

PONTIFÍCIA UNIVERSIDADE CATÓLICA
DO RIO DE JANEIRO



Maurício Machado Fernandes

Ensaio em microeconomia aplicada

Tese de Doutorado

Tese apresentada como requisito parcial para
obtenção do título de Doutor pelo Programa de Pós-
Graduação em Economia da PUC-Rio.

Orientador: Prof. Rodrigo Reis Soares

Co-orientador: Prof. Claudio Abramovay Ferraz do Amaral

Rio de Janeiro
Setembro de 2013



Maurício Machado Fernandes

Ensaio em microeconomia aplicada

Tese apresentada como requisito parcial para obtenção do título de Doutor pelo Programa de Pós-Graduação em Economia da PUC-Rio. Aprovada pela Comissão Examinadora abaixo assinada.

Prof. Rodrigo Reis Soares

Orientador
Departamento de Economia - PUC-Rio

Prof. Claudio Abramovay Ferraz do Amaral

Co-orientador
Departamento de Economia - PUC-Rio

Prof. Gustavo Maurício Gonzaga

Departamento de Economia - PUC-Rio

Prof. Juliano Junqueira Assunção

Departamento de Economia - PUC-Rio

Profa. Elaine Toldo Pazello

FEA-RP/USP

Prof. Miguel Nathan Foguel

IPEA

Profa. Monica Herz

Coordenador(a) Setorial do Centro de Ciências Sociais - PUC-Rio

Rio de Janeiro, 03 de setembro de 2013

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

Maurício Machado Fernandes

Graduou-se em Engenharia Agrônômica na ESALQ/USP (Escola Superior de Agricultura “Luiz de Queiroz” – Universidade de São Paulo), Mestre em Economia Aplicada pela FEA-RP/USP (Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto – Universidade de São Paulo). É aluno de pós-doutorado do Programa de Pós Graduação em Economia da FEA-RP/USP.

Ficha Catalográfica

Fernandes, Maurício Machado

Ensaio em microeconomia aplicada / Maurício Machado Fernandes; orientador: Rodrigo Reis Soares; co-orientador: Claudio Abramovay Ferraz do Amaral – Rio de Janeiro: PUC-Rio, Departamento de Economia, 2013.

196 f. : il. ; 30 cm

Tese (doutorado) – Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Administração.

Inclui referências bibliográficas.

1. Economia - Teses. 2. Discriminação entre gêneros. 3. Qualidade do professor. 4. Identidade Partidária I. Soares, Rodrigo Reis. II. Amaral, Claudio Abramovay Ferraz do III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Administração. IV. Título.

CDD:330

Agradecimentos

Ao meu orientador Professor Rodrigo Reis Soares e meu co-orientador Claudio Ferraz pelos ensinamentos, estímulos e compreensão.

Ao IPEA e à Secretaria da Educação do Estado de São Paulo pela disponibilização do acesso às bases de dados fundamentais para a elaboração de dois dos três capítulos dessa tese.

Ao CNPq, à Capes, ao Departamento de Economia da PUC-Rio e à PUC-Rio pelos auxílios concedidos, sem os quais esta tese não poderia ter sido realizada.

Aos professores que participaram da banca examinadora pelas significativas sugestões e comentários.

Aos professores e funcionários do Departamento de Economia da PUC-Rio pelos ensinamentos e ajuda.

Aos meus amigos da PUC-Rio pelo companheirismo ao longo desses últimos anos e aos demais por entenderem minhas ausências e ainda assim não desistirem de nossa amizade.

À Elen Barreto pela paciência e carinho.

À minha família, em especial aos meus pais Maria Inez e Saint Clair, pelo apoio e afeto nos momentos mais difíceis.

Resumo

Fernandes, Maurício Machado; Soares, Rodrigo Reis (Orientador); Amaral, Claudio Abramovay Ferraz do (Co-orientador). **Ensaio em microeconomia aplicada**. Rio de Janeiro, 2013. 196p. Tese de Doutorado - Departamento de Economia, Pontifícia Universidade Católica do Rio de Janeiro.

Essa tese é composta por três artigos empíricos independentes. No primeiro capítulo é avaliado em que medida diferenças no histórico profissional entre os gêneros influenciam o diferencial de salários observado no mercado de trabalho formal brasileiro. Para isto, utiliza-se uma amostra aleatória e representativa de 1% dos trabalhadores presentes na RAIS / MTE entre os anos de 1994 e 2009. A partir dessas informações é reconstruída a trajetória profissional dos indivíduos pertencentes à amostra. As estratégias empíricas exploram a característica longitudinal dessa base de dados para gerar informações complementares acerca do diferencial de salários entre gêneros. Os resultados revelam que as medidas de histórico profissional têm impactos economicamente relevantes sobre os rendimentos individuais. Períodos de ausência no mercado de trabalho reduzem em média os rendimentos e um maior engajamento dos trabalhadores implica salários maiores. Entretanto, a inserção dessas medidas mais fidedignas de histórico profissional dos trabalhadores acarreta uma diminuição de no máximo 10% na magnitude do coeficiente associado ao diferencial de salários entre os gêneros, ou seja, um impacto bastante reduzido. O segundo capítulo investiga a importância relativa de duas dimensões da qualidade dos professores para a aprendizagem em matemática e língua portuguesa dos alunos da oitava série do ensino fundamental na rede de ensino paulista. Com este propósito, adota-se uma abordagem de função de produção educacional e a principal especificação utiliza um modelo de valor adicionado com controle para o desempenho passado dos estudantes. Os resultados mostram que tanto o conhecimento quanto as atividades pedagógicas dos professores em sala de aula têm impacto positivo e estatisticamente significativo sobre a aquisição de habilidades cognitivas. Entretanto, o efeito do conhecimento dos docentes apresenta uma magnitude pequena em termos econômicos. Já os impactos associados à adoção frequente de práticas pedagógicas eficazes tem magnitude bastante relevante. Por exemplo, a intervenção de substituir um professor de matemática que não passa lição de casa

sempre por outro que o faz aumenta a proficiência dos alunos em aproximadamente 12% de um desvio padrão da distribuição de notas. O terceiro capítulo analisa a relação entre identidade partidária e as escolhas políticas para o contexto das municipalidades brasileiras no ciclo político entre 2004 e 2008. Para isto, utiliza-se o arcabouço de regressão com descontinuidade para estimar o efeito causal local de um município ser governado por um partido de esquerda ao invés de um de direita sobre as políticas públicas. Os resultados apontam que governos de esquerda gastam proporcionalmente menos com urbanismo e saúde e mais com administração. No entanto, esses maiores gastos administrativos não estão associados a um inchaço da máquina pública com servidores.

Palavras-chave

Discriminação entre gêneros; qualidade do professor; identidade partidária.

Abstract

Fernandes, Maurício Machado; Soares, Rodrigo Reis (Advisor); Amaral, Claudio Abramovay Ferraz do (Co-advisor). **Essays in applied microeconomics**. Rio de Janeiro, 2013. 196p. PhD Thesis - Departamento de Economia, Pontifícia Universidade Católica do Rio de Janeiro.

This thesis is composed of three independent empirical articles. In the first chapter is evaluated to what extent differences in labor supply factors and careers by gender influence the wage gap observed in the Brazilian formal labor market. For this, we use a 1% representative random sample of the workers in RAIS / MTE between the years 1994 and 2009. From this information is retrieved the career path of individuals in the sample. The empirical strategies exploit the longitudinal feature of this database to generate complementary information about the gender wage gap. The results show that the labor market history measures have economically relevant impacts on individual incomes. Career interruptions reduce average earnings and workers with continuous labor market attachment have higher wages. However, the inclusion into the analysis of these more reliable job experience variables results in a reduction of up to 10% in the magnitude of the gender wage gap estimates. This represents a quite reduced influence. The second chapter investigates the relative importance of two dimensions of teacher quality for the learning in mathematics and Portuguese of eighth graders of the elementary school in São Paulo state. For this purpose, we adopt an approach based on the educational production function and the main specification uses a value added model with control for the students' past grades. The results show that both the teachers' knowledge and pedagogical activities inside the classroom have a positive and statistically significant impact on the acquisition of cognitive skills. However, the teachers' knowledge effect has a small economic magnitude. Yet the impacts associated with the frequent application of effective teaching practices are quite large. For instance, the intervention defined by the replacing a math teacher who does not always give homework for another that does it, increases the students' proficiency in approximately 12% of a standard deviation of the grades distribution. The third chapter examines the relationship between political partisanship and government size for the context of the Brazilian municipalities after 2004 local election. In order to achieve this, we use a

regression discontinuity research design to estimate the local causal effect on political choices of a municipality being governed by a left-wing party instead of a right-wing one. The results show that left-wing governments spend proportionately less on urbanism and health, and more on administration. Nevertheless, this higher administrative spending is not associated with an excessive hiring of public employees.

Keywords

Gender wage gap; teacher quality; political partisanship.

Sumário

1. Diferencial de salário por gênero: discriminação ou história profissional? Uma análise a partir dos dados da RAIS	18
1.1 Introdução	18
1.2 Dados	29
1.3 Estratégias empíricas	36
1.3.1 Medida de experiência profissional real	38
1.3.2 Comparação entre trabalhadores com elevado engajamento no mercado de trabalho	40
1.3.3 Decomposição do diferencial entre gêneros	42
1.4 Resultados	45
1.4.1 O diferencial de salário entre gêneros e a experiência profissional prévia	45
1.4.2 O diferencial de salários entre gêneros e sua evolução ao longo do ciclo de vida para trabalhadores com históricos profissionais similares e alto comprometimento no mercado de trabalho	57
1.4.3 Decomposição do diferencial de salários entre gêneros: componente explicado e não explicado	68
1.5 Conclusões	73
2 Medindo os efeitos do professor na sala de aula: evidências a partir da prova de promoção para professores em São Paulo	75
2.1 Introdução	75
2.2 Estratégica empírica	82
2.2.1 Aspectos que definem a qualidade do professor	83
2.2.2 O modelo de valor adicionado	86
2.2.3 Controle para capacidade de gestão do professor	88
2.3 O sistema de ensino público estadual paulista	90
2.3.1 O sistema de avaliação externa da SEE/SP (SARESP)	90

2.3.2 O Sistema de Promoção para os integrantes do Quadro do Magistério	92
2.4 Dados	94
2.4.1 As bases de dados	94
2.4.2 Amostras e estatísticas descritivas	98
2.4.3 Seleção e <i>sorting</i>	101
2.4.4 Comparação entre distribuições de desempenho	105
2.5 Resultados	111
2.5.1 O efeito do conhecimento dos professores	111
2.5.2 A decomposição da qualidade do professor: efeitos do conhecimento vs. das práticas pedagógicas	114
2.5.3 Condicionando para qualidade de gestão da escola	127
2.6 Conclusões	130
3 Ideologia e tamanho de governo no municípios brasileiros	133
3.1 Introdução	133
3.2 Estratégia de identificação	136
3.3 Dados	139
3.4 Contexto político brasileiro	142
3.5 Resultados	147
3.6 Conclusão	163
4 Referências Bibliográficas	165
Apêndices	170
Anexo	196

Siglas

BHPS – *British Household Panel Study*.
CBO – Classificação Brasileira de Ocupações.
CPS – *Current Population Survey*.
FINBRA – Finanças do Brasil.
FUNDEF – Fundo de Manutenção e Desenvolvimento do Ensino Fundamental e de Valorização do Magistério.
IBGE – Instituto Brasileiro de Geografia e Estatísticas.
INEP – Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira.
INPC – Índice Nacional de Preços ao Consumidor.
IPEA – Instituto de Pesquisa Econômica Aplicada.
IPTU – Imposto sobre a Propriedade Predial e Territorial Urbana.
ISS – Imposto sobre Serviços.
MBA – *Master of Business and Administration*.
MEC – Ministério da Educação.
MQO – Mínimos Quadrados Ordinários.
Munic – Pesquisa de Informações Básicas Municipais.
MTE – Ministério do Trabalho e Emprego.
PASEP – Programa de Formação do Patrimônio do Servidor Público.
PBF – Programa Bolsa Família.
PIS – Programa de Integração Social.
PNAD – Pesquisa Nacional por Amostra de Domicílios.
PSID – *Michigan Panel Study of Income Dynamics*.
RAIS – Relação Anual de Informações Sociais.
Saeb – Sistema de Avaliação da Educação Básica.
SARESP – Sistema de Avaliação do Rendimento Escolar do Estado de São Paulo.
SSE-SP – Secretaria da Educação do Estado de São Paulo.
STN – Secretaria do Tesouro Nacional.
SUS – Sistema Único de Saúde.
TSE – Tribunal Superior Eleitoral.

Lista de figuras

Figura 1.1 – Evolução do diferencial de salários entre gêneros para algumas coortes	59
Figura 1.2 – Evolução dos diferenciais de salários entre gêneros para as coortes 1972 a 1976, 1967 a 1971, 1962 a 1966 e 1957 a 1961	65
Figura 1.3 – Evolução do diferencial de salários para coortes jovens: 1974 a 1976, 1977 a 1979 e 1980 a 1982	67
Figura 2.1 – Distribuições de desempenho em matemática e língua portuguesa entre os grupos de alunos com professores identificados e não identificados a partir dos dados de atribuição de aulas da SEE/SP	103
Figura 2.2 – Diferenças na distribuição de desempenho em matemática e língua portuguesa entre os alunos associados a professores com alto e baixo desempenho na prova de promoção	107
Figura 2.3 – Diferenças na distribuição de desempenho em matemática entre os alunos associados a professores que sempre adotam uma das práticas pedagógicas investigadas e os demais estudantes	109
Figura 2.4 – Diferenças na distribuição de desempenho em língua portuguesa entre os alunos associados a professores que sempre adotam uma das práticas pedagógicas investigadas e os demais estudantes	110
Figura 3.1 – Histograma da população dos municípios brasileiros	143
Figura 3.2 – Efeito de partido de esquerda sobre tamanho de governo local	152
Figura 3.3 – Efeito de partido de esquerda sobre fontes de receitas tributárias, despesas com encargos sociais e investimentos	155
Figura 3.4 – Efeito de partido de esquerda sobre a composição de gastos públicos	161

Figura A1 – Distribuição das notas na parte objetiva da prova de promoção para os professores de matemática e língua portuguesa	192
Figura A2 – Relação entre a nota média em matemática e a nota na parte objetiva da prova de promoção dos professores	193
Figura A3 – Relação entre a nota média em língua portuguesa e a nota na parte objetiva da prova de promoção dos professores	194
Figura A4 – Efeito de partido de esquerda sobre o total e composição do quadro de funcionários públicos	195

Lista de tabelas

Tabela 1.1 – Estatística descritiva	35
Tabela 1.2 – Equação de determinação de salários para os trabalhadores entre 18 e 55 anos em 2009	47
Tabela 1.3 – Equação de determinação de salários para os trabalhadores entre 18 e 55 anos por faixa de escolaridade no ano de 2009	50
Tabela 1.4 – Equação de determinação de salários no ano de 2009 para seis cortes de trabalhadores: nascidos entre 1987 e 1991, 1982 e 1986, 1977 e 1981, 1972 e 1976, 1967 e 1971; e 1962 e 1966	54
Tabela 1.5 – Equação de determinação de salários com modelo de efeito-fixos para trabalhadores estimados separadamente para homens e mulheres: período 1996 a 2009	70
Tabela 1.6 – Decomposição do diferencial de salários entre homens e mulheres para modelo de painel com efeito fixo de trabalhadores: período 1996 a 2009	72
Tabela 2.1 – Distribuição da frequência de escolas quanto ao número de professores identificados na amostra principal	99
Tabela 2.2 – Estatísticas descritivas para a amostra principal: disciplinas de matemática e língua portuguesa	100
Tabela 2.3 – Teste para verificar a existência de <i>sorting</i> de professores às turmas com melhores alunos em média	105
Tabela 2.4 – Efeito do conhecimento do professor sobre o ganho de desempenho dos alunos no SARESP 2009: matemática e língua portuguesa	112
Tabela 2.5 – Efeito conjunto do conhecimento e das práticas pedagógicas do professor sobre o ganho de desempenho dos alunos em matemática	115
Tabela 2.6 – Efeito conjunto do conhecimento e práticas pedagógicas do professor sobre o ganho de desempenho dos alunos em língua portuguesa	120

Tabela 2.7 – Efeito conjunto do conhecimento e das práticas pedagógicas do professor sobre o ganho de desempenho dos alunos em matemática: inclusão de interações entre as notas e atividades dos docentes	123
Tabela 2.8 – Efeito conjunto do conhecimento e das práticas pedagógicas do professor sobre o ganho de desempenho dos alunos em língua portuguesa: inclusão de interações entre as notas e atividades dos docentes	125
Tabela 2.9 – Efeito conjunto do conhecimento e das práticas pedagógicas do professor – e interações – sobre o ganho de desempenho dos alunos em matemática: condicional a qualidade do diretor	128
Tabela 2.10 – Efeito conjunto do conhecimento e das práticas pedagógicas do professor – e interações – sobre o ganho de desempenho dos alunos em língua portuguesa: condicional a qualidade do diretor	130
Tabela 3.1 – Resultados eleições municipais 2004	140
Tabela 3.2 – Eleitorado governado por partido 2005-2008	141
Tabela 3.3 – Estatísticas descritivas dos municípios em geral e especificamente com eleições direita vs. esquerda	146
Tabela 3.4 – Efeito de governos de esquerda sobre receitas orçamentárias	148
Tabela 3.5 – Efeito de governos de esquerda sobre receitas tributárias	149
Tabela 3.6 – Efeito de governos de esquerda sobre despesas orçamentárias	150
Tabela 3.7 – Efeito governos de esquerda sobre despesas com pessoal e encargos sociais	151
Tabela 3.8 – Efeito de governos de esquerda sobre impostos/receitas tributárias	153
Tabela 3.9 – Efeito de governos de esquerda sobre taxas/receitas tributárias	154

Tabela 3.10 – Efeito de governos de esquerda sobre IPTU/receitas tributárias	156
Tabela 3.11 – Efeito de governos de esquerda sobre gastos com saúde/despesas orçamentárias	157
Tabela 3.12 – Efeito governos esquerda sobre gastos com educação/despesas orçamentárias	158
Tabela 3.13 – Efeito governos esquerda sobre gastos com urbanismo/despesas orçamentárias	159
Tabela 3.14 – Efeito governo esquerda sobre despesas com administração/despesas orçamentárias	160
Tabela 3.15 – Efeito governo esquerda sobre despesas com pessoal/despesas orçamentárias	162
Tabela 3.16 – Efeito governo esquerda sobre número funcionários públicos per capita	163
Tabela A1 – Comparação das distribuições de ganho de desempenho entre os SARESP 2007 e 2009: grupo de estudantes para os quais foi possível identificar os professores vs. os demais alunos	174
Tabela A2 – Equação de determinação de salários para os trabalhadores entre 18 e 55 anos por faixa de escolaridade com a inclusão de controles para a ocupação no ano 2009	175
Tabela A3 – Equação de determinação de salários para seis coortes de trabalhadores com inclusão de controles para ocupação no ano de 2009	176
Tabela A4 – Estimativas do diferencial de salários entre gêneros por ano para a coorte 1972 - 1976	177
Tabela A5 – Estimativas do diferencial de salários entre gêneros por ano para a coorte 1967 - 1971	178
Tabela A6 – Estimativas do diferencial de salários entre gêneros por ano para a coorte 1962 - 1966	179
Tabela A7 – Estimativas do diferencial de salários entre gêneros por ano para a coorte 1957 – 1961	180

Tabela A8 – Estimativas do diferencial de salários entre os gêneros para a amostra geral e coortes jovens: nascidos entre 1974-76, 1977-79 e 1980-82	181
Tabela A9 – Estimativas do diferencial de salários entre os gêneros para a amostra de trabalhadores continuamente empregados e coortes jovens: nascidos entre 1974-76, 1977-79 e 1980-82	182
Tabela A10 – Efeito governo esquerda sobre impostos reais per capita	183
Tabela A11 – Efeito governo esquerda sobre investimentos reais per capita	184
Tabela A12 – Efeito governo esquerda sobre superávit per capita	185
Tabela A13 – Efeito governo esquerda sobre imposto sobre serviços/receitas tributárias	186
Tabela A14 – Efeito governo esquerda despesas com transporte/despesas orçamentárias	187
Tabela A15 – Efeito governo esquerda despesas com assistência social/despesas orçamentárias	188
Tabela A16 – Efeito governo esquerda despesas diretas com pessoal /despesas orçamentárias	189
Tabela A17 – Efeito governo esquerda sobre número de funcionários estatutários per capita	190
Tabela A18 – Efeito governo esquerda sobre funcionários temporários/total funcionários	191

1

Diferencial de salário por gênero: discriminação ou história profissional? Uma análise a partir dos dados da RAIS.

1.1

Introdução

No modelo tradicional de mercado de trabalho sob competição perfeita, o salário recebido por um trabalhador deve ser igual à sua produtividade marginal. No entanto, situações nas quais os salários de trabalhadores são correlacionados com características não produtivas, tais como gênero ou raça, são comuns na literatura empírica (Cahuc & Zylberberg, 2004 e Altonji & Blank, 1999). Muitas vezes esse padrão de correlações é identificado como uma manifestação da existência de discriminação no mercado de trabalho.

Altonji & Blank (1999) definem a discriminação no mercado de trabalho como uma situação na qual trabalhadores igualmente produtivos são tratados de maneira desigual, por questões relacionadas a uma característica observável. Ou seja, o comportamento discriminatório existe quando os indivíduos são tratados de maneira diferenciada apenas por pertencerem a diferentes grupos demográficos, dado características produtivas equivalentes.

Entre as possíveis explicações para a discriminação contra as mulheres no mercado de trabalho, a teoria econômica consolidou duas doutrinas mais influentes. Sob o primeiro ponto de vista, os empregadores (ou os trabalhadores do grupo majoritário) têm uma desutilidade em contratar (ou interagir com) mulheres e devem ser compensados para trabalhar com estas. Em equilíbrio, isso leva a um menor salário para as mulheres. Esse arcabouço teórico, inicialmente desenvolvido por Becker (1957), denomina-se discriminação por preferências e o comportamento discriminatório requer a existência de alguma renda econômica pura para os empregadores.

Já uma segunda vertente da teoria, cujos artigos seminais são Phelps (1972) e Arrow (1973), adota o pressuposto que os empregadores têm informação imperfeita sobre as habilidades ou comportamento profissional dos indivíduos. As firmas têm um conjunto de informação sobre os trabalhadores que é apenas

correlacionado com a produtividade real desses e, por conseguinte a explica de maneira imperfeita. Por esse motivo, as firmas procuram utilizar informações complementares – ou crenças – sobre a qualidade média dos trabalhadores pertencentes a um determinado grupo demográfico, nesse caso as mulheres. Dessa forma, os empregadores têm incentivos a ‘discriminar estatisticamente’ trabalhadores com produtividade similar a partir de características facilmente observáveis. Esses indivíduos acabam apresentando trajetórias profissionais distintas devido à produtividade média (real ou imaginada) do grupo demográfico a que pertencem. De maneira complementar, essas distinções na carreira entre trabalhadores de diferentes grupos demográficos podem se tornar persistentes ao longo do tempo. Isso ocorre quando as crenças dos contratantes influenciam as decisões dos trabalhadores sobre investimentos em capital humano e escolha de profissões, anteriores à entrada no mercado de trabalho e/ou o comportamento e comprometimento destes nos períodos subsequentes (Cahuc & Zylberberg, 2004). No caso de gênero, um motivo natural para a discriminação estatística seria a perspectiva de saída do mercado de trabalho associado à gravidez e aspectos relacionados à família.

Inicialmente, os estudos empíricos sobre diferenças salariais entre os gêneros utilizavam uma equação minceriana e apresentavam estimativas associadas a uma variável *dummy* para o sexo como indicador de discriminação. Para o contexto dos EUA no ano de 1995, Altonji & Blank (1999) estimam o diferencial bruto contra as mulheres em aproximadamente 27,9% e a adição de controles para educação e experiência tem apenas um pequeno efeito sobre a estimativa desse diferencial, passando para 27,2%. No entanto, ao se controlar para ocupação e indústria observa-se uma redução mais significativa na magnitude do diferencial, que passa a ser de 22,1%. Giuberti & Menezes-Filho (2005) comparam o diferencial de rendimentos entre gêneros para o Brasil e os EUA e incluem controles para escolaridade, idade (*proxy* para experiência), variáveis indicadoras de região e de jornada de trabalho. Os resultados encontrados para os EUA são muito semelhantes aos obtidos por Altonji & Blank (1999) – ambos os trabalhos utilizam a base de dados da CPS. Já as evidências para o Brasil, obtidas a partir da PNAD, mostram um diferencial desfavorável às mulheres de 39,6%, 33,9% e 17,1%, respectivamente, para os anos de 1981, 1988 e 1996, com os mesmos controles listados acima.

Os estudos empíricos sobre os diferenciais salariais para o Brasil vem de longa data e em sua maioria utilizam a metodologia tradicionalmente adotada na literatura internacional. Entre os pioneiros destacam-se Langoni (1973) e Camargo e Serrano (1983), esse último tendo como particularidade o uso dos dados da RAIS para o ano de 1976. Na década de 1990 surgiram novos estudos, entre os quais se destacam Cavalieri & Fernandes (1998) e Kassouf (1998). Em particular Cavalieri & Fernandes (1998) obtém estimativas do diferencial de salários contrário às mulheres – 59% aproximadamente – para regiões metropolitanas em 1989, que destoam das obtidas por Giuberti & Menezes (2005) para o mesmo período. Leme & Wajnman (2000) analisam a evolução do diferencial de rendimentos entre os anos de 1977 e 1997 e observam uma forte redução de 70% para 25% no período. Scorzafave & Pazello (2007) corroboram esse padrão de redução do diferencial entre gêneros no período mais recente. Em 1988, os trabalhadores do sexo masculino recebiam 47,5% a mais do que as mulheres e essa diferença se reduziu para 21,6% em 2004. Scorzafave & Pazello (2007) também realizam uma decomposição do diferencial observado entre gêneros e concluem que o principal fator para a diminuição do diferencial ao longo do tempo foi o declínio da contribuição dos retornos às características¹. Sobretudo, na diferença entre as estimativas associadas às constantes das regressões de equações mincerianas de homens e mulheres, isto é, redução mais forte nos componentes usualmente relacionados à discriminação. Por último, Madalozzo (2010) também observa uma redução significativa do componente não explicado do diferencial de rendimentos entre gêneros para o Brasil entre 1978 e 2007. No início do período as mulheres recebiam 33% a menos do que os homens, diferença essa que se reduziu a 23% em 1988, 19% em 1998 e 16% no ano de 2007. Dessa forma, Madalozzo (2010) ratifica a tendência previamente observada nos estudos anteriores de importante redução na magnitude da discriminação percebida contra as mulheres no mercado de trabalho brasileiro ao longo das últimas décadas.

¹ Scorzafave e Pazello (2007) utilizam um método de decomposição de diferenciais proposto por Yun (2005), semelhante à metodologia de Oaxaca-Blinder, que permite solucionar o problema de indeterminação relacionado a variáveis explicativas indicadoras. No entanto, as frações totais do diferencial associadas ao componente explicado e às estimativas dos retornos às características produtivas não se alteram com a definição do grupo base escolhido para a variável indicadora. Para maiores detalhes vide Scorzafave e Pazello (2007).

Esse conjunto de evidências fundamenta a ideia de que existe um importante diferencial de salários entre gêneros. Entretanto, isso não é suficiente para garantir a existência de discriminação. De acordo com Altonji & Blank (1999), esses diferenciais negativos enfrentados pelas mulheres podem ser oriundos, sobretudo, de duas fontes. A primeira refere-se a uma diferença não explicada entre os retornos de mercado para as características produtivas entre os sexos, usualmente interpretada como discriminação. A outra possibilidade é a existência de variáveis relevantes para a determinação da produtividade no mercado de trabalho que são omitidas nas especificações adotadas por esses estudos mais tradicionais e que, ao mesmo tempo, são correlacionadas com o gênero.

Devido a essa questão de variáveis omitidas e/ou erro de medida em variáveis determinantes da produtividade individual é importante ressaltar que, ao longo do ciclo de vida, as mulheres passam por um período crítico no qual devem tomar decisões relativas à constituição de suas famílias, fecundidade e demandas diferenciadas entre os cônjuges acerca dos cuidados com os filhos. Muitas vezes essas decisões fazem com que as mulheres precisem se ausentar um maior número de vezes do mercado de trabalho e por períodos mais longos, ou reduzam suas jornadas de trabalho e até mesmo optem por empregos/carreiras menos hostis à família.

Devido a essas escolhas, as mulheres tendem em média a acumular uma menor experiência no mercado de trabalho condicional a sua coorte e, por conseguinte, potencialmente acabam por realizar um menor nível de investimento em capital humano específico à firma ou ao mercado (Blau & Kahn, 2011). Esses fatores podem, a priori, ser responsáveis por um diferencial de salários motivado por características produtivas individuais associadas em média ao sexo dos trabalhadores. Todavia, em geral, essas informações não estão presentes nos bancos de dados utilizados pelos estudos empíricos que procuram investigar o diferencial de rendimentos entre gêneros. Assim, as medidas de experiência no mercado de trabalho tradicionalmente utilizadas (experiência potencial, definida como a idade dos indivíduos menos o número de anos de escolaridade menos seis anos) não são fidedignas da história profissional das mulheres em especial (Blau & Kahn, 2011).

A incapacidade de levar em consideração a história profissional efetiva dos indivíduos na análise empírica levanta preocupações acerca da adequação das

estimativas que são geralmente associadas à discriminação das mulheres no mercado de trabalho.

Recentemente, uma nova linha de pesquisa tem examinado mais atentamente a influência da história profissional dos indivíduos sobre o diferencial de rendimentos entre gêneros e a evolução desse diferencial ao longo da carreira profissional. Para isto, a maioria desses estudos baseia sua análise em informações provenientes de pesquisas biográficas retrospectivas para reconstruir o histórico dos indivíduos no mercado de trabalho e obter medidas mais fidedignas da experiência profissional e interrupções na carreira.

Alguns trabalhos concentram a análise em indivíduos provenientes das melhores faculdades (*colleges* ou MBAs) dos EUA. Goldin & Katz (2008) obtêm informações sobre vários momentos do ciclo de vida de indivíduos de doze turmas de Harvard/Radcliffe (registros administrativos, casamento, filhos, emprego, ocupações, períodos de não participação ou desemprego). Apesar desse amplo conjunto informações, Goldin & Katz (2008) constatam que persiste um substancial diferencial de salários entre gêneros. Por exemplo, na especificação básica o diferencial contra as mulheres é de 0,359 pontos de log e após a inclusão de controle para ocupação este se reduz para 0,301 pontos de log. Esses resultados são similares para as três diferentes coortes analisadas de egressos de Harvard/Radcliffe.

Já Bertrand, Goldin & Katz (2010) estudam a evolução do diferencial de rendimentos e da oferta de trabalho para jovens profissionais egressos do curso de MBA da *Booth School of Business* da Universidade de Chicago. Para essa amostra, os rendimentos laborais são quase idênticos entre homens e mulheres no início da carreira. Mas, com o passar do tempo, o padrão salarial diverge e cinco anos após a conclusão do MBA, o diferencial se aproxima de 0,30 pontos de log. Entre 10 e 16 anos, a desigualdade salarial é de quase 0,60 pontos de log. Contudo, Bertrand, Goldin & Katz (2010) identificam três determinantes que são capazes de explicar 84% do diferencial de salários observado. Estes são diferenças nas notas e cursos realizados, nas jornadas de trabalho e no número e duração das interrupções na carreira.

Tanto Goldin & Katz (2008) quanto Bertrand, Goldin & Katz (2010) observam que a interrupção na oferta de trabalho acarreta uma importante penalização sobre rendimentos. Complementarmente, a presença de crianças é o

principal fator para a maior descontinuidade na carreira, menor experiência e jornadas de trabalho mais curtas para as mulheres.

A partir de um contexto mais geral para os E.U.A., Blau & Kahn (2011) demonstram a importância de se possuir medidas reais de experiência para investigar o diferencial de salários entre gêneros. Para isto, os autores utilizam base de dados longitudinal da *Michigan Panel Study of Income Dynamics* (PSID). Os resultados apontam que o componente não explicado do diferencial de salários entre gêneros é menor ao se utilizar a experiência real ao invés da potencial. Essa redução foi de aproximadamente 7,7 pontos de log em 1980 e 2,8 pontos de log em 1999. Dessa forma, Blau & Kahn (2011) concluem que uma análise *ceteris paribus* do diferencial entre gêneros utilizando a experiência real é claramente superior ao uso da medida de experiência potencial. Ademais, a inclusão do histórico profissional real aumentou de forma considerável o poder explicativo das equações de salários para as mulheres, mesmo ao se condicionar a análise à experiência potencial e a duração do vínculo empregatício atual.

Nordman & Roubaud (2009) utilizam metodologia e informações semelhantes no contexto de um país subdesenvolvido (Madagascar). Os resultados também mostram que uma parcela significativamente maior do diferencial de salários entre gêneros pode ser explicada pelas características produtivas ao se utilizar a medida real de experiência. De maneira complementar, Nordman & Roubaud (2009) também constata um aumento (redução) no retorno à experiência para os (as) homens (mulheres) quando se utiliza a medida real de experiência.

Uma alternativa nessa linha de estudos é a utilização de bases de dados longitudinais sobre os trabalhadores para recuperar suas histórias profissionais. Essa é a estratégia utilizada por Manning & Swaffield (2008), a partir da *British Household Panel Study* (BHPS), para os doze anos entre 1991 e 2002 no Reino Unido. De acordo com os autores, os diferenciais de salários entre gêneros na entrada no mercado de trabalho são aproximadamente nulos, entretanto após dez anos existe um diferencial de quase 0,25 pontos de log. Essa diferença atinge o seu máximo cerca de vinte anos após o início da carreira profissional e se reduz marginalmente nos anos posteriores. Aproximadamente metade desse diferencial pode ser explicada por uma acumulação distinta de capital humano entre os sexos. Entre as dimensões de capital humanas investigadas por Manning & Swaffield

(2008) está a acumulação de experiência real, derivada de informações sobre as interrupções na oferta de trabalho dos indivíduos. As estimativas indicam que cada ano fora do mercado de trabalho reduz os rendimentos em 4,4% para homens e 4,7% para mulheres. Todavia, após dez anos da entrada no mercado de trabalho persiste sem explicação uma considerável parcela do diferencial entre gêneros. Assim, conforme destacado pelos autores, uma mulher que sempre trabalhou em período integral, não teve filhos e afirma o desejo de não tê-los no futuro ainda recebe em média 8% a menos do que um homem com características semelhantes.

Este artigo procura contribuir para com essa literatura e avalia em que medida o diferencial entre gêneros mensurado no mercado de trabalho formal brasileiro pode ser atribuído a diferentes histórias profissionais de homens e mulheres. Para isto, nós utilizamos uma amostra aleatória e representativa de 1% dos trabalhadores presentes na base de dados da ‘Relação Anual de Informações Sociais’ – RAIS – do Ministério do Trabalho e Emprego – MTE, entre 1994 e 2009. As informações prestadas pelas firmas que compõem esse banco de dados têm caráter censitário e representam o mercado de trabalho formal no Brasil. Dadas as características particulares dessa base de dados, nós podemos identificar os trabalhadores e recuperar as informações sobre seus respectivos vínculos empregatícios ao longo de todos os anos utilizados. Isto significa que a RAIS/MTE nos permite reconstruir o histórico profissional dos trabalhadores – no mercado de trabalho formal – ao longo dos dezesseis anos disponíveis, obtendo assim uma medida mais fidedigna da sua história profissional (ou períodos de afastamento).

Assim, a nossa principal contribuição é contornar um dos principais problemas presentes na maioria dos estudos empíricos sobre discriminação entre gêneros. Ou seja, a questão de erro de medida não clássico na variável de experiência profissional potencial que culmina com viés nas estimativas dos coeficientes de interesse (Reagan & Oaxaca, 2009). Para isto, nós exploramos a característica longitudinal da nossa base de dados (RAIS/MTE) e recuperamos a história profissional dos indivíduos no mercado de trabalho formal a partir de registros realizados em cada um dos dezesseis anos que constituem nosso intervalo de análise. Portanto, ao contrário de outros artigos que investigam o diferencial de salários entre gêneros, tais como Goldin & Katz (2008), Nordman & Roubaud (2009) e Bertrand, Goldin & Katz (2010), nosso estudo não depende

de pesquisas biográficas e/ou retrospectivas para ter acesso às informações sobre as carreiras profissionais dos trabalhadores. Apesar de Blau & Kahn (2011) obterem evidências de que as informações sobre experiência profissional obtidas de maneira retrospectiva se adequam bem aos dados derivados de pesquisas anuais dentro do painel da PSID. Os mesmos autores reconhecem que as informações retrospectivas de curto prazo (anual) parecem ser um pouco mais precisas do que as recordações de um prazo mais longo. E que existe um ganho de precisão sobre as informações de experiência no mercado de trabalho ao se utilizar informações anuais coletadas em cada ano em comparação com os dados derivados de questões retrospectivas realizadas em um único ponto no tempo. Dessa forma, dado a ampla dimensão temporal que utilizamos para reconstruir a história profissional dos trabalhadores, o uso de informações retrospectivas poderia influenciar a qualidade dos dados, sobretudo devido a problemas de memória por parte dos respondentes.

Deste modo, nesse trabalho, nós investigamos detalhadamente o papel que o histórico profissional dos trabalhadores tem sobre o diferencial de salários observado entre gêneros. O objetivo é verificar o quanto da discriminação percebida contra as mulheres decorre de decisões acerca da participação no mercado e interrupções na oferta de trabalho tomadas ao longo do ciclo de vida produtivo.

Com este propósito, nós empregamos três estratégias empíricas que têm em comum a utilização do aspecto longitudinal da nossa base de dados – RAIS/MTE. Cada uma dessas estratégias apresenta vantagens e desvantagens em relação às demais e por isso os resultados obtidos a partir de cada uma podem ser vistos como complementares. Inicialmente, restringimos nossa análise às informações sobre desempenho no mercado de trabalho (salários) para a *cross-section* no último período observado – ano de 2009. Assim, para o conjunto de trabalhadores com vínculo empregatício ativo em 2009, nós podemos reconstruir o histórico profissional no mercado de trabalho formal brasileiro para os quinze anos anteriores – entre 1994 e 2008. Esse aspecto permite o uso de medidas mais realistas sobre a experiência profissional como controle nas equações de determinação de salários. Deste modo, nós procuramos minorar as preocupações sobre a possibilidade de viés nas estimativas do diferencial de salários entre gêneros decorrentes de erros de medida na variável de experiência.

Em seguida, nós comparamos os salários ao longo do tempo para uma subpopulação de homens e mulheres bastante semelhantes em relação a seu comprometimento com o mercado de trabalho formal. Esses trabalhadores são denominados ‘continuamente empregados’ e apresentaram vínculo empregatício ativo para todos os anos que compõem nosso intervalo de observação. Com essa estratégia, nós reduzimos a influência de distintas dinâmicas de emprego entre os sexos sobre as estimativas obtidas para a estatística de interesse. Esse benefício informacional se dá apesar do custo de restringirmos os resultados a uma medida local do diferencial de rendimentos entre gêneros, tal como realçado por Machado (2011)².

Por último, nós decompos a diferença média de salários entre homens e mulheres, usando o procedimento de Oaxaca-Blinder, nos efeitos composição (explicado) e não explicado, e verificamos como esses componentes se alteram com a inclusão das medidas de histórico profissional dos indivíduos. Por estarmos interessados na influência das medidas de experiência profissional sobre o componente associado à discriminação contra as mulheres, nós utilizamos uma amostra de trabalhadores jovens e com experiência nula ou reduzida no mercado de trabalho formal no início do painel. De maneira complementar, nós incorporamos efeitos-fixos individuais para controlar para as características intrínsecas dos indivíduos constantes no tempo. Essa estratégia, por utilizar efeitos-fixos, nos permite levar em consideração a seleção de indivíduos no mercado de trabalho ao longo do ciclo da vida.

Os nossos resultados mostram que as medidas de histórico profissional, construídas a partir dos dados da RAIS/MTE, têm impactos estatisticamente significantes e economicamente relevantes sobre os rendimentos individuais. Adicionalmente, os efeitos dessas medidas de experiência têm os sinais esperados, ou seja, períodos de ausência no mercado de trabalho reduzem em média os rendimentos e um maior engajamento dos trabalhadores implica salários maiores. Além disso, a influência da história profissional dos indivíduos sobre os salários se reduz, em termos absolutos, com o passar do tempo. Essas evidências corroboram que as variáveis de histórico profissional usadas neste estudo expressam de maneira fidedigna e realista a experiência dos indivíduos no

² O termo ‘continuamente empregado’ é uma tradução livre a partir da expressão *always employed* utilizada por Machado (2011).

mercado de trabalho formal brasileiro. Por exemplo, no intervalo entre 2004 e 2008, cada ano ausente do mercado de trabalho implica uma redução de aproximadamente 3% nos salários médios. Já a permanência durante todo um ano no mesmo emprego entre 2004 e 2008 aumenta em 5,5% os rendimentos individuais.

Todavia, a inclusão dessas medidas de histórico profissional dos trabalhadores tem uma pequena influência sobre o diferencial estimado entre gêneros. Esse impacto é aproximadamente igual a uma redução de 10% na estimativa do coeficiente associado ao respectivo diferencial de salários, em termos absolutos. Outro aspecto interessante, e que segue no sentido contrário ao observado por estudos em outros países, refere-se ao aumento no diferencial após a adoção de controles para as variáveis de escolaridade dos indivíduos. Para o caso do mercado formal brasileiro, a diferença média de salários entre os sexos sofre um incremento de um terço, em valores absolutos, após a inclusão das variáveis indicadoras para as faixas de escolaridade – passando de aproximadamente 21,1% para 30%. Ou seja, ao restringirmos a comparação entre indivíduos com níveis de escolaridade semelhantes, a discriminação mensurada contra as mulheres é ainda maior que a diferença média de salários para a amostra como um todo.

Sob um ponto de vista mais descritivo, nós observamos uma clara tendência ascendente em termos absolutos do diferencial de salários entre gêneros na fase inicial do ciclo de vida produtivo. O padrão do diferencial é monotonicamente crescente até aproximadamente os 35 - 40 anos de idade. Nas idades posteriores, se constata uma estabilização desse diferencial entre gêneros e um suave decaimento no final do ciclo de vida. Essas evidências sugerem a existência de um claro efeito idade no aumento do diferencial de rendimentos entre os sexos na fase inicial da carreira profissional.

Esse padrão de evolução do diferencial também está presente para a amostra de trabalhadores ‘continuamente empregados’. O nível de penalização, em termos de rendimentos, sofrido pelas mulheres desse grupo é apenas marginalmente inferior em relação à população como um todo. Isto sugere que, apesar de um maior comprometimento com o mercado de trabalho, as mulheres ‘continuamente empregadas’ não derivam benefícios relevantes a partir desse comportamento em relação à média das mulheres no mercado de trabalho formal. Um fator que

possivelmente ajudaria a explicar as diferenças de desempenho observadas entre os sexos, para esse grupo de trabalhadores em particular, é o menor acúmulo médio de capital humano das mulheres ao longo do ciclo de vida produtivo. Esse aspecto se deve, sobretudo pelo menor comprometimento de tempo das mulheres, apesar de permanecerem ativas no mercado de trabalho (Blau & Kahn, 2011). No entanto, parece pouco provável que esse único fator seja responsável pela grande persistência do diferencial de salários contrário às mulheres ‘continuamente empregadas’ constatada para o Brasil.

Já o exercício de decomposição do diferencial de rendimentos entre gêneros revela que as mulheres deveriam receber um salário 2% superior aos dos homens, caso o retorno às características produtivas fossem iguais entre os sexos. Esse diferencial a favor das mulheres decorre principalmente de uma maior escolaridade. No entanto, a diferença entre os retornos às características produtivas é favorável aos homens e supera o efeito composição. Na especificação sem a inclusão de medidas do histórico profissional, as mulheres recebem 32% menos do que homens com a mesma idade e essa variável é a principal fonte de discriminação contra as mulheres. Entretanto, depois de adicionarmos as variáveis relativas à história no mercado de trabalho, a contribuição da idade para o componente não explicado se reduz bastante. Os resultados mostram que mulheres com o mesmo número de interrupções na oferta de trabalho e experiência similar recebem, respectivamente, 1,2% e 3,7% a menos do que os homens. A contribuição dessas duas variáveis ao componente não explicado (discriminação) do diferencial mais do que anula o ganho médio de rendimentos que as mulheres obtêm devido às diferenças nas características produtivas.

