832/1996 da Presidência da República.	Resoluções nº 3.694/2011, 3.695/2011 e 3.696/2011 da ANTT.
Tráfego mútuo	Havia possibilidade de compartilhamento de uma mesma via mediante pagamento, mas as condições de operações seriam estabelecidas pelas próprias concessionárias.	Os critérios para o tráfego mútuo são baseados na capacidade ociosa da via e nas metas de produção por trecho pactuado entre a concessionária e a ANTT. Mesmo assim, ocorrerá pagamento da requerente para a cedente.
Direito de passagem	Caso o tráfego mútuo não fosse possível, havia possibilidade de cessão da via a outra concessionária mediante pagamento. Novamente as condições de operações seriam estabelecidas pelas próprias concessionárias.	Os critérios para direito de passagem são baseados na capacidade ociosa da via e nas metas de produção por trecho pactuado entre a concessionária e a ANTT. Ocorrerá o pagamento da requerente para a cedente.
Impossibilidade de tráfego mútuo	As situações em que o tráfego mútuo não fosse possível não eram estabelecidas pela legislação.	Quando houver desacordo comercial entre as partes ou as características operacionais inerentes ao tráfego mútuo comprometem o atendimento eficiente da necessidade de transporte do usuário.
Condições de operação	Eram estabelecidas entre as concessionárias.	Baseiam-se em critérios claros estabelecidos pela legislação e informações pactuadas entre concessionárias e ANTT sobre a capacidade ociosa e metas de produção.
Inventário de capacidade	Não existia.	Anualmente, cada concessionária deve entregar à ANTT a Declaração de Rede, informando a capacidade instalada e ociosa.
Metas de produção por trecho	Não existia.	Anualmente, cada concessionária deve entregar Proposta de Pactuação de Metas de Produção e Metas de Segurança por Trecho. As metas apresentadas serão acordadas entre ANTT e concessionária.

Fonte: CNT (2013). Elaboração dos autores.

Dando continuidade ao processo de mudança do setor, em agosto de 2012, o governo federal anunciou o Programa de Investimentos em Logística (PIL). Entre as ações previstas no programa para o setor ferroviário, há a implantação de

um novo modelo de concessão em que haverá uma separação entre a atividade de fornecimento de infraestrutura ferroviária e a prestação do serviço de transporte ferroviário. Segundo CNT (2013), o modelo apresentado pelo governo federal é significativamente mais complexo que o atualmente em vigor no país. A iniciativa privada participará tanto da oferta de infraestrutura quanto do serviço de transporte ferroviário, mas em ambos haverá uma participação direta do ente público. Para a expansão e manutenção da malha ferroviária, serão firmadas parcerias públicoprivadas (PPPs) visando financiar o investimento e dar celeridade à execução das obras de engenharia, sendo a pessoa jurídica fruto dessa parceria com o gestor da infraestrutura ferroviária (GIF). Para garantir o retorno das PPPs e a concorrência no serviço de transporte ferroviário, a Engenharia, Construções e Ferrovias S/A (Valec) comprará a capacidade integral de transporte das ferrovias. Por meio de oferta pública, essa empresa venderá o direito de passagem às empresas de transporte ferroviário, chamadas de operadores ferroviários (OFs), nas malhas submetidas a esse modelo de concessão, para garantir a manutenção da modicidade tarifária e assegurar a oferta do serviço em trechos de menor demanda. Técnicos do governo aventam a relação de melhora de produtividade com o novo modelo, mas como seria este ganho de eficiência? Abordar-se-á tal tema no próximo item.

3 TÉCNICAS DE ANÁLISE DE FEICIÊNCIA

A produção pode ser definida como um processo pelo qual os insumos (*inputs*) são combinados e transformados em bens e serviços (*output*) (Case e Fair, 1999), bem como em unidades tomadoras de decisão (DMU, da sigla em inglês *decision making units*). Estudar a produção é de grande importância devido à escassez de recursos e ao desejo humano de utilizá-los plenamente. A noção microeconômica presume que as empresas (DMUs), no âmbito do mercado livre, devem alocar os *inputs* de forma eficiente com o objetivo de obter o lucro máximo e/ou custo mínimo a partir da produção de certos *outputs* (Erkoc, 2012).

Dyson (2001) afirma que a mensuração do desempenho tem um papel essencial na avaliação da produção, pois definirá o futuro da empresa. Thanassoulis (2001) identifica as seguintes informações que podem ser obtidas pela mensuração do desempenho: *i*) identificação de boas práticas operacionais para divulgação; *ii*) escala operacional mais produtiva; *iii*) espaço para ganhos de eficiência no uso dos recursos e/ou para o aumento da produção; *iv*) modelo mais adequado para uma unidade ineficiente melhorar seu desempenho; *v*) taxas marginais de substituição entre os fatores de produção; e *vi*) alteração ao longo do tempo da produtividade por cada unidade de operação e da forma mais eficiente dessas unidades em cada ponto no tempo. Produtividade e eficiência são dois conceitos importantes na medição do desempenho. A produtividade de uma firma pode ser definida como a relação entre o(s) *output*(s) e o(s) *input*(s) (Coelli, Prasada Rao e Battese, 1998). A

eficiência pode ser definida como a produtividade em relação ao tempo ou espaço, ou ambos (Lansink, Silva e Stefanou, 2001). Entretanto, para a formalização da análise por DEA (data envelopment analysis), devem-se considerar outros aspectos discutidos na próxima seção.

3.1 Análise envoltória de dados

Charnes *et al.* (1978) descrevem o DEA como um modelo matemático de programação aplicada a dados observáveis, que fornecem uma nova maneira de obter estimativas empíricas das relações, como as funções de fronteira das possibilidades de produção, que são pilares da economia moderna. A abordagem da DEA tem a vantagem de evitar a necessidade de atribuir uma medida *a priori* da importância relativa de qualquer *input* ou *output* comum e outras abordagens (como modelagem de regressão estatística) e resultou em seu uso em vários estudos nos setores governamental e sem fins lucrativos, regulatório e privado.

No DEA, a eficiência é calculada obtendo-se uma relação entre o vetor de *outputs* e de *inputs*, estando situada entre zero e um, sendo representada pela equação (1):

$$\theta = \frac{u_0 + u_1 y_1 + \dots + u_n y_n}{v_0 + v_1 x_1 + \dots + v_n x_n} \le 1,0$$
(1)

em que $\vec{y} = (y_1, y_2, ..., y_n)$ é o vetor de *inputs*; $\vec{x} = (x_1, x_2, ..., x_n)$ é o vetor de *outputs*; u_i são os pesos dos *inputs*; e v_i os pesos dos *outputs*. Os primeiros modelos DEA desenvolvidos na década de 1960 foram: o modelo CRS (constant return scale – retornos constantes de escala) ou CCR (abreviatura dos nomes de seus criadores Charnes, Cooper e Rhodes, 1978), e o modelo VRS (variable return scale – retornos variáveis de escala) ou BCC (abreviatura de Banker, Charnes e Cooper, 1984). Toda a teoria que fomenta essa abordagem de tratamento de dados surgiu conforme as definições a seguir.

Definição 1.1 (eficiência – extensão da definição de Pareto-Koopmans):

 a eficiência completa (100%) é alcançada por qualquer DMU se, e somente se, nenhum de seus *inputs* e *outputs* podem ser melhorados sem o agravamento de alguns dos seus outros *inputs* ou *outputs*.

Essa definição apresenta um problema na maioria das aplicações de gestão ou ciências sociais, pois os possíveis níveis teóricos de eficiência não são conhecidos. Assim, a definição 1.1 é, portanto, substituída, destacando seus usos apenas com a informação que é empiricamente conforme a definição 1.2.

2) Definição 1.2 (eficiência relativa): a DMU deve ser classificada como totalmente eficiente (100%) com base na evidência disponível se, e somente se, os desempenhos das outras DMUs não mostrarem que

alguns de seus *inputs* ou *outputs* podem ser melhorados sem agravamento de alguns dos seus outros *inputs* ou *outputs*.

Essa definição evita a necessidade de recorrer a preços ou outras hipóteses de pesos, que são selecionados *a priori* e, também, a necessidade de especificar explicitamente as relações formais que deveriam existir entre *inputs* e *outputs*. Este tipo básico de eficiência, referida como "eficiência técnica" em economia, pode ser estendido a outros tipos de eficiência quando os dados, tais como preços e custos unitários, estão disponíveis para uso na DEA. Com base na definição 1.2., os modelos DEA podem ser subdivididos em termos de retornos de escala. Charnes, Cooper e Rhodes (1978) propuseram, inicialmente, a medição da eficiência da DMU para CRS, em que todas as DMUs estão operando em sua escala ótima; Banker, Charnes e Cooper (1984) introduziram as variáveis de retornos de escala (VRS), permitindo a quebra de eficiência na eficiência técnica e de escala na DEA. Assim, sendo Θ^* a função que descreve a relação entre *inputs* e *outputs* e a folga *si*, o quanto de *input i* precisa ser adicionado ou retirado para se obter o máximo de eficiência desejado cabe às definições 1.3 e 1.4, mais formal, para eficiência DEA.

- 3) Definição 1.3 (eficiência DEA): o desempenho do DMUo é totalmente eficiente (100%) se, e somente se, ambos (1) $\Theta^* = 1$ e (2) todas as folgas $si \neq 0$ e $si + \neq 0$.
- 4) Definição 1.4 (fracamente eficiente): o desempenho do DMUo é fracamente eficiente se, e somente se, ambos (1) $\Theta^* = 1$ e (2) si $\neq 0$ e/ ou si+ $\neq 0$ para algum i ou r em alguma alternativa ótima.

A seguir é apresentado o teorema para identificação do tipo de retorno de escala obtido de Banker e Thrall (1994) que identifica o sinal de Θ^* no modelo BCC.

1) Teorema 1.1: as condições de identificação dos retornos de escala para o modelo BCC são: IRS prevalece em (\hat{x}_0, \hat{y}_0) se, e somente se, $\Theta^* < 0$ para todas as soluções ótimas; DRS prevalece em (\hat{x}_0, \hat{y}_0) se e somente se, $\Theta^* > 0$ para todas as soluções ótimas; CRS prevalece em (\hat{x}_0, \hat{y}_0) se e somente se, $\Theta^* = 0$ para pelo menos uma solução ótima.

Aqui, pode-se notar, (\hat{x}_0, \hat{y}_0) são as coordenadas do ponto na fronteira de eficiência que é obtido a partir da projeção do modelo BCC sobre a fronteira na avaliação do DMUo através da solução do modelo BCC.

3.1.1 Aplicação da DEA em transportes

Lan e Lin (2003), por meio de modelos DEA, investigam a eficiência técnica do serviço de 76 ferrovias no mundo durante o período de 1999 a 2001 e mostram que a maioria das ferrovias apresentaram retornos variáveis à escala. Complementarmente, os autores realizaram a detecção de *outliers* e análise de sensibilidade para a DMU

eficiente. Por seu turno, Malhotra, Malhotra e Lermack (2009) analisam o desempenho financeiro de sete ferrovias norte-americanas de carga utilizando a técnica DEA e indicam as empresas que operam de forma mais eficiente relativamente a outras do setor, diagnosticando pontos de melhora. Já Santos (2011) aplica a metodologia DEA para avaliar a eficiência produtiva das ferrovias de carga brasileiras com dados da ANTT de 2006 a 2009, objetivando avaliar se ferrovias de bitola larga seriam mais eficientes do que as ferrovias de bitola métrica, porém os resultados sugerem que esta hipótese deve ser refutada.

Em Caldas *et al.* (2012), analisou-se o desempenho operacional do transporte ferroviário de carga dos Estados Unidos e do Brasil com a ferramenta DEA e o resultado da modelagem DEA, utilizando CCR e algumas premissas simplificadoras, apontaram que os trechos americanos são mais eficientes que os brasileiros. Petrovic *et al.* (2012), com a mesma técnica DEA, mensuram a eficiência relativa do transporte ferroviário de cargas em países europeus durante o período de 2005 a 2009, concluindo que a melhoria dos indicadores de investimento (*inputs*) podem aumentar as receitas (*output*) das ferrovias europeias de carga.

Nag (2013) utilizou a DEA para determinar a eficiência das ferrovias indianas e compará-las com outras grandes ferrovias no mundo, bem como avaliar o desempenho relativo e o desempenho ao longo do tempo dessas ferrovias. O autor conclui que as ferrovias chinesas, indianas, japonesas, russas, italianas e norteamericanas são as mais eficientes do mundo. Complementarmente, o estudo revela que todos os inputs utilizados na análise são críticos e o desempenho das ferrovias indianas vem aumentando ao longo do tempo. Por fim, Bhanot e Singh (2014) aferem o desempenho das ferrovias indianas tendo em vista a quebra de monopólio decorrente das alterações da legislação do setor. Os autores analisam o segmento de transporte ferroviário de contêineres comparando dados secundários de terminais de contêineres de duas companhias privadas, Adani e Gateway Logistics, e uma estatal indiana, a Container Corporation of India (CONCOR). Os resultados apontam para um maior nível de eficiência para a CONCOR quando comparado com os agentes privados. Outro resultado apresentado é que o nível de eficiência da CONCOR é mais estável ao longo do tempo quando comparado com as grandes flutuações do nível de eficiência dos agentes privados.

3.2 Modelo Tobit

Tobin (1958) propôs o modelo, denominado Tobit, para descrever a relação entre uma variável dependente y_i não negativa e uma variável independente (ou vetor) x_i . O modelo supõe que há uma variável latente não observável, y_i^* . Esta variável depende linearmente de x_i pelo parâmetro (ou vetor) β que determina a relação entre a variável independente (ou vetor) x_i e a variável latente y_i^* , tal como no modelo linear. Na proposição, o termo de erro, u_i deve ser normalmente distribuído para

capturar influências aleatórias. A variável observável, *y_i*, é definida para ser igual à variável latente (Tobin, 1958; Amemiya, 1984; 1985). A equação estrutural do modelo é:

$$y_{i}^{*} = X_{i}\beta + u_{i}, \tag{2}$$

onde $u_i N(0,\sigma 2)$. y_i^* é uma variável latente observada com os valores maiores do que τ e censuradas caso contrário. O y_i observado é definido pela equação 3. No modelo Tobit típico, assume-se que τ = 0, ou seja, os dados são censurados em 0. Assim, tem-se:

$$y_i = \begin{cases} y^* & \text{se } y^* \le 0. \\ 0 & \text{se } y^* \le 0. \end{cases}$$
 (3)

A censura de dados ocorre quando os dados da variável dependente são perdidos (ou limitados), mas não os dados dos regressores. Quando uma distribuição é censurada no lado esquerdo, as observações com valores iguais ou inferiores a τ são definidos para τ_{ν} (Amemiya, 1973):

$$y_i = \begin{cases} y^* se \ y^* \le \tau. \\ \tau se \ y^* \le \tau. \end{cases} \tag{4}$$

O uso de τ e τ_y são apenas uma generalização de ter τ e τ_y igual a 0. Se uma variável contínua y tem uma função densidade de probabilidade f(y) e τ é uma constante, então temos:

$$f(y) = [f(y^*)]^{d_i} [F(\tau)]^{1-d_i}.$$
(5)

Em outras palavras, a densidade de y é a mesma que para y^* , para $y > \tau$, e é igual à probabilidade de observar $y^* < \tau$ se $y = \tau$. d é uma variável indicadora que é igual a 1 se $y > \tau$, ou seja, a observação é censurada e é igual a 0 se $y = \tau$. A função de verossimilhança pode ser escrita como (Schnedler, 2005):

$$L = \prod_{i}^{N} \left[\frac{1}{\sigma} \phi(\frac{y - \mu}{\sigma}) \right]^{d_{i}} \left[1 - \phi(\frac{y - \mu}{\sigma}) \right]^{1 - d_{i}}.$$
 (6)

O valor esperado de uma variável censurada é:

$$E[y]$$
 - $(P(n\tilde{a}o\ censurado) \times E[y/y > \tau]) + (P(n\tilde{a}o\ censurado) \times E[y/y = \tau]).$

$$E[y] = \{\phi(\frac{y-\mu}{\sigma})[\mu + \sigma\lambda(\alpha)]\} + \phi(\frac{y-\mu}{\sigma})\tau_{y}. \tag{7}$$

Para o caso particular de quando $\tau = 0$, temos:

$$E[y] = \phi(\frac{\mu}{\sigma})[\mu + \sigma\lambda]$$
 (8)

onde,

$$\lambda = \frac{\phi(\frac{\mu}{\sigma})}{\Phi(\frac{\mu}{\sigma})} \tag{9}$$

3.2.1 Aplicação de regressão Tobit em transportes

Li e Hu (2010) analisam o sistema ferroviário da China, identificando os principais fatores que afetam a eficiência por meio de regressão Tobit. Após avaliarem a eficiência ferroviária, aplicam a metodologia DEA e Índice Malmquist de Produtividade Total de Fatores (MPI). Eles concluem que o sistema ferroviário chinês apresenta baixos níveis de eficiência técnica global e evidencia o desequilíbrio no desenvolvimento do sistema, mostrando que as ferrovias que operam no leste do país possuem melhores níveis de eficiência e as que operam no oeste apresentaram maior potencial de crescimento da produção ferroviária. Eles apontam como fatores dessa ineficiência o desenvolvimento da economia regional, a população, a densidade ferroviária e a densidade rodovia.

Kutlar, Ali Kabasakal e Sarikaya (2013) mediram o desempenho das ferrovias de passageiro e carga em todo o mundo com dados de 2000 a 2009, correlacionando os *inputs* da DEA por regressão Tobit, utilizando uma amostra de 31 empresas ferroviárias. Os resultados mostram que a mesma composição de *inputs* usados na análise Tobit dá resultados mais compatíveis com os escores de eficiência alocativa, quando comparado com os escores de eficiência técnica. Por seu turno, Merkert, Smith e Nash (2009) objetivavam determinar o impacto da estrutura institucional, o tamanho das empresas, grau de concorrência, tipo de operação e o nível de custos de transação sobre a eficiência técnica e global e, para tanto, aplicam regressões Tobit para aferir o grau de influência dos *inputs* utilizados nos modelos DEA nas eficiências calculadas, utilizando 43 empresas ferroviárias de passageiros e cargas da Inglaterra, Alemanha e Suécia. Os resultados mostram que a separação vertical não tem efeito sobre a eficiência técnica e de custo, mas apresenta um efeito marginal negativo sobre a eficiência alocativa e que a concorrência de operação de trens de carga diminui a eficiência e os custos de transação possuem

uma associação negativa com todas as eficiências. Chen (2010) explora a relação entre implantação do Taiwan High Speed Rail (THSR) e a eficiência econômica regional de Taiwan, usando dados anuais de 2004 a 2010 e aplicando o método DEA e a regressão Tobit. Os resultados indicam que as eficiências médias para as regiões analisadas estão se deteriorando desde o início da operação do THSR em 2007. De acordo com os resultados da regressão Tobit, as eficiências médias, para as regiões, tiveram uma melhora no curto prazo, mas uma piora no longo prazo, indicando que a construção de infraestrutura de transporte não estimulou o desenvolvimento econômico regional.