Um aspecto que deve ficar claro em nosso estudo é que todos os resultados obtidos devem ser analisados sob a ótica do mercado de trabalho formal brasileiro. Isso, pois a base de dados que nós utilizamos nos exercícios empíricos – RAIS/MTE – refere-se apenas ao setor formalizado da economia, ou seja, exclui os trabalhadores que estão na informalidade, desempregados ou fora da força de trabalho. Sobretudo no Brasil a questão da importância relativa do setor informal se faz presente. De maneira geral, a informalidade no Brasil se manteve estável em um período muito próximo ao investigado por este estudo (1992 a 2005). Ramos & Ferreira (2006) observam que a informalidade era igual a 51,9% em 1992 e apresentou crescimento discreto até o final da década de 1990s, com

posterior redução nos anos 2000s, sobretudo após 2002, atingindo um patamar de aproximadamente 50% em 2005. Apesar de padrões distintos de comportamento ocorre uma convergência das taxas de informalidade entre as regiões metropolitanas e não metropolitanas.

Dessa forma, dada a magnitude da relevância do setor informal para a economia brasileira, as nossas estimativas acerca do diferencial de salários entre gêneros podem ser uma medida bastante equivocada da discriminação percebida contra as mulheres na economia como um todo. Especialmente caso o processo de seleção dos indivíduos entre os setores formal e informal da economia sejam distintos entre os sexos, ou evoluam de maneira distinta ao longo do tempo. Por exemplo, se supormos que a informalidade passa a ser a principal opção de colocação no mercado de trabalho para os homens menos qualificados, enquanto uma importante fração das mulheres opta por empregos por conta própria e/ou mais amigáveis do ponto de vista da dedicação à família. E, adicionalmente, existe um *matching* de habilidades positivo no mercado de casamentos. Então, os nossos resultados para o diferencial entre gêneros no setor formal estariam superestimando o verdadeiro valor dessa estatística para a economia como um todo. Essa ilustração reforça a importância de restringirmos os nossos resultados para o mercado de trabalho formal apenas.

Além dessa seção introdutória, o capítulo está organizado da seguinte forma. A seção 1.2 descreve a base de dados utilizada – RAIS/MTE. As estratégias empíricas adotadas são discutidas na seção 1.3. Os resultados são apresentados e interpretados na seção 1.4. Por último, a seção 1.5 conclui este capítulo.

1.2

Dados

O trabalho faz uso de uma base de dados longitudinal derivada da ‘Relação Anual de Informações Sociais’ (RAIS), pertencente ao Ministério do Trabalho e Emprego (MTE). Nesta seção nós apresentamos a RAIS/MTE, suas principais características e limitações. Em seguida, descrevemos os procedimentos adotados para definir as amostras e as principais variáveis utilizadas na análise empírica que segue. Por último, nós apresentamos estatísticas descritivas das principais

variáveis para cada uma das amostras utilizadas nos três exercícios empíricos realizados neste trabalho.

Nossos dados são oriundos da RAIS/MTE. Esses dados têm caráter administrativo e contêm informações sobre todos os vínculos empregatícios que uma firma teve com seus respectivos empregados para cada ano³. Tais informações devem ser prestadas por todas as firmas devidamente registradas ao governo federal brasileiro⁴. Dessa forma, a RAIS/MTE tem um caráter censitário a respeito do mercado de trabalho formal no Brasil.

A unidade de observação é o vínculo empregatício definido como cada combinação entre uma firma e um trabalhador para cada ano; e cada um desses elementos – a firma e o trabalhador – têm códigos que os identificam de maneira única. Complementarmente aos códigos identificadores, para cada empregado estão disponíveis informações sobre sexo, idade, escolaridade definida por faixas, ocupação, rendimento mensal médio e jornada semanal de trabalho contratual. Para as empresas, nós temos informações sobre a principal atividade econômica (classificação de 25 subsetores do IBGE), natureza jurídica, município e, por conseguinte, o respectivo Estado da Federação. Por último, temos também informações específicas ao vínculo empregatício, tais como: datas de admissão e desligamento, tipos de admissão e desligamento e status em 31 de dezembro de cada ano (ativo ou não).

Assim como na maioria das bases de dados administrativas referentes a emprego no mundo, a RAIS/MTE não tem informações sobre as famílias dos trabalhadores. Isto é, não há relato sobre o estado civil dos indivíduos; sobre o

³ Estes são: empregados contratados sob o regime da CLT, por prazo indeterminado ou determinado; trabalhadores avulsos (Lei nº 8.630, de 25/02/1993); trabalhadores temporários (Lei nº 6.019, de 3/1/1974); trabalhadores com Contrato de Trabalho por Prazo Determinado (Lei nº 9.601, de 21/1/1998 ou Lei nº 8.745, de 9/12/1993, com a redação dada pela Lei nº 9.849, de 26/10/1999 ou Lei Estadual ou Lei Municipal); trabalhadores regidos pelo Estatuto do Trabalhador Rural (Lei nº 5.889, de 8/6/1973); aprendiz (contratado nos termos do art. 428 da CLT, Decreto nº 5.598, de 1/12/2005); servidores da administração pública direta ou indireta; servidores públicos não-efetivos; servidores e trabalhadores licenciados; servidores públicos cedidos e requisitados; empregados de cartórios extrajudiciais; dirigentes sindicais; e diretores sem vínculo empregatício (desde que opte pelo recolhimento de FGTS).

⁴ Entre os entes obrigados a entregar a declaração da RAIS tem-se: todos os empregadores, conforme definidos na CLT; os estabelecimentos inscritos no CNPJ com ou sem empregados (estes últimos são obrigados a entregar a RAIS Negativa); todas as pessoas jurídicas de direito privado; empresas individuais, inclusive as que não possuem empregados; empregadores urbanos e rurais pessoas físicas, que mantiveram empregados no ano-base; órgãos da administração direta e indireta dos governos federal, estadual ou municipal; condomínios e sociedades civis; cartórios extrajudiciais e consórcios de empresas; e quaisquer entidades vinculadas à pessoa jurídica domiciliada no exterior.

cônjuge – quando for o caso; sobre fecundidade ou sobre a distribuição de responsabilidades relativas aos cuidados com as crianças etc. Outra questão padrão nas bases de dados administrativas é a ausência de medidas de proficiência em testes para os trabalhadores. Essa informação possibilitaria a adoção do desempenho nos testes como uma variável *proxy* para a habilidade intrínseca dos indivíduos ou o conhecimento acumulado condicional a escolaridade. Também não estão disponíveis informações sobre raça ou cor de pele dos empregados.

Nós tivemos acesso à base de dados completa da RAIS/MTE para todos os anos no intervalo entre 1994 e 2009. Ao longo desse período de tempo, existem mais de 80 milhões de trabalhadores presentes em pelo menos um ano e no máximo em dezesseis anos no mercado de trabalho formal brasileiro. Devido a problemas de dimensionalidade da base de dados e viabilidade na implantação da pesquisa empírica, nós optamos por utilizar uma amostra aleatória de aproximadamente 1% do total de trabalhadores presentes no banco de dados original da RAIS entre 1994 e 2009. Essa amostra inicial tinha 1.364.646 trabalhadores. Posteriormente à definição da amostra de trabalhadores, nós utilizamos os seus respectivos códigos identificadores (números PIS/PASEP) para recuperar as informações sobre todas as observações (vínculos empregatícios) relativas aos empregados que compõem a nossa amostra inicial. Para que cada vínculo empregatício fosse considerado e incluído na amostra inicial, este deveria permanecer ativo por um período mínimo de quatro meses em cada ano, ou seja, todos os vínculos empregatícios com duração inferior a quatro meses não foram incorporados na amostra inicial. Esse procedimento permitiu reconstruir o histórico profissional dos trabalhadores da nossa amostra no mercado de trabalho formal brasileiro e resultou em um banco de dados em formato de painel não balanceado.

A partir dessa amostra inicial, nós realizamos alguns procedimentos com o intuito de tornar as informações contidas na base de dados mais homogêneas e próximas do que julgamos ideal para a estimação das equações de determinação de salários que seguem. O nosso interesse recai especialmente sobre o vínculo empregatício principal de cada trabalhador em cada ano, definido como a relação entre trabalhador e firma com o maior valor para o produto entre o número de semanas nas quais a relação de trabalho permaneceu ativa e o número de horas de trabalho (contratuais) por semana. Os demais vínculos empregatícios secundários

e simultâneos de cada trabalhador em um mesmo ano foram eliminados quando existentes. Adicionalmente, nós também excluimos todas as relações de emprego com jornadas de trabalho declaradas inferiores a cinco horas ou superiores a sessenta horas semanais, assim como os vínculos empregatícios com rendimentos médios não positivos. Ainda, devido a questões relacionadas a diferentes dinâmicas do mercado de trabalho, nós optamos por eliminar as observações referentes aos trabalhadores empregados na administração pública direta e autárquica ou agricultura, floresta e criação de animais. Esse último procedimento foi adotado para todos os anos utilizados na estimação das equações de determinação de salários. Portanto, os resultados obtidos, a partir das análises que seguem, devem ser restritos ao contexto dos trabalhadores do setor privado não agrícola.

As informações que constituem a RAIS/MTE são declaradas pelas firmas e, por conseguinte existe a possibilidade de inconsistências, sobretudo em relação aos dados dos trabalhadores. Dada à característica longitudinal da base de dados, as informações referentes aos empregados foram testadas e os algoritmos utilizados nesses testes são descritos no apêndice 1. Sempre que possível realizamos a correção das informações com inconsistências e para os casos restantes as informações foram excluídas. Após esse conjunto de testes e correções, nós restringimos a amostra aos indivíduos com idade entre dezoito e sessenta anos.

Após essas restrições a amostra básica permaneceu com 930.151 indivíduos, dos quais 561.515 (60,37%) eram do sexo masculino e 368.636 (39,63%) do feminino. Do total de dezesseis anos observados no painel, os homens estavam presentes em média em 6,48 anos (com desvio-padrão de 4,67 anos) e as mulheres 5,68 anos em média (com desvio-padrão de 4,30 anos). Para os homens, aproximadamente 5,84% estavam presentes em todos os anos observados – trabalhadores definidos como ‘continuamente empregados’ no mercado de trabalho formal brasileiro – enquanto apenas 3,39% das mulheres na amostra estavam presentes em todo o período.

A partir dessa amostra básica, nós elaboramos três amostras finais de trabalhadores do mercado de trabalho formal brasileiro. Cada uma dessas amostras foi criada com o intuito de melhor se ajustar às estratégias empíricas – apresentadas na próxima seção – que nós adotamos neste estudo para investigar o

papel da história profissional sobre o diferencial de salários entre gêneros. A primeira amostra (final) é composta pelos trabalhadores com vínculo empregatício ativo na *cross-section* referente ao ano de 2009. E para cada indivíduo nós reconstruímos seu histórico no mercado de trabalho formal entre 1994 e 2008. A amostra final dois refere-se aos trabalhadores ‘continuamente empregados’, isto é, o conjunto de indivíduos que permaneceram com vínculo empregatício ativo nos dezesseis anos entre 1994 e 2009. Já a amostra três é constituída de trabalhadores jovens que tinham vínculo empregatício ativo em 1996, mas não possuíam experiência profissional no setor formal da economia nos dois anos anteriores.

Em todos os exercícios empíricos deste trabalho a variável dependente utilizada é o logaritmo natural do salário mensal real médio para cada ano. Originalmente, a RAIS/MTE disponibiliza dados sobre rendas laborais como a média mensal do salário recebido (não o valor contratual) em cada ano em termos do correspondente número de salários mínimos da época. Posteriormente, essa informação foi multiplicada pelo valor monetário do salário mínimo real para cada ano, deflacionado pelo ‘Índice Nacional de Preços ao Consumidor’ (INPC/IBGE), com o período base sendo janeiro de 2012. Em seguida, o valor real do salário médio de cada ano é transformado com o auxílio da função logaritmo natural. Na análise empírica que segue a nossa principal variável explicativa de interesse é o gênero (ou sexo) do trabalhador.

Além disso, usamos algumas variáveis de controle nas equações de salários. Um primeiro conjunto de covariadas é a idade dos trabalhadores e variáveis dicotômicas derivadas dessa informação que identificam a coorte à qual cada indivíduo pertence. A educação dos trabalhadores é definida por variáveis indicadoras que representam o intervalo de número de anos de escolaridade. Essa informação está disponível na RAIS/MTE nos seguintes intervalos: analfabeto e até a quarta série incompleta do ensino fundamental; quarta série completa; quinta a oitava série incompleta do ensino fundamental; ensino fundamental completo; ensino médio incompleto; ensino médio completo; ensino superior incompleto e ensino superior completo ou níveis de educação superiores (mestrado ou doutorado). Diferentes dinâmicas e particularidades dos mercados de trabalho locais são levadas em consideração pela inclusão de variáveis indicadoras para os Estados da Federação. Outro conjunto de covariadas está associado à informação sobre a jornada de trabalho contratual de cada vínculo empregatício. Nesse caso,

nós controlamos nas equações de salário para o número de horas de trabalho mensal informado pelas firmas. Além disso, também condicionamos a análise ao fato do trabalhador manter vínculo empregatício em tempo parcial, ou seja, jornada semanal inferior a trinta horas.

Por último, temos as variáveis referentes à história profissional dos indivíduos no mercado de trabalho formal brasileiro. Essas têm diferentes definições de acordo com a respectiva amostra final e a estratégia empírica correspondente. Na primeira amostra, as variáveis referentes ao histórico profissional são: número de anos ausente do mercado formal separadamente para os intervalos entre 2004-2008, 1999-2003 e 1994-1998; número de anos nos quais o indivíduo ficou empregado o ‘ano inteiro’, ou seja, teve vínculo empregatício ativo por nove meses ou mais em um único ano, também nos períodos de 2004-2008, 1999-2003 e 1994-1998; e número de anos com atuação no subsetor de administração pública direta ou autárquica entre 1994 e 2008. A amostra dois utiliza a variável denominada de ‘ano inteiro’ para identificar o maior grau de empenho para com o mercado de trabalho. Já na terceira amostra, o histórico no mercado de trabalho é captado por uma variável contínua relativa ao número de anos de experiência entre 1996 e 2009, além de um vetor de três variáveis indicadoras que tomam valor igual a um quando o trabalhador apresenta uma, duas e três ou mais interrupções na oferta de trabalho entre 1996-2009, e zero em caso contrário.

A tabela 1.1 apresenta as estatísticas descritivas para cada uma dessas três amostras finais. O painel A refere-se à primeira amostra, o painel B lista as informações para as amostras usadas no segundo exercício empírico deste trabalho e o painel C exibe a média e o desvio-padrão das variáveis para a amostra três. Em cada um desses painéis, as duas primeiras colunas referem-se ao total dos indivíduos na amostra. Já a terceira e quarta colunas são relativas aos trabalhadores do sexo masculino e as duas últimas colunas mostram as estatísticas para as mulheres. As informações presentes nessa primeira tabela se referem a média e desvio-padrão da remuneração mensal média real dos trabalhadores, idade, escolaridade e para as variáveis mais importantes associadas ao histórico profissional dos indivíduos no mercado de trabalho formal brasileiro.

Tabela 1.1 – Estatísticas descritivas

Painel A: Amostra 1						
Variáveis	Cross section de trabalhadores entre 18 e 55 anos em 2009					
	Total		Homens		Mulheres	
	Média	Desvio-padrão	Média	Desvio-padrão	Média	Desvio-padrão
Salário médio real	1.675,05	2.540,08	1.840,37	2.807,83	1.414,36	2.020,65
'Ano inteiro' em 2009	0,4670	0,4989	0,4664	0,4989	0,4679	0,4990
Número anos experiência	8,98	5,34	9,43	5,35	8,27	5,26
Número anos ausente	1,18	2,24	1,21	2,23	1,15	2,24
Número interrupções 2004-2008	0,49	1,06	0,48	1,05	0,50	1,08
Número interrupções 1999-2003	0,50	1,15	0,52	1,16	0,47	1,14
Jornada semanal	42,7	4,2	43,1	3,4	42,1	5,1
Tempo parcial	0,0195	0,1381	0,0110	0,1045	0,0327	0,1779
Idade	32,0	9,3	32,3	9,4	31,5	9,0
Ensino por faixas de escolaridade:						
Superior Completo	0,1278	0,3339	0,0990	0,2987	0,1732	0,3784
Médio completo ou superior incompleto	0,5629	0,4960	0,5297	0,4991	0,6153	0,4865
Médio incompleto	0,2917	0,4545	0,3473	0,4761	0,2040	0,4029
Analfabeto ou básico incompleto	0,0176	0,1314	0,0240	0,1530	0,0075	0,0861
Número trabalhadores	511.130		312.777		198.353	
Fração por sexo			0,6119		0,3881	

Painel B: Amostra 2						
Variáveis	Trabalhadores em geral entre 1994 e 2009					
	Total		Homens		Mulheres	
	Média	Desvio-padrão	Média	Desvio-padrão	Média	Desvio-padrão
Salário médio real	1.671,30	2.507,30	1.844,25	2.761,53	1.370,56	1.953,17
Número anos experiência	6,2	4,5	6,5	4,7	5,7	4,3
Ano inteiro	0,5720	0,4948	0,5709	0,4949	0,5739	0,4945
Jornada semanal	42,5	4,6	42,9	3,8	41,7	5,7
Tempo parcial	0,0242	0,1537	0,0135	0,1155	0,0428	0,2025
Número observações	5.731.861		3.639.066		2.092.795	
Número trabalhadores	930.151		561.515		368.636	
Fração trabalhadores por sexo			0,6037		0,3963	

Variáveis	Trabalhadores 'continuamente empregados entre 1994 e 2009					
	Total		Homens		Mulheres	
	Média	Desvio-padrão	Média	Desvio-padrão	Média	Desvio-padrão
Salário médio real	2.774,56	3.294,97	2.942,09	3.466,37	2.334,76	2.747,23
Ano inteiro	0,8127	0,3901	0,8070	0,3947	0,8279	0,3775
Jornada semanal	42,2	4,2	42,6	3,5	41,3	5,4
Tempo parcial	0,0132	0,1140	0,0069	0,0831	0,0295	0,1691
Número observações	725.104		525.088		200.016	
Número trabalhadores	45.319		32.818		12.501	
Fração trabalhadores por sexo			0,7242		0,2758	

Painel C: Amostra 3						
Variáveis	Trabalhadores jovens (a anos) em 1996 e sem experiência prévia no mercado de trabalho					
	Total		Homens		Mulheres	
	Média	Desvio-padrão	Média	Desvio-padrão	Média	Desvio-padrão
Salário médio real	1.327,93	1.806,38	1.385,11	1.890,11	1.233,51	1.654,58
Ano inteiro	0,55	0,50	0,55	0,50	0,55	0,50
Jornada semanal	42,8	4,0	43,2	3,2	42,2	5,0
Tempo parcial	0,0178	0,1324	0,0089	0,0942	0,0325	0,1774
Informações ao final do painel:						
Exposição mercado trabalho	10,9	4,5	11,2	4,4	10,6	4,6
Número anos experiência	8,8	4,4	9,0	4,4	8,3	4,3
Número interrupções	0,80	0,85	0,82	0,86	0,79	0,83
Idade	30,4	4,7	30,6	4,7	30,0	4,8
Ensino por faixas de escolaridade:						
Superior Completo	0,1247	0,3304	0,0893	0,2852	0,1787	0,3832
Médio completo ou superior incompleto	0,4942	0,5000	0,4527	0,4978	0,5575	0,4967
Médio incompleto	0,3444	0,4752	0,4065	0,4912	0,2498	0,4329
Analfabeto ou básico incompleto	0,0367	0,1879	0,0515	0,2211	0,0140	0,1174
Número observações	212.089		132.095		79.994	
Fração observações por sexo			0,6228		0,3772	
Número trabalhadores	24.195		14.611		9.584	
Fração trabalhadores por sexo			0,6039		0,3961	

Fonte: Elaboração própria, dados da RAIS/MTE 1994 a 2009.

1.3 Estratégias empíricas

Diferentemente da discriminação relacionada à raça/cor de pele/etnia, para o qual a questão da correlação existente entre essa dimensão e as características socioeconômicas familiares é bastante relevante para os estudos empíricos. Pois, em geral, os negros e pardos (latinos no caso dos EUA) tendem a pertencer a famílias mais pobres, com menor escolaridade, com menores chances de acesso a boas escolas etc. A questão da correlação entre o sexo dos indivíduos e sua habilidade intrínseca ou condições socioeconômicas parece ser de segunda ordem na investigação dos diferenciais de resultados no mercado de trabalho entre gêneros⁵. Este aspecto minora as preocupações derivadas da restrição metodológica de não ser possível a utilização de modelos de painel com efeito-fixo para a obtenção direta de estimativas de discriminação entre gêneros.

Por outro lado, é de primeira ordem para os trabalhos empíricos sobre discriminação o fato de que o sexo dos indivíduos é correlacionado com o padrão de histórico profissional observado nas sociedades ocidentais, no nosso caso particular o contexto brasileiro. Por histórico no mercado de trabalho estamos nos referindo às consequências das decisões diferenciadas entre gêneros quanto a: participar ou não no mercado de trabalho; interrupções na carreira – muitas vezes motivadas por questões relacionadas à família e/ou fecundidade; investimento em capital humano geral e específico às firmas ou setores da economia; variações na carga horária de trabalho ao longo do ciclo de vida; tipos de ocupações no mercado de trabalho etc.

Posto essa questão, nosso estudo procura elucidar essa lacuna de conhecimento ao examinar o papel que o histórico profissional dos trabalhadores tem para explicar o diferencial de rendimentos entre gêneros observado empiricamente. No entanto, a simples inclusão de uma variável indicadora de gênero do trabalhador em uma equação de determinação de salários clássica, mesmo com a adoção de controle para a experiência potencial, não nos fornece

⁵ No entanto, existem alguns exemplos hipotéticos nos quais a questão da correlação entre as características de *background* familiar e o sexo dos filhos pode ser mais importante. Entre esses podemos citar o caso de famílias restritas a crédito e que decidam investir mais em capital humano para o filho homem em detrimento da mulher visando uma maior segurança financeira futura, ou a existência de uma maior correlação entre criminalidade e famílias pobres, e a maior incidência da primeira em média sobre os homens etc.

uma medida isenta de viés acerca da magnitude da discriminação que as mulheres sofrem no mercado de trabalho⁶.

Assim as informações detalhadas sobre a história dos indivíduos no mercado de trabalho formal, que temos a partir da RAIS/MTE, nos permite superar essa limitação usualmente encontrada nos estudos empíricos e testar a questão da discriminação entre gêneros em mais de uma dimensão e de formas complementares. Em primeiro lugar, nós utilizamos o procedimento padrão da literatura e estimamos o diferencial de rendimentos entre gêneros para uma *cross-section* (ano de 2009). Com a diferença de que reconstruímos a experiência real dos trabalhadores no mercado de trabalho formal a partir da dimensão longitudinal do painel (entre 1994 e 2008). Dessa forma, nós procuramos enfrentar um dos principais problemas presentes na maioria dos estudos empíricos sobre discriminação entre gêneros, que é o de viés nas estimativas obtidas decorrente de erro de medida não clássico na variável de experiência potencial (Reagan e Oaxaca, 2009). Sendo esta a principal contribuição do presente trabalho.

Entretanto, essa primeira abordagem pode apresentar outros aspectos potencialmente problemáticos para a correta estimação do diferencial. Esses são a existência de variáveis não observáveis que são relevantes para o sucesso dos trabalhadores no mercado de trabalho (salários) e que, portanto não podem ser incluídas como controles nas equações de salário. Entre essas variáveis omitidas podemos destacar aspectos relacionados à motivação e ao comprometimento dos indivíduos para com a sua atividade laboral. Assim como o *trade-off* enfrentado, sobretudo pelas mulheres, em relação a permanência no mercado de trabalho e as consequências associadas às decisões relacionadas a família e a fecundidade. Com o intuito de investigar a importância desse conjunto de variáveis omitidas, nós optamos por comparar homens e mulheres muito semelhantes quanto ao seu alto grau de engajamento – *attachment* – no mercado de trabalho formal durante o seu ciclo de vida produtiva. Em seguida, nós confrontamos a evolução do diferencial de rendimentos entre gêneros desse grupo de indivíduos ‘continuamente empregados’ com o observado para a amostra geral de trabalhadores entre 1994 e 2009.

⁶ A experiência potencial é definida como a idade dos indivíduos menos sua escolaridade menos cinco ou seis anos (Altonji & Blank, 1999, p. 3151).

Nas duas estratégias empíricas citadas acima existe o aspecto da seleção dos indivíduos no mercado de trabalho. Se existir uma diferenciação entre os sexos na forma de seleção dos trabalhadores no mercado de trabalho formal, isso refletirá sobre as estimativas de interesse a respeito do diferencial de salários. Ademais, no segundo exercício empírico, nós estamos intencionalmente selecionando os indivíduos que constituem nossa amostra. Isto é, estamos restringindo aos trabalhadores que permanecem no mercado de trabalho independente de fatores externos, usualmente associados com a constituição de famílias e criação de filhos. A questão da seleção é abordada neste trabalho com a utilização de modelos de efeitos-fixos, separadamente para trabalhadores do sexo masculino e feminino, que permitem controlar para as características intrínsecas dos indivíduos (constantes no tempo) que influenciam o padrão de seleção no mercado de trabalho. Essas equações de salários são estimadas inicialmente sem as variáveis de controle para experiência e posteriormente nós as adicionamos. A partir desses resultados, realizamos exercícios de decomposição, segundo o método Oaxaca-Blinder, do diferencial de salários entre gêneros entre os componentes explicado e não explicado, usualmente associado à discriminação.

Apesar deste trabalho não estar baseado em uma única estratégia empírica fundamentada em hipóteses de identificação claras, nós acreditamos que a conjunção dessas três estratégias empíricas revela informações úteis sobre o diferencial de salários entre gêneros no mercado de trabalho formal brasileiro e o papel do histórico profissional dos trabalhadores sobre este. Assim, as estratégias empíricas empregadas no presente trabalho podem ser vista como complementares. Nas subseções que seguem, nós explicitaremos em maiores detalhes cada uma dessas.

1.3.1 Medida de experiência profissional real

O tópico de erros de medida em variáveis explicativas e suas consequências estão usualmente presentes no contexto de trabalhos empíricos. No caso do estudo dos diferenciais de rendimentos entre gêneros esse tipo de preocupação é realçado, sobretudo para a variável referente à experiência dos indivíduos no

mercado de trabalho. Isso se deve ao fato de que a medida usualmente adotada – experiência potencial – mostra-se especialmente problemática para o caso das mulheres. De tal magnitude são os problemas derivados dessa medida potencial e das consequências associadas aos pressupostos da abordagem tradicional de erros de medida, que Regan & Oaxaca (2009) consideram mais adequada a interpretação do problema em questão sob o contexto de erro de especificação da equação a ser estimada.

Dada a relevância desse problema, torna-se um ponto fundamental na nossa análise a reconstrução da maneira mais precisa possível do histórico profissional dos indivíduos e a avaliação do modo pelo qual a incorporação dessas medidas influencia as estimativas do diferencial de rendimentos entre gêneros. Com esse intuito, em uma primeira etapa, nós restringimos nossa análise aos trabalhadores com vínculo empregatício ativo em 2009, para os quais podemos recuperar as informações referentes à sua permanência/ausência no mercado de trabalho formal brasileiro ao longo dos quinze anos anteriores. Dessa forma, por meio da reconstrução da trajetória profissional dos trabalhadores podemos incorporar na equação de determinação de salários uma medida de experiência real no mercado de trabalho.

Logo, nós limitamos a investigação empírica aos indivíduos que tinham entre 18 e 55 anos de idade em 2009, independentemente da ocupação e faixa de escolaridade⁷. Em seguida, a partir dessa amostra inicial, nós restringimos a análise aos trabalhadores de acordo com as suas faixas de escolaridade. Posteriormente, os procedimentos de estimação abaixo descritos serão reproduzidos para algumas coortes específicas de trabalhadores.

As variáveis referentes ao histórico profissional dos trabalhadores foram construídas de tal maneira a levar em consideração possíveis não linearidades do impacto da experiência prévia de acordo com a distância temporal em relação ao período específico utilizado nas estimações (ano de 2009). Dessa forma, nós criamos três variáveis para captar o número de anos ausente no mercado de trabalho, de acordo com o critério: entre 2004 e 2008; entre 1999 e 2003 e entre 1994 e 1998. Outras três variáveis, que utilizam o mesmo padrão acima, foram

⁷ Conforme já explicitado na seção de dados, excluímos os trabalhadores com rendimentos não positivos e que em 2009 estavam empregados nos subsetores de administração pública direta ou autárquica ou na agricultura, silvicultura e criação de animais.

criadas para captar o número de anos que os indivíduos trabalharam o ‘ano inteiro’. Essa variável é definida como um caso o trabalhador tenha apresentado vínculo empregatício ativo por nove meses ou mais em um único ano e zero caso contrário. Por último, criamos uma variável que expressa o número de anos que o empregado trabalhou em órgãos da administração pública direta ou autárquica anteriormente a 2009.

Essas medidas relativas à experiência e histórico profissional dos trabalhadores compõem o vetor denominado de ‘História’ na equação de salários abaixo que será utilizada para obter as estimativas do diferencial médio de rendimentos entre gêneros para a *cross-section* referente ao ano de 2009:

$$\begin{aligned} \ln(Y_{i,2009}) = & \beta \cdot Sexo_i + \lambda \cdot História_{i,2009} + \theta \cdot Escolaridade_{i,2009} \\ & + \alpha \cdot AnoNasc_i + \kappa \cdot Estado_{i,2009} + \omega \cdot Jornada_{i,2009} \\ & + \phi \cdot TempoParcial_{i,2009} + \rho \cdot Ocupação_{i,2009} + \varepsilon_{i,2009}. \end{aligned} \quad (1.1)$$

Onde $\ln(Y_{i,2009})$ denota o logaritmo natural do salário médio anual real em termos de janeiro de 2012 do trabalhador i para o ano de 2009. A nossa variável de interesse é $Sexo_i$, que assume valor igual a um se o indivíduo é do sexo feminino e zero caso contrário. O coeficiente associado a esta variável indicadora de gênero – β – reflete a redução salarial percentual (aproximadamente) média que as mulheres sofrem no mercado de trabalho formal brasileiro. Os demais termos da equação são vetores de variáveis (indicadoras em alguns casos) usados como controles nas equações de salários, tal como explicado acima e na seção de dados.

1.3.2

Comparação entre trabalhadores com elevado engajamento no mercado de trabalho

No segundo exercício empírico deste artigo, nós procuramos comparar subgrupos de trabalhadores que são semelhantes em características não observáveis determinantes para a evolução e sucesso da carreira profissional e, por conseguinte para os patamares de salários obtidos. Essas características,

conforme ressaltado por Machado (2011), são elevado grau de comprometimento e motivação em relação ao mercado de trabalho. As mulheres pertencentes a esse grupo de trabalhadores são particularmente interessantes de serem comparadas com os homens, visto que seu grau de permanência e comprometimento em relação ao mercado de trabalho é elevado e estável, assim como é tradicionalmente verificado para uma grande parcela dos homens.

Para isto, nós restringimos atenção ao conjunto de indivíduos que permaneceram empregados em todos os dezesseis anos da nossa amostra básica. Nesse caso, nós não limitamos a amostra aos trabalhadores que permaneceram empregados por nove meses ou mais em um mesmo ano (definição associada a variável denominada de ‘ano inteiro’). Para ser considerado um trabalhador ‘continuamente empregado’ bastava que o indivíduo apresentasse vínculo empregatício ativo em todos os anos analisados e que esse vínculo perdurasse por pelo menos quatro meses em cada ano – o mesmo critério utilizado para o sorteio da amostra inicial composta por 1% dos trabalhadores presentes na RAIS/MTE.

Assim, para cada ano em particular – entre 1994 e 2009 – nós comparamos os salários de homens e mulheres ‘continuamente empregados’. Esse procedimento nos permite estimar o diferencial de rendimentos entre gêneros para esse subgrupo de trabalhadores, excluindo o efeito de diferentes experiências de mercado de trabalho. Pois o conjunto de trabalhadores investigados se mantém constante ao longo do período analisado, ou seja, sobre essa estimativa de diferencial não há influência de uma possível taxa de saída e/ou entrada diferenciada entre gêneros no mercado de trabalho. Ademais, podemos avaliar a evolução desse diferencial ao longo do ciclo produtivo para diferentes coortes. Por último, comparamos as estimativas do diferencial entre gêneros dessa amostra de trabalhadores ‘continuamente empregados’, e suas evoluções ao longo do tempo, com os resultados obtidos para a amostra completa. Essa última inclui todos os trabalhadores, independentemente da variação na composição dos indivíduos dentro da amostra entre os diferentes anos utilizados na análise.

Entre as coortes estudadas estão os indivíduos nascidos entre 1972 e 1976 e que tinham 18 a 22 anos de idade em 1994. As coortes nascidas de 1967 a 1971 e 1962 a 1966, que tinham respectivamente 23 a 27 anos e 28 a 32 anos em 1994. E por último os trabalhadores nascidos entre 1957 e 1961, com 33 a 37 anos de idade no início do painel. Essa segmentação por coortes se justifica pela

possibilidade de avaliar a evolução das estimativas do diferencial de salários entre gêneros para as populações com diferentes graus de maturidade em relação ao ciclo de vida profissional.

Para a realização das comparações descritas nessa subseção foram estimadas para cada ano – t – do painel e para cada coorte – j – as duas equações abaixo e que se distinguem pela inclusão de um maior conjunto de variáveis de controle, principalmente pela presença de variáveis indicadoras de faixa de escolaridade dos trabalhadores:

$$\ln(Y_{ijt}) = \beta \cdot \text{Sexo}_{ij} + \alpha \cdot \text{Idade}_{ijt} + \varepsilon_{ijt}. \quad (1.2)$$

$$\begin{aligned} \ln(Y_{ijt}) = & \beta \cdot \text{Sexo}_{ij} + \theta \cdot \text{Escolaridade}_{ijt} + \alpha \cdot \text{Idade}_{ijt} \\ & + \kappa \cdot \text{Estado}_{ijt} + \omega \cdot \text{Jornada}_{ijt} + \phi \cdot \text{TempoParcial}_{ijt} \\ & + \delta \cdot \text{AnoInteiro}_{ijt} + \varepsilon_{ijt}. \end{aligned} \quad (1.3)$$

As variáveis utilizadas nessas equações já foram discutidas acima e na seção de dados do presente artigo. Novamente, nosso interesse se concentra nas estimativas do coeficiente β para a amostra completa e de indivíduos ‘continuamente empregados’, para as diferentes coortes e sua evolução no período entre 1994 e 2009.

1.3.3 Decomposição do diferencial entre gêneros

Por último, nós utilizamos o método de decomposição de Oaxaca-Blinder para dividir a diferença média de salários entre homens e mulheres nos efeitos composição (explicado) e não explicado. Além de verificar a contribuição das variáveis usualmente adotadas nas equações de determinação de salários para cada um desses efeitos. Complementarmente, nós examinamos se a inclusão de variáveis relativas ao histórico profissional dos trabalhadores – número de anos de experiência e de interrupções na oferta de trabalho – no modelo de determinação de salários tem algum impacto sobre esses efeitos. Este aspecto torna-se especialmente relevante dado à possibilidade de, a partir da RAIS/MTE, nós

recuperarmos informações críveis sobre a experiência real dos trabalhadores no mercado de trabalho formal brasileiro.

Conforme realçado anteriormente, neste trabalho temos especial interesse sobre o papel que as medidas reais de histórico profissional têm sobre os resultados usualmente associados à discriminação entre gêneros. Por esse motivo, ao decompor o diferencial de rendimentos entre gêneros entre os componentes explicado e não explicado, nós limitamos explicitamente nossa análise a um grupo de trabalhadores jovens e para os quais existe forte evidência de que estes não estavam inseridos, pelo menos de maneira estável, no mercado de trabalho formal brasileiro no início do período analisado. Para isso, nós definimos uma terceira amostra composta por trabalhadores com vínculo empregatício ativo em 1996 e faixa etária entre 18 e 25 anos. Adicionalmente, esses trabalhadores não estavam presentes na amostra da RAIS para os anos de 1994 e 1995, ou seja, estavam fora do mercado de trabalho formal nesse intervalo. Assim, com este procedimento nós pretendemos excluir os trabalhadores que já tinham experiência prévia substancial no mercado de trabalho.

Essa terceira amostra é composta por 24.195 trabalhadores, sendo aproximadamente 60% homens e 40% mulheres. Para esse conjunto de trabalhadores recuperamos todas as informações disponíveis na base de dados da RAIS/MTE para o período entre 1996 e 2009. E estimamos as equações de salários com efeitos-fixos – γ_i – por trabalhadores, separadamente para homens e mulheres, tal como apresentado abaixo nas equações (1.4) e (1.5):

$$\begin{aligned} \ln(Y_{it}) = & \alpha \cdot Idade_{it} + \theta \cdot Escolaridade_{it} + \omega \cdot Jornada_{it} \\ & + \phi \cdot TempoParcial_{it} + \delta \cdot AnoInteiro_{it} \\ & + \kappa \cdot Estado_{it} + \gamma_i + \varepsilon_{it}. \end{aligned} \quad (1.4)$$

$$\begin{aligned} \ln(Y_{it}) = & \alpha \cdot Idade_{it} + \sigma \cdot Experiência_{it} + \zeta \cdot ExperiênciaQuadrado_{it} \\ & + \eta \cdot Interrupções_{it} + \theta \cdot Escolaridade_{it} + \omega \cdot Jornada_{it} \\ & + \phi \cdot TempoParcial_{it} + \delta \cdot AnoInteiro_{it} \\ & + \kappa \cdot Estado_{it} + \gamma_i + \varepsilon_{it}. \end{aligned} \quad (1.5)$$

Na equação (1.5), $Experiência_{it}$ é o número total de anos que o trabalhador ‘ i ’ esteve presente com vínculo ativo no mercado de trabalho até o período ‘ t ’, no intervalo de anos entre 1996 e 2009. O termo seguinte é o quadrado dessa medida de número de anos de experiência no mercado de trabalho e tem o objetivo de captar possíveis não linearidades no efeito da experiência sobre os rendimentos. Já $Interrupções_{it}$ é um vetor constituído por três variáveis indicadoras que tomam valor igual a um quando o indivíduo apresenta, respectivamente, uma, duas e três ou mais interrupções na sua oferta de trabalho no mercado formal ao longo do intervalo de anos entre 1996 e 2009.

A diferença total entre os valores médios para a variável dependente – nas equações (1.4) e (1.5) – entre o grupo de trabalhadores do sexo feminino e masculino é dada por:

$$\Delta_{fm} = \overline{\ln(Y_f)} - \overline{\ln(Y_m)}.$$

E pode ser reescrita da forma que segue⁸:

$$\Delta_{fm} = \sum_{k=1}^K (\bar{X}_{fk} - \bar{X}_{mk}) \cdot \hat{\beta}_{mk} + \sum_{k=1}^K \bar{X}_{fk} \cdot (\hat{\beta}_{fk} - \hat{\beta}_{mk}). \quad (1.6)$$

O primeiro termo do lado direito da equação (1.6) refere-se ao efeito composição (explicado), ou seja, a parcela do diferencial de salários entre gêneros devido às diferenças médias nas características entre os trabalhadores do sexo feminino e masculino, quando utilizamos os homens como grupo base. Já o segundo termo é o efeito não explicado e representa a parcela do diferencial de rendimentos decorrente da diferença nos retornos estimados para as características dos trabalhadores entre os sexos. Esse último termo é usualmente denominado na literatura como a fração do diferencial de rendimentos devido à discriminação entre gêneros. Além disso, podemos subdividir ambos os efeitos nas contribuições

⁸ Com o intuito de simplificar a notação, especialmente para esta equação referente a decomposição de salários entre os gêneros, as variáveis explicativas do modelo são representadas por X_k e as estimativas dos coeficientes associados são dados por $\hat{\beta}_k$, $k = 1, 2, \dots, K$. O subscrito ‘ m ’ denota o grupo de trabalhadores do sexo masculino e o subscrito ‘ f ’ o grupo de indivíduos do sexo feminino.

de cada uma das variáveis explicativas presentes no modelo de determinação de salários⁹.

1.4 Resultados

Nessa seção apresentamos os resultados obtidos a partir dos procedimentos descritos anteriormente. As diferentes estratégias e, por conseguinte as distintas amostras utilizadas tem o intuito de explorar as várias dimensões da influência do histórico profissional dos trabalhadores sobre o diferencial de rendimentos entre gêneros e sua evolução ao longo do ciclo de vida.

1.4.1 O diferencial de salário entre gêneros e a experiência profissional prévia

Inicialmente, com o intuito de mensurar o diferencial de salários entre gêneros para o contexto do mercado de trabalho formal brasileiro e melhor entender a influência da experiência profissional sobre este, nós restringimos a amostra aos trabalhadores presentes na RAIS em 2009 e nascidos entre os anos de 1954 e 1991. As estimativas obtidas a partir da equação (1.1) para essa primeira amostra são apresentadas na tabela 1.2. A base de comparação a partir do qual as estimativas apresentadas nessa tabela devem ser interpretadas são os homens que permaneceram no mercado de trabalho formal em todos os anos entre 1994 e 2008, mas que não mantiveram vínculo empregatício ativo por nove meses ou mais em cada um desses períodos e sempre estiveram empregados no setor privado no mesmo período. Ademais, no ano de 2009 esses trabalhadores do sexo masculino tiveram jornadas de trabalho semanais superiores a trinta horas e vínculo empregatício ativo inferior a nove meses. A princípio, essa definição de grupo base pode parecer rebuscado, mas por meio das variáveis utilizadas nas

⁹ Para maiores detalhes acerca da decomposição de Oaxaca-Blinder, pressupostos implícitos e possíveis limitações na interpretação dos efeitos ver Altonji & Blank (1999) e Fortin, Lemieux & Firpo (2011, seção 3).

regressões podemos mensurar especificamente os efeitos de interrupções na oferta de trabalho e de um maior comprometimento para com o mercado formal.

O diferencial médio desfavorável às mulheres é de aproximadamente 21% condicional apenas a variáveis indicadoras de coorte e estados da federação, tal como apontado na coluna I. Esses controles nos permitem condicionar a análise apenas a idade e aspectos particulares dos mercados de trabalho locais. A inclusão das variáveis relativas à experiência anterior no mercado de trabalho formal, reconstruídas a partir da observação da oferta de trabalho dos indivíduos nos quinze anos precedentes a 2009, tem influência marginal sobre a redução na magnitude do diferencial de salários entre gêneros – colunas II da tabela 1.2. Isso ocorre apesar das medidas de experiência ter impactos significativos e relevantes sobre os rendimentos individuais. Por exemplo, durante o intervalo de anos entre 2004 e 2008, a cada ano fora do mercado de trabalho os indivíduos sofrem uma redução próxima a 3% nos seus rendimentos médios. O efeito da ausência no mercado de trabalho sobre os salários, no entanto, reduz-se em valor absoluto conforme o tempo transcorrido entre o período no qual a não participação é observada e as informações sobre salários são obtidas (2009). E para a amostra investigada, a ausência no mercado durante o intervalo de tempo mais distante (1994 a 1998) tem um efeito positivo sobre os rendimentos médios em 2009, acréscimos estes de aproximadamente 1,5%. Uma possível explicação para essa alteração de padrão no período mais distante pode ser uma correlação entre as decisões de investimento em capital humano e de não participação no mercado de trabalho, sobretudo para alguns grupos de indivíduos com maior nível de escolaridade, e que acaba por refletir em ganhos na remuneração futura.