Tsai, Mulley e Merkert (2014) avaliam os fatores que impulsionam a eficiência operacional dos sistemas ferroviários urbanos por meio de regressão Tobit. Eles avaliam os determinantes das eficiências calculadas pelos modelos DEA, utilizando dados de sistemas ferroviários urbanos internacionais entre 2009 e 2011. Os resultados identificam sistemas com eficiência alocativa e de custos (por exemplo, Hong Kong) e sistemas com um melhor desempenho em termos de eficiência (por exemplo, Sydney) e que o número de estações influencia a eficiência técnica e o fator determinante da eficiência alocativa é o custo e a densidade populacional. Por fim, Sameni e Preston (2012) analisam a eficiência dos operadores de passageiros da Grã-Bretanha na utilização da capacidade. Com os escores de eficiência resultantes da aplicação do modelo DEA, avalia-se por modelo de regressão Tobit os fatores da eficiência. Os resultados mostram que existe correlação positiva entre operar na região de Londres e os escores de eficiência. Além disso, também é verificada a existência de correlação negativa entre a oferta de serviços regionais e os escores de eficiência.

4 METODOLOGIA E ESPECIFICAÇÃO DOS MODELOS

Esta seção apresentará a proposta metodológica, a definição dos modelos e será aplicado o procedimento para as concessionárias ferroviárias em operação no Brasil.

4.1 Procedimento metodológico

Com base na revisão bibliográfica, realizada nas seções anteriores, foram selecionadas as seguintes variáveis. Estas serão utilizadas como *inputs* nos modelos DEA: número de trens formados; consumo de combustível; utilização de locomotiva; utilização de vagão; e número de empregados.

Cenário	Modelo	Objetivo	Inputs	Output
Retornos constantes de escala	CCR	Medir a eficiência da produção do transporte de cargas.	Número de trens formados; consumo de combustível; utilização de locomotiva; utilização de vagão; número de empregados.	Toneladas por quilômetros úteis.
Retornos variáveis de escala	ВСС	Medir a eficiência da produção do transporte de cargas.	Número de trens formados; consumo de combustível; utilização de locomotiva; utilização de vagão; número de empregados.	Toneladas por quilômetros úteis.

QUADRO 2
Especificação dos modelos DEA

Fonte: Macambira Filho (2014). Elaboração dos autores.

Para o *output* dos modelos DEA, foi selecionada a variável *toneladas por quilômetro úteis*. De posse das variáveis, foram testados dois modelos DEA – apresentados no quadro 2 –, estabelecidos conforme as equações 10 e 11. Ambas utilizam onze DMUs, cinco *inputs* e um *output*.

$$\max z = \mu y_{0}$$

$$\sup z = u y_{0}$$

$$\mu y_{j} - \sum_{i=1}^{5} v_{i} x_{ij} \leq 0 \quad j = 1,...,11;$$

$$\sum_{i=1}^{5} v_{i} x_{i0} = 1;$$

$$\mu_{r}, v_{i} \geq 0;$$

$$\max z = u y_{0} - u_{0}.$$

$$\sup z = u y_{0} - u_{0}.$$

$$\sup z = u y_{0} - u_{0}.$$

$$\sup z = u y_{0} - u_{0}.$$

$$\lim z = v_{0} x_{0} - u_{0}.$$

$$\lim z = v_{0} x_{0}$$

Após obtenção das eficiências técnicas, θ_i^* , pelos modelos DEA, estas serão correlacionadas por meio de regressão Tobit, com os *inputs* utilizados nos modelos DEA para determinar o nível de influência desses *inputs* na eficiência técnica, conforme equação 12, dada a restrição 13.

$$\theta_{i}^{*} = X_{i}\beta + u_{i}, \tag{12}$$

onde,

$$\theta_i = \begin{cases} \theta^* & \text{se } \theta^* \le 0. \\ 0 & \text{se } \theta^* \le 0. \end{cases} \tag{13}$$

Onde θ representa o vetor dos escores de eficiência e $z^* = min \theta = \theta^*$.

Complementarmente, executa-se o algorítimo de Simar e Wilson (1998) e o teste de retornos conforme apontado em Bogetoft e Otto (2011). O bootstrap parte da hipótese de que se a distribuição empírica de uma variável (no caso deste trabalho, o escore de eficiência) corresponder mais ou menos para a verdadeira distribuição, então a distribuição empírica da estatística de t de student corresponderá mais ou menos à verdadeira distribuição de t de student da eficiência. Com a amostra de bootstrap, determinam-se diretamente os intervalos de confiança para a estatística, sendo os resultados desta mais precisos do que a construção de intervalos de confiança baseados no desvio-padrão estimado, uma vez que esta última se baseia no pressuposto de que a distribuição em teste é simétrica e pode ser razoavelmente aproximada usando uma distribuição normal.

Tecnicamente o algorítimo funciona da seguinte forma: tome a definição 1.1, $\vec{y} = (y_1, y_2, ..., y_n)$ é o vetor de *outputs*, $\vec{x} = (x_1, x_2, ..., x_n)$ é o vetor de *inputs*, calcula-se Θ^* , conforme equação 11. Assume-se que os *inputs* e *outputs* das amostras são uma realização de variáveis aleatórias distribuídas de forma idêntica e independentemente com probabilidade P(.). Gera-se $P^*(.)$ como estimativa de P(.) com base em Θ^* obtidos. A relação entre $P^*(.)$, P(.) e Θ^* são complexas e gerarão estimadores Θ^{k*} . Dadas as replicações, calcula-se a média e a variância das eficiências com *bootstrap* conforme equações 14 e 15.

$$\overline{\theta}^{k^*} = \frac{1}{B} * \sum_{b=1}^{B} \theta^{kb} \tag{14}$$

$$Var(\theta^{k^*}) = \frac{1}{B} * \sum_{b=1}^{B} (\theta^{kb} - \overline{\theta}^{k^*})^2$$
(15)

Os resultados obtidos são utilizados para calcular intervalos de confiança e aplicar teste de hipótese de modelo mais adequado aos dados conforme apresentado em Simar e Wilson (1998; 1999; 2000) e Bogetoft e Otto (2011).

4.2 Dados descritivos

As informações utilizadas neste estudo proveem do banco de dados do Sistema de Acompanhamento e Fiscalização do Transporte Ferroviário (Saff) da ANTT. Na análise, não foram considerados os dados da concessionária Ferrovia Norte-Sul (FNSTN), pois o início de sua operação deu-se a partir de 2008. As variáveis utilizadas para aplicação dos modelos estão descritas a seguir.

- tku toneladas por quilômetro útil, que corresponde à produção ferroviária, ou seja, ao somatório do produto entre o volume transportado, em toneladas úteis (tu), e a distância percorrida, em quilômetros (km), de cada composição transitada na malha concedida ou na malha de outras ferrovias, na realização do transporte de carga própria ou de terceiros.
- 2) ntf número de trens formados, que corresponde ao número total de trens de carga formados no período.
- ccl consumo de combustível, que representa o somatório da quantidade de combustível, em litros (l), consumido pelas locomotivas dieselelétricas utilizadas para o desempenho das operações de transporte, manobra e serviço.
- 4) ul utilização de locomotiva de frota própria e de outras ferrovias, que corresponde ao somatório dos tempos de locomotivas utilizadas no transporte de carga, inclusive em carga própria, em locomotiva.hora, por modelo de locomotiva própria, arrendada, alugada, de outras concessionárias e de terceiros no período considerado.
- 5) uv utilização de vagão de frota própria e de outras ferrovias, que corresponde ao somatório dos tempos de vagões carregados no serviço remunerado e carga própria, em vagão.dia, por tipo de vagões próprios, arrendados, alugados e de terceiros (exceto de outras ferrovias), no período considerado.
- 6) ne número de empregados por lotação, que corresponde ao efetivo total de empregados por área de atuação e por tipo de vínculo empregatício com a concessionária, sendo *próprios* para os empregados com vínculo empregatício com a concessionária ou *terceirizados* para os trabalhadores de empresas contratadas para a prestação de serviços nas instalações da concessionária. Em ambos os casos, considera-se a quantidade no último dia do mês informado.

A seguir são apresentadas, para cada concessionária, algumas estatísticas descritivas dos dados utilizados na tabela 2.

TABELA 2
Estatísticas descritivas das variáveis por concessionárias

	Améri	ca Latina Logística Malha	a Norte (ALLMN-RMN) –	rumo malha norte		
Variável	Média	Mediana	Desvio-padrão	Mínimo	Máximo	
TKU	1.256.560.097	1.224.930.144	482.156.056	179.203.788	2.300.403.035	
NTF	316	295	125,7327471	64	792	
CCL	7	7	2	2	10	
UL	268	237	97	153	416	
UV	4.683	4.992	2.061	250	7.711	
NE	1.677	1.604	300,9789409	1142	2.438	
	Amério	a Latina Logística Malha	a Oeste (ALLMO-RMO) –	rumo malha oeste		
Variável	Média	Mediana	Desvio-padrão	Mínimo	Máximo	
TKU	125.357.119	124.712.524	26.408.767	48.721.115	173.738.563	
NTF	814,0462963	661,5	369,8907741	126	1692	
CCL	13	12	2	9	18	
UL	68	64	16	37	94	
UV	2.007	1.755	1.035	37	3.534	
NE	910	898	139	224	1.228	
América Latina Logística Malha Paulista (ALLMP-RMP) — rumo malha paulista						
Variável	Média	Mediana	Desvio-padrão	Mínimo	Máximo	
TKU	286.584.230	280.685.435	101.628.999	62.934.882	570.742.930	
NTF	1.315	799	859,6448281	246	3.167	
CCL	13	13	3	8	20	
UL	174	159	74	79	328	
UV	5.567	3.463	3.586	138	11.251	
NE	1.700	1.926	545,7489494	598	2.332	
	Am	érica Latina Logística Ma	alha Sul (ALLMS-RMS) — ı	rumo malha sul		
Variável	Média	Mediana	Desvio-padrão	Mínimo	Máximo	
TKU	1.412.203.451	1.406.180.133	210.715.453	1.014.634.862	1.938.774.169	
NTF	4500,342593	4090	1438,358476	2499	7507	
CCL	9	9	1	7	20	
UL	449	459	61	336	530	
UV	12.610	14.241	4.477	472	16.620	
NE	5.868	5.528	1.297	3.395	8.189	

(Continua)

		Estrada de	Ferro Carajás (EFC)		
Variável	Média	Mediana	Desvio-padrão	Mínimo	Máximo
TKU	7.680.735.232	7.500.472.033	1.155.203.464	5.245.464.071	10.789.157.756
NTF	1.298	1.340	258,7299799	718	1.735
CCL	2	2	0	2	3
UL	212	215	34	128	257
UV	11.119	11.846	4.497	205	16.739
NE	4.645	4.377	1078,71942	2.642	8.181
		Estrada	de Ferro (EFPO)		
Variável	Média	Mediana	Desvio-padrão	Mínimo	Máximo
TKU	34.846.534	24.566.559	28.000.877	1.251.220	125.702.883
NTF	93	89	28,4866644	42	180
CCL	4.399	15	8.893	8	67.664
UL	9	10	5	1	16
UV	175	100	205	14	976
NE	150	149	13,59851813	114	176
		Estrada de Fer	ro Vitória Minas (EFVM)		
Variável	Média	Mediana	Desvio-padrão	Mínimo	Máximo
TKU	5.977.037.513	6.208.123.352	755.226.636	3.526.142.811	7.042.718.668
NTF	4.424	4.536	531,4719318	2.915	5.210
CCL	3	3	2	3	19
UL	317	316	23	132	377
UV	16.609	19.459	6.293	132	20.833
NE	6.435	6.252	1215,897832	4.740	8.978
		Ferrovia Ce	entro-Atlântica (FCA)		
Variável	Média	Mediana	Desvio-padrão	Mínimo	Máximo
TKU	1.246.833.092	1.263.970.756	278.166.333	586.852.505	1.795.421.835
NTF	4.297	4.265	632,0899509	2.270	5.602
CCL	210	12	1.449	4	10.751
UL	562	530	160	113	1.702
UV	11.116	11.696	4.061	113	16.195
NE	6.370	6.221	961,0311795	4.234	8.948

(Continua)

(Continuação

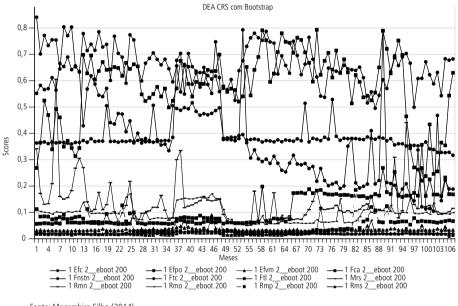
		Ferrovia	Tereza Cristina (FTC)		
Variável	Média	Mediana	Desvio-padrão	Mínimo	Máximo
TKU	17.251.327	16.549.073	3.399.771	8.477.503	26.738.189
NTF	1.033	1.027	138,5686319	482	1.341
CCL	7	7	1	6	9
UL	11	11	1	9	15
UV	335	370	116	11	404
NE	278	288	36,09368219	224	352
		N	MRS Logística		
Variável	Média	Mediana	Desvio-padrão	Mínimo	Máximo
TKU	4.766.993.275	4.813.306.084	676.000.256	2.985.415.224	5.921.861.220
NTF	9.006	9.460	1718,080384	5.902	11.777
CCL	5	5	0	4	7
UL	671	682	114	451	874
UV	15.783	18.026	5.701	663	20.235
NE	7.062	7.465	2834,77348	3.657	11.564
		Transnorde	stina Logística SA (FTL)		
Variável	Média	Mediana	Desvio-padrão	Mínimo	Máximo
TKU	60.562.467	58.440.305	13.648.133	36.614.498	104.878.560
NTF	565	714	362,3391166	105	1.051
CCL	15	15	2	9	19
UL	119	123	15,76333656	84	160
UV	1.834	2.098	653	104	2.442
NE	1747,342593	1750	227,5372481	1320	2344
		Ferrovia	Norte-Sul (FNTSN)		
Variável	Média	Mediana	Desvio-padrão	Mínimo	Máximo
TKU	215.812.401	197.613.488	154.178.479	0	699.839.709
NTF	144	149	61,44731271	6	324
CCL	5	4	2	2	18
UL	22	15	20	0	105
UV	833	560	767	4	3.301
NE	285	231	309,1439003	0	1.655

Fonte: Macambira Filho (2014). Elaboração dos autores.

4.3 Modelos e resultados

Os resultados para cada DMU (concessionária) foram obtidos para modelos DEA, considerando retornos constantes à escala e retornos variáveis à escala. No gráfico 1, são apresentadas as séries temporais com a posição calculada (ordem dos escores classificados do maior para o menor) da DMU em cada período de análise para estimativas de DEA CRS sem *bootstrap*, seguindo-se com a análise dos resultados.





Fonte: Macambira Filho (2014). Elaboração dos autores.

Analisando os resultados da DMU da ALLMN-RMN, pode-se observar que os escores das eficiências técnicas calculadas, a partir dos dois modelos DEA utilizados, apresentaram patamares elevados, porém distintos entre 2006 e 2008. Sendo que, a partir de 2009, houve uma convergência de valores acompanhada por uma tendência de crescimento. Durante o período de análise, a eficiência técnica do modelo DEA-VRS apresentou níveis mais elevados quando comparado com o modelo DEA-CRS. Os resultados sugerem significativas melhoras dos níveis de eficiências técnicas da ALLMN-RMN. Em janeiro de 2006, os escores, para ambas as eficiências, eram 0,498, e em dezembro de 2011, passaram a ser 0,883 para a TE-CRS e 1,0 para TE-VRS. A posição mais observada pela DMU da ALLMN-RMN foi a terceira colocação, com 32 vezes (44% do total de períodos).

Na maior parte do período de análise, a ALLMN-RMN apresentou, predominantemente, retornos crescentes à escala (72% do total de períodos), o que se traduz em eficiência alocativa, ou seja, seus níveis de eficiência podem ser aumentados por meios de investimentos nos fatores de produção (insumos). A partir de 2009, a ALLMN-RMN passou a apresentar ineficiência alocativa, com o surgimento de retornos decrescentes à escala (14% do total de períodos), possivelmente, consequência do início da operação da FNSTN que passou a competir com a ALLMN-RMN no transporte de granéis agrícolas da região Centro-Oeste. Os resultados da DMU da ALLMN-RMN mostram baixos níveis de eficiência técnica, TE-CRS e TE-VRS, patamares semelhantes e uma tendência de melhora para as eficiências técnicas.

Os resultados também revelam que houve um aumento nos níveis das eficiências técnicas da ALLMO-RMO, comportando-se de forma semelhante à ALLMO-RMN. Em relação ao *ranking* de classificação, a posição mais observada pela DMU da ALLMO-RMO foi a décima colocação, com 28 vezes (39% do total de períodos). Sua pior classificação foi a 11ª posição, a qual ocupou por dez vezes (14% do total de períodos). Em agosto de 2008 e dezembro de 2009, ela obteve sua melhor classificação, a sexta posição. Observando os retornos à escala, pode-se perceber que a DMU da ALLMO-RMO não opera em escala ótima. Na maior parte do período de análise, a ALLMO-RMO apresentou, predominantemente, retornos crescentes à escala (94% do total de períodos), o que mostra eficiência alocativa dos fatores de produção (insumos).

Os resultados da DMU da ALLMP-RMP revelam os baixos níveis de eficiência técnica, TE-CRS e TE-VRS, além de patamares semelhantes e uma tendência de melhora para as eficiências técnicas. A posição mais ocupada pela DMU da ALLMP-RMP foi a sexta colocação, com 27 vezes (38% do total de períodos). Em novembro de 2011, ela ocupou a 11ª posição, sua pior classificação no período de análise. Em dezembro de 2010 e maio de 2011, ela obteve sua melhor classificação, quinta posição. Analisando os retornos à escala, pode-se perceber que a DMU da ALLMN-RMN não opera em escala ótima. Durante o período de análise, a ALLMP-RMP apresentou eficiência alocativa em 2006 e 2007; e ineficiência alocativa de 2008 a 2011. Os retornos crescentes à escala corresponderam a 64% do total de períodos e os retornos decrescentes à escala 35% do total de períodos. A ALLMS-RMS obteve baixos níveis de eficiência técnica, TE-CRS e TE-VRS, e um comportamento semelhante em que os escores das eficiências técnicas flutuam em torno do valor médio de 0,258. Os resultados também evidenciam estagnação dos níveis de eficiências técnicas da ALLMS-RMS. Em relação ao ranking de eficiência, a posição mais ocupada pela DMU da ALLMS-RMS foi a sétima colocação, 18 (25% do total de períodos). Em janeiro de 2007, ela ocupou a 11ª posição, sua pior classificação no período de análise, e, de agosto de 2011 a novembro de 2011, ela obteve sua melhor classificação, sexta posição. Já em relação à natureza dos retornos à escala, pode-se perceber que a DMU da ALLMS-RMS não opera em escala ótima. Durante o período de análise, a ALLMP-RMP apresentou eficiência alocativa de 2006 a 2008, e ineficiência alocativa de 2009 a 2011. Os retornos crescentes à escala corresponderam a 79% do total de períodos e os retornos decrescentes à escala, 21%.