O fato dos trabalhadores permanecerem empregados o ‘ano inteiro’, ao longo dos anos observados, tem um efeito positivo e significativo sobre os seus rendimentos médios. A magnitude desse impacto situa-se entre 2,5% e 6%, por ano completo, para os anos mais próximos ao período analisado. O conjunto de resultados relacionados à experiência prévia dos trabalhadores no mercado de trabalho formal brasileiro, conforme apresentado na tabela 1.2 coluna II, indicam a relevância que o grau de comprometimento dos indivíduos para com suas carreiras profissionais tem para o patamar médio de salários atingido no futuro. Entretanto, vale ressaltar que esse papel não se faz presente por meio da redução da discriminação percebida contra as mulheres no mercado de trabalho. Isto seria

o esperado devido aos períodos críticos relacionados com decisões acerca da família e fecundidade, e que estão associados com um menor acúmulo de experiência real no mercado de trabalho por parte das mulheres.

Tabela 1.2 – Equação de determinação de salários para os trabalhadores entre 18 e 55 anos em 2009

Ln salário médio real	Amostra Geral - trabalhadores entre 18 e 55 anos em 2009					
	I	II	III	IV	V	VI
Mulher	-0.213	-0.211	-0.182	-0.330	-0.300	-0.191
	(0.00187)***	(0.00175)***	(0.00183)***	(0.00159)***	(0.00152)***	(0.00171)***
Num. anos fora mercado 2004-08		-0.0323	-0.0274		-0.0297	-0.0259
		(0.000968)***	(0.000809)***		(0.000849)***	(0.000777)***
Num. anos fora mercado 1999-03		-0.00749	-0.00609		-0.00488	-0.00507
		(0.00108)***	(0.000882)***		(0.000919)***	(0.000838)***
Num. anos fora mercado 1994-98		0.0172	0.0128		0.0146	0.0117
		(0.00175)***	(0.00143)***		(0.00148)***	(0.00135)***
Ano inteiro em 2009		0.0436	0.0112		0.0300	0.0171
		(0.00192)***	(0.00160)***		(0.00164)***	(0.00151)***
Num. anos inteiros 2004-08		0.0616	0.0465		0.0557	0.0490
		(0.000751)***	(0.000621)***		(0.000639)***	(0.000587)***
Num. anos inteiros 1999-03		0.0397	0.0283		0.0290	0.0252
		(0.000979)***	(0.000789)***		(0.000816)***	(0.000740)***
Num. anos inteiros 1994-98		0.0774	0.0532		0.0534	0.0459
		(0.00112)***	(0.000895)***		(0.000923)***	(0.000838)***
Num anos setor público		0.0324	0.0122		-0.00514	0.00311
		(0.00114)***	(0.00101)***		(0.00103)***	(0.000992)***
Dummies Idade	Sim	Sim	Sim	Sim	Sim	Sim
Estados	Sim	Sim	Sim	Sim	Sim	Sim
Jornada de trabalho	Não	Sim	Sim	Não	Sim	Sim
Tempo parcial	Não	Sim	Sim	Não	Sim	Sim
Faixas escolaridade	Não	Não	Não	Sim	Sim	Sim
Ocupações	Não	Não	Sim	Não	Não	Sim
Constante	6.598	8.212	7.756	6.399	7.042	7.212
	(0.00524)***	(0.0188)***	(0.0173)***	(0.00687)***	(0.0172)***	(0.0171)***
Observações	511,130	511,130	511,130	511,130	511,130	511,130
R-quadrado	0.164	0.284	0.507	0.411	0.476	0.560

Fonte: Elaboração própria, dados da RAIS/MTE 1994 a 2009.

Erros padrão entre parênteses - *** $p < 0,01$; ** $p < 0,05$; * $p < 0,1$.

Um ponto interessante presente nas colunas IV e V da tabela 1.2 refere-se ao incremento que ocorre no diferencial de salários entre gêneros, em termos absolutos, após a inclusão de variáveis indicadoras para faixas de escolaridade na equação de determinação de salários. No caso brasileiro o diferencial, em termos absolutos, sofre um incremento de mais do que dez pontos percentuais após a inclusão desses controles para escolaridade, passando para 33%. Adicionalmente, ao compararmos as colunas IV e V, a inclusão das variáveis de experiência no mercado de trabalho é responsável por uma redução de aproximadamente 10% na magnitude do coeficiente associado ao diferencial de salários observado contra as mulheres (30%). Apesar de persistirem os impactos relevantes sobre os

rendimentos médios individuais das variáveis relativas ao histórico profissional. De maneira complementar, ao compararmos as colunas II e V da tabela 1.2 constatamos que as estimativas associadas à experiência no mercado de trabalho sofrem uma pequena redução (em termos absolutos) após a inserção de controles para educação dos empregados. Essa evidência segue no sentido contrário ao observado para a magnitude das estimativas do coeficiente associado à discriminação entre gêneros, que se amplia.

Este resultado relacionado à associação entre escolaridade e o diferencial de rendimentos entre os sexos é distinto do obtido na literatura internacional, na qual o diferencial entre gêneros se reduz ou permanece constante condicional a escolaridade¹⁰. Se adotarmos o pressuposto de que a educação dos indivíduos é uma boa variável *proxy* para capital humano geral, então temos que as mulheres brasileiras enfrentam em média uma maior penalização em termos salariais em comparação aos homens quando restringimos nossa análise apenas aos indivíduos pertencentes a um mesmo grupo definido pelo nível de produtividade. Ou seja, condicional a sua produtividade as mulheres são ainda mais discriminadas no Brasil.

Entretanto, uma parcela significativa desse elevado patamar de discriminação contra as mulheres no mercado de trabalho formal brasileiro parece estar fortemente relacionada com as decisões ocupacionais diferenciadas entre gêneros e os seus respectivos retornos. A coluna VI da tabela 1.2 apresenta as estimativas obtidas para a equação de determinação de salários com o acréscimo de variáveis indicadoras para cada uma das famílias ocupacionais definidas na Classificação Brasileira de Ocupações – CBO2002. As famílias ocupacionais são a unidade do sistema de classificação e agrupam as situações de emprego ou ocupações similares, representando, por conseguinte uma medida bastante desagregada das ocupações observadas no mercado de trabalho brasileiro. Assim, ao controlarmos para as diferentes ocupações dos indivíduos, o diferencial entre os sexos sofre uma redução de aproximadamente 11 pontos percentuais, o que representa um terço do diferencial previamente estimado para nossa amostra – comparar colunas IV e VI da tabela 1.2. Dada a magnitude desse declínio, torna-

¹⁰ Esse padrão de redução ou constância para o diferencial entre gêneros após a inclusão de controles para escolaridade é observado por Altonji & Blank (1999) para duas diferentes bases de dados dos E.U.A. (CPS e NYLS).

se imprescindível que estudos futuros investiguem de maneira mais detalhada a discriminação entre gêneros e sua evolução ao longo do histórico profissional para ocupações particulares, tais como já realizados para outros países. O estudo pormenorizado do diferencial de rendimentos entre os sexos e sua relação com as escolhas ocupacionais dos indivíduos está além do escopo do presente trabalho, por isso nas tabelas e gráficos que seguem nós não incluímos as estimativas da equação de salário com controles para ocupações¹¹. Entretanto, é importante destacar que mesmo após a adição de variáveis de controle para ocupação, o diferencial entre gêneros persiste em um patamar economicamente relevante, 19,1% (ver coluna VI da tabela 1.2). Esse resultado é muito próximo do verificado inicialmente quando incluímos na análise apenas variáveis de controle para coortes e estados da federação.

O grande incremento no diferencial de rendimentos entre gêneros após a inclusão da escolaridade dos trabalhadores motivou uma investigação mais detalhada sobre a existência de algum grau de heterogeneidade desse diferencial ao restringirmos a amostra por faixas de escolaridade. Os resultados ao se utilizar a mesma abordagem acima e segregando a amostra apenas pelo nível de educação dos trabalhadores no ano de 2009 são apresentados na tabela 1.3. Nas colunas I e II temos os trabalhadores com ensino superior completo ou maior nível de escolaridade – mestrado e/ou doutorado. Nas colunas III e IV são apresentados os resultados para os indivíduos com ensino médio completo ou ensino superior incompleto. Os empregados com nível de escolaridade entre a quarta série do ensino fundamental e o ensino médio incompleto tem suas estimativas apresentadas nas colunas V e VI. Por último, os resultados referentes aos trabalhadores analfabetos ou com escolaridade até a primeira metade do ensino fundamental incompleto estão nas colunas VII e VIII. As colunas ímpares contêm as estimativas dos diferenciais entre gêneros para as respectivas amostras com controles apenas para coorte e estados da federação. Já as colunas pares também incluem as variáveis relativas à jornada de trabalho, experiência e particularidades sobre o histórico profissional dos trabalhadores no mercado de trabalho formal nos quinze anos anteriores.

¹¹ Em alguns casos essas estimativas com a inclusão de controles para ocupações estarão disponíveis nos anexos do presente trabalho apenas com o intuito de corroborar a robustez dos resultados.

Tabela 1.3 – Equação de determinação de salários para os trabalhadores entre 18 e 55 anos por faixa de escolaridade no ano 2009

Ln salário médio real	Amostra Geral - trabalhadores entre 18 e 55 anos em 2009							
	Ensino Superior		Ensino Médio		Ensino Fundamental		Semi-Analfabeto	
	I	II	III	IV	V	VI	VII	VIII
Mulher	-0.416	-0.389	-0.277	-0.255	-0.351	-0.305	-0.368	-0.331
	(0.00646)***	(0.00620)***	(0.00194)***	(0.00184)***	(0.00229)***	(0.00215)***	(0.0111)***	(0.0106)***
Num. anos fora mercado 2004-08		-0.0639		-0.0255		-0.0199		-0.0253
		(0.00395)***		(0.00109)***		(0.00118)***		(0.00479)***
Num. anos fora mercado 1999-03		-0.0317		-0.00150		0.00149		0.0100
		(0.00386)***		(0.00123)		(0.00125)		(0.00490)**
Num. anos fora mercado 1994-98		0.00409		0.0183		0.0125		-0.00209
		(0.00604)		(0.00210)***		(0.00189)***		(0.00755)
Ano inteiro em 2009		0.127		0.0375		-0.0122		-0.0801
		(0.00699)***		(0.00204)***		(0.00242)***		(0.0109)***
Num. anos inteiros 2004-08		0.0623		0.0682		0.0383		0.0226
		(0.00259)***		(0.000836)***		(0.000921)***		(0.00394)***
Num. anos inteiros 1999-03		0.00764		0.0399		0.0267		0.0277
		(0.00280)***		(0.00109)***		(0.00117)***		(0.00484)***
Num. anos inteiros 1994-98		0.0359		0.0602		0.0471		0.0343
		(0.00291)***		(0.00131)***		(0.00131)***		(0.00491)***
Num anos setor público		-0.0188		-0.00388		0.00582		0.0165
		(0.00167)***		(0.00137)***		(0.00246)**		(0.00769)**
Dummies Idade	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Estados	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Jornada de trabalho	Não	Sim	Não	Sim	Não	Sim	Não	Sim
Tempo parcial	Não	Sim	Não	Sim	Não	Sim	Não	Sim
Constante	7.021	7.461	6.665	7.414	6.523	6.955	6.685	7.385
	(0.0953)***	(0.102)***	(0.00658)***	(0.0207)***	(0.00703)***	(0.0379)***	(0.0481)***	(0.203)***
Observações	65,326	65,326	287,738	287,738	149,085	149,085	8,981	8,981
R-quadrado	0.214	0.283	0.196	0.301	0.230	0.318	0.228	0.293

Fonte: Elaboração própria, dados da RAIS/MTE 1994 a 2009.

Erros padrão entre parênteses - *** p<0,01; ** p<0,05; * p<0,1.

A tabela 1.3 revela que o diferencial de rendimentos entre gêneros tem magnitude ainda mais relevante, em comparação com a amostra como um todo, para o grupo de trabalhadores com maior escolaridade e que por hipótese tem um elevado nível de produtividade. Para esse grupo, o diferencial de salários contrários às mulheres perfaz um valor de aproximadamente 39%. Já para os trabalhadores com ensino médio e/ou nível técnico, o diferencial entre os sexos apresenta o seu menor valor, em termos absolutos, entre 25% e 28%. Abaixo inclusive ao valor médio obtido para a amostra geral. Esses trabalhadores tem um nível de produtividade intermediário e, no entanto sofrem o limite inferior de discriminação contra as mulheres, mesmo quando comparados aos grupos de indivíduos com menor escolaridade, que apresentam estimativas do diferencial próximas às da amostra geral – ao redor de 30% a 33%. O padrão acima descrito mostra uma importante não monotonicidade do diferencial de rendimentos entre gêneros em relação à produtividade/escolaridade dos trabalhadores no contexto brasileiro. Onde se destaca o fato de que as mulheres no topo da distribuição de

habilidades (escolaridade) sofrem de maneira mais acentuada com o comportamento discriminatório dos empregadores, padrão este que está em consonância ao que se costuma denominar de *glass ceiling* na literatura sobre discriminação entre gêneros.

Mais uma vez, a inclusão de variáveis para a experiência real no mercado de trabalho formal não tem papel relevante para explicar a discriminação percebida contra as mulheres. A inclusão das variáveis relativas ao histórico no mercado de trabalho implica a redução de dois a três pontos percentuais na estimativa do diferencial salarial entre os sexos. Padrão este similar ao observado na tabela 1.2 entre as colunas IV e V. Contudo, é importante destacar que os efeitos de períodos de ausência dos indivíduos no mercado de trabalho são de grande magnitude, sobretudo para aqueles pertencentes ao grupo com maior escolaridade. As estimativas de redução do nível de rendimentos de 6,39% e 3,17% a cada ano de ausência, respectivamente, para o período 2004 – 2008 e 1999 – 2003, são superiores ao dobro dos impactos observados para a amostra geral (tabela 1.2). Além disso, cada ano que os indivíduos permanecem empregados o ‘ano inteiro’ tem um impacto positivo e economicamente relevante, especialmente para os trabalhadores com maior escolaridade – colunas II e IV da tabela 1.3. Dessa forma, os nossos resultados corroboram, principalmente para o grupo de trabalhadores com maior nível de educação, as fortes penalizações associadas a interrupções na oferta de trabalho e os prêmios em termos salariais associados a uma elevada permanência no mercado de trabalho formal.

A tabela A2 no apêndice apresenta as estimativas do diferencial de salários entre gêneros dos modelos análogos às colunas pares da tabela 1.3 com a inclusão de controles para ocupação dos empregados. Esses resultados reafirmam a grande redução em magnitude da discriminação contra as mulheres quando se leva em consideração suas decisões ocupacionais e esse padrão se faz presente para os trabalhadores de todas as classes de escolaridade. Entretanto, os diferenciais entre gêneros permanecem economicamente relevantes, principalmente para os empregados com maior escolaridade. Também persiste o impacto diferenciado do grau de comprometimento dos indivíduos no mercado de trabalho – penalizações relacionadas a interrupções e acréscimos salariais associados à permanência em ‘anos inteiros’ no mercado – para os trabalhadores com maiores níveis de instrução.

Outro aspecto que investigamos refere-se à existência de distintos padrões de diferencial de rendimentos entre gêneros de acordo com a maturidade dos trabalhadores, ou seja, por faixa etária. Conforme os indivíduos envelhecem, estes ganham experiência no mercado de trabalho. Esse atributo produtivo em geral é bastante valorizado, pois pode ser interpretado segundo a teoria econômica como uma acumulação de capital humano específico à firma. Devido a decisões de fecundidade e composição de família pela qual passam os indivíduos, e que tem efeitos mais importantes na alteração do padrão de oferta de trabalho, rotatividade e comprometimento das mulheres em comparação com os homens. Esse acúmulo de experiência profissional no mercado e, por conseguinte de capital humano, apresenta uma evolução média diferenciada entre gêneros conforme a coorte envelhece. Por outro lado, conforme a idade dos trabalhadores avança muitas vezes os empregadores precisam tomar decisões sobre promoções e investimentos em capital humano específico às firmas que refletem por sua vez nas remunerações individuais futuras. Assim, à medida que as mulheres se aproximam das idades comumente associadas ao matrimônio, gestação e criação dos filhos – o que ocorre ainda no início do ciclo de vida produtivo dos trabalhadores – os empregadores, dado o problema informacional, podem apresentar comportamento discriminatório contra as mulheres. Isso ocorre apenas pelo fato das mulheres pertencerem ao grupo que em média terá um menor grau de comprometimento com o mercado de trabalho no futuro próximo, o que caracteriza um processo bastante similar à discriminação estatística.

Na tabela 1.4 abaixo, seguimos a abordagem metodológica utilizada nessa subseção. Dessa forma, nós restringimos a comparação de salários a uma *cross section* (2009), isto é, mantemos constante a dimensão tempo e temos tanto os efeitos idade quanto coorte presentes na análise da evolução do diferencial entre gêneros¹². O painel A dessa tabela apresenta as estimativas dos coeficientes da equação de salários para as amostras formadas pelos indivíduos que compõem as coortes de indivíduos nascidos entre os anos de 1987 a 1991, 1982 a 1986 e 1977 a 1981. Esses trabalhadores tinham em 2009, respectivamente, 18 a 22 anos, 23 a 27 anos e 28 a 32 anos. Já o painel B da mesma tabela mostra os resultados

¹² Para maiores detalhes sobre a descrição dos efeitos tempo, coorte e idade na análise da evolução do padrão de comportamento dos indivíduos, em especial para o caso de diferenças de participação no mercado de trabalho entre homens e mulheres no Brasil, ver Scorzafave e Menezes-Filho (2001).

relativos às coortes de trabalhadores que nasceram de 1972 a 1976, 1967 a 1971 e 1962 a 1966. No ano de 2009, os membros dessas coortes tinham respectivamente 33 a 37 anos, 38 a 42 anos e 43 a 47 anos.

Para cada uma dessas coortes nós estimamos três especificações do modelo de determinação de salários que diferem pelo conjunto de variáveis de controle incluídas em cada equação. Na primeira coluna, para cada coorte, temos o diferencial de salários entre gêneros quando condicionamos a análise a variáveis indicadoras para idades e Estados da Federação, além de um vetor de variáveis referentes ao histórico profissional dos indivíduos. A segunda coluna exhibe as estimativas para a especificação que inclui variáveis indicadoras de escolaridade, porém na ausência das variáveis de história profissional. Por último, para cada coorte, a terceira coluna expõe os resultados da nossa especificação mais completa, ou seja, que condiciona para as variáveis de idade, Estados da Federação, escolaridade e de histórico profissional no mercado de trabalho formal. Com o objetivo de exemplificar a disposição dos resultados presentes na tabela 1.4, para a coorte de trabalhadores nascidos entre 1982 e 1986 as estimativas são apresentadas no painel A, entre as colunas IV e VI. De tal forma que os resultados da primeira especificação estão na coluna IV, da segunda na coluna V e da nossa especificação mais completa na coluna VI.

A partir do conjunto de resultados listados na tabela 1.4, nós constatamos que o diferencial de rendimentos entre gêneros é bastante inferior nas idades mais tenras, sendo aproximadamente igual a 12% na faixa etária entre 18 a 22 anos. Além disso, essa medida de discriminação contra as mulheres cresce de maneira monotônica, em termos absolutos, com a faixa etária de cada coorte no ano de 2009 para as três distintas especificações utilizadas. Ao restringirmos nossa atenção apenas às estimativas associadas à especificação mais completa da equação de salários, verificamos que o módulo do diferencial aumenta em 9,3 pontos percentuais entre a primeira faixa etária e a amostra que tinha 23 a 27 anos em 2009. Entre essa última faixa de idade e a seguinte – 28 a 32 anos em 2009 – o incremento no diferencial é de aproximadamente seis pontos percentuais, mantendo o mesmo patamar entre a terceira e a quarta faixas etárias (de 28 a 32 anos para 33 a 37 anos em 2009). Por último, as duas transições na variação do diferencial entre gêneros associadas as três últimas faixas de idade revelam um

crescimento, em módulo, de aproximadamente 3,5 pontos percentuais para cada uma destas.

Tabela 1.4 – Equação de determinação de salários no ano de 2009 para seis cortes de trabalhadores: nascidos entre 1987 e 1991, 1982 e 1986, 1977 e 1981, 1972 e 1976, 1967 e 1971; e 1962 e 1966

Ln salário médio real	PAINEL A								
	Coorte 1987-91 18 a 22 anos em 2009			Coorte 1982-86 23 a 27 anos em 2009			Coorte 1977-81 28 a 32 anos em 2009		
	I	II	III	IV	V	VI	VII	VIII	IX
Mulher	-0.100	-0.134	-0.122	-0.162	-0.240	-0.233	-0.192	-0.320	-0.298
	(0.00262)***	(0.00268)***	(0.00261)***	(0.00301)***	(0.00280)***	(0.00275)***	(0.00395)***	(0.00348)***	(0.00338)***
Num. anos fora mercado 2004-08	0.000152		-0.00837	-0.0205		-0.0231	-0.0345		-0.0302
	(0.00553)		(0.00533)	(0.00191)***		(0.00170)***	(0.00183)***		(0.00160)***
Num. anos fora mercado 1999-03				0.0234		0.0204	0.00164		0.00141
				(0.00512)***		(0.00453)***	(0.00200)		(0.00170)
Num. anos fora mercado 1994-98							-0.00615		0.00435
							(0.00638)		(0.00533)
Ano inteiro em 2009	0.0108		0.0119	0.0618		0.0457	0.0784		0.0512
	(0.00291)***		(0.00284)***	(0.00320)***		(0.00283)***	(0.00418)***		(0.00353)***
Num. anos inteiros 2004-08	0.0597		0.0601	0.0466		0.0477	0.0640		0.0561
	(0.00246)***		(0.00242)***	(0.00124)***		(0.00112)***	(0.00148)***		(0.00126)***
Num. anos inteiros 1999-03				0.0305		0.0212	0.0262		0.0203
				(0.00302)***		(0.00273)***	(0.00170)***		(0.00146)***
Num. anos inteiros 1994-98							0.0142		0.00963
							(0.00393)***		(0.00340)***
Num anos setor público	0.0181		-0.00880	0.0199		-0.0212	0.0310		-0.0154
	(0.0133)		(0.0132)	(0.00438)***		(0.00402)***	(0.00291)***		(0.00267)***
Observações	80,157	80,157	80,157	117,715	117,715	117,715	103,702	103,702	103,702
R-quadrado	0.197	0.179	0.227	0.152	0.278	0.312	0.179	0.357	0.408

Ln salário médio real	PAINEL B								
	Coorte 1972-76 33 a 37 anos em 2009			Coorte 1967-71 38 a 42 anos em 2009			Coorte 1962-66 43 a 47 anos em 2009		
	X	XI	XII	XIII	XIV	XV	XVI	XVII	XVIII
Mulher	-0.240	-0.396	-0.358	-0.284	-0.438	-0.388	-0.326	-0.471	-0.417
	(0.00524)***	(0.00447)***	(0.00435)***	(0.00616)***	(0.00536)***	(0.00513)***	(0.00733)***	(0.00639)***	(0.00602)***
Num. anos fora mercado 2004-08	-0.0369		-0.0331	-0.0366		-0.0280	-0.0345		-0.0285
	(0.00234)***		(0.00203)***	(0.00280)***		(0.00247)***	(0.00335)***		(0.00292)***
Num. anos fora mercado 1999-03	-0.00566		-0.00126	-0.00419		-0.00482	-0.0112		-0.00853
	(0.00219)***		(0.00186)	(0.00256)		(0.00222)***	(0.00317)***		(0.00267)***
Num. anos fora mercado 1994-98	0.00198		0.00269	0.0149		0.0109	0.0306		0.0168
	(0.00307)		(0.00259)	(0.00351)***		(0.00297)***	(0.00430)***		(0.00364)***
Ano inteiro em 2009	0.0566		0.0392	0.0331		0.0308	0.00938		0.0126
	(0.00562)***		(0.00464)***	(0.00684)***		(0.00571)***	(0.00818)		(0.00677)***
Num. anos inteiros 2004-08	0.0689		0.0549	0.0617		0.0561	0.0638		0.0619
	(0.00198)***		(0.00163)***	(0.00237)***		(0.00197)***	(0.00283)***		(0.00233)***
Num. anos inteiros 1999-03	0.0393		0.0241	0.0490		0.0297	0.0429		0.0310
	(0.00214)***		(0.00175)***	(0.00253)***		(0.00209)***	(0.00296)***		(0.00240)***
Num. anos inteiros 1994-98	0.0278		0.0215	0.0636		0.0358	0.0918		0.0500
	(0.00227)***		(0.00187)***	(0.00233)***		(0.00192)***	(0.00262)***		(0.00215)***
Num anos setor público	0.0196		-0.0194	0.0327		-0.0119	0.0293		-0.00888
	(0.00254)***		(0.00226)***	(0.00293)***		(0.00255)***	(0.00261)***		(0.00234)***
Observações	71,630	71,630	71,630	54,278	54,278	54,278	42,779	42,779	42,779
R-quadrado	0.214	0.407	0.466	0.267	0.427	0.501	0.303	0.438	0.534
Controles:									
Dummies Idade	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Estados	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Jornada de trabalho	Sim	Não	Sim	Sim	Não	Sim	Sim	Não	Sim
Tempo parcial	Sim	Não	Sim	Sim	Não	Sim	Sim	Não	Sim
Faixas escolaridade	Não	Sim	Sim	Não	Sim	Sim	Não	Sim	Sim

Fonte: Elaboração própria, dados da RAIS/MTE 1994 a 2009.

Erros padrão entre parênteses - *** p<0,01; ** p<0,05; * p<0,1.

Conforme já destacamos anteriormente, como estamos mantendo a dimensão tempo constante na análise que deriva os resultados apresentados na tabela 1.4, a magnitude desses incrementos nas estimativas do diferencial deve-se

a dois efeitos: idade e coorte. Assim, se as coortes mais antigas de trabalhadores têm um conjunto de características intrínsecas que conferem uma tendência à existência de um maior diferencial entre gêneros, os resultados apresentados e discutidos acima podem estar superestimados para as respectivas faixas etárias. Entretanto, dada a magnitude da variação observada para o diferencial de rendimentos entre as amostras com diferentes idades e se supormos que o efeito coorte é aproximadamente constante – algo em torno de três pontos percentuais entre cada coorte de indivíduos. Então podemos deduzir a existência de um efeito senioridade (idade) negativo sobre a evolução do diferencial de salários entre gêneros entre o grupo de trabalhadores mais jovens (até 27 anos), com idade intermediária (28 a 37 anos) e dos indivíduos mais maduros (acima de 38 anos). Isto é, conforme os trabalhadores envelhecem ocorre um aumento no diferencial de rendimentos médio entre gêneros que está diretamente relacionado com a idade e esses incrementos são decrescentes na medida em que os indivíduos atingem um maior grau de maturidade.

Dessa forma, esse conjunto de evidências acerca do padrão de evolução do diferencial entre gêneros ao longo das faixas de idade, somado ao pressuposto de uma aproximada constância do efeito coorte, nos permite concluir que a maior expansão dessa discrepância salarial entre gêneros se faz presente no início do ciclo de vida produtivo dos trabalhadores. Justamente um intervalo crítico, sobretudo para as mulheres, no qual se apresentam questões que são consequências do período de fertilidade feminina, tais como, decisões sobre matrimônio, fecundidade e maior frequência de interrupções na oferta de trabalho no mercado para se dedicar a assuntos familiares. Essa coincidência entre um forte crescimento do diferencial entre gêneros em um período do ciclo da vida marcado por decisões bastante distintas entre os sexos pode ser interpretada como um indício de comportamento discriminatório dos empregadores contra as mulheres. Uma forma de racionalizar esse comportamento é o fato de que os tomadores de decisão no interior das firmas antecipam possíveis diferenças médias no grau de comprometimento com o mercado de trabalho entre homens e mulheres. E, por conseguinte, tomem decisões que favoreçam a evolução da carreira dos funcionários do sexo masculino em detrimento a dos empregados do sexo feminino. Consoante ao declarado por Bertrand, Goldin & Katz (2010, p. 236) de que “os salários de mercado são uma medida fundamental para resumir o sucesso

e progresso na carreira profissional”, o comportamento discriminatório dos empregadores – sob essa ótica – obrigatoriamente deve se refletir em média no padrão de rendimentos diferenciados entre gêneros.

A partir da comparação das estimativas do diferencial de rendimentos entre gêneros antes e depois da inclusão de variáveis de controle para escolaridade também presentes na tabela 1.4, nós observamos uma ampliação da magnitude desse diferencial que é crescente com a faixa etária das amostras. Por exemplo, para a amostra de trabalhadores mais jovens a inclusão das variáveis de educação implica um aumento de 2,2 pontos percentuais no módulo da estimativa de discriminação. Para os indivíduos entre 23 e 27 anos, esse incremento é de sete pontos percentuais. Já para os trabalhadores com idade superior a 28 anos o acréscimo, em termos absolutos, no diferencial entre gêneros derivado do fato de condicionarmos a análise ao nível de escolaridade dos trabalhadores se estabiliza em aproximadamente dez pontos percentuais¹³.

De maneira complementar ao verificado na tabela 1.2, novamente na tabela 1.4 nós constatamos a pequena influência, quando existente, que as variáveis de história profissional no mercado de trabalho formal têm para explicar o diferencial de salários entre os sexos. No caso dos trabalhadores mais jovens esse efeito é praticamente inexistente e tende a se ampliar apenas marginalmente para os grupos com faixas etárias mais avançadas. Essa redução na discriminação percebida contra as mulheres atinge seu máximo – cinco pontos percentuais – nas faixas etárias mais maduras (comparação entre as colunas XIV e XV; e XVII e XVIII). Esse padrão de redução do diferencial de salários entre gêneros nas faixas etárias mais avançadas, após a inclusão de controles para o histórico profissional, está em consonância com o que seria esperado pelo fato de que a experiência torna-se mais relevante para os trabalhadores com mais tempo de exposição (potencial) ao mercado de trabalho. Mesmo que a magnitude desse efeito seja pequena.

Por último, a tabela A3, no apêndice, lista os resultados do diferencial de rendimentos entre gêneros obtidos a partir das estimações da especificação mais completa das equações de salários por faixas etárias análogas a tabela 1.4. A

¹³ O padrão de variação do diferencial motivado pelo controle para a escolaridade/produktividade dos trabalhadores ao longo das faixas etárias é obtido por meio da comparação entre as colunas I e III, IV e VI, VII e IX no painel A e X e XII, XIII e XV, XVI e XVIII no painel B da tabela 1.4.

diferença é a incorporação de um vetor de covariadas para as famílias ocupacionais dos trabalhadores como controle nas regressões. O principal padrão de resultados, associado à existência de um efeito idade, mantém-se, apesar de uma redução na magnitude desse efeito.

1.4.2

O diferencial de salários entre gêneros e sua evolução ao longo do ciclo de vida para trabalhadores com históricos profissionais similares e alto comprometimento no mercado de trabalho

Conforme discutido na seção de metodologia deste capítulo, o uso de informações mais precisas sobre a história profissional dos indivíduos nos permite superar as deficiências anteriormente encontradas pelos demais estudos empíricos acerca do diferencial de rendimentos entre os sexos. Entretanto, o uso e, por conseguinte os resultados derivados da análise com auxílio dessas informações dependem em grande medida da estratégia empírica utilizada. Na subseção anterior, nós exploramos um desses métodos, o que nos auxiliou a lançar luz sobre alguns pontos importantes relacionados ao tema aqui estudado. Contudo, se o histórico profissional dos indivíduos, reconstruído a partir das informações do mercado de trabalho formal brasileiro, ainda for imperfeito e essas falhas apresentarem padrões médios distintos entre gêneros, então temos uma limitação sobre a estratégia empírica anteriormente utilizada.

Com o intuito de complementar os resultados apresentados acima, na presente subseção nós buscamos comparar trabalhadores tão semelhantes quanto o possível em relação a sua experiência profissional. Para isto, nós definimos a amostra utilizada nessa seção como o conjunto de indivíduos com vínculo empregatício ativo no mercado de trabalho formal em todos os anos entre 1994 e 2009¹⁴. Estes são denominados de trabalhadores ‘continuamente empregados’ e excluímos assim o efeito de diferentes experiências de mercado de trabalho sobre os resultados obtidos. Dessa forma, para cada ano, estimamos o diferencial de rendimentos entre gêneros para esse grupo de trabalhadores para os quais a

¹⁴ Nessa etapa nós não restringimos os vínculos empregatícios quanto a sua duração em cada ano observado. Dessa forma, dada as limitações impostas para a construção da amostra inicial – apresentada na seção de dados – os vínculos empregatícios têm entre quatro e doze meses de duração em cada ano.

motivação e o grau de permanência no mercado de trabalho são bastante alto. Por último, nós confrontamos a evolução do diferencial de rendimentos entre gêneros dessa amostra em particular com o observado para a amostra como um todo de trabalhadores do mercado de trabalho formal no Brasil ao longo dos dezesseis anos.

A figura 1.1 apresenta, por meio de gráficos, o módulo das estimativas do diferencial de salários entre gêneros obtidas a partir da estimação das equações de determinação de salários (1.2) e (1.3) para cada ano da amostra da RAIS e para as quatro coortes descritas na subseção 1.3.2. Essas coortes são definidas pelos trabalhadores nascidos entre 1972 e 1976, 1967 e 1971, 1962 e 1966, além de 1957 e 1961. As linhas pontilhadas representam a evolução do diferencial de rendimentos bruto entre os sexos, ou seja, quando estimamos a equação (1.2) com a inserção de controles apenas para as idades dos trabalhadores. Já as linhas cheias traçam a tendência das estimativas do diferencial de salários entre gêneros ao utilizarmos a equação (1.3) que incorpora adicionalmente controles para as características produtivas dos indivíduos, tais como: escolaridade, jornada de trabalho mensal, se trabalha em jornada parcial (menos do que 30 horas semanais), Estado da Federação e se permaneceu empregado durante o ‘ano inteiro’. A coloração vermelha refere-se à amostra de trabalhadores definidos como ‘continuamente empregados’ no mercado de trabalho formal brasileiro. Cada ponto que constitui essas linhas refere-se ao módulo da estimativa do valor percentual médio aproximado que as mulheres recebem a menos do que os homens, igualmente comparáveis, como salário mensal médio real, nas respectivas faixas etárias associadas a cada uma das quatro coortes para cada um dos dezesseis anos. Esses resultados, assim como o número de observações que constitui cada amostra, são apresentados nas tabelas A4, A5, A6 e A7 do apêndice. De acordo com o apresentado nessas tabelas, para a amostra de trabalhadores ‘continuamente empregados’ a composição do grupo de indivíduos presentes em cada ano é mantida constante, o que exclui problemas decorrentes de possíveis taxas diferenciadas de entrada e saída no mercado de trabalho entre os sexos. De maneira complementar, as linhas azuis nos gráficos da figura 1.1 representam as estimativas para a amostra geral de trabalhadores pertencentes à respectiva coorte, por faixa etária, em cada ano.

Figura 1.1 – Evolução do diferencial de salários entre os gêneros para algumas coortes

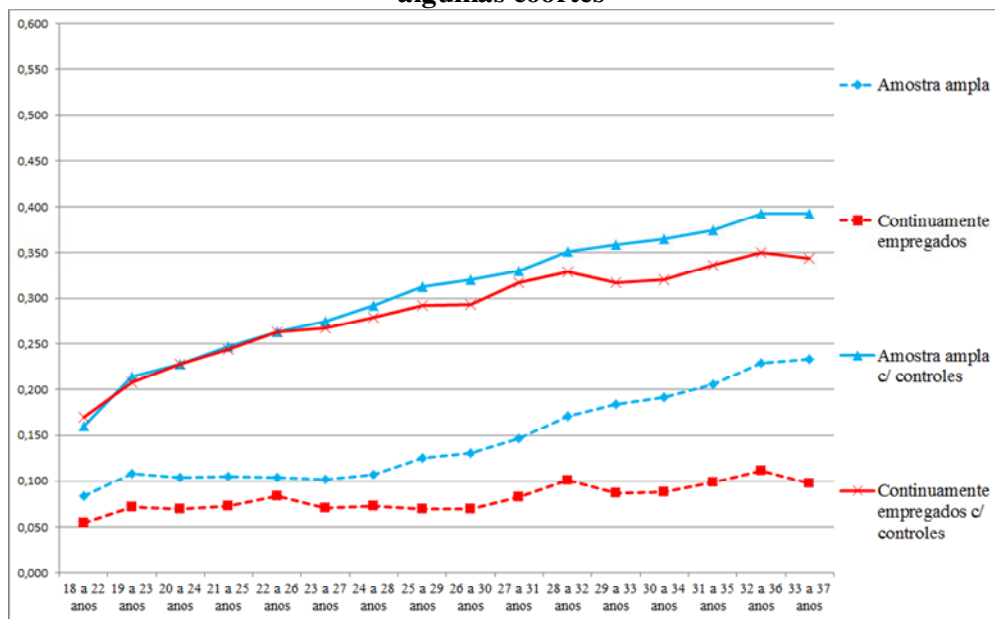


Gráfico 1.1.1: Coorte 1972 – 1976.

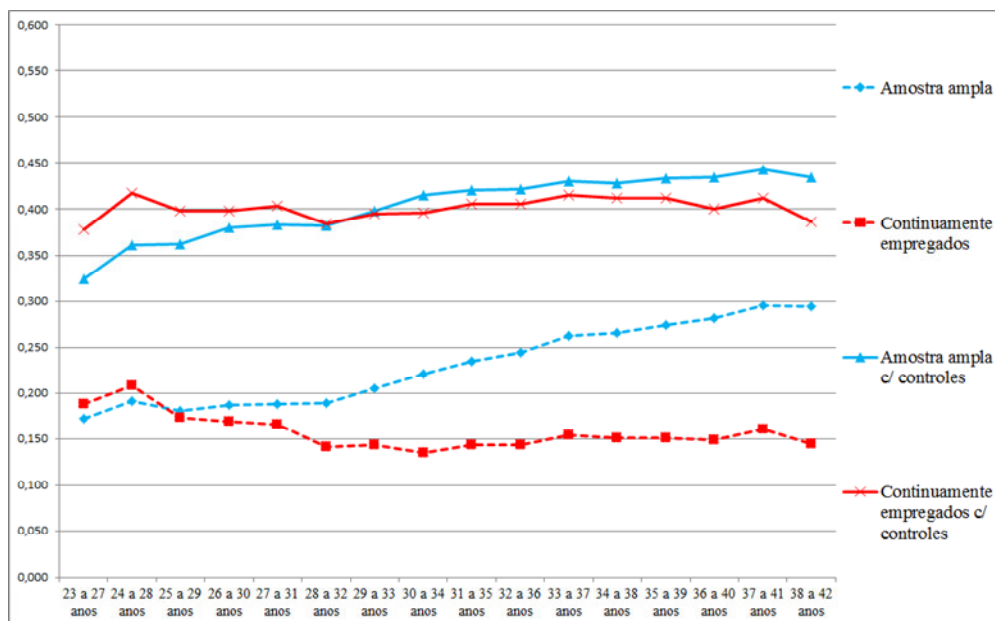


Gráfico 1.1.2: Coorte 1967 – 1971.

Fonte: Elaboração própria, dados da RAIS/MTE 1994 a 2009.

Figura 1.1 – Evolução do diferencial de salários entre os gêneros para algumas coortes (continuação)

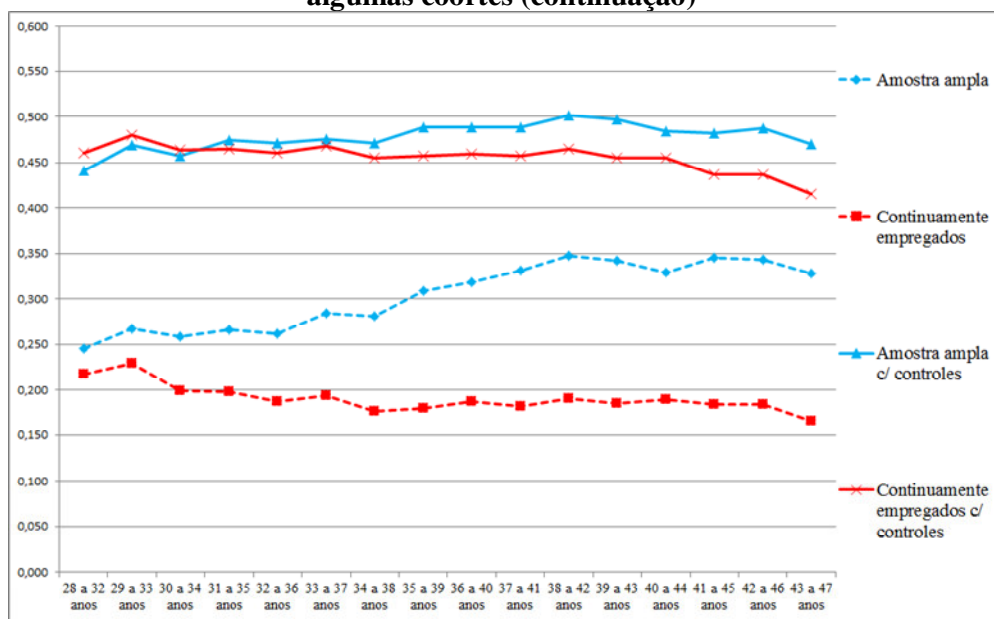


Gráfico 1.1.3: Coorte 1962 – 1966.

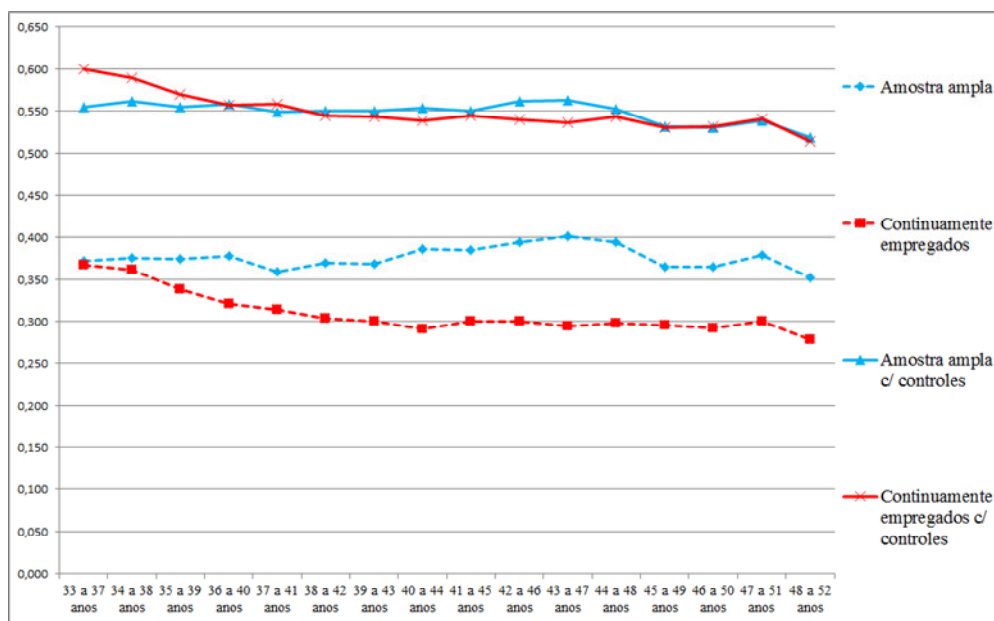


Gráfico 1.1.4: Coorte 1957 – 1961.

Fonte: Elaboração própria, dados da RAIS/MTE 1994 a 2009.

A comparação da evolução dos diferenciais de rendimentos brutos (linhas pontilhadas) para as coortes mais novas – 1972 a 1976 e 1967 a 1971 – mostra que a discriminação percebida contra as mulheres para a amostra dos indivíduos ‘continuamente empregados’ é inferior. Além disso, esta se mantém em patamares

reduzidos ao longo do período observado, enquanto que para a amostra geral ocorre um crescimento do diferencial com o envelhecimento das coortes. Esse padrão pode ser observado por meio do descolamento entre as linhas pontilhadas vermelha ('continuamente empregados') e azul (amostra geral) ao longo das faixas etárias. Isso poderia ser uma evidência de que para o grupo de trabalhadores com alto grau de comprometimento e permanência no mercado de trabalho o diferencial de salários entre gêneros não refletiria um comportamento discriminatório dos empregadores, mas sim diferenças de produtividade entre os grupos. Visto que esse diferencial (bruto) já se fazia presente no início do período observado e não sofre variação relevante em comparação com a evolução observada para a amostra geral. Já para o grupo de trabalhadores pertencentes às coortes mais antigas a conclusão acima exposta a respeito do comportamento discriminatório não é válida, pois não observamos possíveis diferenças de experiência prévia dos trabalhadores no início do ciclo de vida.

Todavia, após a inclusão das variáveis de controle para as características produtivas individuais, em especial as faixas de escolaridade, o padrão de evolução das estimativas do diferencial de salários entre gêneros converge entre a amostra geral e a dos indivíduos 'continuamente empregados'. A comparação agora deve ser feita entre as linhas vermelha e azul cheias. Complementarmente, assim como documentado nas tabelas 1.2 e 1.4, os diferenciais de rendimentos apresentam um grande incremento condicional à escolaridade/produtividade dos trabalhadores comparados às estimativas do diferencial bruto.

Assim, ao restringirmos nossa análise às linhas cheias dos gráficos da figura 1.1, destacamos uma tendência fortemente crescente do valor absoluto do diferencial de rendimentos entre gêneros para a coorte de indivíduos nascidos entre 1972 e 1976 (gráfico 1.1.1 da figura 1.1). Esses são os trabalhadores com as idades mais tenras para os quais é possível analisar o comportamento no mercado de trabalho ao longo dos dezesseis anos compreendidos pela amostra da RAIS. Justamente por esse motivo, essa coorte de indivíduos é especialmente importante, pois a questão da experiência prévia não observada, anterior ao ano de 1994, deve ter pequena relevância (se esta não for nula). Deste modo, no início do painel, nós observamos os trabalhadores da coorte 1972-1976 iniciando sua história profissional no mercado de trabalho formal brasileiro.

O padrão ascendente do diferencial de rendimentos é muito similar tanto para a amostra geral quanto para a amostra de trabalhadores ‘continuamente empregados’. Para a coorte de 1972-1976, o diferencial adverso às mulheres é de aproximadamente 16% entre os 18 e 22 anos e cresce, em termos absolutos, monotonicamente com o envelhecimento desse grupo de trabalhadores. Entre os 20 e 24 anos, o diferencial já atingiu o patamar de 22%, ultrapassa os 26% na faixa etária entre 22 e 26 anos e chega a aproximadamente 30% antes dos 30 anos de idade para as duas amostras analisadas. Vale ressaltar que o nível do diferencial entre gêneros é um pouco inferior para os trabalhadores ‘continuamente empregados’ em comparação com a amostra geral, porém a evolução das estimativas é muito semelhante até aproximadamente os 30 anos de idade. Apenas na segunda metade do período avaliado para a coorte 1972 a 1976, se observa uma discrepância no padrão de evolução das estimativas. De tal maneira que o diferencial para os trabalhadores da amostra geral cresce até os 39% e as estimativas para os indivíduos ‘continuamente empregados’ permanecem aproximadamente quatro pontos percentuais inferiores ao do primeiro grupo. Esse resultado nos permite concluir que apesar de um maior engajamento com o mercado de trabalho, este comportamento não garante às mulheres ‘continuamente empregadas’ benefícios significativos em relação à média das mulheres presentes no mercado de trabalho formal brasileiro.

Para as coortes mais antigas os diferenciais de salários entre gêneros tendem a se estabilizar após os 35 anos de idade. Além disso, o padrão de evolução das estimativas das diferenças de rendimentos entre homens e mulheres são bastante similares para ambas as amostras, no caso da especificação que condiciona para as características produtivas individuais – equação (1.3). Para a coorte de indivíduos nascidos entre 1967 e 1971 (gráfico 1.1.2 da figura 1.1) o diferencial entre gêneros (contrário às mulheres) estabiliza-se em torno de 40% a 43%. Para os trabalhadores nascidos entre 1962 e 1966 (gráfico 1.1.3), a estabilidade ocorre entre 45% e 50%. Já para a coorte de 1957 a 1961 (gráfico 1.1.4), o patamar de estabilidade ocorre ao redor dos 55%. Novamente nós verificamos uma tendência de diferenciais entre gêneros maiores para as coortes mais antigas, pelo menos na faixa de estabilidade desses diferenciais ao longo do ciclo de vida produtivo. Contudo a magnitude dessa variação no diferencial entre as coortes mais velhas mostrou-se apenas moderada. De maneira distinta aos exercícios empíricos

realizados na seção 1.4.1 desse capítulo, nos gráficos da figura 1.1, nós mantemos constante o efeito coorte e a evolução do diferencial de rendimentos entre gêneros decorre dos efeitos idade e tempo.

Sobretudo, para as coortes mais novas analisadas, fica evidente a existência de um importante efeito idade para explicar o crescimento da discriminação percebida contra as mulheres no mercado de trabalho formal na fase inicial do ciclo de vida produtivo. E esse efeito se faz presente mesmo para a amostra de trabalhadores ‘continuamente empregados’ que por hipótese são mais comprometidos com o mercado de trabalho. Portanto, essas evidências para a coorte mais nova corroboram o padrão verificado na seção 1.4.1 de coincidência entre o forte crescimento dos diferenciais entre gêneros e um período crítico para as mulheres acerca de suas decisões sobre família e fecundidade, ou seja, o início do ciclo da vida produtivo. Além disso, o padrão ascendente do diferencial se mantém mesmo para o grupo de indivíduos ‘continuamente empregados’, isto é, composto por mulheres que apesar de atravessarem esse período crítico de sua vida produtiva não apresentam interrupções significativas na sua oferta de trabalho no mercado formal. O que reforça a ideia de existência de um componente discriminatório desfavorável às mulheres por parte dos empregadores no diferencial de rendimento entre gêneros estimado para o Brasil.

A princípio, para obtermos conclusões precisas sobre o padrão de evolução dos diferenciais de salários entre gêneros ao longo de todo o ciclo de vida dos trabalhadores, seria ideal acompanhar cada coorte desde suas idades mais tenras até o período em que se observam as decisões sobre aposentadoria. No entanto, nossa amostra não dispõe de uma dimensão longitudinal tão extensa, pois observamos a participação dos indivíduos no mercado de trabalho apenas entre 1994 e 2009. Dessa forma, devemos restringir nossa análise a frações da trajetória da estimativa de interesse para cada coorte. Adicionalmente, se adotarmos o pressuposto de uma parcial estabilidade do padrão de diferenciais de salários entre gêneros para diferentes coortes, nós podemos realizar uma análise aproximada da evolução desse diferencial ao longo do ciclo de vida. A figura 1.2 explicita o esforço de colocar em um mesmo gráfico a evolução do diferencial de salários entre gêneros por faixa etária para as quatro diferentes coortes de trabalhadores anteriormente descritas. Os resultados apresentados no gráfico 1.2.1 da figura 1.2 referem-se ao módulo das estimativas dos coeficientes associados a variável

indicadora de sexo em regressões para a amostra geral que adotam a especificação da equação (1.3), ou seja, condicionando para as características produtivas individuais. Já os resultados para a amostra de indivíduos ‘continuamente empregados’ é exibido no gráfico 1.2.2 da mesma figura.

A partir do gráfico 1.2.1 da figura 1.2, nós constatamos uma clara tendência ascendente, em termos absolutos, do diferencial de salários entre homens e mulheres na fase inicial do ciclo de vida produtivo dos trabalhadores da amostra geral. Esse padrão é monotônico até aproximadamente os 35 anos de idade. Posteriormente, nós percebemos uma estabilização dessa medida média de discriminação no mercado de trabalho e um suave decaimento no final do ciclo de vida dos trabalhadores. Os resultados que constituem o gráfico 1.2.1, entretanto, estão sujeitos a variações na composição da amostra de trabalhadores em cada coorte ao longo dos anos. Esse ponto deve ser especialmente problemático na comparação entre homens e mulheres no final do ciclo de vida, ou seja, nas idades mais avançadas. Pois uma possível seleção positiva das mulheres mais produtivas e, por conseguinte com maiores salários, nessa fase do ciclo de vida poderia guiar as evidências de suave redução do diferencial entre os sexos nessas faixas etárias. Enquanto a taxa de permanência no mercado de trabalho permanece elevada para os homens de maneira geral.

Analogamente, o gráfico 1.2.2 da figura 1.2 representa graficamente os resultados para a amostra dos trabalhadores ‘continuamente empregados’ e a especificação do modelo de determinação de salários dado pela equação (2.3). Nesse caso em particular, o efeito composição da amostra não existe e se mantém a tendência de suave queda no diferencial de salários entre gêneros no final do ciclo de vida dos indivíduos – após os quarenta anos de idade.

Figura 1.2 – Evolução dos diferenciais de salários entre os gêneros para as coortes 1972 a 1976, 1967 a 1971, 1962 a 1966 e 1957 a 1961

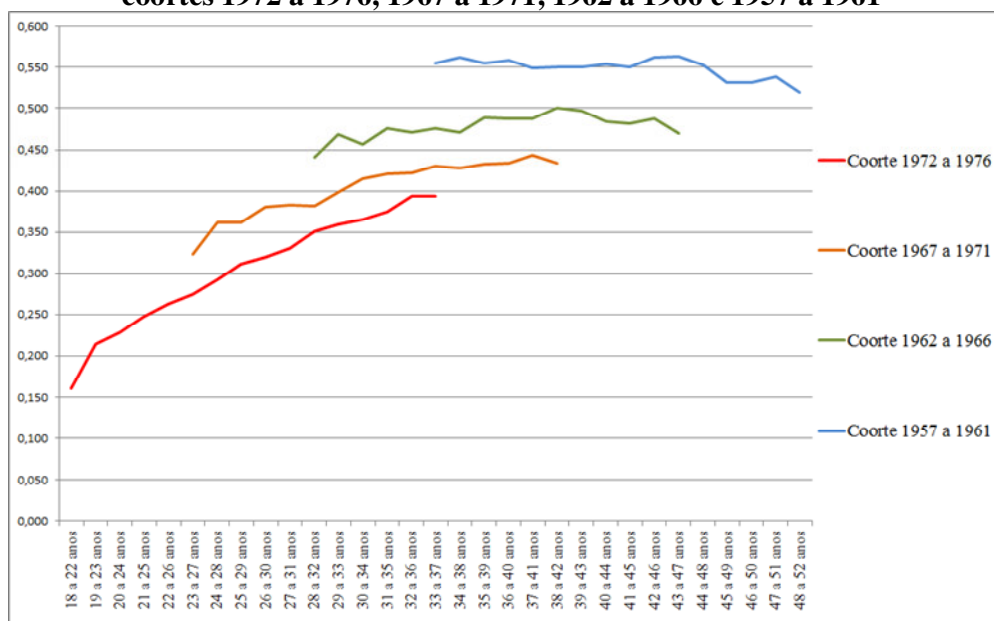


Gráfico 1.2.1 – Amostra geral: modelos de determinação de salários condicionando para as características produtivas individuais.

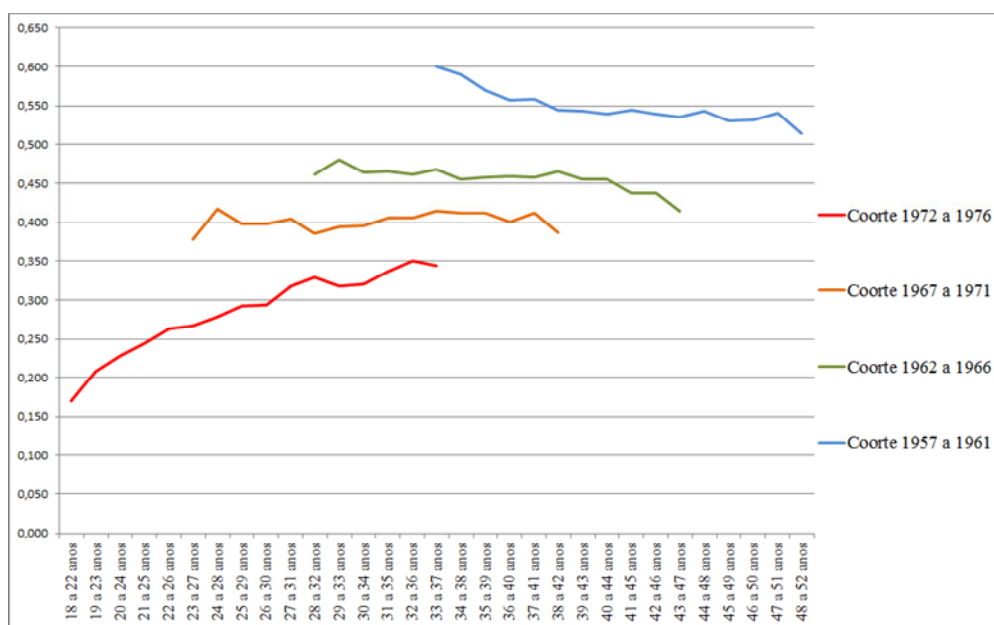


Gráfico 1.2.2 – Amostra de indivíduos 'continuamente empregados': modelos de determinação de salários condicionando para as características produtivas individuais.

Fonte: Elaboração própria, dados da RAIS/MTE 1994 a 2009.

Para esse grupo de trabalhadores com elevado engajamento no mercado de trabalho – 'continuamente empregados', a tendência ascendente do diferencial até os 35 anos é verificada apenas para os trabalhadores pertencentes à coorte de

1972-1976. Já para as amostra de indivíduos das coortes nascidas entre os anos de 1967 a 1971 e 1962 a 1966, esse padrão crescente das diferenças salariais entre gêneros não se faz mais presente. Os trabalhadores dessas duas últimas coortes começam a ser observados na amostra da RAIS, respectivamente com faixas etárias entre 23 e 27 anos e 28 e 32 anos. Essas são fases do ciclo de vida profissional, nas quais a experiência previamente acumulada no mercado pode ter um papel importante para explicar a trajetória profissional diferenciada em média entre os sexos. Dada a restrição na dimensão temporal da base de dados da RAIS/MTE, nós não temos como condicionar a análise ao histórico profissional desses indivíduos anteriormente a 1994.

De maneira complementar, com o propósito de investigar detalhadamente o padrão de evolução do diferencial de salários entre gêneros no início do ciclo de vida, nós reproduzimos os procedimentos de estimação para as amostras geral e de trabalhadores ‘continuamente empregados’ com base nas equações (1.2) e (1.3) para coortes mais novas. Essas coortes são compostas por indivíduos nascidos entre os anos de 1974 a 1976 e que tinham 18 a 20 anos em 1994; de 1977 a 1979 e são observados com a mesma faixa etária a partir de 1997; e por último de 1980 a 1982, inicialmente observados na amostra da RAIS apenas em 2000. Assim, essa análise complementar é baseada em coortes de indivíduos definidas por um menor intervalo de número de anos e para duas dessas coortes nós observamos sua história profissional em intervalos mais curtos (treze e dez anos). A figura 1.3 ilustra a evolução do valor absoluto do diferencial médio de rendimentos entre gêneros para essas três coortes separadamente para a amostra geral e dos ‘continuamente empregados’. A importância dessa análise refere-se a restringirmos sempre a grupos de trabalhadores que são observados inicialmente com pouca idade e, por conseguinte reduzida ou inexistente experiência no mercado de trabalho formal. Logo, podemos observar a evolução do diferencial conforme o grupo de trabalhadores passa pela exposição potencial ao mercado de trabalho e acumula – de maneira diferenciada ou não entre os sexos – experiência. Sem a influência da experiência profissional passada sobre o histórico profissional analisado.

Figura 1.3 – Evolução do diferencial de salários entre os gêneros para coortes jovens: 1974 a 1976, 1977 a 1979 e 1980 a 1982

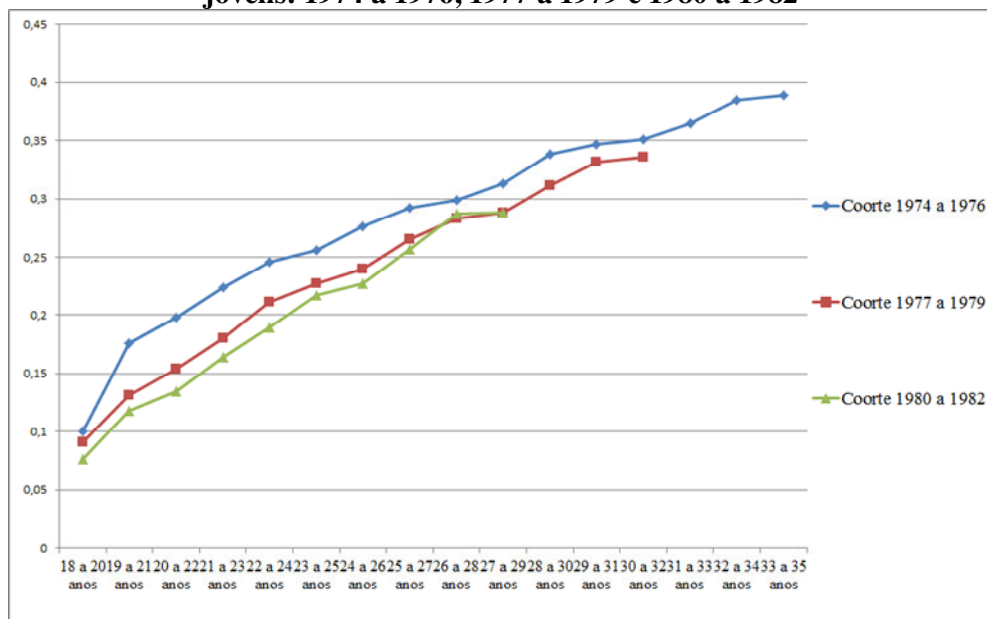


Gráfico 1.3.1: Amostra geral: modelo de equação de determinação de salários condicionando para as características produtivas.

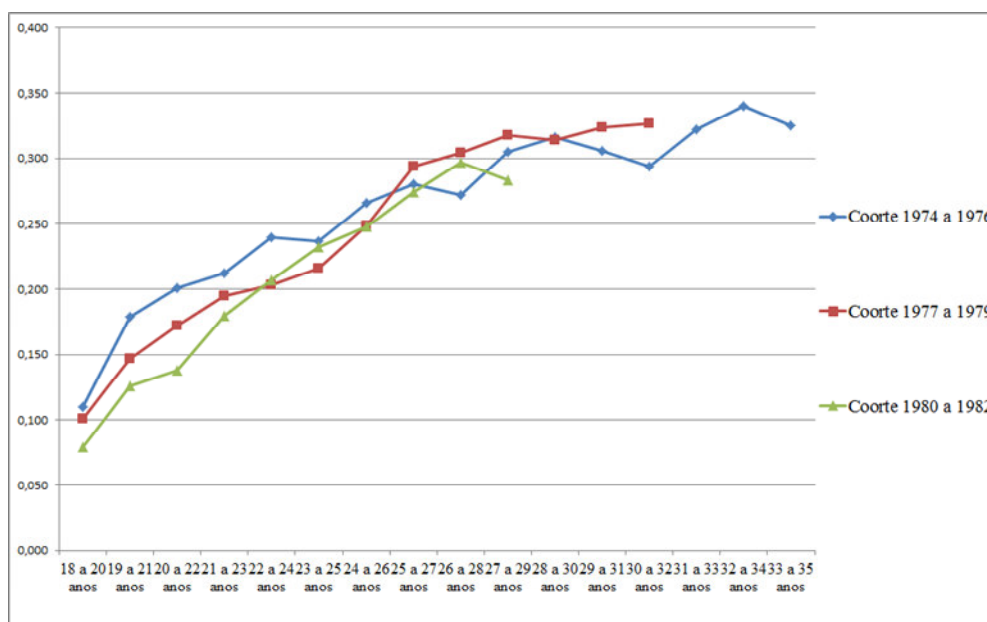


Gráfico 1.3.2: Amostra de trabalhadores 'continuamente empregados': modelo de equação de determinação de salários condicionando para as características produtivas.

Fonte: Elaboração própria, dados da RAIS/MTE 1994 a 2009.

O gráfico 1.3.1, da figura 1.3, ilustra, para a amostra geral, que o diferencial de salários entre gêneros é monotonicamente ascendente ao longo de todas as faixas etárias observadas no início do ciclo de vida produtivo, ou seja, entre os 18

e 35 anos de idade¹⁵. Esse padrão é bastante semelhante para as três coortes estudadas. No entanto, essa amostra está sujeita a questões de alterações na sua composição e acúmulo diferenciado de experiência média entre os sexos, o que poderia induzir os resultados observados.

O padrão ascendente na evolução do diferencial de rendimentos entre gêneros no início do ciclo de vida é mantido mesmo quando restringimos a amostra a um grupo de trabalhadores constante entre os diferentes anos e muito similares quanto ao elevado grau de comprometimento e estabilidade no mercado de trabalho formal – gráfico 1.3.2 da figura 1.3. Assim, tanto as questões referentes à composição da amostra quanto às relacionadas ao acúmulo diferenciado de experiência entre os sexos são minoradas. Isto gera uma maior robustez para a evidência de crescimento da discriminação percebida contra as mulheres na fase inicial e crítica de suas carreiras profissionais, em consequência a aspectos relacionados a escolhas pessoais e de família. Além disso, como estamos realizando uma comparação intra-coortes e o padrão persiste para diferentes coortes de indivíduos jovens e com pouca experiência no mercado de trabalho, podemos concluir que existe um efeito idade importante no aumento do diferencial de salários entre gêneros na fase inicial do ciclo de vida. E esse padrão de comportamento não se deve exclusivamente aos efeitos coorte e/ou tempo.

1.4.3

Decomposição do diferencial de salários entre gêneros: componente explicado e não explicado.

A partir dos resultados anteriores de crescente importância da discriminação contra as mulheres no início do ciclo de vida dos trabalhadores, nós buscamos separar a parcela do diferencial de salários decorrente de diferenças nas características médias entre trabalhadores, da fração oriunda de diferenças nos retornos a esses atributos. Esse procedimento de decomposição de Oaxaca-Blinder foi realizado para uma amostra de trabalhadores jovens e com reduzida experiência prévia no mercado de trabalho.

¹⁵ As estimativas pontuais, desvios-padrão e número de observações referentes aos gráficos da figura 3 são listados nas tabelas A8 e A9 do Apêndice.

Para isto, conforme descrito na seção de estratégia empírica, nós utilizamos uma amostra de trabalhadores que em 1996 tinham entre 18 e 25 anos de idade e que não estavam presentes no mercado de trabalho formal brasileiro nos dois anos anteriores. Para esses trabalhadores, nós estimamos separadamente para homens e mulheres equações de determinação de salários com efeitos-fixo individuais de acordo com (1.4) e (1.5). Ou seja, primeiramente não incluímos as variáveis de histórico profissional e posteriormente nós condicionamos a análise a estas. As informações utilizadas nesse exercício empírico foram recuperadas para cada indivíduo pertencente à amostra em questão no intervalo entre 1996 e 2009. Isto significa que acompanhamos trabalhadores com reduzida (ou inexistente) experiência prévia no mercado de trabalho ao longo dos quatorze anos seguintes e utilizamos apenas a variação intra-trabalhador para recuperar as estimativas dos coeficientes presentes nas equações (1.4) e (1.5). A tabela 1.5 apresenta os resultados obtidos.

Em linhas gerais, cada ano extra de experiência no mercado de trabalho formal tem um impacto positivo e estatisticamente significativo sobre os rendimentos individuais, apesar desse efeito não ser linear com a experiência acumulada. Na realidade, conforme demonstrado nas estimativas negativas e estatisticamente significativas do coeficiente da variável de experiência ao quadrado, os retornos são crescentes a uma taxa decrescente. A magnitude desse impacto é muito próxima tanto para homens quanto para mulheres – incrementos de rendimento de aproximadamente 6,5% por ano extra de experiência. Já no caso das interrupções na oferta de trabalho dos indivíduos no mercado formal as mulheres sofrem em média penalizações de magnitude superior às experimentadas pelos homens. Para o caso de trabalhadores jovens e sem experiência prévia no mercado de trabalho formal brasileiro, uma interrupção na sua oferta de trabalho no período posterior a 1996 implica uma redução de 9,6% nos rendimentos dos homens, enquanto para as mulheres esse impacto é de 12,2%. Já a ocorrência de duas interrupções leva a um incremento, em módulo, de aproximadamente cinco e nove pontos percentuais sobre essas estimativas de penalização no mercado de trabalho, respectivamente para trabalhadores do sexo masculino e feminino. Isto é, além de as mulheres sofrerem um maior patamar de redução salarial média decorrente de interrupções na sua oferta de trabalho, essa ainda é crescente com o número de interrupções para o sexo feminino.

Tabela 1.5 – Equação de determinação de salários com modelo de efeito-fixos para trabalhadores estimados separadamente para homens e mulheres: período 1996 a 2009

Ln salário médio horário real	Efeito Fixo por trabalhador			
	Especificação (4)		Especificação (5)	
	Homem	Mulher	Homem	Mulher
Idade	0.0491 (0.000625)***	0.0367 (0.000758)***	0.0260 (0.00252)***	0.0225 (0.00285)***
Experiencia			0.0681 (0.00312)***	0.0635 (0.00371)***
Experiência ao quadrado			-0.00245 (0.000118)***	-0.00266 (0.000158)***
1 Interrupção			-0.0959 (0.00944)***	-0.122 (0.0119)***
2 Interrupções			-0.146 (0.0157)***	-0.208 (0.0194)***
3 ou mais Interrupções			-0.217 (0.0252)***	-0.266 (0.0305)***
Quarta série completa	-0.0638 (0.0123)***	0.0307 (0.0258)	-0.0508 (0.0124)***	0.0301 (0.0256)
Quinta a oitava série incompleta	-0.0839 (0.0122)***	0.0207 (0.0242)	-0.0683 (0.0124)***	0.0212 (0.0239)
Oitava série completa	-0.124 (0.0128)***	-0.0458 (0.0239)*	-0.108 (0.0131)***	-0.0428 (0.0238)*
Ensino médio incompleto	-0.120 (0.0146)***	-0.0394 (0.0248)	-0.101 (0.0149)***	-0.0316 (0.0248)
Ensino médio completo	-0.0912 (0.0144)***	-0.0271 (0.0247)	-0.0785 (0.0146)***	-0.0259 (0.0248)
Ensino superior incompleto	0.0727 (0.0222)***	0.141 (0.0297)***	0.0807 (0.0224)***	0.133 (0.0297)***
Ensino superior completo	0.370 (0.0248)***	0.416 (0.0309)***	0.381 (0.0248)***	0.414 (0.0309)***
Jornada	-0.00272 (0.000348)***	-0.000374 (0.000288)	-0.00271 (0.000344)***	-0.000466 (0.000285)
Tempo parcial	-0.654 (0.0450)***	-0.278 (0.0324)***	-0.645 (0.0446)***	-0.273 (0.0322)***
Ano Integral	-0.0436 (0.00305)***	-0.0316 (0.00375)***	-0.0528 (0.00305)***	-0.0416 (0.00377)***
Controles:				
Estados	Sim	Sim	Sim	Sim
Intercepto	6.208 (0.0659)***	5.940 (0.0644)***	6.558 (0.0797)***	6.137 (0.0828)***
Numero observações	132,095	79,994	132,095	79,994
R-quadrado	0.253	0.210	0.267	0.227
Número trabalhadores	14,611	9,584	14,611	9,584

Fonte: Elaboração própria, dados da RAIS/MTE 1994 a 2009.
Erros padrão entre parênteses - *** $p < 0,01$; ** $p < 0,05$; * $p < 0,1$.

A tabela 1.5 ainda revela outro aspecto interessante, após a inclusão de variáveis relativas ao histórico dos trabalhadores no mercado de trabalho nas equações de salários, as estimativas associadas ao efeito da idade sobre os

rendimentos individuais sofre uma redução pela metade em comparação com os resultados derivados do uso da equação (1.4).

A amostra de jovens trabalhadores com reduzida ou inexistente experiência profissional, a que se refere essa seção, apresenta um diferencial médio de salários desfavorável às mulheres de aproximadamente 10,7%, conforme apresentado na linha 1 da tabela 1.6. Essa tabela exhibe os resultados da decomposição de Oaxaca-Blinder entre os componentes composição (explicado) e não explicado – linhas 2 e 3 – de acordo com a equação (1.6). Além disso, a tabela 1.6 também apresenta a contribuição das variáveis explicativas dos modelos de determinação de salários para cada um desses efeitos.

A linha 2 da tabela 1.6 mostra que, tanto na especificação sem medidas para o histórico profissional quanto com a inclusão dessas variáveis, as mulheres deveriam receber um salário em média 2% superior aos dos homens na nossa amostra, caso o retorno às características produtivas fossem iguais entre homens e mulheres. Em média, as mulheres têm um maior nível de escolaridade e a diferença nessa variável a favor do sexo feminino corresponde a mais do que 100% do componente explicado do diferencial de salários entre gêneros, que é favorável aos trabalhadores do sexo feminino – linha 5 da tabela 1.6. A contribuição das diferenças nas características pouco se altera com a inclusão das variáveis de experiência profissional na equação de determinação de salários individuais, com efeito-fixo para o trabalhador. A exceção é a redução pela metade na contribuição da variável idade (linha 4 da tabela 1.6) sobre o componente explicado do diferencial, que é compensada principalmente pela contribuição de um menor acúmulo de número de anos de experiência no mercado de trabalho por parte das mulheres – linha 10 da tabela 1.6.

Tabela 1.6 – Decomposição do diferencial de salários entre homens e mulheres para modelo de painel com efeito fixo de trabalhadores: período 1996 a 2009

		Mulheres vs. Homens	
1	Diferença no Ln salário horário	-0,1069	
		Sem experiência	Com experiência
Parcela da diferença devido:			
2	Características Médias	0,0243	0,0212
3	Coefficientes	-0,1301	-0,1272
Diferenças associadas às características			
4	Idade	-0,0120	-0,0064
5	Escolaridade	0,0291	0,0291
6	Estados (mercados locais)	0,0115	0,0113
7	Jornada mensal trabalho	0,0112	0,0112
8	Tempo parcial	-0,0154	-0,0152
9	Ano Integral	-0,000056	-0,000068
10	Experiência	-----	-0,0091
11	Interrupções	-----	0,0004
Diferenças associadas aos coeficientes			
12	Idade	-0,3240	-0,0915
13	Escolaridade	0,0687	0,0565
14	Estados (mercados locais)	-0,0216	-0,0198
15	Jornada mensal trabalho	0,3959	0,3787
16	Tempo parcial	0,0122	0,0121
17	Ano Integral	0,0066	0,0061
18	Experiência	-----	-0,0365
19	Interrupções	-----	-0,0120
20	Intercepto	-0,2680	-0,4210

Fonte: Elaboração própria, dados da RAIS/MTE 1994 a 2009.

No entanto, a diferença nas estimativas dos retornos às características produtivas entre os sexos é favorável aos homens e de magnitude muito superior ao efeito composição (explicado). Esse fato resulta em um diferencial de salários final desfavorável às mulheres. Na nossa amostra a diferença de rendimentos associado ao componente não explicado, ou seja, associado à discriminação contra as mulheres, é de aproximadamente 13% para as duas especificações consideradas, linha 3 da tabela 1.6. As mulheres com a mesma escolaridade dos homens recebem em média 5,5% a 7% a mais do que estes. É justamente na contribuição das variáveis para o componente não explicado do diferencial de

salários entre gêneros onde observamos as maiores alterações decorrentes da inclusão das medidas da história no mercado de trabalho formal dos indivíduos. Inicialmente, na ausência do histórico profissional na análise, as mulheres com mesma idade dos homens recebem 32,4% a menos que estes. Juntamente com as diferenças no intercepto, essas são as duas principais fontes de discriminação contra as mulheres a partir dos resultados obtidos com a equação de salários (1.4). Todavia após a inclusão da experiência acumulada e número de interrupções na oferta de trabalho entre 1996 e 2009, a diferença não explicada associada à idade permanece contrária às mulheres, mas se reduz para 9,15% (linha 12 da tabela 1.6).

De forma complementar, as mulheres com mesmo número de anos de experiência no mercado de trabalho e número de interrupções na oferta de trabalho recebem, respectivamente, 3,65% e 1,2% a menos que os homens (linhas 18 e 19 da tabela 6). As contribuições dessas duas últimas variáveis correspondem a aproximadamente 29% e 9%, respectivamente, da magnitude do efeito não explicado para o diferencial de rendimentos entre gêneros verificado empiricamente. Assim, a discriminação contra as mulheres diretamente associada ao acúmulo de anos de experiência no mercado de trabalho formal e ao número de interrupções na oferta de trabalho, somadas (-4,85%), mais do que anulam o ganho médio de rendimentos (2%) que as mulheres deveriam receber devido às diferenças nas características produtivas que são favoráveis a estas. Mais uma vez, observa-se uma forte evidência de comportamento discriminatório associado ao padrão de oferta de trabalho diferenciado entre homens e mulheres no início de seu ciclo de vida produtivo.

1.5 Conclusões

Os nossos resultados mostram a existência de um importante diferencial de salários entre gêneros, desfavorável às mulheres, no mercado de trabalho formal brasileiro. Esse diferencial de rendimentos é monotonicamente crescente em valor absoluto na fase inicial do ciclo de vida produtivo dos trabalhadores, até 35 – 40 anos. Justamente, um período crítico no qual as mulheres se defrontam com

decisões relacionadas a seu ciclo de fertilidade e enfrentam as consequências de suas escolhas sobre a oferta de trabalho. Além disso, o histórico profissional dos indivíduos – reconstruído a partir das informações da RAIS/MTE – tem uma pequena influência para explicar esse diferencial observado. A redução do diferencial de rendimentos contra as mulheres é de 10% após a inclusão de controles para a experiência prévia. Por outro lado, as evidências sugerem que as mulheres no mercado formal brasileiro são ainda mais discriminadas, condicional a sua produtividade, medida pela escolaridade. Após a inclusão de controles para o nível de escolaridade dos indivíduos, o diferencial observado sofre um incremento de um terço, em termos absolutos, passando aproximadamente de 21% para 33%. Esse padrão é dissemelhante ao observado na literatura internacional para o contexto de outros países.

A reduzida capacidade das medidas de experiência profissional real em explicar o diferencial de salários entre gêneros para o Brasil se dá apesar destas apresentarem um padrão de comportamento em consonância com o esperado. Isto é, os efeitos do histórico profissional são estatisticamente significantes e economicamente relevantes sobre os rendimentos dos trabalhadores. Ademais, períodos de ausência no mercado de trabalho formal reduzem os rendimentos e um maior comprometimento dos trabalhadores está relacionado a maiores salários. A magnitude desses efeitos se reduz com o passar do tempo.

Por último, uma fração significativa do diferencial de rendimentos entre gêneros no mercado de trabalho formal brasileiro está relacionada com as decisões ocupacionais diferenciadas entre homens e mulheres. As evidências mostram que a inclusão de controles para as famílias ocupacionais (CBO2002) é responsável por uma redução, em valores absolutos, de aproximadamente um terço da estimativa do diferencial de salários entre gêneros. Esses resultados ressaltam a relevância de estudos futuros que procurem investigar com maior nível de detalhe as relações existentes entre a escolha ocupacional dos indivíduos e, por conseguinte suas carreiras profissionais e as diferenças de rendimentos decorrentes. Além disso, também é válida a análise pormenorizada do diferencial de rendimentos entre gêneros e sua evolução ao longo do ciclo da vida produtivo para carreiras profissionais particulares, tais como para: profissionais da saúde, advogados, no setor público, cargos de direção e gerência de firmas etc.

2

Medindo os efeitos do professor na sala de aula: evidências a partir da prova de promoção para professores em São Paulo

2.1

Introdução

Diversos estudos recentes têm mostrado a importância dos professores e a existência de substancial variação na qualidade destes mensurada pelo valor adicionado ao desempenho dos alunos (Rockoff, 2004; Rivkin, Hanushek & Kain, 2005; Aaronson, Barrow & Sander, 2007; Jacob & Lefgren, 2008; Kane & Staiger, 2008; Rothstein, 2010 e Chetty, Friedman & Rockoff, 2011). Em geral, esses estudos têm como contexto os ensinos fundamental e médio, e os resultados apontam que, em média, um aumento de um desvio-padrão na qualidade dos docentes está associado a um incremento de 11% de um desvio-padrão da distribuição de desempenho em leitura e 15% para matemática (Hanushek & Rivkin, 2010)¹⁶.

Mas quais são os atributos dos professores responsáveis por essa ampla variação na eficácia observada? Em geral, as características facilmente observáveis tem pouco poder explicativo sobre a variação na qualidade dos docentes. Entre essas características destacam-se: escolaridade, características demográficas, certificação por meio de diplomas, especialização, mestrado, doutorado, experiência, salários, etc. (Wayne & Youngs, 2003; Aaronson, Barrow & Sander, 2007; Hanushek & Rivkin, 2010; Rockoff & Speroni, 2010;). Ultimamente, graças à disponibilização de bases de dados com novas informações, alguns estudos tem procurado investigar o papel de algumas características intrínsecas dos docentes – que não são facilmente observáveis – para explicar essas diferenças de qualidade. Entre essas medidas destacamos o conhecimento específico às disciplinas (Santibañez, 2006; Metzler & Woessmann,

¹⁶ Fabian (2010) ao investigar as demissões de professores de matemática judeus na Alemanha nazista, como uma variação exógena na qualidade das universidades, encontra evidências de que a qualidade da instrução também é bastante importante mesmo em mercados de educação altamente seletivos.

2012), avaliações subjetivas dos docentes (Rockoff & Speroni, 2010) e adoção de práticas pedagógicas nas salas de aula (Lavy, 2011).

Dessa forma, apesar da existência de forte indício de que a qualidade do professor importa para o aprendizado, ainda existe uma lacuna importante acerca de evidências robustas sobre quais atributos ou práticas dos professores são responsáveis pelo aprendizado dos alunos, sobretudo em países em desenvolvimento.

O nosso artigo tem como objetivo preencher esse hiato de conhecimento. Para isto, nós investigamos o impacto de duas dimensões da qualidade dos professores sobre a aprendizagem de alunos da oitava série do ensino fundamental em matemática e língua portuguesa. Em particular, nós focamos no conhecimento específico à disciplina ministrada pelo docente e o conjunto de habilidades deste no interior da sala de aula que fomenta o processo de aprendizagem dos alunos.

Para examinar a relação entre essas duas características dos docentes e a aprendizagem dos alunos, nós utilizamos uma estratégia empírica baseada em um modelo de valor adicionado da função de produção educacional com controle para o desempenho passado¹⁷. Essa análise foi possível pela conjunção de cinco bancos de dados. Assim, temos informações sobre os alunos (notas em testes contemporâneos e passados, etc.), sobre a adoção de práticas pedagógicas pelos docentes na sala de aula, dados administrativos sobre a atribuição dos professores às turmas no ano letivo de 2009 e o desempenho dos professores na avaliação do Sistema de Promoção instituído pela Secretaria de Educação do Estado de São Paulo¹⁸.

Os nossos resultados mostram que tanto o conhecimento específico à disciplina quanto as atividades pedagógicas empregadas pelos docentes nas salas de aula têm um impacto positivo e estatisticamente significativo sobre a aquisição de habilidades cognitivas dos alunos em ambas as disciplinas. Entretanto, o efeito do conhecimento do docente mostrou-se bastante limitado em termos da magnitude dos ganhos de desempenho dos estudantes. Por outro lado, os impactos associados aos professores adotarem sempre práticas pedagógicas eficazes são economicamente importantes para a produção de habilidades cognitivas dos

¹⁷ A nossa especificação principal é bastante semelhante à apresentada e discutida por Rothstein (2010) como VAM2 (*Value-added model 2*).

¹⁸ No restante desse capítulo nós faremos referência a Secretaria de Educação do Estado de São Paulo pela sua sigla SEE/SP.

alunos. Ademais, as evidências indicam que os efeitos dessas práticas pedagógicas realizadas em sala de aula sobre o ganho de proficiência dos alunos são, em grande parte, independentes do nível de conhecimento do professor na disciplina. Os resultados ainda são robustos a diferentes definições da medida de adoção das atividades pedagógicas pelos professores e a inclusão de controles para a qualidade de gestão da unidade escolar.

A interpretação causal dos nossos resultados está fundamentada no arcabouço teórico da função de produção educacional que investiga as relações entre os níveis de insumos educacionais e o desempenho dos alunos (notas em testes). Sob esse ponto de vista, a produção de habilidades cognitivas é um processo cumulativo e que depende de toda a história de insumos educacionais (escolares e familiares) e das dotações individuais herdadas pelas crianças, assim como a interação entre esses fatores (Todd & Wolpin, 2003 e 2007; Cunha & Heckman, 2003; Rothstein, 2010; e Boardman & Murnane, 1979).

De acordo com Todd & Wolpin (2003) para o caso de estudos baseados em dados observacionais sobre alunos e professores é razoável admitir que os insumos produtivos empregados no processo educacional estão sujeitos às escolhas feitas pelos agentes envolvidos (os pais e as escolas). Isso torna potencialmente endógeno o nível e a qualidade dos insumos utilizados. Sob o ponto de vista estatístico, este aspecto não deveria ser motivo de preocupações caso todos os fatores relevantes no processo de produção de habilidades cognitivas fossem observáveis¹⁹.

Dessa maneira, o sucesso de estudos empíricos em responder questões relacionadas à função de produção educacional depende, em grande escala, do acesso a informações sobre todos os insumos escolares e familiares, presentes e passados, assim como informações sobre o conjunto de habilidades intrínsecas de cada criança. Entretanto, em geral, as bases de dados apresentam restrições em relação a essas informações, sobretudo para o histórico de fatores produtivos ao longo da vida escolar dos alunos.

Outra questão presente nos estudos que utilizam a abordagem de função de produção educacional diz respeito à existência de atribuição não aleatória de

¹⁹ Para maiores detalhes acerca dos pressupostos implícitos relacionados com a abordagem de função de produção educacional e as limitações impostas aos estudos não experimentais ver o trabalho de Todd & Wolpin (2003).

estudantes a turmas (e professores) e a possibilidade de que esse padrão leve a distorções nas medidas de eficácia dos professores. O grau de importância desses problemas na estimação dos efeitos dos docentes também dependerá da escolha da especificação da função de produção educacional com valor adicionado a ser utilizada nos procedimentos de estimação.

A principal especificação do nosso estudo é um caso particular do modelo de valor adicionado da função de produção educacional com controle para desempenho passado discutido por Rothstein (2010). Essa especificação tem sido a mais utilizada na literatura de economia da educação e consiste de uma regressão das notas dos alunos nas variáveis de interesse, de controle e no desempenho passado dos alunos²⁰. A sua estimação por Mínimos Quadrados Ordinários permite identificar os parâmetros de interesse (efeito do professor ou dos atributos deste) somente se a atribuição dos alunos na turma, para a série avaliada, for aleatória condicional ao desempenho passado dos estudantes (Rothstein, 2010). Assim, pelos motivos acima discutidos a respeito da função de produção educacional, nós identificamos o efeito causal do conhecimento e da adoção frequente de práticas pedagógicas na sala de aula pelo professor sobre o aprendizado dos alunos, caso sejam válidos dois pressupostos. Esses são que a medida de desempenho defasado é uma estatística suficiente para o conjunto de insumos escolares e familiares passados e para a habilidade inata dos alunos; e que a atribuição dos estudantes às turmas (e, por conseguinte aos professores) é aproximadamente aleatória condicional no desempenho passado dos alunos.

Em relação aos estudos empíricos que exploram a conexão entre o conhecimento dos docentes e a aprendizagem de seus alunos, a literatura dispõe de poucos estudos com resultados satisfatórios. Santibañez (2006) investiga a relação entre as notas dos professores em testes e o desempenho dos estudantes para escolas primárias e secundárias da Cidade do México. Os resultados apontam que existe uma pequena relação positiva entre desempenho médio dos estudantes e as notas dos professores nos testes²¹. Entretanto, a análise de Santibañez (2006) apresenta um importante problema, pois é elaborada com dados agregados ao

²⁰ Exemplos de estudos que utilizam esta especificação do modelo de valor adicionado da função de produção educacional são: Aaronson, Barrow & Sander (2007); Goldhaber (2007); Jacob & Lefgren (2008); e Kane, Rockoff & Staiger (2008).