A Estrada de Ferro Carajás (EFC) mostra que ela pode ser considerada benchmarking. Durante todo o período de análise ela operou em escala ótima, apresentando apenas retornos constantes à escala e escores 1,000 para TE-CRS e TE-VRS. Além disso, a EFC só não ocupou primeira colocação no ranking de eficiência em março de 2008. Já a Estrada de Ferro (EFPO) apresentou elevados níveis de eficiência técnica, TE-CRS e TE-VRS, e pequena flutuação no final do período de análise, mas apresentou estabilidade dos níveis de eficiência, ou seja, pouca variabilidade no indicador de eficiência. Em relação ao ranking de eficiência, a posição mais ocupada pela DMU da EFPO foi a primeira colocação, 44 vezes (61% do total de períodos). Em maio de 2011, ela ocupou a sétima posição, sua pior classificação no período de análise. Os retornos à escala para EFPO operaram em escala ótima, apresentando retornos constantes à escala. A partir de setembro de 2008, ela passou a apresentar eficiência alocativa com retornos crescentes à escala. Durante o período de análise, os retornos constantes à escala corresponderam a 74% do total de períodos e os retornos crescentes à escala, 26%.

A Estrada de Ferro Vitória Minas (EFVM) apresentou patamares distintos e elevados durante o período de análise e uma tendência de redução nos níveis de eficiência. No período de análise, os escores da TE-VRS apresentaram níveis mais elevados quando comparado aos escores da TE-CRS. O valor médio dos escores das variáveis TE-VRS e TE-CRS foram, respectivamente, 0,841 e 0,711. A posição mais ocupada da EFVM foi a quarta colocação, 35 vezes (63% do total de períodos). Em março de 2009, ela obteve sua pior classificação, sexta posição, e ocupou a primeira localização em seis períodos distintos. Com relação à natureza dos retornos à escala, pode-se perceber que a DMU da EFVM não está operando em escala ótima. Na maior parte do período de análise, a EFVM apresentou eficiência alocativa, obtendo retornos crescentes à escala em 83% do total de períodos. Os retornos decrescentes à escala, que correspondem a 15% do total de períodos, estão concentrados entre 2006 e 2008, sendo que houve apenas uma ocorrência depois de 2009.

A Ferrovia Centro-Atlântica (FCA) pode constatar os baixos níveis de eficiência técnica, TE-CRS e TE-VRS, além de patamares semelhantes e uma tendência de aumento dos escores das eficiências técnicas. Os resultados revelam uma elevação dos níveis de eficiências técnicas da FCA. Em janeiro de 2006, os escores eram 0,175 para TE-CRS e 0,176 para TE-VRS; e, em dezembro de 2011, passaram a ser 0,307 para a TE-CRS e 0,310 para TE-VRS. A posição mais ocupada pela

FCA foi a oitava, com 27 vezes (28% do total de períodos). Em junho de 2006 e janeiro de 2008, ela ocupou a 11ª posição, sua pior classificação no período de análise. A melhor posição ocupada pela FCA foi a sexta colocação, por seis vezes. Analisando os retornos de escala, pode-se perceber que a FCA não opera em escala ótima. Durante o período de análise, a FCA apresentou eficiência alocativa nos anos de 2006 até meados de 2009, e ineficiência alocativa a partir desta última data. Os retornos crescentes à escala corresponderam a 77% do total de períodos e os retornos decrescentes à escala, 25%.

Analisando os resultados da DMU da FTC pode-se constatar os baixos níveis de eficiência técnica, TE-CRS e TE-VRS, e um comportamento semelhante em que os escores das eficiências técnicas flutuam em torno do valor médio de 0,288. Durante o período de análise, os escores médios da TE-CRS e TE-VRS foram, respectivamente, 0,283 e 0,293. Os resultados também evidenciam uma estagnação dos níveis de eficiências técnicas da FTC. Em janeiro de 2006, os escores eram 0,247 para TE-CRS e TE-VRS; e, em dezembro de 2011, passaram a ser 0,299 para a TE-CRS e TE-VRS. A FTC foi a sexta colocação, com dezoito vezes (31% do total de períodos). A pior colocação da FTC foi a décima posição, ocupada por nove vezes. Em fevereiro de 2008, ela obteve sua melhor classificação, terceira posição. Já em relação à natureza dos retornos à escala, pode-se perceber que a DMU da FTC está operando em escala ótima, uma vez que ela obteve apenas retornos constantes à escala.

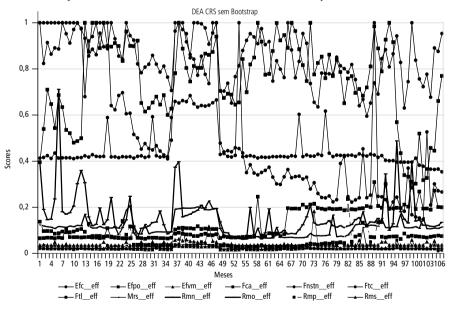
A MRS apresentou patamares elevados e distintos, no início e no fim do período de análise. Além disso, é evidente a estagnação dos níveis de eficiência. Durante o período de análise, os escores da TE-VRS apresentaram níveis mais elevados quando comparados aos escores da TE-CRS. O valor médio dos escores das variáveis TE-VRS e TE-CRS foram, respectivamente, 0,630 e 0,690. A MRS foi a quinta colocação, com vinte vezes (28% do total de períodos). Seus retornos à escala não se revelaram em escala ótima, mas, na maior parte do período de análise, a MRS apresentou eficiência alocativa obtendo retornos crescentes à escala em 85% do total de períodos. Os retornos decrescentes à escala, que correspondem a 13% do total de períodos, estão concentrados entre 2010 e 2011.

A FTL-FTNL apresentou baixos níveis de eficiência técnica, TE-CRS e TE-VRS, e um comportamento semelhante em que os escores das eficiências técnicas flutuam em torno do valor médio de 0,1618. Durante o período de análise, os escores médios da TE-CRS e TE-VRS foram, respectivamente, 0,158 e 0,166. Em julho de 2011, as eficiências técnicas obtiveram seus menores escores, 0,104 para TE-CRS e 0,105 para TE-CRS. Os resultados evidenciam uma estagnação dos níveis de eficiências técnicas da FTL. A posição mais ocupada pela FTL foi a 11ª, com 58 vezes (81% do total de períodos). A sua melhor colocação foi a sexta, ocupada em novembro de 2008. Já em relação à natureza dos retornos à escala,

pode-se perceber que a DMU-FTL está operando em escala ótima. Ela apresentou eficiência alocativa na maior parte do período de análise (88% do total de períodos) e retornos constantes à escala nos outros 12%.

Executou-se bootstrap com duzentas, quinhentas e mil replicações do modelo DEA com retornos constantes baseados nos trabalhos de Simar e Wilson (1998; 1999; 2000) e aplicados no pacote benchmarking do R Project indicado no livro do Bogetoft e Otto (2011). Os resultados com aumento do número de replicações não mudaram consideravelmente os resultados, portanto, apresenta-se no gráfico 2 apenas o bootstrap com duzentas replicações. São apresentadas as séries temporais com a posição calculada (ordem dos escores classificados do maior para o menor) da DMU em cada período de análise para estimativas de DEA CRS e com bootstrap (gráfico 2). Na sequência, apresenta-se a tabela 3 com os valores médios de diferença entre a estimativa com e sem bootstrap em termos percentuais e a estatística do intervalo de confiança.

GRÁFICO 2
Estimação de eficiência das concessionárias com *bootstrap*



Fonte: Macambira Filho (2014). Elaboração dos autores.

O teste sugerido por Simar e Wilson (2002; 2011) foi realizado. Nele obtiveram-se resultados do *bootstrap* quanto a retornos de escala com a hipótese nula de retornos constantes de escala frente à alternativa de retornos variáveis, concluindo por não rejeição da hipótese nula – portanto, o modelo de retornos

constantes de escala é aceita. Assim, o modelo CRS é o adequado e os resultados do *bootstrap*, embora resultem em valores menores do que o do modelo DEA CRS, replicam o comportamento encontrado no modelo. Valendo as discussões de resultado anterior.

Nas tabelas 3 e 4, são apresentadas as estimações do modelo Tobit, aplicado nos *inputs* utilizados no modelo DEA, considerando retornos constantes e variáveis de escala para as onze concessionárias ferroviárias selecionadas, para avaliar se estes *inputs* estão correlacionados com as eficiências técnicas resultantes da aplicação do modelo DEA. Embora o teste utilizando o *bootstrap* sugira retornos constantes de escala, decidiu-se fazer dois modelos Tobit para verificar se haveria divergência dos resultados ao adotar especificação diferente da sugerida no teste de Simar e Wilson (2002; 2011).

TABELA 3
Resultados da regressão Tobit para a variável dependente eficiência técnica considerando retornos constantes à escala

Coeficiente	Estimativa	Erro-padrão	Teste-z	p-valor
Constante	0,374435	0,010516	35,6	<0,0001
NTF	-2,929* 10 ⁻⁰⁵	3,97 _* 10 ⁻⁰⁶	-7,377	<0,0001
CCL	1,60* 10-05	4,58* 10 ⁻⁰⁶	3,492	0,0005
UL	-0,000641300	8,01* 10 ⁻⁰⁵	-8,011	<0,0001
UV	4,50E-06	2,02* 10 ⁻⁰⁶	2,226	0,026
NE	1,94* 10-05	6,62* 10 ⁻⁰⁶	2,935	0,0033
χ^2	705,0188			<0,0001

Elaboração dos autores.