²¹ Santibañez (2006) também conclui que a senioridade e os títulos dos professores não são significativamente relacionados com o desempenho médio dos alunos.

nível das séries e essa restrição é motivada pela impossibilidade de utilização das informações ao nível dos alunos. Guimarães et al (2013) investigam a mesma relação para alunos da quarta série em seis estados do Brasil em 1999, para a disciplina de matemática. Na especificação mais completa, Guimarães et al (2013) encontram que um aumento de um desvio-padrão no conhecimento dos professores eleva o desempenho dos alunos em 4% de um desvio-padrão da distribuição de notas em matemática²². Novamente, o estudo para a realidade brasileira avalia a relação apenas ao nível da escola, pois não identifica o professor que está associado a cada turma no interior da escola. Dessa forma, os resultados desses dois artigos dependem fortemente da hipótese de que os professores são aleatoriamente atribuídos às turmas dentro das escolas. No nosso caso, graças as informações sobre a atribuição dos docentes às salas de aula, nós não estamos sujeitos as mesmas restrições dos trabalhos de Santibañez (2006) e Guimarães et al (2013).

Já Metzler & Woessmann (2012) estudam o contexto de alunos da sexta série em escolas pequenas (um único professor por série) do Peru. Nesse caso, os autores utilizam a variação intra-professor e intra-estudante para estimar o efeito causal do conhecimento do professor específico à disciplina sobre o desempenho dos alunos. Os resultados mostram que um aumento de um desvio-padrão no desempenho em matemática dos professores eleva o desempenho dos alunos em 9% de um desvio padrão²³. Já os efeitos sobre leitura são muito inferiores e estatisticamente não distintos de zero. Apesar das diversas peculiaridades do estudo de Metzler & Woessmann (2012), a nossa investigação tem um importante aspecto que não é contemplado no contexto peruano. No caso dos professores da rede pública estadual paulista que fizeram a prova de promoção havia um incentivo financeiro associado às notas obtidas na avaliação. Ademais, somente foram promovidos, ou seja, receberam incremento salarial, os docentes pertencentes ao vigésimo percentil superior na distribuição de notas. Isto é, os candidatos não sabiam *ex-ante* a nota mínima necessária para garantir a promoção e, por conseguinte o incremento salarial. Essa característica incentivava os

²² Esse efeito é significativo ao nível de 5%. Além disso, para as demais especificações utilizadas pelos autores as estimativas não eram estatisticamente significantes.

²³ A magnitude desse impacto é reduzida para 6,4% de um desvio-padrão sem a adoção de uma metodologia para a correção de erros de medida, que sob o modelo clássico tem um papel de atenuar as estimativas obtidas.

candidatos a se esforçarem nos testes, com o intuito de alcançar a maior nota possível e assim obter a recompensa financeira. Por este motivo, as notas derivadas da prova de promoção devem refletir melhor o conhecimento dos docentes.

Outros estudos procuram investigar os efeitos de medidas alternativas da qualidade dos docentes. Esse é o caso de Rockoff & Speroni (2010) que encontram que medidas subjetivas da qualidade do professor – notas em entrevista prévia a contratação e em avaliação por mentores no primeiro ano da carreira – apresentam relações significativas com o desempenho dos alunos futuros dos docentes. Para o contexto da cidade de Nova York, um aumento de um desvio-padrão na avaliação subjetiva dos professores eleva em aproximadamente 1,5% de um desvio-padrão o desempenho dos alunos em matemática.

Os nossos resultados para o impacto da melhoria da qualidade do professor por meio de uma maior erudição desse profissional são bastante próximos aos obtidos por Rockoff & Speroni (2010). Para a realidade dos alunos da oitava série da rede de ensino pública estadual paulista, um aumento de um desvio-padrão na medida de conhecimento dos professores tem um efeito de elevar em 1,4% de um desvio-padrão da distribuição de notas dos alunos para matemática e de 1,2% para língua portuguesa²⁴.

Uma vertente ainda mais recente de pesquisa focaliza a análise das atividades que os professores realizam no interior da sala de aula com o propósito de identificar quais são as práticas pedagógicas eficazes na produção de habilidades cognitivas dos alunos. Complementarmente, procuram obter estimativas consistentes da magnitude dos efeitos dessas práticas. Lavy (2011) utiliza uma categorização conceitual das práticas pedagógicas dos docentes, de acordo com o paradigma na literatura de psicologia educacional (Bloom, 1956), e examina a importância de cinco medidas agregadas de práticas de ensino sobre o aprendizado de alunos na quinta e oitava séries em Israel. Os resultados de Lavy (2011) mostram que dois elementos de práticas pedagógicas causam incrementos no desempenho dos alunos. Atividades dos professores em sala de aula que estimulam o conhecimento e a melhora na compreensão (denominado de estilo tradicional de lecionar) tem um efeito positivo e forte sobre as notas em testes de

²⁴ Essas estimativas são derivadas da especificação mais completa da função de produção educacional com controle para proficiência passada.

maneira geral. O uso de técnicas de ensino que dotam os alunos com habilidades analíticas, aplicativas e críticas do conhecimento (estilo moderno de lecionar) também apresenta um efeito positivo relevante. Em acréscimo, Lavy (2011) ressalta a grande magnitude do impacto dessas duas medidas de práticas pedagógicas. Por exemplo, uma intervenção realista baseada na melhora dessas duas atividades pedagógicas do nível médio para o máximo observado na amostra implica um aumento de 50% de um desvio-padrão da distribuição de notas em cada disciplina.

Para matemática, os nossos resultados mostram que a intervenção de trocar um professor que não passa lição de casa em todas ou quase todas as aulas por outro que o faz sempre tem um impacto de 12,6% de um desvio-padrão da distribuição de desempenho dos alunos. Já os efeitos *ceteris paribus* de alterar o professor de matemática por outro que sempre corrige lição de casa é igual a 10,7% de um desvio-padrão e a mudança para um docente que sempre explica a matéria é de 7,9%. Os efeitos para professores que sempre adotam as práticas de relacionar os conteúdos da disciplina às situações do cotidiano ou propor a resolução de problemas variados são de aproximadamente 14% de um desvio-padrão. Esses efeitos referem-se à estimação isolada do impacto de cada uma das práticas pedagógicas, apenas condicional ao conhecimento do docente. Já para língua portuguesa, os resultados são qualitativamente similares aos de matemática, com a principal diferença relacionada a uma menor magnitude dos impactos estimados. Além disso, a atividade de sempre indicar aos alunos livros de literatura para ler mostrou-se eficaz para o aprendizado dos alunos e essa prática pedagógica é comumente relacionada ao estímulo da capacidade de estudo individual dos estudantes.

Uma visão alternativa aos nossos resultados sobre a eficácia das práticas pedagógicas é a hipótese de que essas estão, na realidade, refletindo características não observáveis dos professores, tais como motivação e o comprometimento destes em relação a sua atividade profissional. Isso é possível caso as variáveis que resumem a intensidade da adoção das atividades pedagógicas pelos docentes forem positivamente correlacionadas com esses aspectos não observáveis. Assim, se esta hipótese alternativa for verdadeira temos que mudanças *ceteris paribus* nas atividades realizadas em sala de aula por um dado docente não necessariamente teriam os efeitos estimados e apresentados no nosso estudo. Contudo, as

evidências de que os efeitos das práticas pedagógicas permanecem significantes quando essas são incluídas de maneira simultânea nas estimações sugere que os resultados observados são oriundos pelo menos em grande parte das atividades pedagógicas, e não apenas por fatores não observáveis. Pois ao condicionarmos a análise às demais medidas de práticas de ensino, que são potencialmente correlacionadas com o comprometimento dos professores, a adoção em todas ou quase todas as aulas de cada atividade pedagógica investigada continua impactando positivamente e de forma relevante o ganho de proficiência dos alunos. Apesar da redução na magnitude desses efeitos na especificação que inclui as práticas pedagógicas conjuntamente.

Além dessa introdução, este capítulo está organizado da seguinte forma. A segunda seção apresenta e discute a estratégia empírica adotada para a estimação das especificações da função de produção educacional. A terceira seção apresenta uma visão geral do arcabouço institucional no qual o sistema de ensino público estadual paulista se insere. Na quarta seção, nós descrevemos as bases de dados utilizadas na análise empírica. Já os resultados são apresentados e discutidos na seção cinco. Por último, a sexta seção conclui o capítulo.

2.2 Estratégia empírica

O objetivo do presente trabalho é estimar os efeitos da qualidade dos professores sobre a aprendizagem dos seus respectivos alunos. A princípio nós desejaríamos que, mantido constante os demais insumos escolares, ocorresse uma variação exógena na qualidade dos professores na disciplina lecionada e que nós fossemos capazes de medir o impacto dessas mudanças sobre a proficiência dos estudantes.

Para isto, inicialmente nós adotamos o pressuposto de que o desempenho do aluno i na disciplina s no final da série g é determinado pela seguinte equação:

$$A_{isg,2009} = \alpha_s + \beta_s T_{isg} + \rho Z_{sg} + \theta X_{ig} + \varphi_s W_{isg} + \kappa_s E_{isg} + \varepsilon_{isg}. \quad (2.1)$$

Onde, $A_{isg,2009}$ é a nota padronizada do aluno i na disciplina s e na série g no ano de 2009²⁵; T_{isg} é a qualidade intrínseca do professor da disciplina s ; Z_{sg} é um vetor composto pelas características médias dos professores, por disciplina s e por escola, na oitava série e de insumos escolares disponibilizados pela escolas; X_{ig} é um vetor de características socioeconômicas pré-determinadas dos estudantes e de suas famílias²⁶; W_{isg} é um vetor de variáveis indicadoras para o tamanho das classes²⁷; E_{isg} é uma variável *proxy* para a experiência profissional dos docentes, ou seja, a sua idade em 2009. O termo de erro aleatório é denotado por ε_{isg} e nos procedimentos de estimação adotamos *cluster* ao nível das turmas/classes.

Diferentemente de Dee (2005 e 2007) e Metzler & Woessmann (2012), nós não impomos o pressuposto implícito de que os impactos dos professores sobre o ganho de desempenho dos alunos são iguais entre as diferentes disciplinas analisadas²⁸. No caso de disciplinas tão distintas quanto matemática e língua portuguesa (leitura, vocabulário etc.), não parece trivial supor que os mecanismos pelos quais os professores influenciem a aprendizagem dos alunos sejam idênticos. Assim, nós realizamos os procedimentos de estimação separadamente para as duas matérias e obtemos estimativas específicas dos efeitos dos professores sobre o ganho de desempenho dos alunos em matemática e língua portuguesa.

2.2.1 Aspectos que definem a qualidade do professor

Uma primeira preocupação que surge a partir da especificação da função de produção educacional em (2.1) é a definição de quais dimensões compreendem a

²⁵ As disciplinas s avaliadas nesse estudo são matemática e língua portuguesa. Já a série g refere-se sempre a oitava série do ensino fundamental.

²⁶ As variáveis referentes às características socioeconômicas e insumos educacionais familiares são: gênero dos alunos, escolaridade da mãe (por faixas de anos de estudo completos), escolaridade do pai, raça (cor da pele) da mãe e raça (cor da pele) do aluno, frequência com que os pais ajudam a fazer a lição de casa, disponibilidade de jornais e/ou revistas em casa, disponibilidade de livros educativos e/ou romance, se tinha computador e internet em casa.

²⁷ Foram utilizadas quatro variáveis indicadoras de tamanho de turma, definidas da seguinte forma: turmas com 30 ou menos alunos, 31 a 35 alunos, 36 a 40 alunos ou com mais de 40 alunos.

²⁸ Apesar de Metzler & Woessmann (2012) testarem a hipótese de igualdade dos coeficientes associados a tais efeitos para o contexto dos estudantes e professores de sexta série no Peru.

qualidade de um determinado professor. Inicialmente, nós temos o conhecimento intrínseco do professor específico à disciplina lecionada. Além disso, outro aspecto relevante são as habilidades do professor no interior da sala de aula que contribuem para a transmissão desse conhecimento e influenciam a aquisição de habilidades cognitivas por parte dos alunos.

Com o intuito de avaliar a importância relativa de cada uma dessas dimensões, nós separamos o conceito de qualidade do professor em dois conjuntos de variáveis. O primeiro é a nota padronizada dos docentes na parte objetiva da prova de promoção. Já o segundo grupo é um vetor de variáveis *proxies* para a habilidade dos professores, definidas a partir da intensidade de adoção de práticas pedagógicas em sala de aula. Dessa forma, a função de produção de habilidades cognitivas dos alunos passa a ser expressa por:

$$A_{isg,2009} = \alpha_s + \beta_{ks}K_{isg} + \beta_{ps}P_{isg} + \rho Z_{sg} + \theta X_{ig} + \varphi_s W_{isg} + \kappa_s E_{isg} + \varepsilon_{isg} . \quad (2.2)$$

Onde, K_{isg} é o conhecimento específico à disciplina s do professor, mensurado pela nota padronizada na parte objetiva da prova do Sistema de Promoção da SEE/SP²⁹. P_{isg} denota um vetor de práticas pedagógicas adotadas pelo docente responsável pela disciplina s na série g . O nosso interesse se dá sobre os coeficientes associados a essas duas medidas de qualidade dos professores. As demais variáveis presentes na equação (2.2) têm a mesma definição apresentada na equação (2.1) acima.

As práticas pedagógicas dos docentes utilizadas são definidas como variáveis indicadoras que adotam valor igual a um caso mais do que 75% dos alunos das turmas afirmam que o professor de matemática, em todas ou quase todas as aulas: (i) “passa lição de casa”; (ii) “corrige as lições de casa”; (iii) “explica a matéria até que todos os alunos entendam”; (iv) “propõe atividades de resolução de problemas variados” e (v) “relaciona os conteúdos de matemática às situações do cotidiano”. Já para o docente de língua portuguesa, além das três primeiras atividades acima listadas, também é utilizada a informação se este

²⁹ Maiores detalhes sobre o Sistema de Promoção para os integrantes do Quadro do Magistério da SEE/SP, assim como a respectiva avaliação, são fornecidas na seção três deste capítulo.

“indica livros de literatura para ler” em todas ou quase todas as aulas. Nós também empregamos definições alternativas para as variáveis referentes às práticas de ensino, em alguns exercícios empíricos, com o intuito de verificar a robustez dos resultados. A primeira medida alternativa indica que os docentes realizam sempre a atividade pedagógica quando 50% ou mais dos alunos da turma afirmam que isto ocorre. Já a segunda definição alternativa exige que 90% ou mais dos alunos de cada turma informem que o professor realiza em todas ou quase todas as aulas a prática pedagógica.

Nós entendemos que essas variáveis *proxies* representam uma forte evidência de que os respectivos professores realmente realizavam essas práticas de ensino em sala de aula. Complementarmente, acreditamos que essas atividades pedagógicas são mais factíveis de serem respondidas de maneira objetiva pelos alunos da oitava série do ensino fundamental. Dessa forma, nós procuramos evitar medidas que poderiam estar captando outros aspectos da relação entre alunos e professores³⁰. Por último, nós utilizamos três práticas – “passar lição de casa”, “corrigir lição” e “explicar a matéria até que todos os alunos entendam” – para as quais tínhamos informações para ambas as disciplinas. Assim, podemos comparar a importância relativa do impacto dessas práticas entre as duas matérias investigadas.

Ademais, sob o ponto de vista da psicologia educacional, segundo Bloom (1956) apud Lavy (2011), as práticas pedagógicas investigadas no presente trabalho são comumente associadas ao desenvolvimento de importantes categorias no domínio cognitivo. O hábito do docente se esforçar para explicar o conteúdo à maioria dos estudantes é essencial para fomentar o entendimento/compreensão do significado de cada tópico abordado. Complementarmente, o fato dos professores passarem lições para casa que ajudam a entender o material visto em sala de aula contribui para a construção do conhecimento e aprimorar a compreensão do significado do conteúdo estudado³¹.

Por outro lado, se os professores dão lições e/ou propõem atividades de resolução de problemas que demandam a elaboração de relações com temas anteriormente estudados, cujas respostas ainda não foram vistas. Então essas

³⁰ Exemplos dessas medidas são: “o professor incentiva os alunos a melhorarem o seu desempenho”, “o professor é atencioso e auxilia os alunos a realizarem suas tarefas” etc.

³¹ Sob a óptica da psicologia educacional, o conhecimento é definido como o processo de lembrar/recordar adequadamente das informações previamente aprendidas.

atividades estão fomentando as habilidades aplicativas, analíticas e críticas dos estudantes. Aspectos estes também explorados quando os docentes relacionam os conteúdos às situações do cotidiano, especialmente para matemática, que é caracterizada por um maior grau de abstração.

No caso da disciplina de língua portuguesa, a indicação frequente de livros de literatura para serem lidos pelos alunos tem um papel de destaque no estímulo da capacidade de estudo individual dos alunos, além de também contribuir para a construção do conhecimento. Por último, a correção das tarefas de casa em sala de aula é um mecanismo de retroalimentação (*feedback*) sobre o processo de aprendizagem dos alunos³².

2.2.2

O modelo de valor adicionado

Um segundo conjunto de problemas que se apresenta frequentemente no contexto dos estudos sobre a função de produção educacional refere-se aos problemas de variáveis omitidas e atribuição não aleatória dos professores aos alunos. O conceito cumulativo do processo de produção de habilidades cognitivas requer que a análise do impacto de fatores contemporâneos sobre o desempenho dos alunos seja condicionada a toda a história de insumos escolares, familiares e habilidade inata dos alunos (Todd & Wolpin, 2003). A incapacidade de incorporar essas informações implica a existência de variáveis omitidas relevantes. Por outro lado, observa-se de maneira geral, que melhores professores são associados em média a melhores alunos (vice-versa). Esses dois aspectos podem levar a sério viés nas estimativas de interesse.

O nosso estudo reconhece explicitamente esses dois aspectos recorrentes nas investigações empíricas e adota uma especificação de função de produção educacional com valor adicionado condicional à proficiência passada dos alunos³³. A partir da definição da função de produção de habilidades cognitivas expressa em (2.1), a especificação seguindo o modelo de valor adicionado passa a ser determinada pela seguinte equação:

³² Uma descrição pormenorizada das categorias no domínio cognitivo está disponível no site: <http://www.krummefamily.org/guides/bloom.html>

³³ Nós utilizamos uma especificação da função de produção educacional, com valor adicionado, muito próxima da estrutura descrita por Rothstein (2010) como o modelo de valor adicionado “VAM2”.

$$A_{isg,2009} = \alpha_s + \beta_s T_{isg} + \lambda_s A_{isg-2,2007} + \rho Z_{sg} + \theta X_{ig} + \varphi_s W_{isg} + \kappa_s E_{isg} + \varepsilon_{isg} . \quad (2.3)$$

No caso da equação (2.3), o desempenho passado do aluno i na disciplina s e na série $g-2$ é representado pelo termo $A_{isg-2,2007}$. Essa informação é a nota padronizada dos alunos na sexta série do ensino fundamental obtida a partir do SARESP 2007.

Usualmente esse modelo de valor adicionado utiliza como variável de controle a nota em testes dos alunos no final da série imediatamente anterior, ou seja, $g - 1$. No nosso estudo em particular, isto não é possível devido à estrutura de avaliação externa dos alunos empregada pela SEE/SP. No sistema de ensino paulista os estudantes são avaliados pelo SARESP somente nas séries pares do ensino fundamental (quarta, sexta e oitava séries). Por esse motivo, para os alunos que observamos na oitava série em 2009, nós podemos utilizar somente a sua nota na sexta série do ensino fundamental no ano de 2007 como uma medida para controlar para a proficiência passada. Os demais termos presentes na equação (2.3) são idênticos aos apresentados na equação (2.1).

Com o propósito de tornar mais completa a análise, nós incorporamos a discussão a respeito das dimensões da qualidade dos docentes – apresentada na subseção 2.2.1 – no modelo de valor adicionado da função de produção educacional. Dessa forma, a expressão (2.3) passa a ser reescrita como:

$$A_{isg,2009} = \alpha_s + \beta_{ks} K_{isg} + \beta_{ps} P_{isg} + \lambda_s A_{isg-2,2007} + \rho Z_{sg} + \theta X_{ig} + \varphi_s W_{isg} + \kappa_s E_{isg} + \varepsilon_{isg} . \quad (2.4)$$

A equação (2.4) é a principal especificação da função de produção educacional a ser utilizada em nossa investigação empírica. A sua estimação por mínimos quadrados ordinários (MQO) permite a identificação consistente dos impactos causais do conhecimento e das práticas pedagógicas dos professores

sobre a aquisição de habilidades cognitivas dos estudantes, se adotarmos como válidos alguns pressupostos importantes (Rothstein, 2010)³⁴.

Sob o modelo de valor adicionado com controle para o desempenho passado dos alunos, nós admitimos o pressuposto de que uma medida de proficiência defasada dos alunos na disciplina analisada é uma estatística suficiente para todo o conjunto de informações não observáveis anteriores (Todd & Wolpin, 2003 e 2007). Ademais, o modelo que inspira a equação (2.4) implica a hipótese de que a atribuição dos alunos e professores às turmas é aproximadamente aleatória condicional nessa medida de desempenho passado (Rothstein, 2010). Ou seja, devem valer as seguintes restrições:

$$E(K_{isg} \cdot \varepsilon_{isg} | A_{isg-2,2007}) = 0 \quad (2.5)$$

$$E(P_{isg} \cdot \varepsilon_{isg} | A_{isg-2,2007}) = 0 \quad (2.6)$$

Para uma amostra pequena de professores em Los Angeles, Kane & Staiger (2008) comparam as estimativas experimentais e observacionais do efeito do professor sobre o ganho de desempenho dos alunos e constatam que condicionar na nota defasada dos estudantes é suficiente para eliminar o viés decorrente da atribuição não aleatória de professores às turmas.

2.2.3

Controle para a capacidade de gestão do diretor

De acordo com Clark, Martorell & Rockoff (2009), na sua maioria, as decisões tomadas no âmbito da escola são de responsabilidade dos seus respectivos diretores. Esses tem (ou deveriam ter) uma importante ascendência sobre a força de trabalho da escola – em especial sobre os professores, muitas

³⁴ De acordo com Rothstein (2010) a estimação por mínimos quadrados ordinários da equação (2.4) gera estimativas inconsistentes para a taxa de decaimento (ou persistência) uniforme – λ_s – que é o coeficiente associado ao desempenho defasado dos alunos na disciplina s na sexta série do ensino fundamental. Isto ocorre mesmo quando são válidos os pressupostos explicitados no texto. Entretanto, é importante destacar que o nosso interesse é específico sobre os efeitos do conhecimento e habilidades pedagógicas dos professores sobre a proficiência dos alunos e não sobre o coeficiente associado à taxa de decaimento.

vezes sobre o conteúdo das disciplinas ministradas e são os principais responsáveis pelo monitoramento da qualidade do ensino que os alunos recebem.

Como exemplos da influência que os diretores têm sobre a gestão das escolas podemos citar o controle sobre a assiduidade dos docentes (um dos mais sérios problemas nos diversos sistemas de ensino no Brasil), a capacidade de motivar e ensinar o nível de esforço adequado dos professores na instrução dos alunos, a competência para a resolução de conflitos de interesse entre os diferentes atores presentes nas escolas (professores, alunos, pais etc.). Além de serem em última instância os responsáveis pela formação das turmas e alocação dos respectivos professores³⁵.

Todos esses aspectos, que a priori não são observáveis ao econometrista, compõem as diversas atribuições descritas como responsabilidade dos diretores e podem influenciar tanto o desempenho dos estudantes quanto as atividades dos professores. Por esses motivos, torna-se relevante ao menos condicionar a nossa análise sobre o processo de produção de habilidades cognitivas à influência da qualidade de gestão da escola. Para isso, nós utilizamos o desempenho do diretor na prova de promoção como uma variável *proxy* para a sua capacidade de gestão e liderança da unidade de ensino que está sob sua responsabilidade.

Ao controlarmos para a qualidade dos diretores é bastante plausível que nós estaremos condicionando, mesmo que de maneira aproximada, para diversas características não observáveis que são correlacionadas com a questão da administração das escolas. Por conseguinte, com a gestão dos professores e também são determinantes do desempenho dos alunos nos testes. Assim, o modelo de valor adicionado para o processo de produção de habilidades cognitivas dos alunos dado pela equação (2.4), após a inserção de controle para a qualidade de gestão dos diretores, passa a ser representado por:

$$A_{isg,2009} = \alpha_s + \beta_{ks}K_{isg} + \beta_{ps}P_{isg} + \tau_s D_{isg} + \lambda_s A_{isg-2,2007} + \rho Z_{sg} + \theta X_{ig} + \varphi_s W_{isg} + \kappa_s E_{isg} + \varepsilon_{isg} \cdot \quad (2.7)$$

Onde, a variável D_{isg} é a nota padronizada dos diretores associados à escola na qual o aluno i cursa a série g em 2009. Exatamente como realizado para os

³⁵ Dessa maneira, sendo decisivos sobre as questões a respeito de *sorting* de alunos nas turmas e *tracking* dos estudantes ao longo das séries.

docentes, nós optamos por utilizar apenas a nota dos diretores na parte objetiva da avaliação do Sistema de Promoção da SEE/SP.

2.3

O sistema de ensino público estadual paulista

A rede de ensino pública estadual paulista é a maior do país. Em 2009, as escolas sob administração da SEE/SP tinham 2.720.685 alunos matriculados no ensino fundamental e 1.449.782 estudantes no ensino médio – excluindo as escolas estaduais mantidas e administradas pelas universidades estaduais paulistas. De forma conjunta, são 4.170.467 alunos alocados em mais de 5.000 unidades de ensino e aproximadamente 230.000 professores. No restante dessa seção nós apresentamos dois importantes programas da SEE/SP a partir dos quais as informações associadas são fundamentais para a análise empírica deste estudo. Essas são o Sistema de Avaliação do Rendimento Escolar do Estado de São Paulo (SARESP) e Sistema de Promoção para os integrantes do Quadro do Magistério da SEE/SP.

2.3.1

O sistema de avaliação externa da SEE/SP (SARESP)

O Sistema de Avaliação do Rendimento Escolar do Estado de São Paulo – SARESP – é uma avaliação externa e foi implantado pela SEE/SP em 1996. Apesar de importantes mudanças metodológicas desde sua criação, o SARESP objetiva mensurar o desempenho escolar dos alunos com o intuito de fornecer informações sobre a situação da escolaridade na rede pública de ensino paulista e subsidiar a SEE/SP nas tomadas de decisão quanto à política educacional. No seu início, o SARESP previa a participação das equipes escolares na aplicação e correção das provas, análise dos resultados e na elaboração de propostas para melhoria do ensino (ARCAS, 2009). De acordo com Arcas (2009) nos três primeiros anos do SARESP (1996 a 1998) ocorreu essa interação das escolas, professores e diretores nas avaliações. Além disso, ao longo desses anos as séries e disciplinas avaliadas não permaneceram constantes, sendo que em 1997 houve a

inclusão do ensino médio no sistema de avaliação³⁶. A partir de 2001 as avaliações relacionadas ao SARESP passaram a ser aplicadas aos alunos nas séries finais dos ciclos, ou seja, a quarta e oitava séries do ensino fundamental e a terceira série do ensino médio. Entre 2003 e 2005 o SARESP passou a ser censitário a todos os estudantes e séries do ensino fundamental e médio da rede estadual pública paulista. A avaliação de 2006 foi suspensa com o intuito de permitir reflexão sobre o sistema de avaliação externa até então vigente, suas características, objetivos e a relação com as demais avaliações existentes³⁷.

Após um processo de reformulação, o SARESP a partir de 2007 passou a mensurar o desempenho dos alunos na primeira, segunda, quarta, sexta e oitava séries do ensino fundamental, assim como a terceira série do ensino médio. A participação das escolas estaduais tornou-se obrigatória e para as unidades de ensino municipais foi permitida a possibilidade de adesão ao sistema. Em 2007 o SARESP avaliou as disciplinas de matemática e língua portuguesa. Estas foram mantidas para os anos posteriores e a partir de 2008 ocorreu à incorporação, de maneira rotativa, de outras disciplinas³⁸. Adicionalmente, os resultados das avaliações do SARESP tornaram-se comparáveis às avaliações nacionais³⁹ – Saeb e Prova Brasil – e entre si para os anos posteriores. Isto foi possível devido à adoção, por parte da SEE/SP, de um conjunto de técnicas estatísticas denominado de Teoria de Resposta ao Item – TRI. Essa permite expressar os resultados de proficiência dos alunos em uma mesma métrica das demais avaliações existentes.

A avaliação referente ao SARESP 2009 foi aplicada entre os dias 17 e 19 de novembro do mesmo ano. As provas para a oitava série do ensino fundamental – a nossa análise sobre o ganho de proficiência dos alunos restringe-se aos estudantes matriculados nesta série em 2009 – foram compostas por itens de múltipla escolha. E, de acordo com a SEE/SP eram compostas por questões cognitivas que avaliavam competências, habilidades e conteúdos nas áreas de conhecimento avaliadas. Além das provas também foram aplicados questionários aos pais e alunos com o propósito de obter informações sobre o contexto socioeconômico e

³⁶ A avaliação do SARESP não foi aplicada em 1999.

³⁷ Para uma exposição detalhada do panorama histórico do Sistema de Avaliação do Rendimento Escolar do Estado de São Paulo ver o capítulo 2 da tese de Arcas (2009).

³⁸ Para o ano de 2008 foram incorporadas provas de ciências nas séries do ensino fundamental; além de biologia, física e química para o ensino médio. Já para 2009, além de matemática e língua portuguesa, também foram avaliados os conhecimentos dos alunos em geografia e história.

³⁹ Para língua portuguesa e matemática.

cultural das famílias, dados sobre a percepção dos alunos acerca das atividades de ensino utilizadas pelos seus respectivos professores em cada uma das disciplinas avaliadas etc.

2.3.2

O Sistema de Promoção para os integrantes do Quadro do Magistério

O Sistema de Promoção para os integrantes do Quadro do Magistério foi um programa de valorização da carreira docente instituído pela SEE/SP e consistia na passagem do titular de determinado cargo para faixa de classe imediatamente superior no plano de carreira vigente⁴⁰. Concomitantemente, a SEE/SP instituiu a existência de cinco faixas em cada uma das carreiras, inclusive para as de docente e diretor, no Sistema de Promoção no Quadro do Magistério⁴¹. A promoção dos candidatos representava uma evolução salarial de 25%, 50%, 75% e 100% sobre o salário inicial, respectivamente para as transições entre as faixas 1 e 2, da faixa 2 para a 3, da 3 para a 4 e por último entre as faixas 4 e 5. O primeiro processo de promoção, referente ao ano de 2010 e que nós utilizamos nesse estudo, permitiu que os profissionais concorressem apenas da faixa inicial (1) para a faixa dois.

Os funcionários da SEE/SP para participarem do programa de promoção e prestarem as provas deviam cumprir alguns requisitos de elegibilidade. Esses eram: (i) estar em efetivo exercício no dia 30 de novembro de 2009; (ii) ter vínculo com a rede estadual de ensino por no mínimo quatro anos ou 1.460 dias; (iii) ter permanecido em uma mesma unidade de ensino (escola) por no mínimo 80% deste período e (iv) ser considerado assíduo de acordo com critério da SEE/SP. Nesse primeiro processo de promoção, 88.630 professores realizaram as provas, sendo 76% de professores de educação básica II (PEB II) e 24% de professores de educação básica I (PEB I). Adicionalmente, 2.647 diretores também fizeram as provas do programa de promoção.

Os candidatos foram avaliados por uma prova composta de duas partes, uma primeira etapa denominada de objetiva – composta por sessenta questões – e uma

⁴⁰ No ano de 2009 as principais carreiras na SEE/SP eram: professor educação básica II (PEB II), professor educação básica I (PEB I), diretor de escola, assistente de diretor de escola, coordenador pedagógico e supervisor de ensino.

⁴¹ Adicionalmente, dentro de cada faixa foram mantidas as evoluções por níveis previstas na legislação vigente no momento em questão.

segunda parte chamada de dissertativa, ambas avaliadas entre zero e dez. Em particular, nosso interesse se dá na parcela objetiva da prova do Sistema de Promoção da SEESP e nas respectivas notas obtidas pelos professores de educação básica II (PEB II) e diretores. Para o caso dos docentes PEB II, as provas avaliaram conhecimentos específicos à disciplina que o professor ministrava aula na rede de ensino pública estadual paulista, além de conhecimento pedagógico.

A nota final do candidato na prova de promoção foi definida como a média entre a parte objetiva e a dissertativa, e a nota mínima requerida para concorrer à promoção era igual a seis. De maneira complementar, apenas foram promovidos os profissionais pertencentes ao vigésimo percentil superior na distribuição de notas finais dos integrantes de cada carreira do magistério da SEE/SP. Essa regra torna endógeno o desempenho mínimo necessário e suficiente para que o docente fosse considerado promovido⁴². Essa característica incentivava os candidatos a se esforçarem nos testes, com o intuito de alcançar a maior nota possível e assim garantir o respectivo incremento de remuneração.

O cronograma do Sistema de Promoção iniciou-se com a Lei Complementar nº 1.097 de 27 de outubro de 2009. A resolução SE 80 foi promulgada em 03 de novembro de 2009 e dispunha sobre o perfil de habilidades e competências específicas a cada disciplina e à parte geral exigidas na avaliação, além de uma bibliografia básica⁴³. Já as provas foram realizadas nos dias 29 de janeiro, 01 e 02 de fevereiro de 2010. Sob essas circunstâncias, os candidatos tiveram aproximadamente três meses entre o início do processo e a realização das provas. Em conjunto com o ineditismo do Sistema de Promoção, esse aspecto ressalta o pequeno intervalo de tempo que os docentes tiveram para se preparar especificamente para as provas nesse primeiro ano do programa.

Pelos motivos acima listados, nós acreditamos que as notas dos docentes PEB II na parte objetiva da prova de promoção refletem de maneira fidedigna o conhecimento que estes tinham das matérias quando lecionaram aos seus respectivos alunos no ano de 2009.

⁴² Em outras palavras, os candidatos não sabiam *ex-ante* a nota mínima necessária para garantir a promoção e, por conseguinte o incremento salarial.

⁴³ A Resolução SEE 80 está disponível no seguinte sítio eletrônico: http://siau.edunet.sp.gov.br/ItemLise/arquivos/80_09.HTM#_Toc242183494.

2.4 Dados

Nossa análise empírica é possível pela reunião de cinco bases de dados. Primeiramente, nós utilizamos as notas dos professores na parte objetiva da prova do Sistema de Promoção da SEE/SP, o que nos permite obter uma medida fidedigna do conhecimento dos professores na disciplina lecionada⁴⁴. Em seguida, por meio de informações administrativas da SEE/SP sobre a atribuição dos docentes em 2009, conseguimos associar os professores às suas respectivas turmas. As informações referentes às notas dos alunos em matemática e língua portuguesa, as características socioeconômicas dos alunos e as medidas a respeito da adoção de práticas pedagógicas pelos professores em sala de aula são oriundas do SARESP 2009. O desempenho passado dos estudantes em ambas as disciplinas na sexta série do ensino fundamental foi obtido a partir do SARESP 2007. A associação entre as notas de 2007 e 2009 para cada aluno foi realizada por meio de um código da SEE/SP – denominado ‘registro de aluno’ – que identifica unicamente os alunos da rede de ensino ao longo do tempo. Por último, nós utilizamos o Censo Escolar 2009 para reconstruir as características médias contemporâneas dos professores e insumos escolares disponibilizados pela escola para os alunos.

2.4.1 As bases de dados

As notas dos estudantes da oitava série do ensino fundamental no ano letivo de 2009 para as disciplinas de matemática e língua portuguesa são provenientes do SARESP 2009. Essa base de dados também contém as respostas dos alunos e dos pais aos questionários presentes na avaliação externa de ensino da SEE/SP. A partir dessas respostas, nós utilizamos as seguintes informações sobre as características socioeconômicas dos alunos: escolaridade da mãe, escolaridade do pai, raça (cor da pele) da mãe, raça (cor da pele) do aluno e gênero do aluno. E também sobre insumos educacionais fornecidos pela família: frequência com que

⁴⁴ Aproximadamente 40% dos professores da rede de ensino pública estadual paulista fizeram a avaliação do Sistema de Promoção.

os pais ajudam a fazer a lição de casa, disponibilidade de jornais e/ou revistas em casa, disponibilidade de livros educativos e/ou romance em casa, se tinha computador e internet em casa.

Com relação ao questionário dos alunos, nós dedicamos especial atenção a dois conjuntos de questões relativas ao comportamento e atividades pedagógicas dos professores de matemática e língua portuguesa em sala de aula⁴⁵. Aos alunos foram explicitadas algumas afirmações a respeito de cada professor e estes deveriam responder com qual frequência o docente em questão realizava a respectiva prática: ‘nunca’, ‘em algumas aulas’ ou ‘em todas ou quase todas as aulas’. Alguns desses itens, especificamente relacionados às práticas de ensino dos professores em sala de aula, foram usados na criação de variáveis dicotômicas para identificar quando havia forte evidência de que o profissional realizava a atividade pedagógica sempre. Dessa forma, na análise empírica nós utilizamos variáveis indicadoras que tomam valor igual a um caso mais do que 75% dos estudantes em cada turma reportam que o professor realiza a prática pedagógica em questão ‘em todas ou quase todas as aulas’⁴⁶.

De acordo com o SARESP 2009, 482.005 alunos estavam matriculados na oitava série do ensino fundamental da rede pública estadual paulista em 2009. Estes estavam alocados em 3.736 escolas e 13.874 turmas (ou classes). Cada escola tinha uma média de 4,65 turmas, com desvio-padrão igual a 2,06 e valores mínimo e máximo de 1 e 15 turmas, respectivamente. Já as turmas em média tinham aproximadamente 35 alunos, com desvio-padrão de 5,19 e valores mínimo e máximo, respectivamente, de 7 e 65 alunos. Desses alunos, 420.744 (87,29%) fizeram a prova de matemática e 423.084 (87,78%) realizaram a prova de língua portuguesa do SARESP 2009.

Todos os alunos da rede de ensino pública da SEE/SP são identificados de maneira única e ao longo dos anos por um código denominado de ‘registro do aluno’. Por meio desse número, nós recuperamos informações sobre o desempenho passado em matemática e língua portuguesa na sexta série do ensino fundamental, a partir do SARESP 2007, para uma fração dos alunos da oitava

⁴⁵ No apêndice 2 nós apresentamos as perguntas específicas aos professores de língua portuguesa e matemática presentes nos questionários dos alunos do SARESP 2009.

⁴⁶ Na seção de estratégia empírica nós apresentamos e justificamos a adoção das práticas pedagógicas utilizadas neste estudo. Também explicitamos nessa seção as definições alternativas utilizadas para a medida de intensidade de adoção das atividades de ensino pelos docentes.

série avaliados em 2009. Essa informação é fundamental, pois nos permitiu controlar para o nível de proficiência anterior dos alunos – abordagem de valor adicionado – ao investigarmos o papel da qualidade dos professores contemporâneos no processo de aprendizagem dos estudantes⁴⁷.

Assim, do total de alunos que fizeram a prova de matemática em 2009, recuperamos as notas nessa mesma disciplina no SARESP 2007 para 338.104 alunos, ou seja, 80,36%. Já para língua portuguesa este número foi de 336.898, o que representa 79,63% dos alunos que fizeram a prova do SARESP 2009. Ao longo do intervalo de anos empregado na análise estava em vigor o sistema de promoção automática na rede estadual paulista, o que implica uma menor taxa de reprovação nas séries que não constituem final de ciclo. Esse aspecto certamente contribuiu para a elevada fração de alunos para os quais recuperamos as notas no teste disponível imediatamente anterior e reduz as preocupações relacionadas à possibilidade de viés de seleção na amostra.

Adicionalmente, nós recebemos da SEE/SP os dados administrativos referentes à atribuição de aulas dos professores para o ano letivo de 2009. Essas informações já estavam restritas ao conjunto de docentes que participaram das avaliações relativas ao Sistema de Promoção de 2010⁴⁸. A partir dessas informações conseguimos identificar os docentes que lecionaram as disciplinas de matemática e língua portuguesa para cada turma.

Na base de dados sobre a atribuição de aulas, para uma minoria das turmas havia mais do que um professor associado para a mesma disciplina. Isto representa 5,9% das classes na disciplina de matemática e 6,5% para língua portuguesa. Como não foi possível identificar qual era o professor que permaneceu o maior período de tempo lecionando para as turmas em questão e quais eram os docentes substitutos, optamos por restringir a análise apenas às turmas alocadas a um único professor. Após este recorte dos dados, permaneceram um pequeno número de turmas com professores únicos para as quais existiam casos de docentes que prestaram o programa de promoção como Professor de Educação Básica I (PEB I) ou fizeram a prova para uma disciplina

⁴⁷ Por este motivo, restringimos nossa análise apenas às disciplinas de matemática e língua portuguesa. Não utilizando as informações sobre o desempenho dos alunos em história e geografia, pois para estas disciplinas nós não temos uma medida passada de proficiência dos alunos.

⁴⁸ Vale ressaltar que para participar do Sistema de Promoção os professores deveriam cumprir certos requisitos e por isso temos a atribuição de aulas apenas para uma parcela dos docentes da SEESP.

distinta daquela que era informada pelos dados administrativos da SEE/SP. Esses fatos totalizam 6,7% dos docentes atribuídos para aulas de matemática e 2,8% para língua portuguesa. Novamente, nós mantivemos o padrão de restringir a amostra apenas ao grupo de professores, e turmas associadas, para os quais podemos comparar as informações sobre as suas notas. Isso se justifica, pois o conteúdo avaliado nas provas do Sistema de Promoção para os professores PEB I e PEB II eram distintos, assim como as próprias avaliações. Da mesma maneira, as provas para professores PEB II de diferentes disciplinas, pelo menos em parte, avaliaram diferentes conhecimentos. Tais fatos inviabilizam a comparação das notas entre os professores de diferentes cargos e matérias, o que justifica a nossa opção de restringir a amostra apenas ao conjunto de turmas associadas aos professores com notas que podem ser confrontadas entre si. Dessa forma, dada às limitações informacionais, conseguimos identificar 6.160 turmas da oitava série associadas a 3.035 professores de matemática que fizeram a prova objetiva de promoção e tem notas positivas. Já para a disciplina de língua portuguesa, podemos observar 6.515 turmas atribuídas a 3.161 professores.

Com base no Censo Escolar 2009 (Inep/MEC) recuperamos informações sobre os insumos físicos disponibilizados pelas escolas aos seus alunos do ensino fundamental. Essas características referem-se a: se a escola era urbana ou rural, se possuía biblioteca, sala de leitura, laboratório de informática, laboratório de ciências, salas para os professores e para o diretor, computador, televisão e se disponibilizava alimentação para os estudantes. Além disso, nós calculamos os valores médios das características do corpo docente por escola, separadamente para cada uma das disciplinas avaliadas. As informações sobre os professores são: a idade média; a fração do sexo feminino; a proporção com ensino superior completo, licenciatura, especialização, mestrado ou doutorado; e a proporção que se declararam brancos ou negros. Esse conjunto de informações sobre as escolas e docentes tem como objetivo controlar para os insumos produtivos contemporâneos fornecidos pelas escolas aos seus alunos da oitava série na função de produção de habilidades cognitivas.

Embora o Censo Escolar disponibilize informações individuais sobre os professores, essas não puderam ser utilizadas de maneira desagregada em nossa análise empírica, pois os códigos identificadores dos docentes utilizados pela SEE/SP e pelo Inep/MEC não são compatíveis. Entretanto, a utilização de valores

médios em relação às características dos professores – tais como: atributos demográficos, títulos etc. – é comumente empregada na literatura. Ademais, temos como fato estilizado que “além da experiência de ensino [... característica esta que controlamos na nossa análise empírica por meio de uma variável *proxy*, ou seja, a idade dos professores ...], as características utilizadas para certificar e pagar os professores tem pouca relação com os resultados dos alunos” (Rockoff & Speroni, 2010, p. 261).

2.4.2

Amostras e estatísticas descritivas

A partir do banco de dados resultante da reunião dessas diferentes informações, nós criamos uma amostra básica de estudantes da oitava série do ensino fundamental em 2009. Essa amostra básica é definida pela existência de um conjunto mínimo de informações necessárias para a estimação da função de produção educacional com controle para o desempenho passado. Essas informações são: (i) o desempenho no SARESP 2009; (ii) a nota na parte objetiva da prova de promoção para o professor associado; e (iii) o desempenho passado do aluno na sexta série no SARESP 2007. Essa amostra é constituída por 155.272 alunos e 3.022 professores para a disciplina de matemática. Já para língua portuguesa, essa é formada por 162.892 alunos e 3.151 professores.