TABELA 4
Resultados da regressão Tobit para a variável dependente eficiência técnica considerando retornos variáveis à escala

Coeficiente	Estimativa	Erro-padrão	Teste-z	p-valor
Constante	0,514	0,017	30,15	<0,0001
NTF	-9,73 * 10 ⁻⁰⁵	6,48 * 10 ⁻⁰⁶	-15,01	<0,0001
CCL	7,00 * 10 ⁻⁰⁸	4,16 * 10 ⁻⁰⁹	16,82	<0,0001
UL	-2,47 * 10 ⁻⁰⁶	4,07 * 10 ⁻⁰⁷	-6,055	<0,0001
UV	5,53 * 10 ⁻⁰⁷	1,82 * 10 ⁻⁰⁷	3,034	0,0024
NE	-5,10 * 10 ⁻⁰⁵	8,46 * 10 ⁻⁰⁶	-6,024	<0,0001
χ^2	612,380			<0,0001

Elaboração dos autores.

Observa-se que todos os *inputs* utilizados possuem correlação significativa com a variável dependente, TE-CRS, quando se considera a estatística z, que mede o nível de significância do parâmetro. Pode-se ainda concluir que os coeficientes NTF, CLL, UL e NF são estatisticamente significantes no nível de 1% de significância, e o coeficiente UV é estatisticamente significante no nível de 5% de significância. Esses resultados indicam que quase todos os coeficientes são estatisticamente significantes em um nível de 1% de significância. Os testes qui-quadrado, χ^2 , verificam se a frequência dos valores observados de TE-CRS se desvia significativamente, ou não, da frequência com que ele é esperado. São testadas as seguintes hipóteses: i) hipótese nula: as frequências observadas de TE-CRS são diferentes das esperadas, ou seja, as frequências observadas de TE-CRS não são diferentes das frequências esperadas, ou seja, as frequências observadas de TE-CRS não são diferentes das frequências esperadas, ou seja, as frequências observadas e esperadas de TE-CRS não são diferentes das frequências esperadas, ou seja, as frequências observadas e esperadas de TE-CRS não são diferentes das frequências esperadas, ou seja, as frequências observadas e esperadas de TE-CRS são dependentes.

De acordo com os valores p da estatística χ^2 , de ambos os modelos, a hipótese nula é rejeitada, indicando a existência de dependência entre as frequências observadas de TE-CRS e as frequências esperadas de TE-CRS. Observa-se ainda que os inputs CCL e UV têm uma relação direta com a variável TE-CRS. Isso significa dizer que qualquer aumento no valor destes inputs gera um aumento da eficiência técnica. Analisando o sentido da correlação do input UV, é razoável afirmar que a alocação do fator de produção vagão está diretamente relacionada à produção de transportes, ou seja, quanto mais vagão.dia é empregado na operação, mais carga é transportada. Em relação ao insumo CCL, pode-se afirmar que a quantidade de combustível consumido pelas locomotivas – utilizadas nas operações de transporte, serviço e manobra – está diretamente relacionada à produção de transporte, uma vez que, para a locomotiva desenvolver esforço de tração, é necessária a queima de diesel. Já os inputs NTF, UL e NF possuem uma relação inversa com a variável TE-VRS, ou seja, um aumento no valor dessas variáveis gera uma redução da eficiência técnica. Observado o sentido da correlação de NTF e UL, pode-se afirmar que uma ferrovia a qual opera com muitos trens, empregados e, consequentemente, mais locomotivas, tende a ser menos eficiente que ferrovias as quais operam com um menor número de trens formados. Essa afirmação pode ser reforçada quando analisamos as ferrovias que transportam essencialmente minério de ferro e granéis agrícolas (soja e milho). O transporte dessas commodities envolve um número menor de trens, porém com mais vagões, quando comparado com a movimentação de outros tipos de carga que envolva trens menores em maior quantidade. Como a consolidação dos trens de minério e granéis agrícolas é mais simples e o número de trens formados é menor, infere-se que o quadro de empregados nas ferrovias tende a ser menor. Em 2011, 98% do transporte ferroviário de minério de ferro foi realizado pelas concessionárias EFC, MRS e EFVM, sendo que o minério de ferro

representou, respectivamente, 95%, 84% e 90% da movimentação total de carga, em toneladas úteis, dessas ferrovias. Os escores médios das eficiências técnicas para as DMUs da EFC, da MRS e da EFVM, foram 0,780 para TE-CRS e 0,843 para TE-VRS. Em relação ao transporte de *commodities* agrícolas, a ALLMN-RMN e a EFPO transportaram 40% do total de soja e milho movimentado nas ferrovias brasileiras em 2011. Essas duas *commodities* agrícolas representam, respectivamente, 78% e 75% da movimentação total de carga, em toneladas úteis, dessas ferrovias. Os escores médios das eficiências técnicas TE-CRS e TE-VRS para as DMUs da ALLMN-RMN e da EFPO foram, respectivamente, 0,807 e 0,893.

Esses dois cenários mercadológicos apresentados são bastante distintos do cenário das ferrovias ALLMO-RMO, ALLMP-RMP, ALLMS-RMS, FCA, FTC e FTL. Os segmentos de mercados atendidos por essas concessionárias são bastante diversificados, não havendo a especialização no transporte de produtos de um segmento de mercado. Os escores médios das eficiências técnicas TE-CRS e TE-VRS para essas DMUs foram, respectivamente, 0,251 e 0,258.

Os resultados apresentados indicam que as ferrovias que atuam quase exclusivamente no transporte de minério de ferro, soja e milho, são relativamente mais eficientes do que ferrovias que atuam em mercados mais pulverizados. Essa diferença está associada, principalmente, à gestão dos trens e dos funcionários. O transporte de *commodities* abrange operações mais frequentes e com maior previsibilidade, além de envolver um menor número de trens formados, porém maiores, ou seja, com mais vagões. Por seu turno, o transporte de carga geral e conteinerizada envolve uma gestão operacional mais complexa e com mais empregados, na medida em que há mais atores envolvidos na consolidação da carga, formação dos trens e realização de manobras.

5 CONCLUSÕES

Este trabalho analisou a eficiência produtiva de onze concessionárias do sistema ferroviário brasileiro de carga no período de 2006 a 2011. Utilizou-se a análise envoltória de dados para calcular as medidas dessa eficiência, comumente conhecida como eficiência técnica, considerando retornos constantes e variáveis à escala. Após obtenção das eficiências, investigou-se em que medidas os parâmetros utilizados na produção de transporte influenciam na eficiência produtiva. Nesse contexto, as evidências empíricas dos modelos DEA mostram que as ferrovias que se especializaram no transporte de *commodities* minerais e agrícolas possuem níveis mais elevados de eficiência produtiva do que as ferrovias que atuam no transporte de carga geral ou conteinerizada e de granéis líquidos. As ferrovias EFC, MRS e EFVM, que atuam no transporte de minério de ferro, e as ferrovias EFPO e ALLMN-RMN, que atendem o setor de exportação de soja e milho, são mais eficientes, do ponto

de vista da produção de transportes, do que as ferrovias ALLMS-RMS, ALLMO-RMO, ALLMP-RMP, FTC, FCA e FTL, que operam com uma quantidade mais diversificada de produtos. Os modelos de regressão Tobit mostram que todas as variáveis utilizadas como *inputs* na modelagem DEA: i) número de trens formados; ii) consumo de combustível; iii) utilização de locomotiva; iv) utilização de vagão; e v) número de empregados, possuem correlação significante com as eficiências produtivas. Interpretando os sentidos das correlações, pode-se inferir que a gestão de pessoal e operacional são matérias significantes no desempenho das ferrovias de carga. A análise conjunta dos resultados dos modelos DEA e das regressões Tobit revela dois cenários distintos para o sistema ferroviário brasileiro de carga. No primeiro, temos as ferrovias especialistas com patamares elevados de eficiência produtiva e que atendem exclusivamente aos setores de exportação de minério de ferro e de soja e milho em grão. Essas ferrovias operam com um número menor de trens e funcionários em relação à TKU produzida e, consequentemente, com trens maiores, operação verticalizada, atendimento a uma pequena carteira de clientes e transporte de mercadorias com escala e de fácil consolidação.

A literatura dos modelos DEA aplicada a transportes ferroviários é resumida em Azambuja (2002) e corrobora o encontrado neste artigo. Silva *et al.* (2010) também apresentam resultados no mesmo sentido. Entretanto, havia uma lacuna quanto à aplicação de modelos Tobit para verificação da significância dos insumos na determinação da eficiência de concessionárias ferroviárias. Desta feita, este trabalho avança aplicando um modelo paramétrico, já empregado na literatura internacional, para verificar os resultados de um modelo DEA estatisticamente, bem como contribui cientificamente para os estudos nacionais, uma vez que aplica procedimentos de *bootstrap* para diminuir viés de estimação, como sugerido por Koster, Balk e Van Nus (2009) e Simar e Wilson (1998).

No segundo cenário, temos as ferrovias com baixos níveis de eficiência produtiva e que operam com um maior número de trens e de funcionários em relação à TKU produzida. Essas ferrovias atendem a vários nichos de mercado e, em consequência, possuem muitos clientes, o que dificulta a consolidação das cargas. Cabe ressaltar que o Brasil é um país preponderantemente exportador de matérias-primas e seus segmentos de exportação de *commodities* são favorecidos pelo direcionamento dos recursos para atendê-los. A exploração dessas *commodities* só é viável se o transporte, que é um dos principais componentes da cadeia de produção, for eficiente, dado o baixo valor agregado desses produtos. Assim, será oportuno avaliar em estudos futuros o desempenho da geração de receita proveniente da produção de transporte nas ferrovias, pois transportar um vagão de minério de ferro, soja ou milho pode ser bem menos lucrativo que transportar um contêiner.

Alguns autores sugerem que pode haver sazonalidade nos dados de transporte ferroviário para a estimação feita e que o método deanálise envoltória de dados seja aplicado separadamente para cada um dos períodos sazonais. Seria razoável fazer modelos por período sazonal, meses ou safra, contudo, tal exercício seria laborioso e não se sabe se assertivo. Optou-se por fazer uma análise com variáveis dummies para modelo Tobit e não se verificou grandes mudanças nos resultados do modelo sem as variáveis dummies, portanto, manteve-se o modelo sem correção de sazonalidade. De qualquer forma, sugere-se uma investigação de métodos de correção de sazonalidade para trabalhos semelhantes a este.