Para a realização dos exercícios empíricos que objetivam separar os efeitos da qualidade dos docentes em duas distintas dimensões – o conhecimento intrínseco e as habilidades pedagógicas dos professores – foi necessário restringir a amostra básica. Dessa forma, nós mantivemos apenas os alunos que reportaram sobre as práticas adotadas pelos professores em sala de aula e analisadas neste estudo. Os principais resultados empíricos do nosso trabalho são derivados dessa amostra e por esse motivo nós a denominamos de ‘amostra principal’ no restante do artigo. Essa é composta por 135.913 alunos associados a 3.016 professores em 6.120 turmas de 2.238 escolas da rede estadual paulista para a base de dados referente à matemática. Já para língua portuguesa, a amostra principal é formada por 142.568 estudantes alocados a 3.147 docentes em 6.467 turmas de 2.365 escolas.

Para a amostra principal, após a realização dos filtros acima descritos, nós constatamos que aproximadamente 70% das escolas têm apenas um professor ministrando as disciplinas na oitava série em 2009. Entre 22% e 26% das escolas têm dois professores e apenas 5% das escolas tem mais do que três professores lecionando na mesma série. A tabela 2.1 mostra que essas frequências são bastante próximas tanto para língua portuguesa quanto para matemática. Logo, a nossa principal fonte de variação para identificação dos efeitos das atividades pedagógicas é proveniente da comparação entre escolas. Assim, dado essa restrição na nossa amostra principal não podemos lançar mão de uma estratégia de efeitos-fixos por escola.

Tabela 2.1 – Distribuição da frequência de escolas quanto ao número de professores identificados na amostra principal

Número Professores	Matemática		Língua Portuguesa	
	Frequência escolas	Percentual	Frequência escolas	Percentual
1	1.550	0,6926	1.701	0,7192
2	581	0,2596	539	0,2279
3	90	0,0402	109	0,0461
4	17	0,0076	14	0,0059
5	0	0,0000	2	0,0008
Total	2.238		2.365	

Fonte: Elaboração própria, dados do SARESP 2009, SARESP 2007, Censo Escolar 2009, Notas da prova de promoção e dados administrativos da atribuição de professores da SEESP.

A tabela 2.2 apresenta as estatísticas descritivas referentes às características dos alunos, professores, turmas e escolas na amostra principal. A nota média dos alunos da oitava série em matemática é de 258,12 pontos e em língua portuguesa igual a 242,66; os respectivos desvios-padrão são iguais a 42,27 e 44,65. Um elevado percentual dos alunos presentes nas escolas estaduais de São Paulo se declarou pardo (38%) ou negro (49%) – no total aproximadamente 87% dos estudantes. Mais da metade das mães dos alunos eram analfabetas ou não tinham o ensino fundamental completo (aproximadamente 52%) e apenas 4% das mães dos estudantes tinham ensino superior completo.

Tabela 2.2 – Estatísticas descritivas para a amostra principal: disciplinas de matemática e língua portuguesa

	MATEMÁTICA		LINGUA PORTUGUESA	
	Média	Desvio-padrão	Média	Desvio-padrão
Características dos estudantes				
Nota Saesp 2009	258,12	42,27	242,66	44,65
Nota Saesp 2007	201,12	38,40	216,56	35,51
% feminino	0,52	0,50	0,52	0,50
% alunos brancos	0,09	0,28	0,09	0,28
% alunos pardos	0,38	0,49	0,38	0,49
% alunos negros	0,49	0,50	0,50	0,50
% mãe ensino superior completo	0,04	0,20	0,04	0,20
% mãe ensino médio completo	0,26	0,44	0,25	0,43
% mãe ensino fundamental completo	0,18	0,39	0,18	0,39
% mãe analfabeta ou semi-analfabeta	0,16	0,37	0,16	0,37
% mãe branca	0,57	0,49	0,57	0,49
% mãe parda	0,31	0,46	0,31	0,46
% mãe negra	0,07	0,25	0,07	0,25
% pais ajudam sempre a fazer lição de casa	0,17	0,38	0,17	0,38
% pais ajudam as vezes a fazer lição de casa	0,51	0,50	0,51	0,50
% alunos tem acesso jornais ou revistas em casa	0,37	0,48	0,36	0,48
% alunos tem acesso livros em casa	0,85	0,36	0,85	0,36
% alunos tem computador em casa	0,57	0,50	0,57	0,50
% alunos tem internet em casa	0,43	0,50	0,43	0,50
Práticas pedagógicas adotadas em todas ou quase todas as aulas				
% lição de casa	0,14	0,34	0,05	0,22
% corrige lição de casa	0,38	0,49	0,33	0,47
% explica a matéria até que todos entendam	0,48	0,50	0,40	0,49
% propõe resolução de problemas variados	0,14	0,34	-	-
% relaciona os conteúdos às situações cotidiano	0,06	0,24	-	-
% indica livros literatura para ler	-	-	0,04	0,21
Características dos Professores				
Nota objetiva prova promoção 2009	4,69	1,02	6,70	0,98
Idade	42	8,29	43	7,95
% Professor efetivo	0,86	0,35	0,93	0,26
Características da turma				
Tamanho turma	35	4,96	35	5,02
% Turma com menos de 30 alunos	0,17	0,38	0,17	0,38
% Turma com 31 a 35 alunos	0,35	0,48	0,35	0,48
% Turma com 36 a 40 alunos	0,41	0,49	0,40	0,49
% Turma com mais de 40 alunos	0,08	0,26	0,08	0,26
Características da escola				
% Professores sexo feminino	0,70	0,36	0,91	0,21
% Professores brancos	0,89	0,24	0,88	0,24
% Professores negros	0,09	0,22	0,10	0,23
% Professores com ensino superior	0,99	0,07	0,99	0,06
% Professores com especialização	0,22	0,34	0,26	0,38
% Professores com mestrado	0,01	0,09	0,01	0,08
% Escola urbana	0,97	0,17	0,97	0,18
% Escola com alimentação	0,99	0,08	0,99	0,10
% Escola com laboratório informática	0,96	0,19	0,96	0,19
% Escola com laboratório ciências	0,34	0,47	0,33	0,47
% Escola com biblioteca	0,05	0,23	0,05	0,23
% Escola com sala de leitura	0,87	0,34	0,86	0,35
% Escola com sala diretor	0,98	0,15	0,97	0,16
% Escola com sala de professores	0,98	0,14	0,98	0,15

Fonte: Elaboração própria, dados do SARESP 2009, SARESP 2007, Censo Escolar 2009, Notas da prova de promoção e dados administrativos da atribuição de professores da SEE/SP.

Em relação aos professores, a média das notas na parte objetiva da prova de promoção é bastante inferior em matemática comparativamente a língua portuguesa, com valores iguais a 4,69 e 6,70, respectivamente. Já os desvios

padrão das notas para ambas as matérias são próximos de um ponto⁴⁹. Em relação às práticas pedagógicas dos professores de matemática, de acordo com as declarações dos alunos, as frações de docentes para os quais temos forte evidência de que realizavam as atividades em “todas ou quase todas as aulas” são: 14% passam lição de casa, 38% corrigem as lições, 48% explicam a matéria até que todos os alunos entendam, 14% propõe a resolução de problemas variados e 6% relacionam os conteúdos de matemática às situações do cotidiano. No caso dos professores de língua portuguesa, essas proporções são iguais a: 5% passam lição de casa sempre, 33% corrigem as lições, 40% explicam a matéria até que todos os alunos entendam e 4,5% indicam livros de literatura para ler. A grande maioria dos docentes de língua portuguesa são mulheres (91%), enquanto que para matemática 30% dos docentes são homens. E a proporção média de professores com títulos de especialização ou mestrado é reduzida para ambas as disciplinas.

A terceira amostra utilizada na análise empírica está relacionada à verificação da robustez dos resultados obtidos para os impactos da qualidade dos docentes sobre a aquisição de proficiência dos seus alunos quando controlamos para a capacidade de gestão das escolas por parte de seus diretores. Esse procedimento tornou-se possível pela existência de uma variável *proxy* para a qualidade dos diretores – as notas na parte objetiva da prova de promoção para esse cargo. Entretanto, essa informação não está disponível para todas as escolas presentes na amostra principal. Dessa forma, a amostra utilizada no teste de robustez para a disciplina de matemática é formada por 78.375 alunos, 1.739 professores e 1.257 diretores. Para língua portuguesa, a amostra é composta por 80.591 alunos, 1.769 docentes e 1.298 diretores.

2.4.3 Seleção e *sorting*

Para uma parcela significativa dos estudantes da oitava série em 2009 não é possível identificar os professores responsáveis por lecionar as disciplinas. Assim, do total de estudantes para os quais temos tanto o desempenho no SARESP 2009 quanto em 2007, para aproximadamente 54% desse conjunto de alunos na matéria

⁴⁹ As distribuições das notas na parte objetiva da prova do sistema de promoção para os professores de matemática e língua portuguesa são apresentadas na figura A1 no Apêndice.

de matemática e 51% na de língua portuguesa, nós não temos informações sobre seus respectivos professores e, por conseguinte suas notas na prova de promoção.

Entre os possíveis motivos pelos quais isto ocorre, destacamos que para o docente ser elegível para a promoção na carreira havia algumas condicionalidades⁵⁰. Caso os professores soubessem de antemão que não satisfaziam tais requerimentos não haveria incentivos a estes participarem e se empenharem na realização das provas. Dada essa limitação informacional, caso os docentes que prestaram a prova de promoção fossem diferentes em relação ao conjunto dos que não a fizeram, então os efeitos estimados não poderiam ser estendidos para o grupo de alunos alocados aos professores que não participaram do programa de promoção.

Uma possibilidade para investigar a veracidade desse problema é a comparação entre as distribuições de ganho de aprendizado dos alunos entre as duas séries para as quais nós temos informações sobre a proficiência. Com esse propósito, nós estimamos regressões ao nível do estudante que tem como variável dependente as notas em 2009 e variável explicativa o desempenho passado (2007) dos alunos. Os resíduos dessas regressões são utilizados como medidas do incremento de aprendizado. Posteriormente, confrontamos as distribuições desses resíduos entre os grupos de estudantes para os quais é possível identificar os professores e os demais alunos. Essa comparação é apresentada nos gráficos da figura 2.1, onde os gráficos 2.1.1 e 2.1.2 referem-se, respectivamente, às disciplinas de matemática e língua portuguesa.

A análise visual não aponta diferenças relevantes nas distribuições de ganho de desempenho entre o grupo de alunos para os quais podemos estimar o efeito do conhecimento dos docentes e o restante dos estudantes para os quais isso não é possível. No apêndice 2 dessa tese, nós apresentamos evidências adicionais acerca da avaliação das semelhanças entre as distribuições de ganho de proficiência dos dois grupos de alunos. De maneira geral, os nossos resultados não nos permitem refutar a existência de um efeito, mesmo que minorado, da qualidade (conhecimento) do professor sobre a aquisição de habilidades cognitivas dos alunos para os quais nós não temos informações sobre os professores.

⁵⁰ Esses requerimentos mínimos são apresentados com maiores detalhes na subseção 2.3.2 que descreve o Sistema de Promoção para os integrantes do Quadro do Magistério da SEE/SP.

Figura 2.1 – Distribuições de desempenho em matemática e língua portuguesa entre os grupos de alunos com professores identificados e não identificados a partir dos dados de atribuição de aulas da SEE/SP

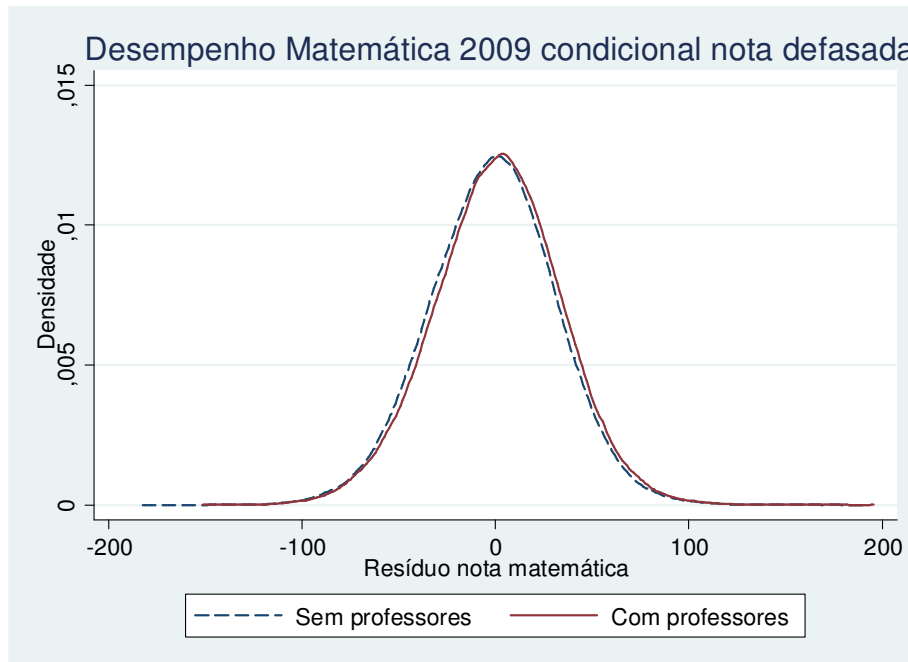


Gráfico 2.1.1. – Densidades do incremento de desempenho em matemática (resíduos) para o conjunto dos alunos para os quais os professores são identificados e os demais alunos.

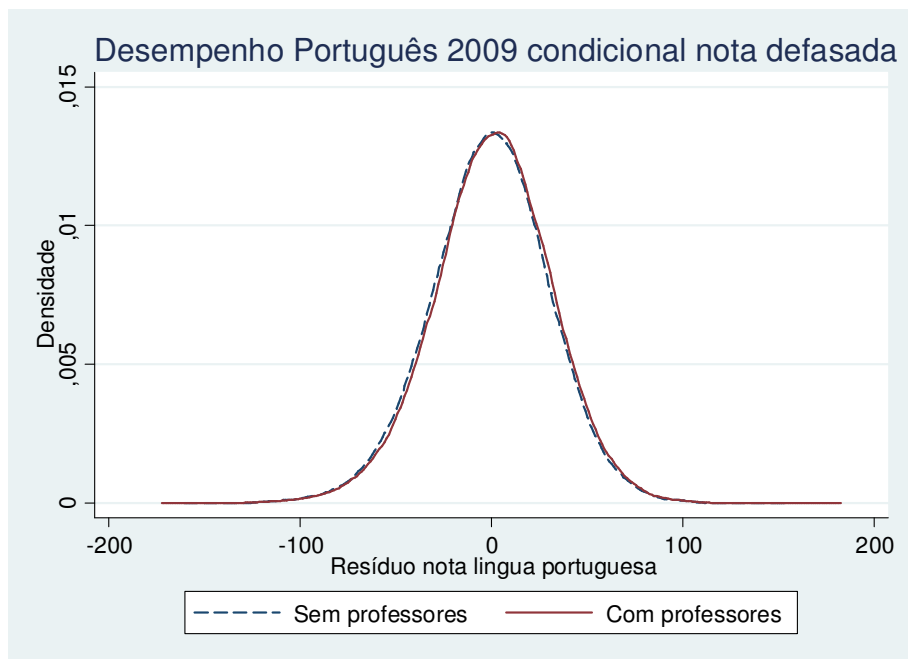


Gráfico 2.1.2. – Densidades do incremento de desempenho em língua portuguesa (resíduos) para o conjunto dos alunos para os quais identificamos os professores e os demais alunos.

Outro aspecto relevante para a estimação correta do impacto da qualidade dos professores é a possível existência de uma segregação de melhores (piores) alunos a turmas específicas e a seleção de determinados professores a essas classes compostas pelos melhores (piores) estudantes⁵¹. Assim, se esses fatores forem relevantes e nós não estivermos tomando as devidas precauções no sentido de controlar para esses aspectos, então as estimativas obtidas poderão ser viesadas e os nossos resultados estarão sendo guiados por outros determinantes do desempenho dos alunos, distintos dos efeitos que desejamos identificar.

A disponibilidade das notas na sexta série em 2007 para os alunos que compõem as turmas atribuídas aos docentes em 2009 nos possibilita a investigação da importância relativa desta questão. Para isso, nós calculamos o desempenho médio passado (2007) dos alunos que formam cada uma das turmas da oitava série em 2009 e examinamos a existência de correlação entre o desempenho defasado médio dos alunos e as notas dos professores na prova de promoção. Se essa correlação for estatisticamente relevante e positiva, então temos evidências de que os professores com maior conhecimento específico à disciplina (com qualidade superior) foram alocados a turmas de alunos que em média eram melhores mesmo antes da exposição destes ao docente em questão. Dessa forma, sob essas circunstâncias os resultados do efeito do professor poderão ser superestimados dependendo do modelo de função de produção educacional adotado.

A tabela 2.3 apresenta as estimativas derivadas de regressões, ao nível das turmas, nas quais a variável dependente é a nota dos professores na prova de promoção e a variável explicativa de interesse é o desempenho passado médio da turma constituída pelos alunos alocados aos respectivos professores em 2009. As evidências da tabela 2.3 apontam que existe uma correlação positiva forte, estatisticamente significativa e robusta entre o desempenho médio passado dos alunos e a qualidade (conhecimento) dos seus professores contemporâneos. Dessa forma, torna-se claro que sistematicamente os docentes com mais conhecimento são alocados a alunos que em média apresentam um patamar de proficiência, antes da interação entre aluno e o professor em questão, bastante superior. Ou seja, os

⁵¹ Neste caso, podemos ter a ocorrência de dois fenômenos usualmente denominados na literatura de economia da educação de *tracking* dos alunos ao longo das séries e *sorting* de alunos a professores.

professores com maior conteúdo nas disciplinas são atribuídos a turmas compostas por alunos melhor preparados e vice-versa. Esses resultados corroboram a necessidade de se condicionar a análise dos efeitos da qualidade (conhecimento e habilidades) dos professores ao desempenho passado dos alunos, com o intuito de levar em consideração essa atribuição não aleatória dos professores aos estudantes.

Tabela 2.3 – Teste para verificar a existência de *sorting* de professores às turmas com melhores alunos em média

Painel A: <i>sorting</i> professores em matemática			
Nota prova objetiva professor	Padronizada: nota prova promoção		
	(1)	(2)	(3)
Nota Saresp 2007 padronizada	0.210 (0.0301)***	0.227 (0.0294)***	0.116 (0.0364)***
Controles:			
Idade professor	Não	Sim	Sim
Tamanho da classe	Não	Sim	Sim
Características background familiar	Não	Não	Sim
Número turmas/classes	6,137	6,137	6,137
R-quadrado	0.008	0.078	0.089
Painel B: <i>sorting</i> Professores em lingua portuguesa			
Nota objetiva professor	Padronizada: nota prova promoção		
	(1)	(2)	(3)
Nota Saresp 2007 padronizada	0.253 (0.0312)***	0.283 (0.0312)***	0.195 (0.0404)***
Controles:			
Idade professor	Não	Sim	Sim
Tamanho da classe	Não	Sim	Sim
Características background familiar	Não	Não	Sim
Número turmas/classes	6,491	6,491	6,491
R-quadrado	0.011	0.081	0.088

Fonte: Elaboração própria, dados do SARESP 2009, SARESP 2007, Censo Escolar 2009, Notas da prova de promoção e dados administrativos da atribuição de professores da SEE/SP.

Erros padrão entre parênteses - *** $p < 0,01$; ** $p < 0,05$; * $p < 0,1$.

2.4.4

Comparação entre distribuições de desempenho

Como uma análise prévia da influência das duas dimensões de qualidade dos professores sobre a aprendizagem dos alunos, nós podemos avaliar graficamente se existem diferenças nas distribuições de ganho de aprendizado

entre os alunos alocados a professores – em 2009 – com diferentes estoques de conhecimento ou que empregam determinada prática pedagógica. Na próxima seção, apresentaremos as evidências dos efeitos do conhecimento e práticas pedagógicas dos docentes sob diferentes especificações da função de produção educacional com controles para insumos familiares, escolares, qualidade de gestão da escola etc. No entanto, uma avaliação mais básica a partir desses gráficos já pode nos fornecer algumas informações sobre a importância das distintas contribuições dos docentes.

Mais uma vez, nós estimamos regressões das notas dos alunos em 2009 sobre o desempenho passado (SARESP 2007) dos alunos e utilizamos os resíduos dessas regressões como as medidas de incremento na aquisição de habilidades pelos alunos. Primeiramente, nós examinamos as diferenças de ganho de aprendizado dos alunos atribuídos a docentes que estavam nos percentis 25 e 75 da distribuição de notas na parte objetiva da prova de promoção⁵². Para ambas as disciplinas, essa variação no desempenho dos docentes é bastante significativa e representa um acréscimo de mais de um desvio-padrão na nota objetiva dos professores. A figura 2.2 apresenta as comparações das densidades dos resíduos acima definidos. O gráfico 2.2.1 refere-se à disciplina de matemática e o gráfico 2.2.2 à língua portuguesa. Nesse caso, apesar da grande diferença de desempenho entre os professores, as distribuições de ganho de aprendizado dos respectivos alunos são bastante semelhantes, especialmente para matemática. Essas evidências podem ser interpretadas como um prenúncio do reduzido impacto do conhecimento específico às matérias dos docentes sobre a aquisição de proficiência de seus alunos.

⁵² As figuras A2 e A3 no Apêndice mostram que não há importantes não linearidades na relação entre as notas dos docentes na parte objetiva da prova de promoção e o aprendizado dos alunos, pelo menos nas faixas de notas dos professores que acumulam a maior quantidade de observações na amostra. Para matemática, essa faixa é entre os valores de dois e seis pontos. Já para língua portuguesa, essa faixa é definida aproximadamente entre quatro e oito pontos.

Figura 2.2 – Diferenças na distribuição de desempenho em matemática e língua portuguesa entre os alunos associados a professores com alto e baixo desempenho na prova de promoção

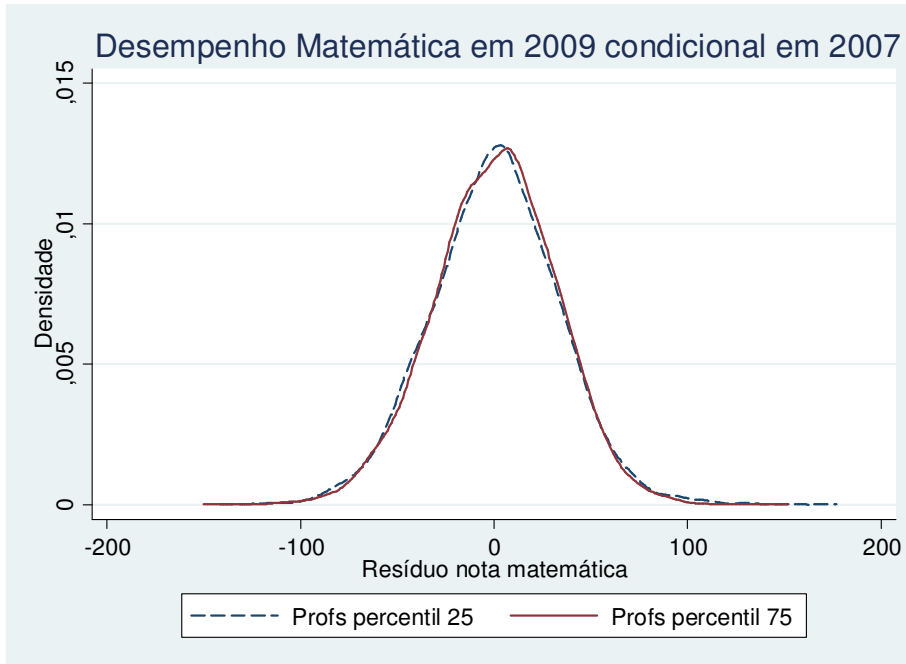


Gráfico 2.2.1. – Densidades do incremento de desempenho em matemática (resíduos) para os alunos associados aos professores no percentil 25 e 75 da distribuição de notas dos professores.

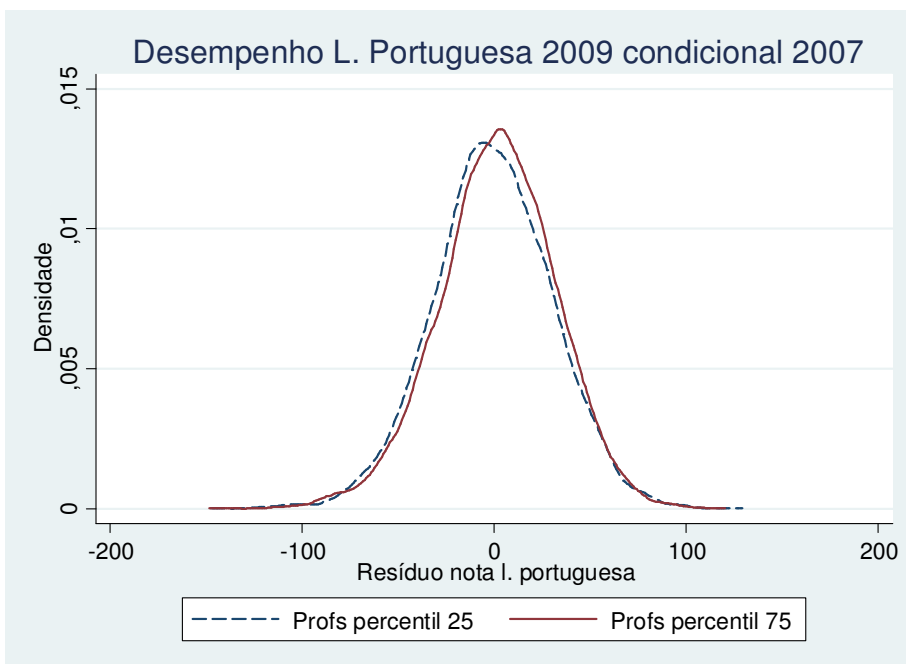


Gráfico 2.2.2. – Densidades do incremento de desempenho em língua portuguesa (resíduos) para os alunos associados aos professores no percentil 25 e 75 da distribuição de notas dos professores.

Exercícios similares ao da figura 2.2 foram elaborados para as densidades dos resíduos dos alunos de acordo com o fato destes pertencerem ou não a turmas nas quais os professores realizam sempre algumas das atividades pedagógicas avaliadas. Para matemática, nós selecionamos as práticas de ensino relativas a sempre “passa lição de casa” e “propõe atividades de resolução de problemas variados”. Já para língua portuguesa, utilizamos as atividades pedagógicas associadas a sempre “passa lição de casa” e “indica livros de literatura para ler”. As figuras 2.3 e 2.4, respectivamente para matemática e língua portuguesa, apresentam os gráficos com a comparação das densidades desses resíduos de acordo com os professores sempre adotarem, ou não, as práticas pedagógicas selecionadas.

Ao contrário do observado na figura 2.2, tanto para matemática quanto para língua portuguesa, nós observamos que as distribuições de ganho de desempenho são deslocadas à direita para os estudantes que constituem as turmas para as quais temos fortes indícios de que os professores adotam cada uma das práticas de ensino. As diferenças entre as densidades para os dois grupos de alunos não são muito grandes, entretanto, a distinção quanto ao deslocamento à direita das densidades associadas aos alunos para os quais os professores realizam sempre as atividades pedagógicas é clara e se faz presente para todos os casos analisados. Esses resultados indicam que as práticas pedagógicas dos professores estão relacionadas a um maior incremento de proficiência dos alunos entre as duas séries observadas para ambas às disciplinas. Na próxima seção, nós investigamos em maiores detalhes a existência e magnitude dessas relações (efeitos) com o auxílio de técnicas de regressão.

Figura 2.3 – Diferenças na distribuição de desempenho em matemática entre os alunos associados a professores que sempre adotam uma das práticas pedagógicas investigadas e os demais estudantes

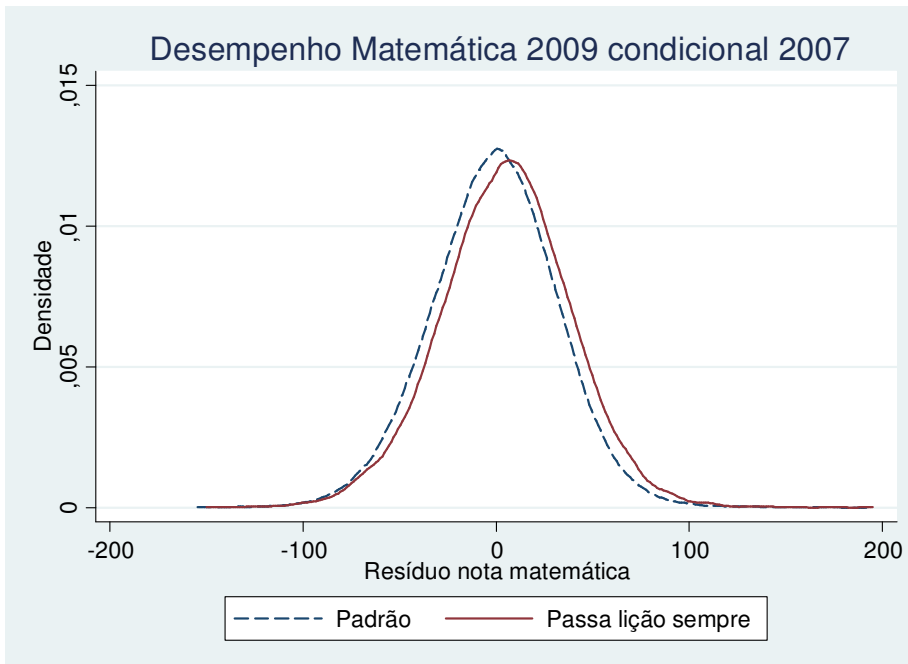


Gráfico 2.3.1 – Densidades do incremento de desempenho em matemática (resíduos) para alunos com professores que sempre passam lição de casa e os demais.

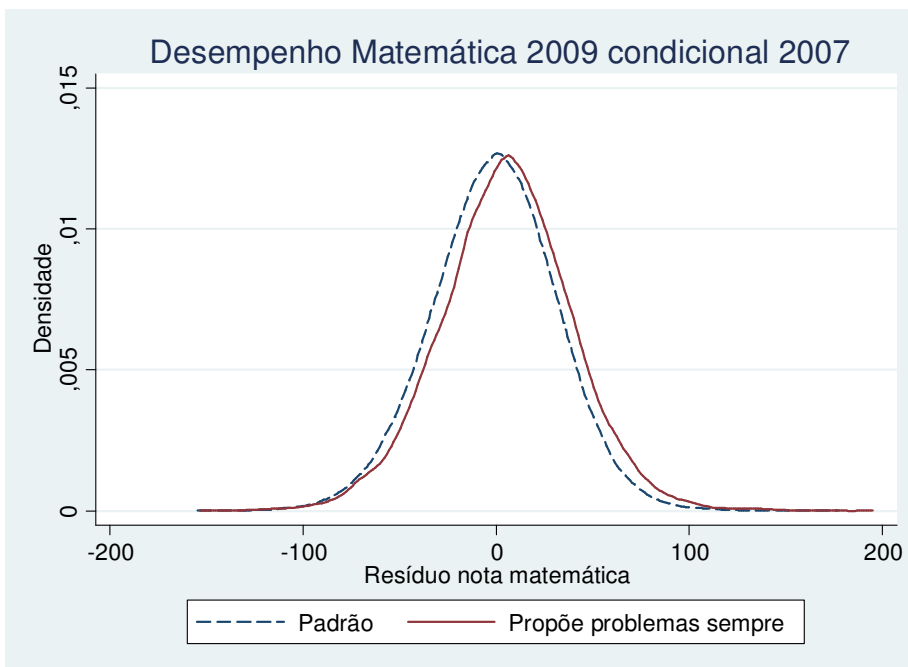


Gráfico 2.3.2 – Densidades do incremento de desempenho em matemática (resíduos) para alunos com professores que sempre propõem resolução de problemas e os demais.

Figura 2.4 – Diferenças na distribuição de desempenho em língua portuguesa entre os alunos associados a professores que sempre adotam uma das práticas pedagógicas investigadas e os demais estudantes

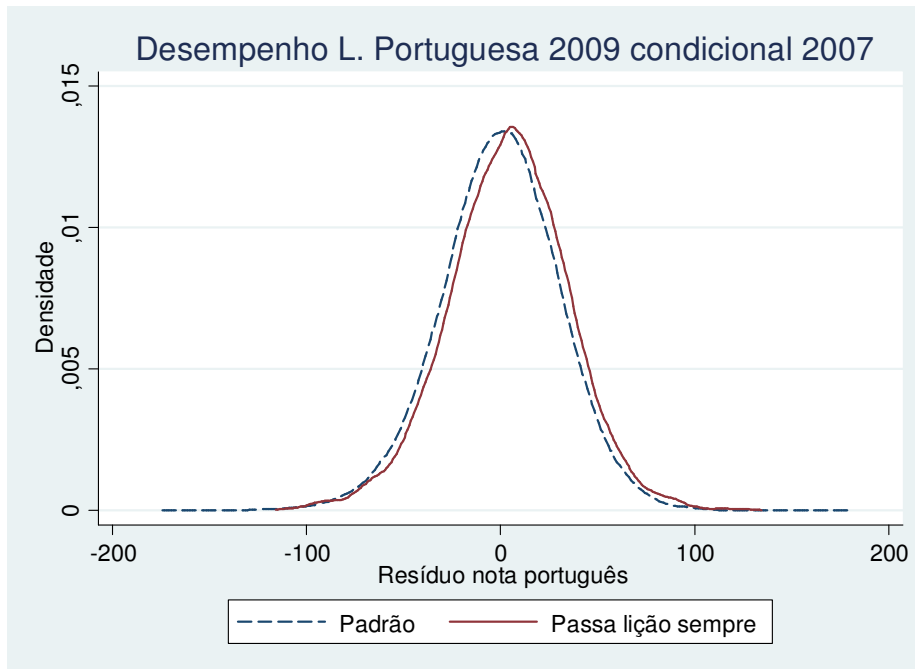


Gráfico 2.4.1 – Densidades do incremento de desempenho em língua portuguesa (resíduos) para alunos com professores que sempre passam lição de casa e os demais.

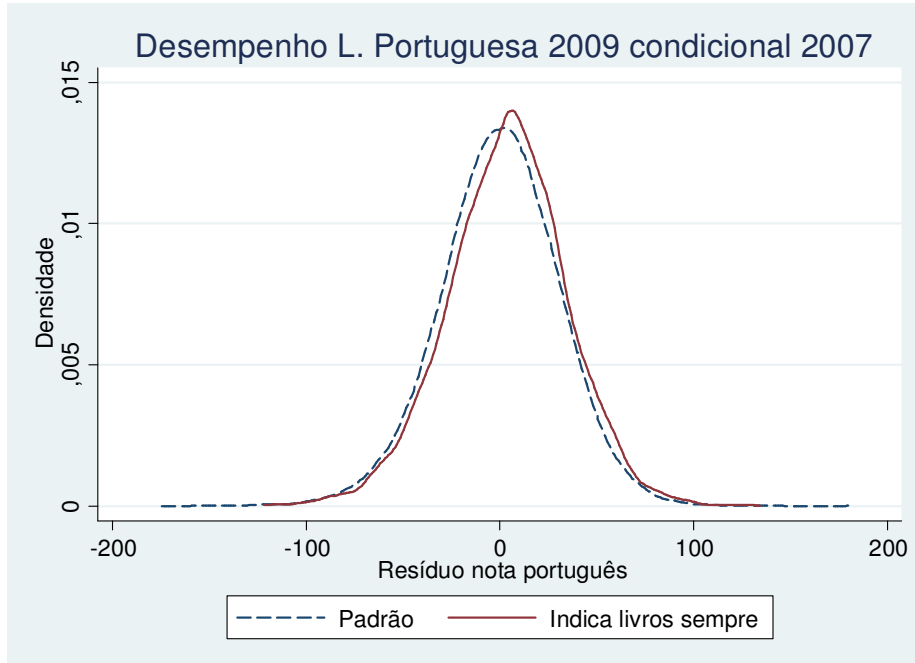


Gráfico 2.4.2 – Densidades do incremento de desempenho em língua portuguesa (resíduos) para alunos com professores que sempre indicam livros de literatura e os demais.

2.5 Resultados

Nessa seção nós apresentamos os resultados do artigo. Inicialmente nós restringimos atenção somente ao impacto do conhecimento do docente sobre o desempenho nos testes dos seus alunos. Em seguida, nós incorporamos as medidas de práticas pedagógicas utilizadas pelos professores em sala de aula e apresentamos as estimativas dos efeitos do conhecimento e dessas atividades dos docentes, de maneira simultânea, sobre a aprendizagem dos estudantes. Isso nos permite separar e comparar a importância relativa dessas duas dimensões da qualidade dos professores. Por último, nós utilizamos as notas dos diretores das respectivas escolas para controlar para aspectos relacionados à qualidade de gestão da escola. Esses fatores influenciam o desempenho educacional dos alunos e podem ser correlacionados com nossas medidas de interesse, o que gera problemas de viés derivado de variáveis omitidas. Isto é, fatores não observáveis que poderiam estar guiando nossos resultados. As evidências mostram-se robustas a essa questão da qualidade da gestão da unidade de ensino, sobretudo para matemática.

2.5.1 O efeito do conhecimento dos professores

A primeira linha da tabela 2.4 apresenta as estimativas da relação entre o conhecimento dos professores e o desempenho dos alunos associados no SARESP 2009 para diferentes especificações da função de produção educacional. Nos procedimentos de estimação nós adotamos ‘cluster’ ao nível das turmas/classes. As quatro primeiras colunas referem-se aos resultados da disciplina de matemática e as últimas quatro colunas são relativas à língua portuguesa. Os resultados para o coeficiente de interesse são positivos e estatisticamente significantes para ambas as disciplinas, distintas especificações e diferentes conjuntos de variáveis de controle.

A nossa especificação mais básica, apresentada nas colunas 1 e 5, avalia a correlação bruta existente entre as notas dos professores e os resultados nos testes dos alunos. As estimativas obtidas indicam que um aumento de um desvio-padrão

na nota dos professores está associado a um incremento de aproximadamente 4,5% de um desvio-padrão da distribuição de notas no SARESP 2009 para a amostra básica de alunos⁵³. Isto corresponde a um acréscimo de pouco mais de dois pontos na escala Saeb tanto para matemática quanto para língua portuguesa. Um impacto de pequena relevância sob o ponto de vista econômico, sobretudo ao levarmos em consideração que estamos captando apenas uma correlação entre as duas variáveis em questão, sem controlar para importantes insumos produtivos da função de produção educacional.

Tabela 2.4 – Efeito do conhecimento do professor sobre o ganho de desempenho dos alunos no SARESP 2009: matemática e língua portuguesa

Nota Saresp 2009 padronizada	Matemática				Lingua portuguesa			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Nota objetiva prova promoção	0.0486 (0.00548)***	0.0220 (0.00352)***	0.0190 (0.00359)***	0.0140 (0.00345)***	0.0448 (0.00514)***	0.0139 (0.00254)***	0.0142 (0.00265)***	0.0116 (0.00256)***
Desempenho passado - Saresp 2007		0.623 (0.00294)***	0.621 (0.00293)***	0.585 (0.00295)***		0.722 (0.00191)***	0.721 (0.00191)***	0.680 (0.00198)***
Controles:								
Desempenho passado alunos	Não	Sim	Sim	Sim	Não	Sim	Sim	Sim
Idade professor	Não	Não	Sim	Sim	Não	Não	Sim	Sim
Tamanho da classe	Não	Não	Sim	Sim	Não	Não	Sim	Sim
Características escola / profs médias	Não	Não	Sim	Sim	Não	Não	Sim	Sim
Características background familiar	Não	Não	Não	Sim	Não	Não	Não	Sim
Média Saresp 2009		256,428				240,916		
Desvio-padrão Saresp 2009		42,373				44,902		
Média prova objetiva promoção		4,691				6,698		
Desvio-padrão prova objetiva		1,017				0,977		
N. Observações	155,272	155,272	155,272	155,272	162,892	162,892	162,892	162,892
R-quadrado	0.002	0.389	0.391	0.406	0.002	0.522	0.522	0.534

Fonte: Elaboração própria, dados do SARESP 2009, SARESP 2007, Censo Escolar 2009, Notas da prova de promoção e dados administrativos da atribuição de professores da SEE/SP.

Erros padrão entre parênteses - *** $p < 0,01$; ** $p < 0,05$; * $p < 0,1$.

Nas colunas 2 e 6 da tabela 2.4, nós incluímos o desempenho passado dos alunos – notas padronizadas em matemática e língua portuguesa na sexta série derivadas do SARESP 2007 – como variável de controle na especificação estimada. A adoção do modelo de valor adicionado implica uma importante redução na magnitude dos efeitos do conhecimento dos professores sobre o desempenho dos alunos nas duas disciplinas. Apesar, de esses efeitos permanecerem positivos e estatisticamente significantes. Para matemática, um

⁵³ A magnitude desse impacto é semelhante à obtida originalmente por Metzler & Woessman (2012) antes da correção para o problema de viés de atenuação do coeficiente derivado de erro de medida (modelo clássico) na variável de interesse – nota dos professores em um teste. Nesse caso, para a realidade de alunos da sexta série do Peru, Metzler & Woessman (2012) concluem que um acréscimo de um desvio-padrão no conhecimento dos professores aumenta a nota de matemática dos alunos em 6,4% de um desvio-padrão da distribuição de notas dos alunos.

incremento de um desvio-padrão na nota do docente passa a ter um impacto de 2,2% de um desvio-padrão na nota dos alunos no SARESP 2009. Já para língua portuguesa, a magnitude dessa queda na estimativa é ainda maior, com um impacto positivo de aproximadamente 1,4% de um desvio-padrão na nota dos estudantes.

A posterior inclusão dos demais controles para características dos professores, das escolas, tamanho de sala (colunas 3 e 7) e por último das características socioeconômicas e de insumos educacionais disponibilizados pelas famílias (colunas 4 e 8) tem um efeito apenas marginal sobre as estimativas de interesse. Assim, na especificação mais completa (colunas 4 e 8 da tabela 2.4) temos que o efeito do conhecimento do professor passa a ser um aumento de 1,4% de um desvio-padrão da distribuição de notas dos alunos para matemática e de 1,2% para língua portuguesa, dado um acréscimo de um desvio-padrão na nota dos docentes. Esses resultados, em conjunto com as evidências apresentadas na seção anterior a respeito de *sorting* de professores e alunos a turmas, atestam a importância da adoção de uma abordagem de valor adicionado no estudo do arcabouço da tecnologia de produção de habilidades cognitivas dos estudantes. Isto, pois as evidências inicialmente obtidas – sem controle para o desempenho passado – podem prestar informações errôneas sobre as estimativas do impacto de interesse.

Ademais, é importante ressaltar que apesar de ter um efeito positivo e estatisticamente distinto de zero, o conhecimento dos docentes específico às disciplinas não parece ser de primeira ordem em termos econômicos para o processo de aprendizagem dos estudantes. Visto que mesmo na especificação inicial, na qual estamos captando apenas uma correlação entre as variáveis, a magnitude dos resultados é pequena. E com a sofisticação da equação que representa o processo por meio do qual ocorre a aprendizagem dos alunos, a influência do conhecimento dos professores passa a ser ainda menos relevante. Por último, nós não observamos efeitos heterogêneos do conhecimento dos docentes sobre o ganho de proficiência dos alunos nos testes ao longo da distribuição de desempenho passado dos alunos nas duas disciplinas avaliadas.

2.5.2

A decomposição da qualidade do professor: efeitos do conhecimento vs. das práticas pedagógicas

A partir de agora nós analisamos os efeitos das diferentes práticas de ensino conjuntamente com a medida de conhecimento dos docentes sobre a aquisição de proficiência dos alunos. Para isto, nas tabelas que seguem, nós apresentamos as estimativas dos coeficientes de interesse sempre empregando a especificação mais completa da função de produção educacional. Essa incorpora o modelo de valor adicionado com a inserção de controle para a proficiência passada individual dos alunos na disciplina avaliada, além de controles para as características médias dos professores, das escolas, tamanho de sala, para as características socioeconômicas e insumos educacionais disponibilizados pelas famílias. Ou seja, os resultados apresentados nessa subseção referem-se à estimação de regressões por mínimos quadrados ordinários que utilizam a especificação associada à equação (2.4). Por conseguinte, nós supomos a validade das restrições expressas nas equações (2.5) e (2.6). Por último, nesses procedimentos nós adotamos ‘cluster’ ao nível das turmas/classes.

A tabela 2.5 apresenta as estimativas para a disciplina de matemática. Nas colunas 1 a 5, as estimativas referem-se a regressões distintas nas quais cada prática de ensino entra como uma variável de interesse acompanhada pela medida padronizada de conhecimento dos docentes. Na última coluna, as estimativas são oriundas de uma única regressão que inclui simultaneamente as cinco medidas de atividades pedagógicas adotadas em sala de aula, além do conhecimento específico dos professores. O painel A exibe os resultados quando utilizamos a nossa definição principal para as variáveis responsáveis por indicar quando o professor sempre adota a atividade pedagógica investigada. Nesse caso, 75% ou mais dos alunos da turma devem afirmar que o docente realiza a práticas em todas ou quase todas as aulas. Já o painel B lista as estimativas para a definição alternativa menos restritiva das práticas pedagógicas (50% ou mais dos alunos na turma) e o painel C mostra os resultados para a medida alternativa mais restritiva

(90% ou mais dos alunos na turma). Os resultados da tabela 2.5 são baseados na amostra principal, anteriormente descrita na seção de dados⁵⁴.

Tabela 2.5 – Efeito conjunto do conhecimento e das práticas pedagógicas do professor sobre o ganho de desempenho dos alunos em matemática

Nota Saresp 2009 padronizada	Matemática					
	(1)	(2)	(3)	(4)	(5)	(6)
Painel A: 75% ou mais dos alunos da turma afirmaram que o professor sempre adotava a prática pedagógica						
Nota objetiva prova promoção	0.0108 (0.00361)***	0.00938 (0.00360)***	0.0107 (0.00361)***	0.0106 (0.00359)***	0.0118 (0.00362)***	0.00729 (0.00355)***
Práticas pedagógicas						
Passa lição sempre (> 75% alunos)	0.126 (0.0111)***					0.0520 (0.0119)***
Corrige lição sempre (> 75% alunos)		0.107 (0.00738)***				0.0597 (0.00856)***
Explica a matéria (> 75% alunos)			0.0795 (0.00709)***			0.0286 (0.00756)***
Propõe resolução problemas (> 75%)				0.134 (0.0120)***		0.0648 (0.0134)***
Relaciona conteúdos ao cotidiano (> 75%)					0.144 (0.0194)***	0.0512 (0.0204)***
Painel B: 50% ou mais dos alunos da turma afirmaram que o professor sempre adotava a prática pedagógica						
Nota objetiva prova promoção	0.0111 (0.00359)***	0.0125 (0.00360)***	0.0117 (0.00364)***	0.0101 (0.00358)***	0.0100 (0.00360)***	0.00825 (0.00354)***
Práticas pedagógicas						
Passa lição sempre (> 50% alunos)	0.0879 (0.00712)***					0.0454 (0.00782)***
Corrige lição sempre (> 50% alunos)		0.0988 (0.00814)***				0.0372 (0.00960)***
Explica a matéria (> 50% alunos)			0.0787 (0.00850)***			0.00394 (0.00962)
Propõe resolução problemas (> 50%)				0.103 (0.00699)***		0.0577 (0.00808)***
Relaciona conteúdos ao cotidiano (> 50%)					0.0880 (0.00749)***	0.0447 (0.00814)***
Painel C: 90% ou mais dos alunos da turma afirmaram que o professor sempre adotava a prática pedagógica						
Nota objetiva prova promoção	0.0121 (0.00365)***	0.0111 (0.00359)***	0.0107 (0.00362)***	0.0128 (0.00363)***	0.0129 (0.00365)***	0.00966 (0.00359)***
Práticas pedagógicas						
Passa lição sempre (> 90% alunos)	0.137 (0.0205)***					0.0196 (0.0233)
Corrige lição sempre (> 90% alunos)		0.127 (0.0126)***				0.0807 (0.0141)***
Explica a matéria (> 90% alunos)			0.105 (0.00998)***			0.0650 (0.00996)***
Propõe resolução problemas (> 90%)				0.207 (0.0293)***		0.0871 (0.0313)***
Relaciona conteúdos ao cotidiano (> 90%)					0.227 (0.0454)***	0.0938 (0.0467)***
Controles:						
Desempenho passado - Saresp 2007	Sim	Sim	Sim	Sim	Sim	Sim
Idade professor	Sim	Sim	Sim	Sim	Sim	Sim
Tamanho da classe	Sim	Sim	Sim	Sim	Sim	Sim
Características escola / profs médias	Sim	Sim	Sim	Sim	Sim	Sim
Características background familiar	Sim	Sim	Sim	Sim	Sim	Sim
N. Observações	135,913	135,913	135,913	135,913	135,913	135,913

Fonte: Elaboração própria, dados do SARESP 2009, SARESP 2007, Censo Escolar 2009, Notas da prova de promoção e dados administrativos da atribuição de professores da SEE/SP.

Erros padrão entre parênteses - *** p<0,01; ** p<0,05; * p<0,1.

⁵⁴ Essa amostra principal é distinta da utilizada para a obtenção das estimativas apresentadas na tabela 2.4. Isto se deve ao fato de não haver respostas para o questionário dos alunos para 19.359 alunos presentes na amostra inicial para a disciplina de matemática.

Nós destacamos que para matemática todos os efeitos do conhecimento e das práticas pedagógicas dos professores são positivos e estatisticamente distintos de zero. O efeito do conhecimento do docente de matemática permanece próximo a um acréscimo de 1% de um desvio-padrão da distribuição de notas da amostra principal, dado um aumento de um desvio-padrão – ou aproximadamente um ponto – no desempenho dos docentes na parte objetiva da prova de promoção. Esse impacto independe da definição da variável de prática pedagógica utilizada entre os três diferentes painéis da tabela 2.5. Além disso, a magnitude desse efeito mantém-se próxima a obtida na coluna 4 da tabela 2.4, ou seja, na nossa especificação mais completa sem a inclusão de medidas para as práticas pedagógicas dos docentes.

Primeiramente, restringimos nossa atenção ao painel A da tabela 2.5. Os resultados apontam que, condicional ao nível de conhecimento do professor de matemática, se trocássemos um professor que não passa lição de casa em todas ou quase todas as aulas por outro para o qual temos forte evidência de que o faz, essa intervenção implica um aumento de 12,6% de um desvio-padrão da distribuição de desempenho em matemática dos alunos no SARESP 2009. Esse acréscimo de desempenho dos alunos representa mais do que cinco pontos na escala Saeb e aproximadamente um aumento de 2% em relação à nota média da amostra principal para matemática. De maneira similar, as colunas 2 a 5, apresentam os efeitos *ceteris paribus* de se alterar o docente para um que sempre realiza cada uma das demais atividades pedagógicas estudadas de maneira isolada – condicional ao conhecimento do professor. Sob tais circunstâncias, um docente que constantemente corrige a lição de casa eleva o desempenho dos seus alunos em matemática em 10,7% de um desvio-padrão. O fato de sempre explicar a matéria até que todos os alunos entendam implica um acréscimo de 7,9% de desvio-padrão na aquisição de habilidades cognitivas em matemática. Já, os professores de matemática que sempre propõem a resolução de problemas variados ou relacionam os conteúdos da disciplina às situações do cotidiano promovem uma ampliação no desempenho dos alunos de aproximadamente 14% de um desvio-padrão da distribuição de notas. É importante ressaltar que esses resultados acima discutidos são os efeitos isolados de substituir um professor que

não adota sempre cada uma das práticas de ensino em questão, por outro que o faz.

Esses impactos da adoção frequente de práticas de ensino no interior das salas de aula apresentam uma magnitude muito superior à estimada para o papel do conhecimento dos docentes nas disciplinas ministradas. A maioria dessas estimativas mostrou-se dez vezes mais eficaz na produção de conhecimento dos alunos em matemática do que uma substancial mudança no nível de conhecimento do professor. A variação na erudição dos docentes que estamos supondo nessa comparação representa aproximadamente o aumento da nota em um ponto ou um desvio-padrão. Isto equivale à passagem de um professor do percentil 30 para o percentil 70 na distribuição de notas na parte objetiva da prova de promoção.

Entretanto, é importante destacar que essas atividades pedagógicas são muitas vezes complementares e provavelmente existe uma alta correlação entre os docentes que realizam cada uma dessas práticas em sala de aula. Complementarmente, tal como destacado na seção introdutória deste capítulo, se as medidas de práticas de ensino estiverem refletindo características não observáveis dos docentes (motivação e/ou comprometimento), então os efeitos estimados podem não ser verificados no caso de intervenções que pretendam incentivar a adoção dessas atividades pedagógicas nas salas de aula. Essa hipótese alternativa para os impactos observados são parcialmente refutadas pelo fato de que as cinco práticas analisadas para os professores de matemática permanecem estatisticamente significantes e com magnitudes relevantes quando incluídas conjuntamente na função de produção educacional. Esses resultados, apresentados na coluna 6 da tabela 2.5, sugerem que o efeito sobre o ganho de desempenho dos alunos é realmente proveniente das práticas pedagógicas e não apenas devido a aspectos não observáveis dos professores. Isso, pois ao condicionarmos a análise às outras variáveis que são potencialmente correlacionadas com o comprometimento dos professores, a adoção frequente das práticas pedagógicas consideradas permanece economicamente importante. Além disso, essas estimativas revelam uma nova informação acerca do impacto de cada uma dessas atividades pedagógicas. Pois na estimação conjunta das cinco práticas pedagógicas, nós estamos condicionando ao contexto de docentes que já tem uma alta qualidade – adotam sempre as demais práticas – e também ao patamar de conhecimento desses professores.

A comparação das estimativas dos efeitos isolados (colunas 1 a 5) com as obtidas de maneira concomitante (coluna 6) revela uma significativa redução na magnitude dos impactos das práticas pedagógicas sobre a aquisição de proficiência dos alunos. Todavia, os efeitos permanecem positivos, estatisticamente distintos de zero e ainda relevantes sob o ponto de vista da magnitude do incremento no aprendizado dos alunos decorrente de cada uma dessas práticas pedagógicas. De acordo com os resultados da coluna 6, condicional ao conhecimento dos docentes, uma intervenção que consiste em alterar a alocação de uma turma de alunos de um professor considerado ruim para outro que sempre realiza essas cinco atividades pedagógicas resulta em um aumento no desempenho desses alunos de 25,6% de um desvio-padrão da distribuição de notas⁵⁵. Essa ação implica um aumento médio na proficiência em matemática dos estudantes de 10,8 pontos na escala Saeb, o que representa um ganho de 4,2% do desempenho médio dos alunos da oitava série no SARESP 2009.

Os resultados relativos ao efeito das práticas pedagógicas são robustos à definição das variáveis que indicam quais são os professores que adotam em todas ou quase todas as aulas as práticas pedagógicas. Isso pode ser observado a partir dos painéis B e C da tabela 2.5. No painel B, utilizamos uma definição menos restritiva de quais são os professores que sempre praticam as atividades – 50% ou mais dos alunos na turma devem afirmar – e as estimativas são marginalmente menores em magnitude. Já o painel C mostra os resultados derivados da adoção de uma definição mais restritiva dos docentes que sempre realizam as práticas de ensino – 90% ou mais dos alunos na turma – e as estimativas são um pouco maiores. No entanto, apesar dessa variação marginal na magnitude dos efeitos, observamos o mesmo padrão de impacto das práticas pedagógicas independente da definição da medida utilizada para a estimação.

De maneira geral, as evidências obtidas para os efeitos das práticas pedagógicas e conhecimento dos docentes sobre o ganho de proficiência em língua portuguesa dos alunos são qualitativamente semelhantes aos verificados para matemática. As estimativas são frequentemente positivas, estatisticamente

⁵⁵ Aqui estamos definindo um professor ruim como aquele que não passa nem corrige lições de casa, não explica a matéria até que todos os alunos entendam, não propõem atividades de resolução de problemas variados e não relaciona os conteúdos de matemática às situações do cotidiano em todas ou quase todas as aulas.

distintas de zero e relevantes sob o ponto de vista do tamanho dos impactos. A tabela 2.6 reproduz para língua portuguesa as mesmas especificações apresentadas na tabela 2.5 para matemática. Os resultados relativos aos coeficientes das práticas de ensino estimados isoladamente são apresentados nas colunas 1 a 4 e as estimativas da especificação que inclui de maneira conjunta as atividades pedagógicas estão listadas na coluna 5 da tabela 2.6. Nesta tabela nós mantemos a estrutura de organização da tabela 2.5 e cada um dos três painéis apresenta os resultados obtidos a partir de diferentes definições da medida de adoção das práticas pedagógicas por parte dos professores de língua portuguesa⁵⁶.

As principais diferenças estão associadas à magnitude dos efeitos em si, com a particularidade de que estes são em geral menores para língua portuguesa em comparação à disciplina de matemática. Para as atividades pedagógicas que são observadas tanto para os professores de língua portuguesa quanto de matemática – passar e corrigir lição de casa e explicar a matéria até que todos os alunos entendam – os efeitos isolados de um docente de língua portuguesa que sempre realiza essas práticas são em termos absolutos menores do que os observados para matemática⁵⁷. Esse padrão de resultados se faz presente para os três painéis apresentados nas tabelas 2.5 e 2.6. Uma hipótese para explicar essa menor importância relativa das práticas de ensino é o fato de que os alunos na oitava série do ensino fundamental provavelmente já têm as habilidades relacionadas à disciplina de língua portuguesa (leitura, vocabulário, interpretação de textos etc.) mais desenvolvidas e conseqüentemente menos sensíveis a influência dos docentes (Metzler & Woessmann, 2012).

Um resultado que merece destaque na tabela 2.6 é o impacto da prática do professor sempre “indicar livros de literatura para ler”. Por exemplo, se nós trocássemos um professor que não adota essa atividade em todas ou quase todas as aulas por outro que sempre indique livros, essa mudança implica um aumento de 7,49% de um desvio-padrão da distribuição de notas em língua portuguesa dos alunos no SARESP 2009 – painel A da tabela 2.6. A adoção isolada dessa prática de ensino é capaz de aumentar o desempenho dos alunos em aproximadamente 3,5 pontos na escala Saeb. Na especificação que também condiciona ao fato dos

⁵⁶ A partir da amostra utilizada para obter os resultados para língua portuguesa na tabela 2.4, para 20.324 alunos nós não recuperamos as respostas destes sobre os itens referentes aos docentes de língua de portuguesa no questionário de alunos do SARESP 2009.

⁵⁷ Isso pode ser constatado a partir da comparação das colunas 1 a 3 entre as tabelas 4 e 5.

professores realizarem sempre as demais práticas de ensino avaliadas (coluna 5), apesar da redução na magnitude do impacto, a estimativa do efeito dessa atividade pedagógica passa a ser igual a 2,9% de um desvio-padrão e permanece estatisticamente significativa (painel A).

Tabela 2.6 – Efeito conjunto do conhecimento e práticas pedagógicas do professor sobre o ganho de desempenho dos alunos em língua portuguesa

Nota Saresp 2009 padronizada	Lingua Portuguesa				
	(1)	(2)	(3)	(4)	(5)
Painel A: 75% ou mais dos alunos da turma afirmaram que o professor sempre adotava a prática pedagógica					
Nota objetiva prova promoção	0.0111 (0.00266)***	0.00821 (0.00266)***	0.00814 (0.00266)***	0.0105 (0.00266)***	0.00657 (0.00266)***
Práticas pedagógicas					
Passa lição sempre (> 75% alunos)	0.0841 (0.0130)***				0.0376 (0.0134)***
Corrige lição sempre (> 75% alunos)		0.0638 (0.00538)***			0.0405 (0.00575)***
Explica a matéria (> 75% alunos)			0.0626 (0.00510)***		0.0438 (0.00540)***
Indica livros literatura (> 75% alunos)				0.0749 (0.0138)***	0.0289 (0.0140)**
Painel B: 50% ou mais dos alunos da turma afirmaram que o professor sempre adotava a prática pedagógica					
Nota objetiva prova promoção	0.0104 (0.00265)***	0.00907 (0.00265)***	0.00924 (0.00265)***	0.00966 (0.00266)***	0.00718 (0.00265)***
Práticas pedagógicas					
Passa lição sempre (> 50% alunos)	0.0518 (0.00582)***				0.0321 (0.00617)***
Corrige lição sempre (> 50% alunos)		0.0574 (0.00573)***			0.0312 (0.00633)***
Explica a matéria (> 50% alunos)			0.0603 (0.00632)***		0.0370 (0.00682)***
Indica livros literatura (> 50% alunos)				0.0432 (0.00659)***	0.0246 (0.00675)***
Painel C: 90% ou mais dos alunos da turma afirmaram que o professor sempre adotava a prática pedagógica					
Nota objetiva prova promoção	0.0111 (0.00265)***	0.0102 (0.00266)***	0.0103 (0.00264)***	0.0113 (0.00266)***	0.00961 (0.00264)***
Práticas pedagógicas					
Passa lição sempre (> 90% alunos)	0.131 (0.0290)***				0.0702 (0.0281)**
Corrige lição sempre (> 90% alunos)		0.0706 (0.00967)***			0.0437 (0.00981)***
Explica a matéria (> 90% alunos)			0.0791 (0.00840)***		0.0630 (0.00855)***
Indica livros literatura (> 90% alunos)				0.0931 (0.0310)***	0.0199 (0.0312)
Controles:					
Desempenho passado - Saresp 2007	Sim	Sim	Sim	Sim	Sim
Idade professor	Sim	Sim	Sim	Sim	Sim
Tamanho da classe	Sim	Sim	Sim	Sim	Sim
Características escola / profs médias	Sim	Sim	Sim	Sim	Sim
Características background familiar	Sim	Sim	Sim	Sim	Sim
N. Observações	142,568	142,568	142,568	142,568	142,568

Fonte: Elaboração própria, dados do SARESP 2009, SARESP 2007, Censo Escolar 2009, Notas da prova de promoção e dados administrativos da atribuição de professores da SEE/SP.

Erros padrão entre parênteses - *** p<0,01; ** p<0,05; * p<0,1.

A prática de indicar livros de literatura para ler está diretamente relacionada ao desenvolvimento de uma importante categoria no domínio cognitivo que é o estímulo da capacidade de estudo individual dos alunos. Assim, os nossos resultados, obtidos para o contexto dos alunos da rede estadual paulista, demonstram que essa prática pedagógica apresenta eficácia na produção de habilidades cognitivas dos alunos pelo menos para língua portuguesa. Esse resultado está em consonância ao observado por Fryer Jr. (2011) entre os alunos que tem como língua mãe o inglês na segunda série da rede de ensino no município de Dallas. Nesse estudo, Fryer Jr. (2011) investiga o impacto de incentivos financeiros sobre o desempenho dos alunos em testes e para os alunos de Dallas, em particular, o tratamento foi pagar US\$ 2,00 por livro lido. Os resultados indicam que essa intervenção para esse subgrupo de estudantes aumenta o desempenho em leitura em 17,3% de um desvio-padrão. Por outro lado, o nosso resultado é contrário à evidência obtida por Lavy (2011) para alunos da terceira e quinta série em Israel. Pois nesse caso os resultados mostram que as medidas agregadas que estimulam a capacidade de estudo individual não têm influência sobre o desempenho dos alunos nos testes.

Assim como para matemática, a estimação da especificação da função de produção educacional com as quatro práticas pedagógicas conjuntamente resulta em uma redução nos efeitos obtidos em comparação com os impactos isolados apresentados nas colunas 1 a 4. Entretanto, essas estimativas permanecem positivas, estatisticamente distintas de zero e com magnitudes relevantes, o que corrobora que as práticas pedagógicas em si são as responsáveis pelo incremento na aprendizagem dos alunos também em língua portuguesa. Ou seja, os impactos estimados não derivam de características não observáveis dos professores, tais como motivação e comprometimento. A partir das estimativas apresentadas na coluna 5, do painel A, da tabela 2.6, nós reproduzimos o exercício hipotético de supor uma intervenção drástica que altere a alocação de uma turma de alunos de um docente considerado ruim para outro que utiliza sempre as quatro práticas de ensino. Para a disciplina de língua portuguesa, essa intervenção implica um ganho de proficiência dos alunos de aproximadamente 15% de um desvio-padrão da distribuição de notas. Esse incremento equivale a 6,7 pontos na escala Saeb e 2,8% do desempenho médio em língua portuguesa da amostra principal de alunos da oitava série.

Dessa forma, os resultados apresentados nas tabelas 2.5 e 2.6 reforçam a importância relativa da adoção frequente de práticas de ensino eficazes no interior das salas de aula. A magnitude dos impactos dessas atividades pedagógicas sobre a produção de habilidades cognitivas pelos alunos mostrou-se bastante superior ao efeito isolado de docentes com um maior estoque de conhecimento específico a cada uma das disciplinas. Além disso, os resultados mostraram-se robustos a diferentes definições da variável que identifica os professores que adotam sempre as práticas de ensino. Motivado por essa robustez dos resultados, nos demais procedimentos de estimação da função de produção de habilidades cognitivas, cujas estimativas são apresentadas nas tabelas abaixo, nós utilizamos apenas a principal definição das práticas pedagógicas. Isto é, quando 75% ou mais dos alunos de uma turma afirmam que o professor realiza a atividade em todas ou quase todas as aulas.

Com o intuito de melhor compreender a importância da qualidade dos professores, nós investigamos a possibilidade de existência de complementaridades entre a medida de conhecimento objetivo dos docentes e as práticas de ensino que estes realizam em sala de aula. O objetivo dessa análise é verificar se o efeito de determinada atividade pedagógica é reforçada quando aplicada por um professor que tem maior erudição na disciplina ou em outros termos verificar se a transmissão de conhecimento do docente para os seus alunos é facilitada por alguma prática em especial.

A tabela 2.7 apresenta as estimativas dos impactos do conhecimento e práticas de ensino dos professores de matemática sobre a aprendizagem dos alunos, assim como os coeficientes associados às interações entre essas variáveis de interesse.

Tabela 2.7 – Efeito conjunto do conhecimento e das práticas pedagógicas do professor sobre o ganho de desempenho dos alunos em matemática: inclusão de interações entre as notas e atividades dos docentes

Nota Saresp 2009 padronizada	Matemática					
	(1)	(2)	(3)	(4)	(5)	(6)
Nota objetiva prova promoção	0.0107 (0.00388)***	0.00564 (0.00438)	0.00222 (0.00477)	0.0122 (0.00375)***	0.0126 (0.00364)***	0.000284 (0.00490)
Práticas pedagógicas						
Passa lição sempre (> 75% alunos)	0.126 (0.0113)***					0.0518 (0.0121)***
Interação nota e lição	0.000686 (0.00991)					-0.00147 (0.0112)
Corrige lição sempre (> 75% alunos)		0.107 (0.00740)***				0.0590 (0.00865)***
Interação nota e corrige lição		0.00898 (0.00723)				0.00605 (0.00888)
Explica a matéria (> 75% alunos)			0.0792 (0.00709)***			0.0285 (0.00756)***
Interação nota e explica matéria			0.0167 (0.00695)**			0.0148 (0.00766)*
Propõe resolução problemas (> 75%)				0.136 (0.0122)***		0.0669 (0.0136)***
Interação nota e resolução problemas				-0.0103 (0.0111)		-0.0175 (0.0126)
Relaciona os conteúdos ao cotidiano (> 75%)					0.147 (0.0199)***	0.0525 (0.0207)**
Interação nota e relaciona ao cotidiano					-0.0133 (0.0192)	-0.00302 (0.0202)
Controles:						
Desempenho passado - Saresp 2007	Sim	Sim	Sim	Sim	Sim	Sim
Idade professor	Sim	Sim	Sim	Sim	Sim	Sim
Tamanho da classe	Sim	Sim	Sim	Sim	Sim	Sim
Características escola / profs médias	Sim	Sim	Sim	Sim	Sim	Sim
Características background familiar	Sim	Sim	Sim	Sim	Sim	Sim
N. Observações	135,913	135,913	135,913	135,913	135,913	135,913
R-quadrado	0.411	0.412	0.411	0.411	0.410	0.413

Fonte: Elaboração própria, dados do SARESP 2009, SARESP 2007, Censo Escolar 2009, Notas da prova de promoção e dados administrativos da atribuição de professores da SEE/SP.

Erros padrão entre parênteses - *** $p < 0,01$; ** $p < 0,05$; * $p < 0,1$.

A comparação das colunas 1 a 5 entre as tabelas 2.5 (painel A) e 2.7 nos permite concluir que a inclusão das interações entre as notas e as medidas de práticas pedagógicas dos professores não tem influência sobre os coeficientes associados ao impacto isolado das atividades pedagógicas. Esses resultados permanecem praticamente idênticos. Além disso, as estimativas referentes às interações são na maioria dos casos estatisticamente não distintas de zero. Exceção feita à interação entre a prática de sempre explicar a matéria até que todos os alunos entendam com o conhecimento específico à matemática do docente, tal como pode ser observado na coluna 3 da tabela 2.7. Nesse caso em particular, o impacto direto do conteúdo do professor na matéria deixa de ser

estatisticamente significativa – primeira linha da coluna 3 – e a transmissão de conhecimento aos alunos se dá por meio da prática frequente de explicar o conteúdo da disciplina aos alunos. Uma possível explicação para esse resultado é o fato de que na oitava série do ensino fundamental as habilidades de matemática dos estudantes ainda estão em franco processo de desenvolvimento (Metzler & Woessmann, 2012). Assim, dentre os professores que optam por sempre explicar a matéria, aqueles que têm um patamar de conhecimento superior obtêm ganhos adicionais no desempenho dos seus alunos. Uma vez que esse aspecto deve ser bastante relevante na tecnologia de produção das habilidades cognitivas para matemática.

O incremento de um desvio-padrão na erudição do docente de matemática que sempre explica a matéria implica um aumento de 1,6% de um desvio-padrão da distribuição de notas dos alunos. Esse efeito é estatisticamente significativo ao nível de 5%. Inúmeros são os exemplos (anedóticos) por meio dos quais uma maior erudição em matemática pode auxiliar os professores no processo de explicar a matéria aos seus alunos, entre eles: (i) uso de exemplos diversos para ilustrar o conteúdo; (ii) maior facilidade em encontrar formas alternativas de explicar a matéria e possíveis dúvidas; (iii) maior confiança e por conseguinte clareza na explanação dos tópicos etc.

A coluna 6 da tabela 2.7 apresenta os resultados relativos ao exercício de estimar conjuntamente os efeitos das práticas pedagógicas dos professores condicional ao nível de proficiência dos docentes. Adicionalmente nós também incluímos as interações entre as práticas e a medida de conhecimento dos docentes. Os resultados para os efeitos diretos das práticas de ensino sobre a produção de habilidades cognitivas dos alunos permanecem muito semelhantes aos observados na especificação sem a inclusão das interações – comparar as estimativas pontuais das colunas 6 entre as tabelas 2.5 (painel A) e 2.7. Ademais, nessa especificação persiste o efeito positivo e de magnitude semelhante à observada para a interação entre o conhecimento em matemática e o fato do professor sempre explicar a matéria, quando estimada isoladamente – na coluna 3 da tabela 2.7⁵⁸. Essa evidência reforça a tese de que, sobretudo para as habilidades relacionadas à matemática, o estoque de conhecimento do docente parece ser

⁵⁸ A única diferença é que a estimativa agora é relevante a um nível de significância de 10% ao invés de 5%.

muito importante para que ele consiga transmitir o conteúdo de uma forma ainda mais eficaz ao explicar a matéria para os seus estudantes.

Já a tabela 2.8 reproduz para língua portuguesa as mesmas especificações – incluindo os termos relativos à interação entre o conhecimento dos docentes de língua portuguesa e as práticas adotadas em sala de aula – apresentadas na tabela 2.7 para matemática.

Tabela 2.8 – Efeito conjunto do conhecimento e das práticas pedagógicas do professor sobre o ganho de desempenho dos alunos em língua portuguesa: inclusão de interações entre as notas e atividades dos docentes

Nota Saresp 2009 padronizada	Lingua Portuguesa				
	(1)	(2)	(3)	(4)	(5)
Nota objetiva prova promoção	0.0106 (0.00269)***	0.00971 (0.00312)***	0.00494 (0.00330)	0.00907 (0.00269)***	0.00562 (0.00344)
Práticas pedagógicas					
Passa lição sempre (> 75% alunos)	0.0829 (0.0133)***				0.0365 (0.0137)***
Interação nota e lição	0.0128 (0.0148)				0.00377 (0.0154)
Corrige lição sempre (> 75% alunos)		0.0644 (0.00547)***			0.0420 (0.00580)***
Interação nota e corrige lição		-0.00503 (0.00571)			-0.0126 (0.00614)**
Explica a matéria (> 75% alunos)			0.0619 (0.00512)***		0.0437 (0.00540)***
Interação nota e explica matéria			0.00860 (0.00526)		0.00790 (0.00551)
Indica livros literatura (> 75% alunos)				0.0629 (0.0147)***	0.0151 (0.0148)
Interação nota e indica livros				0.0411 (0.0154)***	0.0463 (0.0158)***
Controles:					
Desempenho passado - Saresp 2007	Sim	Sim	Sim	Sim	Sim
Idade professor	Sim	Sim	Sim	Sim	Sim
Tamanho da classe	Sim	Sim	Sim	Sim	Sim
Características escola / profs médias	Sim	Sim	Sim	Sim	Sim
Características background familiar	Sim	Sim	Sim	Sim	Sim
N. Observações	142,568	142,568	142,568	142,568	142,568
R-quadrado	0.536	0.536	0.536	0.536	0.537

Fonte: Elaboração própria, dados do SARESP 2009, SARESP 2007, Censo Escolar 2009, Notas da prova de promoção e dados administrativos da atribuição de professores da SEE/SP.

Erros padrão entre parênteses - *** $p < 0,01$; ** $p < 0,05$; * $p < 0,1$.

Assim como para matemática, as estimativas referentes aos efeitos das atividades pedagógicas pouco se alteraram após a inclusão dos termos de interação com a nota dos professores de língua portuguesa. E os resultados para as interações são na maioria dos casos não estatisticamente distintos de zero. A única

prática de ensino que apresentou uma importante complementaridade com o conhecimento do docente foi o fato deste sempre indicar livros de literatura para ler. Em relação ao efeito isolado, tanto a adoção frequente dessa atividade em sala de aula quanto o impacto adicional decorrente de a prática ser utilizada por professores mais eruditos são estatisticamente distintos de zero e com magnitudes importantes (coluna 4 da tabela 2.8). No entanto, na especificação que também condiciona aos professores que adotavam as demais práticas pedagógicas (coluna 5 da tabela 2.8), o canal que persiste estatisticamente significativo é justamente o da interação entre a prática e o conhecimento do docente. Ou seja, em um contexto de professores melhores e/ou mais motivados (adotam sempre as práticas de ensino eficazes) a indicação frequente de livros de literatura para ler apenas influenciará a aquisição de proficiência dos alunos, se os docentes estiverem passando material adequado ao estágio de desenvolvimento intelectual dos alunos. Dessa forma, estimulando o interesse pela leitura e contribuindo para o acúmulo de conhecimento. Esse mecanismo, que estamos racionalizando por meio do qual a transmissão de conhecimento se viabiliza pela atividade do professor em sala de aula, apenas terá fundamento se o professor conhecer profundamente os livros que estiver indicando aos seus alunos. Isto é, conhecer bem a disciplina que leciona.

Dessa forma, as evidências apresentadas nas tabelas 2.7 e 2.8 indicam que os efeitos de grande parte das práticas pedagógicas são independentes do nível de proficiência dos professores tanto de matemática quanto de língua portuguesa. Exceção feita a importância do conhecimento dos docentes de matemática na atividade de sempre explicar a matéria até que todos os alunos entendam e dos professores de língua portuguesa na prática de sempre indicar livros de literatura para ler. Em outras palavras, é praticamente inócua a existência na rede de ensino de docentes com elevado conhecimento nas disciplinas, se estes não souberem e/ou não forem capazes de realizar as atividades em sala de aula que realmente contribuem na produção das habilidades cognitivas dos estudantes.

Por último, assim como verificado para o conhecimento específico à disciplina dos professores, nós não constatamos efeitos heterogêneos das diferentes práticas pedagógicas dos docentes sobre o ganho de proficiência dos alunos ao longo da distribuição de desempenho passado destes, tanto para matemática quanto para língua portuguesa.

2.5.3 Condicionando para a qualidade de gestão da escola

Assim como explicitado na seção de estratégia empírica, a inserção de uma medida da qualidade do diretor na especificação estimada nos permite levar em consideração diferenças em aspectos não observáveis relativos à gestão da escola que podem ter um impacto sobre o comportamento dos professores e o aprendizado dos alunos. No sentido de ilustrar essa questão, supomos que um mesmo professor ministre aulas em duas escolas administradas por diferentes diretores. Caso esse docente não esteja suficientemente motivado devido às condições de trabalho e/ou salariais, o fato de haver em uma das escolas um diretor com melhor capacidade de gestão e/ou liderança em relação ao outro pode ter um papel relevante nas diferenças de desempenho entre os estudantes. Alguns exemplos de mecanismos pelos quais esse diretor mais hábil pode influenciar o esforço/comportamento do docente são o planejamento e orientação sobre os objetivos a serem alcançados, o acompanhamento da execução das atividades de ensino no interior da sala de aula e a fiscalização da qualidade da instrução disponibilizada aos alunos, entre outras ações.

Dessa forma, a partir dos resultados derivados da estimação da equação (2.7) nós podemos averiguar se os impactos das dimensões de qualidade dos professores anteriormente obtidos sofrem alguma influência de características associadas à qualidade da gestão da escola. De maneira complementar, também podemos aferir se a qualidade da gestão da escola – medida pela nota do diretor na parte objetiva da prova de promoção – tem algum efeito direto sobre a aquisição de proficiência dos alunos. As tabelas 2.9 e 2.10 apresentam as evidências dos efeitos do conhecimento e das práticas pedagógicas, isoladas e em conjunto, dos professores sobre o desempenho dos alunos, condicional ao desempenho dos diretores na prova de promoção, respectivamente para matemática e língua portuguesa.

Para matemática, mesmo após a inserção de controle para a qualidade do diretor, as nossas estimativas continuam muito semelhantes às apresentadas nas tabelas 2.5 e 2.7. Isso se dá tanto para os efeitos do conhecimento dos professores quanto para os impactos das atividades pedagógicas empregadas sempre na sala de aula. Nesse último caso, os resultados são muito similares quando estimamos

separadamente o efeito das práticas, mas também para as especificações nas quais condicionamos aos docentes que adotaram as demais práticas de ensino conjuntamente. O fato mais marcante é que as estimativas pontuais dessas variáveis de interesse permanecem bastante próximas entre as duas abordagens, com variações mínimas, apesar da significativa redução da amostra (menos 57.538 alunos)⁵⁹. Dessa forma, as evidências apresentadas sugerem que as nossas estimativas da eficácia do conhecimento e habilidades dos professores sobre a tecnologia de produção de conhecimento em matemática são robustas a fatores não observáveis associados à qualidade da gestão escolar.

Tabela 2.9 – Efeito conjunto do conhecimento e das práticas pedagógicas do professor – e interações – sobre o ganho de desempenho dos alunos em matemática: condicional a qualidade do diretor

Nota Saresp 2009 padronizada	Matemática							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Nota objetiva professor	0.0123 (0.00462)***	0.0127 (0.00484)***	0.00915 (0.00508)*	0.00728 (0.00568)	0.00159 (0.00628)	0.0124 (0.00502)**	0.0122 (0.00477)**	-0.000517 (0.00640)
Nota objetiva diretor	0.0163 (0.00472)***	0.0176 (0.00491)***	0.0166 (0.00486)***	0.0157 (0.00484)***	0.0186 (0.00485)***	0.0183 (0.00491)***	0.0179 (0.00488)***	0.0166 (0.00480)***
Práticas pedagógicas								
Passa lição sempre (> 75% alunos)			0.120 (0.0156)***					0.0564 (0.0158)***
Interação nota e lição			0.00211 (0.0138)					0.00910 (0.0147)
Corrige lição sempre (> 75% alunos)				0.0995 (0.00991)***				0.0581 (0.0115)***
Interação nota e corrige lição				0.00322 (0.00964)				-0.00388 (0.0117)
Explica a matéria (> 75%)					0.0694 (0.00935)***			0.0242 (0.00987)**
Interação nota e explica matéria					0.0154 (0.00920)*			0.0194 (0.0101)*
Propõe resolução problemas (> 75%)						0.117 (0.0160)***		0.0519 (0.0163)***
Interação nota e resolução problemas						-0.0140 (0.0143)		-0.0226 (0.0146)
Relaciona os conteúdos ao cotidiano (> 75%)							0.132 (0.0276)***	0.0469 (0.0278)*
Interação nota e relaciona ao cotidiano							-0.0140 (0.0282)	0.00265 (0.0287)
Controles:								
Desempenho passado - Saresp 2007	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Idade professor	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Tamanho da classe	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Características escola / profs médias	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Características background familiar	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
N. Observações	88.740	78.375	78.375	78.375	78.375	78.375	78.375	78.375
R-quadrado	0.407	0.410	0.411	0.412	0.411	0.411	0.410	0.413

Fonte: Elaboração própria, dados do SARESP 2009, SARESP 2007, Censo Escolar 2009, Notas da prova de promoção e dados administrativos da atribuição de professores da SEE/SP.

Erros padrão entre parênteses - *** $p < 0,01$; ** $p < 0,05$; * $p < 0,1$.

⁵⁹ Essa redução de aproximadamente 40% no número de alunos que compõem a amostra se deve a nós não termos as notas na prova de promoção para todos os diretores das escolas que constituem a amostra principal.

O desempenho do diretor nos testes tem um impacto direto sobre a proficiência em matemática dos alunos. Conforme apresentado na tabela 2.9, este é positivo e estatisticamente distinto de zero. Além disso, em todas as especificações apresentadas entre as colunas 1 a 8 da tabela 2.9, a magnitude desse efeito oscila marginalmente ao redor de 1,6% de um desvio-padrão da distribuição de notas em matemática dado um aumento de um desvio-padrão no desempenho do diretor na prova de promoção.

Entretanto, essa influencia direta da qualidade do diretor sobre o desempenho dos alunos não é verificada para as habilidades cognitivas relacionadas à disciplina de língua portuguesa. Isso é constatado a partir dos resultados apresentados entre as colunas 1 a 7 da tabela 2.10. Essa diferença no papel da qualidade do diretor sobre o desempenho dos alunos entre as duas matérias avaliadas tem um paralelo na escassa literatura a respeito. Clark, Martorell & Rockoff (2009) avaliam a importância de outro fator relacionado à qualidade dos diretores – a experiência no cargo. Nesse estudo, os autores encontram uma relação positiva entre a experiência do diretor e o desempenho escolar, particularmente para as notas dos testes em matemática e absenteísmo dos alunos.

Comparativamente a matemática, os resultados obtidos após a inclusão de controle para a qualidade de gestão dos diretores não são igualmente robustos para língua portuguesa. Primeiramente, o efeito positivo do conhecimento dos professores sobre a proficiência dos alunos em língua portuguesa observado na amostra principal passa a ser estatisticamente não distinto de zero. Isto ocorre, não obstante, nós não termos verificado a existência de uma relação significativa entre as notas dos diretores e o desempenho dos alunos. Além disso, apesar dos impactos isolados das práticas pedagógicas dos professores permanecerem positivos e relevantes sobre o aprendizado dos alunos em língua portuguesa. Na maioria dos casos, as estimativas sofrem uma substancial redução em magnitude condicional na variável *proxy* para a qualidade de gestão das escolas. Já para a especificação que inclui conjuntamente as práticas de ensino na função de produção educacional, os resultados referentes ao docente sempre passar lição de casa e indicar livros de literatura para ler deixam de ser estatisticamente significantes – coluna 7 da tabela 2.10.

Tabela 2.10 – Efeito conjunto do conhecimento e das práticas pedagógicas do professor – e interações – sobre o ganho de desempenho dos alunos em língua portuguesa: condicional a qualidade do diretor

Nota Saresp 2009 padronizada	Lingua Portuguesa						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Nota objetiva professor	0.00444 (0.00338)	0.00624 (0.00352)*	0.00520 (0.00357)	0.00358 (0.00412)	-9.28e-05 (0.00442)	0.00292 (0.00357)	-0.000418 (0.00460)
Nota objetiva diretor	0.00487 (0.00316)	0.00463 (0.00329)	0.00428 (0.00329)	0.00353 (0.00325)	0.00335 (0.00324)	0.00444 (0.00325)	0.00293 (0.00320)
Práticas pedagógicas							
Passa lição sempre (> 75% alunos)			0.0651 (0.0171)***				0.0287 (0.0178)
Interação nota e lição			0.0172 (0.0183)				-9.60e-05 (0.0197)
Corrige lição sempre (> 75% alunos)				0.0538 (0.00722)***			0.0337 (0.00779)***
Interação nota e corrige lição				0.00127 (0.00760)			-0.00616 (0.00809)
Explica a matéria (> 75%)					0.0591 (0.00671)***		0.0456 (0.00715)***
Interação nota e explica matéria					0.00923 (0.00703)		0.00502 (0.00725)
Indica livros literatura (> 75%)						0.0346 (0.0193)*	-0.00872 (0.0195)
Interação nota e indica livros						0.0590 (0.0192)***	0.0644 (0.0195)***
Controles:							
Desempenho passado - Saresp 2007	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Idade professor	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Tamanho da classe	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Características escola / profs médias	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Características background familiar	Sim	Sim	Sim	Sim	Sim	Sim	Sim
N. Observações	91,129	80,591	80,591	80,591	80,591	80,591	80,591
R-quadrado	0.531	0.532	0.532	0.532	0.533	0.532	0.533

Fonte: Elaboração própria, dados do SARESP 2009, SARESP 2007, Censo Escolar 2009, Notas da prova de promoção e dados administrativos da atribuição de professores da SEE/SP.

Erros padrão entre parênteses - *** $p < 0,01$; ** $p < 0,05$; * $p < 0,1$.

No entanto, apesar dessas diferenças, as evidências ainda corroboram algumas das conclusões referentes à eficácia da adoção frequente de certas práticas pedagógicas pelos professores de língua portuguesa no interior das salas de aula, mesmo ao condicionarmos a análise à qualidade da gestão escolar.

2.6 Conclusões

Atualmente existe um consenso na literatura de economia da educação a respeito da importância da qualidade dos professores para o aprendizado dos alunos. Todavia, quais são as características ou habilidades dos professores

eficazes na produção de habilidades cognitivas ainda é uma questão em aberto. Nós utilizamos um conjunto de bases de dados sobre os docentes e alunos da rede estadual paulista que permite investigar os efeitos de duas dimensões de qualidade dos professores sob um arcabouço de função de produção educacional com valor adicionado. Os nossos resultados mostram que tanto o conhecimento da disciplina ministrada quanto a adoção frequente de práticas pedagógicas na sala de aula tem efeitos positivos e estatisticamente significantes sobre a aprendizagem dos alunos.

Em termos relativos, entretanto, os impactos do conhecimento dos professores sobre a aquisição de habilidades cognitivas dos alunos apresentou magnitude reduzida em comparação com os efeitos da realização frequente de algumas atividades pedagógicas pelos professores nas duas disciplinas avaliadas. De acordo com a literatura em psicologia educacional essas práticas estão associadas ao desenvolvimento de importantes categorias no domínio cognitivo. Assim, um aumento em um desvio-padrão na nota dos professores eleva em aproximadamente 1% de um desvio-padrão o desempenho dos alunos. Já uma intervenção que substitui um docente que não passa sempre lição de casa por outro que o faz tem um impacto de 12,6% e 8,4% de um desvio-padrão da distribuição de notas, respectivamente para matemática e língua portuguesa. Esses resultados são robustos a inclusão de controles para a qualidade da gestão escolar pelos diretores. Além disso, há fortes indícios de que os efeitos das práticas pedagógicas são independentes do nível de conhecimento do professor específico à disciplina.