REFERÊNCIAS

AMEMIYA, T. Regression analysis when the dependent variable is truncated normal. Econometrica , v. 41, n. 6, p. 997-1016, 1973.
. Tobit models: a survey. Journal of Econometrics , v. 24, n. 1-2, p. 3-61, 1984.
. Advanced Econometrics. Oxford: Basil Blackwell, 1985.
ANTT – AGÊNCIA NACIONAL DE TRANSPORTES TERRESTRES. Resolução ANTT nº 3.694/2011. Aprova o Regulamento dos usuários dos serviços de transporte ferroviário de cargas. Diário Oficial da União , Brasília, 2011a. Disponível em: https://bit.ly/20SrUo6 .
. Resolução ANTT nº 3.695/2011. Aprova o Regulamento das operações de direito de passagem e tráfego mútuo, visando à integração do sistema ferroviário nacional. Diário Oficial da União , Brasília, 20 jul. 2011b. Disponível em: https://bit.ly/33rnuU8 .
. Sistema de Acompanhamento de Fretes Ferroviários (Saff). Brasília: ANTT, 2012.
. Concessões ferroviárias, 2013a. Disponível em: https://bit.ly/2rcE7Vp .
Manual do sistema de acompanhamento e fiscalização do transporte ferroviário. Brasília: ANTT, 2013b.
ASMILD, M. et al. Railway reforms: do they influence operating efficiency?

ASMILD, M. *et al.* Railway reforms: do they influence operating efficiency? **Discussion Papers**, University of Copenhagen, n. 8-5, 2008. 28p.

AZAMBUJA, A. M. V. Análise de eficiência na gestão do transporte urbano por ônibus em municípios brasileiros. 2002. Tese (Doutorado em Engenharia de Produção) – Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina, Florianópolis, 2002.

BANKER, R.; CHARNES, A.; COOPER, W. W. Some models for estimating technical and scale inefficiencies in data envelopment analysis. **Manage Sci.**, n. 30, p. 1078-1092, 1984.

BANKER, R. D.; THRALL, R. M. Estimation of returns to scale using data envelopment analysis. **Eur J OperRes**., n. 62, p. 74-84, 1994.

BHANOT, N.; SINGH, H. Benchmarking the performance indicators of Indian Railway container business using data envelopment analysis. **Benchmarking: An International Journal**, v. 21, n. 1, p. 101-120, 2014. Disponível em: https://bit.ly/36DYFXn.

BOGETOFT, P.; OTTO, L.; **Benchmarking with DEA, SFA, and R**. New York: Springer, 2011.

BRASIL. Decreto nº 1.832/1996. Aprova o regulamento dos transportes ferroviários. **Planalto.gov**, Brasília, 4 mar. 1996. Disponível em: https://bit.ly/33rLCWX.

BURI, M. R. *et al.* Transporte ferroviário de cargas no Brasil – aproveitamento da malha. *In*: SIMPÓSIO DE ENGENHARIA DE PRODUÇÃO, 13., 2006, Bauru, São Paulo. **Anais...** Bauru: Simpep, 2006.

CALDAS, M. A. F. *et al.* A eficiência do transporte ferroviário de cargas: uma análise do Brasil e dos Estados Unidos. *In*: CONGRESO LATINO-IBEROAMERICANO DE INVESTIGACIÓN OPERATIVA, 16., 2012, Rio de Janeiro; SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL, 44., 2012, Rio de Janeiro. **Anais...** Rio de Janeiro: CLAIO; SBPO, 2012.

CASE, K. E.; FAIR, R. C. **Principles of economics**. 5. ed. Upper Saddle River: Prentice Hall, 1999.

CASTRO, N. Os desafios da regulação do setor de transporte no Brasil. **Revista de Administração Pública**, Rio de Janeiro, v. 34, n. 5, set./nov., 2000. Mimeografado.

CHARNES, A.; COOPER, W. W.; RHODES, E. Measuring the efficiency of decision-making units. **European Journal of Operational Research**, v. 2, p. 429-444, 1978.

CHENG, Y.-H. High-speed rail in Taiwan: new experience and issues for future development. **Transport Policy**, v. 17, n. 2, p. 51-63, 2010.

CNT – CONFEDERAÇÃO NACIONAL DO TRANSPORTE. **O sistema ferroviário brasileiro**. Brasília: CNT, 2013. Disponível em: https://bit.ly/33qxgGn>.

COELLI, T.; PERELMAN, S. Technical efficiency of european railways: a distance function approach. **Applied Economics**, n. 32, p. 1967-1976, 2000.

COELLI, T.; PRASADA RAO, D. S.; BATTESE, G. E. An introduction to efficiency and productivity analysis. Boston; Dordrecht; London: Kluwer Academic Publishers, 1998.

DYSON, R. G. Performance measurement and data envelopment analysis – ranking are ranks! **OR Insight**, v. 13, n. 4, p. 3-8, 2001.

ERKOC, T. E. Estimation methodology of economic efficiency: stochastic frontier analysis vs data envelopment analysis. **International Journal of Academic Research in Economics and Management Sciences**, v. 1, n. 1, Jan. 2012. Disponível em: https://bit.ly/2PVd4Z6.

FRIEBEL, G.; IVALDI, M.; VIBES, C. Railway (de)regulation: a european efficiency comparison. **CEPR Discussion Paper**, n. 4319, London, 2005. 47 p.

KOSTER, M. B. M.; BALK, B. M.; VAN NUS, W. T. I. On using DEA for benchmarking container terminals. **International Journal of Operations & Production Management**, v. 29, p. 1140-1155, 2009.

KUTLAR, A.; ALI KABASAKAL, A.; SARIKAYA, M. Determination of the efficiency of the world railway companies by method of DEA and comparison of their efficiency by Tobit analysis. **Quality & Quantity Journal**, v. 47, n. 6, p. 3575-3602, 2013.

LAN, L. W.; LIN, E. T. J. Technical efficiency and service effectiveness for railways industry: DEA approaches. **Journal of the Eastern Asia Society for Transportation Studies**, v. 5, Oct. 2003.

LANG, A. **As ferrovias no Brasil e avaliação econômica de projetos**: uma aplicação em projetos ferroviários. 2007. Dissertação (Mestrado) – Universidade de Brasília, Brasília, 2007. Disponível em: https://bit.ly/32vJaNX.

LANSINK, A. O.; SILVA, E.; STEFANOU, S. Inter-firm and intra-firm efficiency measures. **Journal of Productivity Analysis**, v. 15, p. 185-199, 2001.

LI, L. B.; HU, J. L. Efficiency analysis of the regional railway in China: An application of DEA-Tobit approach. **Journal of Information and Optimization Sciences**, v. 31, n. 5, p. 1071-1085, 2010.

MACAMBIRA FILHO, J. K. D. Mensuração de eficiência produtiva das ferrovias brasileiras de carga: uma aplicação de modelos DEA e Tobit. Dissertação (Mestrado) — Departamento de Engenharia Civil e Ambiental, Universidade de Brasília, Brasília, 2014, 97p.

MALHOTRA, R.; MALHOTRA, D. K.; LERMACK, H. Using data envelopment analysis to analyze the performance of North American class I freight railroads; Applications of Management Science. *In*: LAWRENCE, K. D.; KLEINMAN, G.

(Eds.) **Financial modeling applications and data envelopment applications**. [s.l.]: Emerald Group Publishing Limited, 2009. v. 13. p. 113-131.

MERKERT, R.; SMITH, A.; NASH, C. The effects of institutional, environmental and transactional factors on train operating company performance? A Tobit regression approach based on DEA efficiency escores. *In*: EUROPEAN TRANSPORT CONFERENCE, 2009, Conference Centre Leeuwenhorst. **Anais**... Disponível em: https://bit.ly/2CsL6vH.

NAG, B. Comparison of efficiency of Indian railways with world railways and across time using data envelopment analysis technique. Calcutá: IIMC, 2013. Disponível em: https://bit.ly/2O2m4t5.

PAVLYUK, D. Efficiency analysis of european countries railways. *In*: INTERNATIONAL CONFERENCE: RELIABILITY AND STATISTICS INTRANSPORTATION AND COMMUNICATION, 8., Riga, 2008. **Anais**... Riga: International Conference, 2008.

PETROVIC, M. *et al.* A DEA based approach for cross-country evaluation of rail freight transport: possibilities and limitations. **Machines, Technologies, Materials Virtual Journal**, v. 22, n. 5, 2012.

SAMENI, M. K.; PRESTON, J. M. Value for rail capacity. Assessing efficiency of operators in Great Britain. **Transportation Research Record Journal of the Transportation Research Board**, n. 2289, p. 134-144, 2012.

SANTOS, M. S. Avaliação da eficiência produtiva das ferrovias de carga no Brasil: uma aplicação da metodologia DEA. 2011. Dissertação (Mestrado) — Programa de Pós-Graduação em Engenharia de Transportes, Universidade Federal do Ceará, Fortaleza, 2011.

SCHNEDLER, W. Likelihood estimation for censored random vectors. **Econometric Reviews**, v. 24, n. 2, p. 195-217, 2005.

SILVA, F. G. F. *et al.* Um estudo exploratório da eficiência produtiva de concessionárias de transporte ferroviário de carga no Brasil. Salvador: Anpet, 2010.

SIMAR, L.; WILSON, P. W. Sensitivity analysis of efficiency escores: how to bootstrap in nonparametric frontier models. **Management Science**, v. 44, n. 1, p. 49-61, 1998.

Estimating and bootstrapping Malmquist indices. European Journal of
Operational Research, v. 115, n. 3, p. 459-471, 1999.
A general methodology for bootstrapping in nonparametric frontier models.
Journal of Applied Statistics, v. 27, n. 6, p. 779-802, 2000.
Non-parametric tests of returns to scale Furguean Journal of Operational

_____. Non-parametric tests of returns to scale. **European Journal of Operational Research**, v. 139, n. 1, p. 115-132, 2002.

___. Inference by the m out of n bootstrap in nonparametric frontier models. Journal of Productivity Analysis, v. 36, p.33-53, 2011. SILVEIRA, M. R. A importância geoeconômica das estradas de ferro no Brasil. Tese (Doutorado) – Universidade Estadual Paulista (UNESP), Presidente Prudente, 2003. . Estradas de ferro no Brasil: das primeiras construções às parcerias públicoprivadas. Rio de Janeiro: Editora Interciência, 2007. 204p. TAVARES, M. P. Concessões ferroviárias de carga: proposta de monitoração dos ativos operacionais arrendados. 2003. Dissertação (Mestrado) - Instituto Militar de Engenharia, Rio de Janeiro, 2003. THANASSOULIS, E. Introduction to theory and application of data envelopment analysis. Norwell: Kluwer Academic Publishers, 2001. TOBIN, J. Estimation of relationships for limited dependent variables. Econometrica, v. 26, n. 1. p. 24-36, Jan. 1958. TSAI, C. H.; MULLEY, C.; MERKERT, R. Measuring the cost efficiency of urban rail systems: an international comparison using DEA and Tobit Models. Journal of Transport Economics and Policy, v. 49, n. 1, 2014. WETZEL, H. European railway deregulation: the influence of regulatory and environmental conditions on efficiency. Working Paper Series in Economics, n. 86, p. 36, 2008. Disponível em: . **BIBLIOGRAFIA COMPLEMENTAR** BRASIL. Lei nº 2.975/1956. Altera a legislação do imposto único sobre combustíveis e lubrificantes líquidos e gasosos, e dá outras providências. Planalto.gov, Rio de Janeiro, 27 nov. 1956. Disponível em: https://bit.ly/33qm5NV>. __. Lei nº 3.115/1957. Determina a transformação das empresas ferroviárias da União em sociedades por ações, autoriza a constituição da Rede Ferroviária S.A., e dá outras providências. **Planalto.gov**, Rio de Janeiro, 16 mar. 1957 Disponível em: <https://bit.ly/34CStg9>. ____. Decreto nº 58.992/1966. Dispõe sobre a implementação da política governamental de supressão de trechos ferroviários antieconômicos de que trata a Lei nº 4.452 de 1964, de 5 de novembro de 1964. **Jusbrasil.com.br**, 4 ago. 1966. Disponível em: https://bit.ly/36E2qvX. _. Decreto-Lei nº 89.396/1984. Autoriza a Rede Ferroviária Federal S.A. – RFFSA a, mantida a condição de subsidiária, mudar a denominação e objeto social

da Empresa de Engenharia Ferroviária S.A. – ENGEFER e dá outras providências. **Diário Oficial da União**, Brasília, 22 fev. 1984. Seção 1. Disponível em: https://

bit.ly/2NrnjTk>.

_____. Decreto-Lei nº 473/1992. Dispõe sobre inclusão, no Programa Nacional de Desestatização - PND, da Rede Ferroviária Federal S.A. – RFFSA, da AGEF – Rede Federal de Armazéns Gerais Ferroviários S.A. e da VALEC – Engenharia, Construções e Ferrovias S.A. **Planalto.gov**, Brasília, 10 mar. 1992. Disponível em: https://bit.ly/2JWUZ96.

_____. Decreto nº 8.134/2013. Estrutura a Valec - Engenharia, Construções e Ferrovias S.A. para a execução das atividades de desenvolvimento dos sistemas de transporte ferroviário e aprova o Estatuto Social da empresa. **Planalto.gov**, Brasília, 28 out. 2013. Disponível em: https://bit.ly/2qz1SqE.

BRASILEIRO, A. **Transportes no Brasil**: história e reflexões. Brasília; Recife: GEIPOT; Editora UFPE, 2001.

CALDEIRA, J. Mauá: empresário do império. São Paulo: Companhia das Letras, 1995.

CHARNES, A.; COOPER, W. W. Programming with linear fractional functionals. **Naval Research Logistics Quartely**, n. 9, p. 181-185, 1962. CRUZ, I. **Gestão ambiental da operação do transporte ferroviário de carga**. 2004. Dissertação (Mestrado) – Instituto Militar de Engenharia, Rio de Janeiro, 2004. Disponível em: https://bit.ly/32q08Np.

DEBREU, G. The coefficient of resource utilization. **Econometrica**, v. 19, p. 273-292, 1951.

FARRELL, M. J. The measurement of productive efficiency. **Journal of the Royal Statistical Society**, n. 120, p. 253-281, 1957.

FONTELLES, C. B.; FROES, M. F. V. A privatização da atividade ferroviária no Brasil. *In*: CONGRESO PANAMERICANO DE FERROCARRILES, 20., 2000, Havana. **Anais**... Havana: Congreso Panamericano de Ferrocarriles, 2000.

HICKS, J. R. A suggestion for simplifying the theory of money. **Economica**, n. 2, p. 1-19, 1935.

LEIBENSTEIN, H. Allocative efficiency vc "x-efficiency". **The American Economic Review**, v. 56, n. 3, p. 392-415, Jun. 1966.

MARCHETTI, D.; WANKE, P. Brazil's rail freight transport: Efficiency analysis using two-stage DEA and cluster-driven public policies. **Socio-Economic Planning Sciences**, v. 59, p. 26-42, 2016. Disponível em: https://bit.ly/2oW5Ust.

MCCULLOUGH, G. J. US railroad efficiency: a brief economic overview. **Transportation Research Board**, 2007 Disponível em: https://bit.ly/36HlwRz.

MELO, A. L. O. Classificação e identificação de trechos ferroviários concedidos passíveis de devolução ao poder concedente: uma proposta metodológica. 2012. Monografia (Especialização) – Centro Interdisciplinar de Estudos em Transportes, Universidade de Brasília, Brasília, 2012.

MERKERT, R.; SMITH, A. S. J.; NASH, C. A. Efficiency analysis of the regional railway in China: an application of DEA-Tobit approach. **Transportation Planning and Technology**, v. 33, n. 1, p. 35-53, 2010.

MICHAELIS. Dicionário da Língua Portuguesa, 2012.

NASH, C. A.; TONER, J. P. **Railways**: structure, regulation and competition policy. [s.l.]: OECD, 1997. Disponível em: https://bit.ly/34CbJKQ.

PAULA, D. A. **As ferrovias no Brasil**: análise do processo de erradicação de ramais. Paraná: Unioeste, 2000. Disponível em: https://bit.ly/2NN5tt3>.

SHOPPA, R. F. **150** anos do trem no Brasil. Rio de Janeiro: Votorantim, 2011. 196p.

SOUSA, R. A.; PRATES, H. F. O processo de desestatização da RFFSA: principais aspectos e primeiros resultados. **Revista do BNDES**, Rio de Janeiro, v. 4, n. 8, p. 119-142, dez. 1997. Disponível em: https://bit.ly/2pPU97p.

TSEKERIS, T. Greek airports: efficiency measurement and analysis of determinants. **Journal of Air Transport Management**, v. 17, p. 140-142, 2011.

VARIAN, H. R. **Microeconomic analysis**. 3. ed. New York; London: W.W. Norton, 1992.

WANKE, P.; BARROS, C. P. Slacks determinants in Brazilian railways: a distance friction minimization approach with fixed factors. **Applied Economics**, v. 47, p. 5103-5120, 2015.

ZHU, J. Quantitative models for performance evaluation and benchmarking: data envelopment analysis with spreadsheets and DEA excel solver. Boston: Kluwer, 2003a.

_____. Quantitative models for performance evaluation and benchmarking: data envelopment analysis with spreadsheets. 2. ed. Boston: Springer Science, 2003b.

Ipea – Instituto de Pesquisa Econômica Aplicada

EDITORIAL

Coordenação

Reginaldo da Silva Domingos

Assistente de Coordenação

Rafael Augusto Ferreira Cardoso

Supervisão

Camilla de Miranda Mariath Gomes Everson da Silva Moura

Revisão

Alice Souza Lopes
Ana Clara Escórcio Xavier
Clícia Silveira Rodrigues
Idalina Barbara de Castro
Luiz Gustavo Campos de Araújo Souza
Olavo Mesquita de Carvalho
Regina Marta de Aguiar
Amanda Ramos Marques (estagiária)
Ana Luíza Araújo Aguiar (estagiária)
Hellen Pereira de Oliveira Fonseca (estagiária)
Ingrid Verena Sampaio Cerqueira Sodré (estagiária)
Isabella Silva Queiroz da Cunha (estagiária)
Lauane Campos Souza (estagiária)

Editoração

Aeromilson Trajano de Mesquita Bernar José Vieira Cristiano Ferreira de Araújo Danilo Leite de Macedo Tavares Herllyson da Silva Souza Jeovah Herculano Szervinsk Junior Leonardo Hideki Higa

Capa

Luís Cláudio Cardoso da Silva

Projeto Gráfico

Renato Rodrigues Bueno

The manuscripts in languages other than Portuguese published herein have not been proofread.

Livraria Ipea

SBS – Quadra 1 – Bloco J – Ed. BNDES, Térreo 70076-900 – Brasília – DF Tel.: (61) 2026-5336 Correio eletrônico: livraria@ipea.gov.br

Missão do Ipea

Aprimorar as políticas públicas essenciais ao desenvolvimento brasileiro por meio da produção e disseminação de conhecimentos e da assessoria ao Estado nas suas decisões estratégicas.