Dessa forma, as nossas evidências apontam no sentido que a dimensão da qualidade do professor que parece ser mais relevante para o acréscimo no aprendizado dos alunos é o emprego de práticas pedagógicas eficazes na sala de aula. Em outras palavras, a seleção e atribuição de professores com alto nível de conhecimento nas disciplinas pode ser inócua, caso não seja acompanhada por outras medidas. Talvez, o mais importante seja identificar as práticas de ensino mais eficazes para cada contexto e treinar os docentes para utilizá-las de maneira adequada e com a frequência necessária. Consequentemente, os nossos resultados tem implicações sobre políticas públicas. A partir das evidências aqui apresentadas, intervenções que tenham como objetivo o treinamento dos professores para a utilização adequada das práticas pedagógicas eficazes devem apresentar uma relação custo-benefício mais favorável que outras políticas

usualmente adotadas como, por exemplo, o pagamento de bônus por meta, redução do tamanho das turmas, alocação de professores adicionais nas salas de aula, ampliação da jornada escolar para período integral, entre outras.

Por último, nós devemos destacar que os nossos resultados dependem em grande parte da capacidade da parte objetiva das provas do Sistema de Promoção da SEE/SP em medir de maneira adequada o conhecimento dos docentes que é realmente relevante para o ensino das disciplinas de matemática e língua portuguesa.

3 Ideologia e tamanho de governo nos municípios brasileiros

3.1 Introdução

Qual é a relação entre identidade partidária e tamanho de governo? Este trabalho oferece evidência sobre esse tema utilizando dados brasileiros. É importante observar que existem visões teóricas distintas sobre essa relação. Modelos clássicos de concorrência política downsiana preveem que, em eleições bipartidárias, a identidade partidária não importará para as políticas públicas. Isso ocorre porque a concorrência política faz as propostas de ambos os partidos convergirem para a proposta preferida pelo eleitor mediano (Black, 1948; Downs, 1947 e Persson & Tabellini, 2000, cap. 3).

Mas os resultados desses modelos são sensíveis às hipóteses que: i) os candidatos são motivados unicamente pelo seu desejo de serem eleitos; e ii) existe comprometimento pleno dos candidatos com as políticas por eles prometidas. O resultado de convergência política desses modelos não será válido sem essas hipóteses. Isto é destacado por Besley & Case (2003) em sua revisão da evidência sobre o tema existente para os EUA. Essa consequência também é evidenciada pelo fato que a ausência de comprometimento dos candidatos com as políticas prometidas aos eleitores pode implicar divergência das plataformas apresentadas pelos partidos políticos mesmo em um contexto bipartidário, tal como nos modelos de “cidadão-candidato” de Besley & Coate (1997) e Osborne & Slivinsky (1997)⁶⁰.

Existe uma longa literatura empírica nos campos de ciência política e economia política que discute a relação entre identidade partidária e escolhas dos políticos em um sentido mais amplo. Mas como a maioria dessa literatura envolve estudos *cross country*, sua capacidade em elaborar afirmações causais sobre essa relação é limitada. Isso ocorre porque a maioria dos resultados dessa literatura

⁶⁰ Ver Persson e Tabellini (2000) para um tratamento simples dos modelos aqui mencionados. Esses autores discutem longamente hipóteses importantes para modelos de concorrência política no capítulo 2, apresentam modelos de concorrência política downsiana no capítulo 3 e um modelo de “cidadão-candidato” no capítulo 5.

pode ser fruto apenas de heterogeneidade não observada das unidades de observação⁶¹.

Foi apenas recentemente que surgiram trabalhos que utilizam estratégias de identificação que tornam mais críveis afirmações causais sobre a relação entre identidade partidária e escolhas políticas. O primeiro estudo dessa literatura é o de Lee, Moretti & Butler (2004). Esses autores utilizam um desenho de regressão com descontinuidade para avaliar se legisladores de partidos diferentes votam leis de forma diferente. A ideia básica desses autores é que distritos que elegem legisladores de esquerda e de direita por margens de vitória pequenas devem ser semelhantes na maioria dos aspectos (inclusive observáveis) relevantes para a determinação dos votos dos legisladores no congresso. Isso permite identificar o efeito causal da identidade partidária sobre os votos dos legisladores comparando os perfis de votos de congressistas de direita e esquerda eleitos por margens pequenas. Os resultados desses autores revelam que os políticos são “partidários”, por exemplo, políticos de diferentes partidos se comportam de maneira bastante diferente. Isso é consistente com modelos de concorrência política sem comprometimento pleno.

Trabalhos subsequentes como de Petersson-Lidbom (2008), Ferreira & Gyourko (2009) e Folke (2011) também utilizam o desenho de regressão com descontinuidade para testarem a influência dos partidos políticos sobre as políticas públicas em governos locais. O trabalho de Petersson-Lidbom (2008) mostra que a identidade ideológica dos governos locais suecos influencia as políticas públicas e também resultados econômicos. Governos de esquerda gastam mais, taxam mais, incentivam menos a imigração e têm menores taxas de desemprego. Já o trabalho de Folke (2011) mostra como pequenos partidos conseguem influenciar as políticas públicas nos governos locais suecos por meio de sua importância para a formação de coalizões.

O trabalho de Ferreira & Gyourko (2009) diverge dos três outros trabalhos apresentados sobre o tema porque reporta que os partidos políticos não são relevantes para a determinação de políticas públicas ao nível local nos EUA. Os autores apresentam evidência de que esse efeito é nulo devido à presença de mecanismos de seleção de *Tiebout*. Mas não é possível descartar a hipótese que

⁶¹ Ver Imbeau, Petry & Lamari (2001) para uma revisão de vários desses estudos.

esse resultado decorre do fato das municipalidades norte-americanas influenciarem um conjunto muito pequeno de políticas públicas. Isso é especialmente verdadeiro se comparamos a liberdade fiscal das municipalidades norte-americanas com as suecas. A maior liberdade fiscal dos municípios desse último país pode ser parcialmente responsável pela oposição dos resultados do trabalho de Ferreira & Gyourko (2009) com os de Petersson-Lidbom (2008) e Folke (2009).

O nosso estudo também utiliza o arcabouço de regressão com descontinuidade para oferecer evidência adicional sobre esse tema utilizando dados brasileiros de eleições e medidas de gastos e receitas locais. Os nossos resultados mostram que existem alguns efeitos da identidade partidária sobre o tamanho de governo nos municípios brasileiros. Em média, as municipalidades governadas pela esquerda gastam proporcionalmente mais com administração e menos com urbanismo e saúde. Entretanto, maiores gastos administrativos por parte de governos de esquerda não ocorrem mediante a contratação de maior pessoal. Em algumas especificações, os resultados também sugerem que a fração do IPTU sobre as receitas tributárias é maior em governos de esquerda.

Este trabalho se relaciona tanto com a literatura discutida acima que estuda se os políticos são “partidários”, quanto com a literatura de ciência política que estuda diversos aspectos do sistema político brasileiro⁶². Acreditamos que este artigo é relevante para ambas as literaturas porque é o primeiro trabalho, que temos conhecimento, que investiga causalmente a relação entre identidade partidária e escolhas políticas utilizando dados de um país em desenvolvimento em que os partidos políticos não são tão fortes ou consolidados como na Europa ou nos EUA⁶³.

O restante do capítulo está organizado da seguinte forma. A seção 3.2 apresenta a estratégia de identificação que adotamos no trabalho. A seção 3.3 descreve os dados e a seção 3.4 discute o contexto político brasileiro e as medidas

⁶² Figueiredo & Limongi (1999), Power (2000) e Power & Zucco Jr. (2009) são exemplos de trabalhos que investigam aspectos diversos do sistema partidário brasileiro.

⁶³ Existe evidência sobre a relação entre identidade (gênero e etnia) sobre políticas públicas na Índia. Ver Pande (2003) e Chattopadhyay & Duflo (2004) sobre esse tema. Mas não conhecemos evidência que investiga a relação entre identidade partidária ou ideológica sobre políticas públicas para países em desenvolvimento utilizando métodos similares aos aqui utilizados.

de ideologia que utilizamos. Já, os resultados são apresentados na seção 3.5. Por último, a seção 3.6 traz as conclusões do trabalho.

3.2 Estratégia de identificação

Nessa seção, nós discutimos a estratégia empírica utilizada para estimar o efeito da identidade partidária sobre as decisões de gastos e receitas nos municípios brasileiros. Suponha que existem apenas dois partidos (A e B) e seja D_i uma variável binária que toma valor zero caso o partido A esteja no poder e um caso o partido B esteja no poder. Defina também y_i como o resultado de uma política pública (gastos, impostos, investimentos públicos etc.) que potencialmente é influenciada pela identidade partidária e ε_i como um termo de erro ($E[\varepsilon_i] = 0$). A maneira mais simples de estimar o efeito da identidade partidária sobre uma escolha política qualquer seria estimar, por Mínimos Quadrados Ordinários (MQO), a seguinte equação:

$$y_i = \alpha + \rho D_i + \varepsilon_i \quad (3.1)$$

Entretanto, o parâmetro ρ apenas representa o efeito causal da identidade partidária (no caso, da presença do partido B no poder) sobre a política pública y_i caso D_i e ε_i não sejam correlacionados. Isso seria verdade caso a identidade partidária fosse determinada aleatoriamente. Mas a identidade do partido no poder não é determinada aleatoriamente. Além disso, é provável que diversos fatores observáveis e não observáveis que afetam as decisões políticas, e estão incluídos no termo de erro ε_i , também influenciem o partido que está no poder. Ideologia, renda, idade, etnia, ocupação etc. Esses são fatores que podem influenciar tanto o partido escolhido pelos eleitores quanto as escolhas de tamanho de governo preferidas por eles.

Esse problema poderia ser superado por meio da comparação de duas unidades políticas idênticas tanto em termos de características observáveis como em termos de características não observáveis. Uma ideia que tem sido amplamente utilizada pela literatura recente sobre o efeito de partidos políticos

sobre políticas públicas é a opção por um desenho de regressão com descontinuidade para recuperar o parâmetro ρ .

A maioria das regras eleitorais envolve descontinuidades. Por exemplo, os candidatos eleitos como prefeitos nos municípios são aqueles que têm mais votos válidos. Isso cria uma descontinuidade em que a identidade partidária do candidato eleito muda a partir de um limite. Com apenas dois candidatos esse limite é dado pela marca de 50% dos votos válidos.

Como a existência desse limite nos ajuda a recuperar estimativas consistentes do parâmetro ρ ? A ideia aqui é que os municípios logo abaixo e logo acima dessa descontinuidade são provavelmente semelhantes na maioria dos seus atributos observáveis e não observáveis. Isso nos permite recuperar ρ apenas comparando os resultados y_i das unidades logo acima e logo abaixo da descontinuidade.

Isso também pode ser realizado com o auxílio de todas as observações e uma estrutura de regressão. Mantenha o pressuposto de que existem apenas dois partidos (A e B) e suponha que x_i seja a proporção de votos do partido B no município i . A regra eleitoral é dada por:

$$D_i = \begin{cases} 1, & \text{se } x_i \geq 0,5 \\ 0, & \text{se } x_i < 0,5 \end{cases}$$

Suponha adicionalmente que a relação entre x_i e y_i é dada pela função $f(\cdot)$. O efeito da identidade partidária sobre as políticas públicas pode então ser consistentemente estimado pelo método de mínimos quadrados ordinários (MQO) para a seguinte equação:

$$y_i = \alpha + \rho D_i + f(x_i) + \varepsilon_i \quad (3.2)$$

É importante observar que o parâmetro ρ estimado acima é o efeito local da identidade partidária sobre as decisões políticas de interesse na proximidade da descontinuidade. Portanto, o que é estimado é basicamente o efeito da identidade partidária em eleições próximas. Podemos também permitir que a função $f(\cdot)$ seja diferente abaixo e acima da descontinuidade e assim estimar alternativamente:

$$y_i = \alpha + \rho D_i + f(x_i) + \tau D_i * f(x_i) + \varepsilon_i \quad (3.3)$$

Também vale destacar que é fundamental estimar de maneira correta a função $f(\cdot)$. Por esse motivo, nós não impomos linearidade a ela e permitimos nas estimações que essa função contenha termos quadráticos e cúbicos.

Uma questão importante que emerge da discussão acima é o fato que na prática o sistema partidário brasileiro é bastante fragmentado e tem mais de dois partidos políticos. Isso não apenas altera o limite necessário de votos para um prefeito ser eleito como também torna esse limite dependente da fragmentação política local. A fragmentação partidária também implica que temos partidos de diversas posições ideológicas (ou mesmo sem ideologia bem definida) e muitas eleições disputadas principalmente por partidos com ideologias similares.

Nas duas próximas seções, nós apresentamos evidências que as eleições municipais brasileiras podem ser tratadas como eleições bipartidárias e apresentamos a classificação ideológica utilizada na análise empírica. Como o nosso intuito é avaliar as diferenças entre escolhas de tamanho de governo implantadas por prefeitos eleitos em eleições próximas e disputadas por partidos de esquerda contra partidos de direita, nós primeiramente excluímos os municípios onde não ocorreu esse confronto ideológico entre os dois primeiros colocados na eleição de 2004. Assim, dos 5.517 municípios nos quais a eleição foi vencida no primeiro turno, em 1.599 o prefeito eleito é de um partido de direita, mas o segundo colocado não pertence a um partido de esquerda. Já para 707 cidades, o vencedor da eleição foi um partido de esquerda, entretanto o segundo colocado não foi um partido de direita. Para 941 municípios, um partido de direita ficou em segundo lugar, no entanto, o prefeito eleito não é de um partido de esquerda. E para 574 municípios, um partido de esquerda ficou em segundo colocado, mas a eleição não foi vencida por um partido de direita. Todos esses municípios descritos acima foram excluídos da análise. De maneira complementar, nós também retiramos da amostra as municipalidades nas quais a eleição teve como primeiro ou segundo colocados partidos de centro ou nãnicos. Estes perfazem 669 cidades. Portanto, após a adequação da amostra em relação ao objetivo principal deste estudo restam 1.027 municípios. Esses estavam distribuídos de tal maneira que 477 elegeram prefeitos ligados a partidos de

esquerda e o segundo colocado pertencia a um partido de direita. E o contrário aconteceu em 550 municípios.

3.3 Dados

Para analisar os efeitos da identidade partidária sobre as decisões políticas de gastos e receitas nos municípios brasileiros utilizamos duas bases de dados principais. A primeira delas é o banco de dados de resultados eleitorais disponibilizado pelo Tribunal Superior Eleitoral (TSE). Essa base de dados contém informações sobre os resultados eleitorais de todas as eleições brasileiras desde 2002 e inclui votação por partido nas eleições para cargos executivos e legislativos. Já a segunda delas é o banco de dados sobre Finanças do Brasil (FINBRA), disponibilizado pela Secretaria do Tesouro Nacional (STN). Essa, por sua vez, inclui informações de receitas e despesas de diversas modalidades dos governos municipais brasileiros. As bases de dados também nos permitem saber o tamanho do eleitorado e da população das unidades de análise. Também utilizamos algumas variáveis do IPEAData e da Pesquisa de Municípios Brasileiros (MUNIC) do IBGE como controles na análise empírica.

Os dados eleitorais utilizados são os referentes às eleições para prefeito ocorridas no ano de 2004. As variáveis relevantes para nossa análise são o partido do candidato vencedor, o partido do segundo colocado, o número de candidatos, a proporção de votos do vencedor, a sua margem de vitória e o número de eleitores. Já os dados fiscais utilizados são referentes ao período de mandato dos prefeitos eleitos em 2004, isto é, os anos entre 2005 e 2008. As variáveis fiscais utilizadas na análise são referentes tanto a receitas (impostos, transferências), quanto a tipos de despesas (educação, saúde, pessoal, saneamento, etc.) e são construídas em duas etapas. Primeiramente, nós utilizamos o Índice Nacional de Preços ao Consumidor Amplo (INPC) para deflacionar as observações de diferentes anos e convertê-las em termos de R\$ de 2008. Posteriormente, nós calculamos os valores médios de cada uma das variáveis para o período completo (2005 a 2008). Essas variáveis fiscais são expressas tanto em termos per capita quanto em termos de fração das receitas ou dos gastos totais. São excluídas da base de dados as

observações de municípios com dados fiscais para menos de dois anos no intervalo considerado.

A Tabela 3.1 apresenta o número de municípios governado por cada um dos principais partidos brasileiros. Os dados mostram que os maiores partidos brasileiros em número de municípios governados são PMDB, PFL (atualmente DEM), PSDB, PP e PT (nessa ordem). Esses cinco partidos governam 66,64% do total de municípios. Esses dados também mostram que 25% das eleições para prefeito tiveram margens de vitória pequenas, ou seja, a diferença entre o primeiro e o segundo colocados na eleição foi de menos de 5% dos votos válidos.

Tabela 3.1 – Resultados eleições municipais 2004

Prefeitos eleitos por partido político				
Partido	Frequência	%	% acumulada	% eleitos por margem pequena
PMDB	913	19,55%	19,55%	24,32%
PSDB	718	15,37%	34,93%	23,12%
PFL	622	13,32%	48,24%	24,60%
PP	487	10,43%	58,67%	28,54%
PT	372	7,97%	66,64%	27,69%
PTB	369	7,90%	74,54%	28,46%
PL	303	6,49%	81,03%	24,75%
PDT	265	5,67%	86,70%	21,51%
PPS	265	5,67%	92,38%	27,17%
Outros partidos	158	3,38%	95,76%	21,73%
PSB	148	3,17%	98,93%	19,59%
PV	44	0,94%	99,87%	22,73%
PCdoB	6	0,13%	100,00%	16,67%
Total	4670			25,01%

Fonte: Elaboração própria a partir de dados do Tribunal Superior Eleitoral (TSE).

Já a Tabela 3.2 apresenta o eleitorado governado pelos principais partidos políticos brasileiros. A medida de tamanho do eleitorado utilizada são os votos válidos nas eleições. Trata-se de uma medida imperfeita, mas que permite observar alguns padrões interessantes. Os principais partidos brasileiros de acordo com o eleitorado por eles governado são PSDB, PT, PMDB e PFL (nessa ordem). Esses quatro partidos governam 63,73% dos eleitores brasileiros. A análise dessa tabela revela que PSDB e PT são particularmente fortes em grandes municípios,

enquanto partidos como PMDB, PFL, PP e PTB são mais fortes em pequenos municípios.

Nós não realizamos uma análise sistemática das informações fiscais porque a variação relevante para a nossa análise desses dados é a referente às eleições próximas disputadas entre partidos de direita e de esquerda. Por isso, deixamos essa análise para a próxima seção após a definição de que partidos são de esquerda e de direita. Mas vale observar que as receitas e os gastos médios per capita dos municípios brasileiros se situam em torno de R\$1.500 no período da análise e que há substantiva variância nessas receitas e gastos.

Tabela 3.2: Eleitorado governado por partido 2005-2008

Eleitorado dos municípios governados por cada partido

Partido	Eleitorado	%	% acumulada
PSDB	18.902.689	21,77%	21,77%
PT	13.103.371	15,09%	36,87%
PMDB	12.550.696	14,46%	51,32%
PFL	10.772.460	12,41%	63,73%
PDT	6.251.461	7,20%	70,93%
PPS	5.028.307	5,79%	76,72%
PP	4.924.419	5,67%	82,39%
PTB	4.704.165	5,42%	87,81%
PSB	4.140.186	4,77%	92,58%
PL	3.221.904	3,71%	96,29%
Outros partidos	1.966.238	2,26%	98,56%
PV	961.492	1,11%	99,66%
PCdoB	291.358	0,34%	100,00%
Total	86.818.746		

Fonte: Elaboração própria a partir de dados do Tribunal Superior Eleitoral (TSE).

Outro aspecto relevante refere-se ao fato de que cerca de 80% das receitas são provenientes de transferências dos governos federal e estadual. Essas transferências incluem as participações dos municípios em impostos administrados e coletados pelo governo federal e estadual, além de transferências referentes à participação dos municípios em fundos constitucionais (Fundo de Participação dos Municípios, FUNDEF, SUS etc.). Parcela substantiva dessas transferências é proveniente do Fundo de Participação dos Municípios. Isso é

importante porque esses recursos podem ser utilizados livremente pelos municípios (ao contrário de recursos para saúde e educação). Portanto, parece existir variação suficiente nesses dados para realização da investigação empírica.

De maneira geral, a análise dessa seção mostra que existe um número relevante de eleições municipais próximas e revela a fragmentação partidária do sistema político brasileiro. Esses dois aspectos são relevantes para a nossa análise empírica uma vez que influenciam diretamente tanto o sucesso de nossa estratégia de identificação quanto a capacidade de realizarmos o estudo. Por esses motivos, esses fatores são analisados com mais detalhe na próxima seção quando descrevermos o contexto político brasileiro e avançarmos com nossas definições de esquerda e direita nesse contexto.

3.4

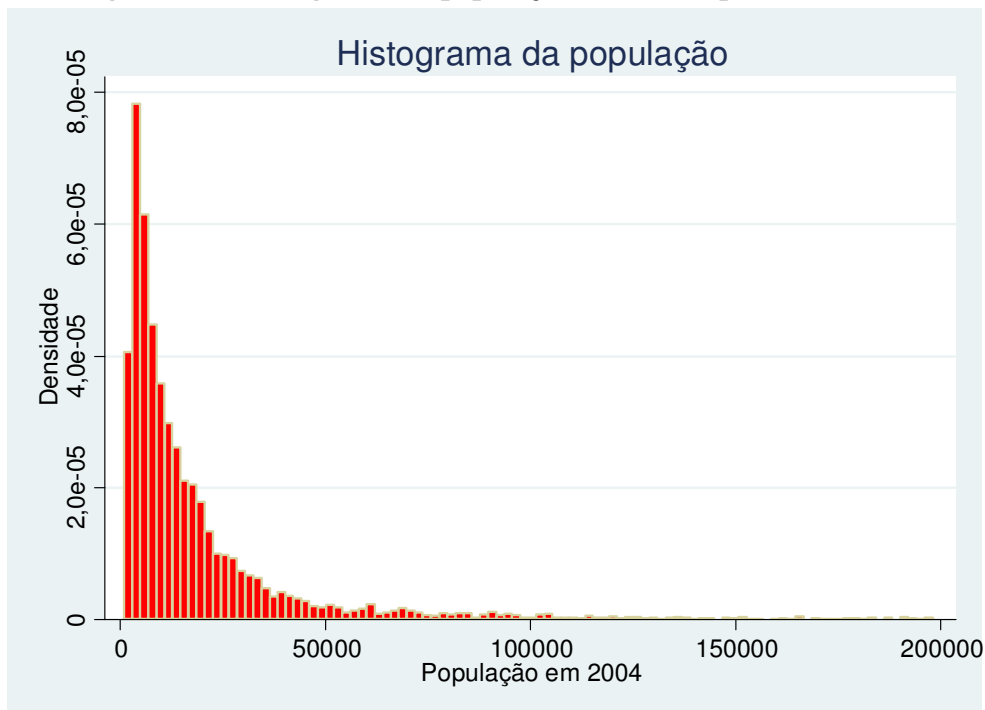
Contexto político brasileiro

Na introdução nós argumentamos que o principal objetivo deste estudo é investigar os efeitos da identidade partidária sobre as escolhas políticas de gastos e receitas nos municípios brasileiros. Isso torna importante discutir alguns aspectos da política local no Brasil. Complementarmente, também se torna relevante apresentar o papel dos municípios na oferta de serviços públicos, considerando tanto a autonomia quanto os limites fiscais que esses municípios devem obedecer.

O município é a menor unidade política existente no Brasil e em 2004 eram 5.561. Estes são governados por prefeitos eleitos por voto direto para mandatos de quatro anos com direito a uma reeleição. Essas eleições ocorrem em turno único e é eleito o candidato com maior número de votos em todos os municípios com menos de 200.000 eleitores. Já nas municipalidades com mais de 200.000 eleitores a eleição ocorre potencialmente em dois turnos. Nesse caso, os dois candidatos mais votados no primeiro turno disputam as eleições no segundo turno. Para a eleição de 2004, 68 municípios tinham mais de 200.000 eleitores, entretanto a disputa em dois turnos ocorreu somente em 44 cidades. Nas outras 24 cidades, o pleito foi decidido no primeiro turno com o candidato vencedor obtendo mais do que 50% dos votos válidos.

A Figura 3.1 apresenta o histograma da distribuição da população em 2004 entre os municípios brasileiros e revela que a grande maioria dos municípios é pequena e tem menos de 50.000 habitantes. De fato, aproximadamente 49% dos municípios do país têm menos de 10.000 habitantes, 73% menos de 20.000 habitantes e 91% menos de 50.000 habitantes⁶⁴. Essa evidência revela que grande parte da variação entre identidade partidária e tamanho de governo, que investigamos neste trabalho, é oriunda de cidades pequenas cujas eleições ocorrem em turno único.

Figura 3.1 – Histograma da população dos municípios brasileiros



Fonte: Elaboração própria a partir de dados do IPEADData.

A combinação de eleições em turno único e multipartidarismo pode invalidar a nossa estratégia empírica. Isto é, para nós estimarmos o efeito da identidade partidária sobre as decisões políticas de gastos e receitas precisamos que as eleições sejam disputadas basicamente por dois candidatos relevantes. Sem

⁶⁴ Nessa análise estamos excluindo os 68 municípios que em 2004 tinham mais do que 200.000 eleitores, ou seja, para os quais existia a possibilidade de eleição em segundo turno. Além disso, no histograma, também estamos restringindo a amostra aos municípios com no máximo 200.000 habitantes para facilitar a visualização do gráfico e a importância relativa dos pequenos municípios.

isso nossos resultados não nos ajudam a avaliar os diversos modelos de concorrência política e votação no contexto das eleições locais brasileiras.

A preocupação expressa acima pode ser mitigada pelo fato da maioria das eleições municipais brasileiras ser disputada por apenas dois candidatos relevantes. Nas eleições de 2004, o número de candidatos efetivos (inverso do índice HH de votos dos candidatos) foi 2,1 para todos os municípios brasileiros. Esse índice foi ainda menor para os municípios com menos de 50.000 habitantes que concentram grande parte da variação relevante para a análise empírica desse trabalho.

Outra preocupação empírica trazida pelo contexto político é a existência de diferenciação ideológica entre os partidos políticos brasileiros. Sem diferenciação ideológica entre os partidos também não é possível estimar o efeito da identidade partidária sobre as decisões políticas. Isso pode ser um problema no nosso contexto, uma vez que a política brasileira é marcada por grande fragmentação partidária e por pouca clareza na diferenciação ideológica dos partidos.

Power & Zucco Jr. (2009) estimam a posição ideológica dos partidos brasileiros para o período 1990 a 2005 utilizando dados de pesquisas de percepções ideológicas dos congressistas brasileiros. Essas pesquisas incluem percepções dos próprios congressistas sobre a sua ideologia (e a ideologia do seu partido), percepções dos congressistas sobre a ideologia dos demais partidos políticos e são especialmente interessantes para serem utilizadas no contexto brasileiro visto que, ao contrário de medidas de *roll call*, não são contaminadas pela barganha existente no processo político.

Dois aspectos da análise desses autores são essenciais para o nosso estudo. Primeiro, a posição ideológica relativa dos principais partidos políticos brasileiros ficou basicamente inalterada no período. Apesar da posição ideológica absoluta dos partidos alterar-se ao longo do tempo e congressistas de partidos de direita se reportam mais à esquerda que seus pares. A ordenação relativa dos partidos no espectro ideológico muda muito pouco ao longo do período. Segundo, existe bastante variação na posição política dos diversos partidos brasileiros.

Utilizando a análise desses autores podemos classificar os partidos brasileiros em quatro grupos distintos: direita, centro, esquerda e nanicos. Os partidos de direita são PFL (hoje DEM), PTB, PP e PL (hoje PR). Os partidos de centro são PSDB e PMDB. Os partidos de esquerda são PCdoB, PT, PDT, PSB e

PPS. Por último, os partidos nanicos são os pequenos partidos e muitas vezes sem representantes no congresso e, por conseguinte não incluídos na análise de Power e Zucco Jr. (2009).

A classificação acima nos permite comparar decisões de gastos e receitas implantadas pelos governantes de esquerda e direita e assim estimar o efeito da ideologia sobre o tamanho de governo nos municípios brasileiros, conforme descrito na estratégia de identificação. Mas quais são as dimensões de escolhas políticas que os municípios brasileiros influenciam? A resposta a essa questão é fundamental para orientar a análise empírica.

Os municípios brasileiros são responsáveis pela oferta de diversos serviços públicos, tais como educação infantil e fundamental, atendimento de saúde, infraestrutura urbana, cultura, esportes e lazer. Os municípios também são parcialmente responsáveis pela oferta de serviços de saneamento básico e pela administração de programas federais de transferência de renda como o Programa Bolsa-Família (PBF). Entretanto, os municípios são constitucionalmente obrigados a gastar frações mínimas das suas receitas com atividades como educação e saúde. Eles também têm limites legais para seus gastos com pessoal e endividamento. Finalmente, os municípios administram a coleta de apenas dois impostos (IPTU e ISS) e uma grande proporção das suas receitas, especialmente dos municípios menores, vêm de repasses do governo federal e, em menor escala, dos governos estaduais.

As restrições legais para gastos e endividamento, e o pequeno espaço fiscal que os governos municipais têm pode limitar a análise deste trabalho. Isso torna extremamente importante uma análise descritiva dos dados utilizados nesse estudo. A Tabela 3.3 apresenta algumas estatísticas descritivas de variáveis eleitorais, socioeconômicas e fiscais do conjunto de municípios com menos de 200.000 eleitores, uma amostra de municipalidades governadas por partidos de esquerda em eleições em que o segundo colocado foi um partido de direita ou vice-versa, e uma subamostra desta última nas quais as eleições tiveram uma margem de vitória menor que 5%.

A análise das variáveis eleitorais mostra que a proporção de votos da esquerda é bastante próxima. Ela também revela que cerca de 20% das eleições da amostra tiveram margens de vitória inferiores a 5% dos votos. Já a análise das variáveis fiscais indica que existe substantiva variação nos dados fiscais tanto de

receita quanto de despesas. Isto, tanto para todas as eleições quanto em eleições próximas. Já a análise das variáveis socioeconômicas demonstra que as estatísticas de municípios em que as eleições foram próximas são similares às dos municípios para os quais isto não ocorreu.

Tabela 3.3 – Estatísticas descritivas dos municípios em geral e especificamente com eleições direita vs. esquerda

	Todas eleições		Eleições esquerda vs. direita		Eleições próximas	
	Média	Desvio-padrão	Média	Desvio-padrão	Média	Desvio-padrão
Variáveis eleitorais						
Margem de vitória	0,1714	0,1870	0,1649	0,1586	0,0235	0,0142
Proporção de votos da esquerda (normalizada em relação a votos da direita)			-0,0119	0,1145	-0,0022	0,0136
Proporção de votos do primeiro colocado	0,5857	0,0935	0,5825	0,0793	0,5117	0,0071
Proporção de votos do segundo colocado	0,4143	0,0935	0,4175	0,0793	0,4883	0,0071
Variáveis socioeconômicas						
População em 2004	21.639	35.428	24.645	39.220	25.203	40.424
IDH em 2000	0,701	0,082	0,702	0,083	0,701	0,082
Proporção com mais de 11 anos de estudo em 2000	0,033	0,028	0,034	0,030	0,035	0,028
Proporção com menos de 4 anos de estudo em 2000	0,4898	0,1698	0,4832	0,1719	0,4857	0,1687
Escolaridade do prefeito	5,3	1,8	5,3	1,8	5,3	1,8
Idade do prefeito	47,4	9,6	47,0	9,4	47,9	10,0
Proporção prefeitos sexo masculino	0,926	0,261	0,927	0,260	0,959	0,200
Variáveis fiscais						
Receitas correntes per capita	1.446,17	742,75	1.392,82	652,32	1.349,44	612,40
Receitas tributárias per capita	83,34	106,18	85,24	115,81	82,19	115,00
Impostos per capita	73,72	97,77	75,18	104,41	72,92	104,82
IPTU per capita	16,72	45,27	16,61	36,79	16,04	30,98
ISS per capita	32,77	61,79	34,19	67,10	32,33	70,91
Transferências da União per capita	744,54	459,42	702,39	393,29	684,82	388,62
Gastos com saúde per capita	283,49	135,20	276,28	122,48	269,37	118,62
Gastos com educação per capita	366,15	148,58	357,16	144,40	338,86	124,07
Gastos com administração per capita	229,23	163,64	224,29	151,54	208,36	136,80
Gastos com pessoal per capita	580,85	268,40	560,26	241,74	541,91	232,83
Gastos com urbanismo per capita	121,13	109,22	110,84	90,64	102,13	88,35
Gastos com transporte per capita	75,96	109,57	67,03	91,85	73,19	101,49
Gastos com infra-estrutura urbana per capita	46,96	73,29	44,57	53,34	37,51	43,01
Investimentos per capita	158,38	142,57	152,74	112,90	129,59	90,73
Superávit per capita	39,60	77,98	40,33	81,14	40,73	67,62
Total de observações	5.136		964		217	

Fonte: Elaboração própria a partir de dados do Tribunal Superior Eleitoral (TSE), da Secretaria do Tesouro Nacional (STN), do IPEADData e da Munic/IBGE.

3.5 Resultados

A presente seção apresenta as evidências obtidas a partir da estratégia empírica escolhida – regressão com descontinuidade – para estimar o efeito da identidade partidária sobre as escolhas políticas de gastos e receitas dos prefeitos nos municípios brasileiros. Assim, conforme ressaltado anteriormente, os resultados abaixo discutidos referem-se a medidas locais desse efeito para a realidade dos municípios que tiveram disputas eleitorais apertadas entre partidos de esquerda e de direita, tal como definidos na seção 3.4.

Nossa análise inicia-se ao investigarmos o efeito de ser governado por um partido de esquerda sobre algumas medidas de tamanho de governo. As tabelas 3.4, 3.5, 3.6 e 3.7 contém estimativas desse efeito sobre, respectivamente, receitas orçamentárias, receitas tributárias, despesas orçamentárias e despesas com pessoal e encargos sociais; todas essas variáveis definidas como valores reais médios, entre 2005 e 2008, per capita. A coluna 1 dessas tabelas apresenta o valor de diferença de médias entre governos de esquerda e de direita. A coluna 2 lista os coeficientes da estimação utilizando mínimos quadrados ordinários com controles para características sociodemográficas (população em 2004, índice de desenvolvimento humano – IDH – no ano 2000 e percentual de habitantes de cada município que tem mais de 11 e menos de 4 anos de estudo) e sobre o prefeito eleito em 2004 (escolaridade). Por último, as colunas 3 a 9 apresentam os efeitos de governos de esquerda sobre cada política pública de interesse utilizando o arcabouço de regressão descontínua. Para essa metodologia, nós utilizamos diversas especificações, desde lineares na margem de vitória da esquerda, a termos quadráticos, cúbicos⁶⁵. Com e sem interações desses polinômios com uma variável indicadora que identifica os municípios nos quais os partidos de esquerda governaram a partir de 2005 e com controles para as características sociodemográficas explicitadas⁶⁶. Vale ressaltar a importância da utilização de covariadas nas regressões estimadas devido à grande heterogeneidade dos municípios brasileiros.

⁶⁵ A margem de vitória do partido de esquerda sobre o partido de direita é a *running variable* na nossa técnica de regressão descontínua.

⁶⁶ Essa descrição das especificações adotadas em cada coluna será útil para todas as demais tabelas apresentadas neste artigo.

A princípio não existe diferença nos valores médios das primeiras quatro variáveis utilizadas como *proxies* para tamanho de governo entre os municípios administrados por partidos de esquerda e direita. Após a inclusão dos controles para as características sociodemográficas, as cidades governadas pela esquerda tendem a ter tamanhos de governo menores (estatisticamente significativa a 10%), pelo menos de acordo com três das variáveis utilizadas – receitas orçamentárias, tributárias e despesas orçamentárias, respectivamente, -4,5%, -0,7% e -4,5% em relação aos valores médios dessas variáveis para a amostra (ver tabela 3.3). Entretanto, essas estimativas não podem ser interpretadas com sentido causal. Para identificar o efeito da identidade partidária sobre tamanho de governo devemos centrar atenção aos resultados obtidos pela técnica de regressão descontínua. As colunas 3 a 9 das tabelas 3.4 a 3.7 indicam que o fato de ser governado por um partido de esquerda não tem impacto sobre o tamanho do governo local.

Tabela 3.4 – Efeito de governos de esquerda sobre receitas orçamentárias

	Variável dependente: Receitas orçamentárias médias (2005 a 2008) per capita em R\$ 2008							
	OLS	OLS	RD	RD	RD	RD	RD	RD
	(1)	controles (2)	Linear (3)	Quadrática (4)	Cúbica (5)	Linear (6)	Quadrática (8)	Cúbica (9)
Partido esquerda	-50.88	-61.15	-18.94	-22.89	-32.43	-25.27	-50.41	-84.70
	(40.05)	(35.75)*	(49.09)	(48.65)	(56.76)	(48.31)	(60.43)	(79.64)
Polinomio em margem vitória:								
Linear			-256.5	-223.8	-135.6	-384.1	-394.7	1,110
			(228.2)	(222.8)	(356.1)	(310.0)	(735.5)	(1,709)
Quadrático				302.8	257.2		-33.97	10,522
				(546.3)	(539.8)		(1,792)	(11,065)
Cúbico					-857.4			17,138
					(1,920)			(16,864)
Interação polinomio margem vitória com dummy para partido esquerda:								
Linear						358.8	1,010	-594.6
						(425.8)	(956.5)	(2,468)
Quadrático							-2,153	-11,942
							(2,387)	(15,762)
Cúbico								-18,394
								(24,549)
Controle:								
População 2004	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% menos de 4 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% mais de 11 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
IDH 2000	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Escolaridade prefeito	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Constante	1,361	-194.2	-217.8	-213.4	-207.1	-229.4	-248.0	-227.5
	(29.15)***	(565.6)	(563.5)	(564.6)	(563.6)	(563.2)	(562.2)	(562.7)
N. observações	964	964	964	964	964	964	964	964
R-quadrado	0.002	0.211	0.212	0.212	0.212	0.212	0.213	0.214

Fonte: Elaboração própria a partir de dados do TSE, FINBRA/STN, IPEAData e Munic/IBGE.

Erros-padrão em parênteses - *** p<0,01, ** p<0,05, * p<0,1.

Tabela 3.5 – Efeito de governos de esquerda sobre receitas tributárias

	Variável dependente: Receitas tributárias médias (2005 a 2008) per capita em R\$ 2008							
	OLS	OLS	RD	RD	RD	RD	RD	RD
	(1)	controles (2)	Linear (3)	Quadrática (4)	Cúbica (5)	Linear (6)	Quadrática (8)	Cúbica (9)
Partido esquerda	-8.378	-11.03	-1.084	-3.502	-5.761	-3.134	-8.469	-16.69
	(7.269)	(5.804)*	(9.721)	(8.903)	(9.926)	(9.013)	(10.59)	(14.51)
Polinomio em margem vitória:								
Linear			-60.45 (55.46)	-40.41 (47.04)	-19.53 (63.27)	-101.8 (77.94)	6.211 (151.4)	333.0 (370.0)
Quadrático				185.4 (144.5)	174.6 (133.0)		327.8 (497.2)	2,620 (2,321)
Cúbico					-203.1 (412.1)			3,721 (3,269)
Interação polinomio margem vitória com dummy para partido esquerda:								
Linear						116.3 (88.54)	28.75 (181.8)	-288.4 (456.0)
Quadrático							-397.2 (521.4)	-2,751 (2,905)
Cúbico								-3,625 (4,453)
Controle:								
População 2004	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% menos de 4 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% mais de 11 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
IDH 2000	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Escolaridade prefeito	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Constante	89.50 (6.034)***	-168.7 (83.03)**	-174.2 (83.64)**	-171.5 (83.72)**	-170.0 (83.44)**	-178.0 (84.04)**	-169.9 (83.75)**	-165.0 (84.55)*
N. observações	964	964	964	964	964	964	964	964
R-quadrado	0.001	0.331	0.333	0.335	0.335	0.334	0.335	0.336

Fonte: elaboração própria a partir de dados do TSE, FINBRA/STN, IPEADData e Munic/IBGE.

Erros-padrão em parênteses - *** p<0,01, ** p<0,05, * p<0,1.

Esse primeiro resultado também foi verificado para outras medidas de tamanho de governo usuais, tais como arrecadação com impostos, arrecadação com IPTU, investimentos e superávit. As estimativas são apresentadas no apêndice desta tese, nas tabelas A10, A11, A12 e A13; e corroboram a inexistência de um efeito da identidade partidária sobre o tamanho médio dos governos. Um detalhe a ser destacado, refere-se à variabilidade da estimativa obtida para o efeito do governo de esquerda de acordo com a alteração da forma funcional adotada pelo desenho de regressão descontínua. Apesar disto, esses efeitos são sempre estatisticamente não significantes.

Tabela 3.6 – Efeito de governos de esquerda sobre despesas orçamentárias

	Variável dependente: Despesas orçamentárias médias (2005 a 2008) per capita em R\$ 2008							
	OLS	OLS	RD	RD	RD	RD	RD	RD
	(1)	controles	Linear	Quadrática	Cúbica	Linear	Quadrática	Cúbica
Partido esquerda	-54.12	-61.81	-23.07	-27.26	-34.63	-30.03	-52.78	-88.67
	(37.91)	(34.14)*	(47.11)	(46.74)	(54.15)	(46.47)	(57.57)	(75.99)
Polinomio em margem vitória:								
Linear			-235.4	-200.8	-132.6	-375.7	-455.2	1,102
			(216.9)	(213.0)	(335.5)	(294.0)	(686.4)	(1,652)
Quadrático				320.8	285.6		-243.3	10,679
				(534.6)	(528.6)		(1,685)	(10,717)
Cúbico					-663.1			17,732
					(1,820)			(16,261)
Interação polinomio margem vitória com dummy para partido esquerda:								
Linear						394.5	1,127	-516.5
						(411.9)	(904.8)	(2,378)
Quadrático							-1,986	-12,238
							(2,260)	(15,186)
Cúbico								-18,833
								(23,761)
Controle:								
População 2004	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% menos de 4 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% mais de 11 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
IDH 2000	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Escolaridade prefeito	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Constante	1,323	-306.8	-328.5	-323.8	-318.9	-341.2	-365.7	-344.2
	(27.52)***	(546.0)	(544.4)	(545.5)	(544.8)	(544.2)	(543.5)	(544.5)
N. observações	964	964	964	964	964	964	964	964
R-quadrado	0.002	0.192	0.193	0.194	0.194	0.194	0.195	0.196

Fonte: Elaboração própria a partir de dados do TSE, FINBRA/STN, IPEADData e Munic/IBGE.

Erros-padrão em parênteses - *** p<0,01, ** p<0,05, * p<0,1.

Tabela 3.7 – Efeito governos de esquerda sobre despesas com pessoal e encargos sociais

Variável dependente: Despesas com pessoal e encargos sociais médias (2005 a 2008) per capita em R\$ 2008								
	OLS	OLS	RD	RD	RD	RD	RD	RD
		controles	Linear	Quadrática	Cúbica	Linear	Quadrática	Cúbica
	(1)	(2)	(3)	(4)	(5)	(6)	(8)	(9)
Partido esquerda	-13.15	-15.10	0.665	-1.996	-8.004	-1.991	-13.65	-20.58
	(15.49)	(14.23)	(20.07)	(19.89)	(22.59)	(19.80)	(24.36)	(32.70)
Polinômio em margem vitória:								
Linear			-95.81	-73.75	-18.23	-149.4	-20.15	564.1
			(95.83)	(93.96)	(139.0)	(128.1)	(279.9)	(692.6)
Quadrático				204.1	175.3		391.9	4,491
				(257.1)	(253.0)		(759.2)	(4,571)
Cúbico					-540.0			6,656
					(831.5)			(6,976)
Interação polinômio margem vitória com dummy para partido esquerda:								
Linear						150.7	178.1	-701.4
						(181.9)	(390.5)	(1,031)
Quadrático							-925.7	-2,857
							(1,082)	(6,619)
Cúbico								-10,164
								(10,423)
Controle:								
População 2004	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% menos de 4 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% mais de 11 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
IDH 2000	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Escolaridade prefeito	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Constante	565.8	-59.62	-68.46	-65.48	-61.49	-73.29	-67.28	-62.89
	(11.20)***	(226.2)	(225.8)	(226.2)	(225.9)	(225.5)	(225.0)	(225.2)
N. observações	964	964	964	964	964	964	964	964
R-quadrado	0.001	0.169	0.170	0.171	0.171	0.171	0.171	0.173

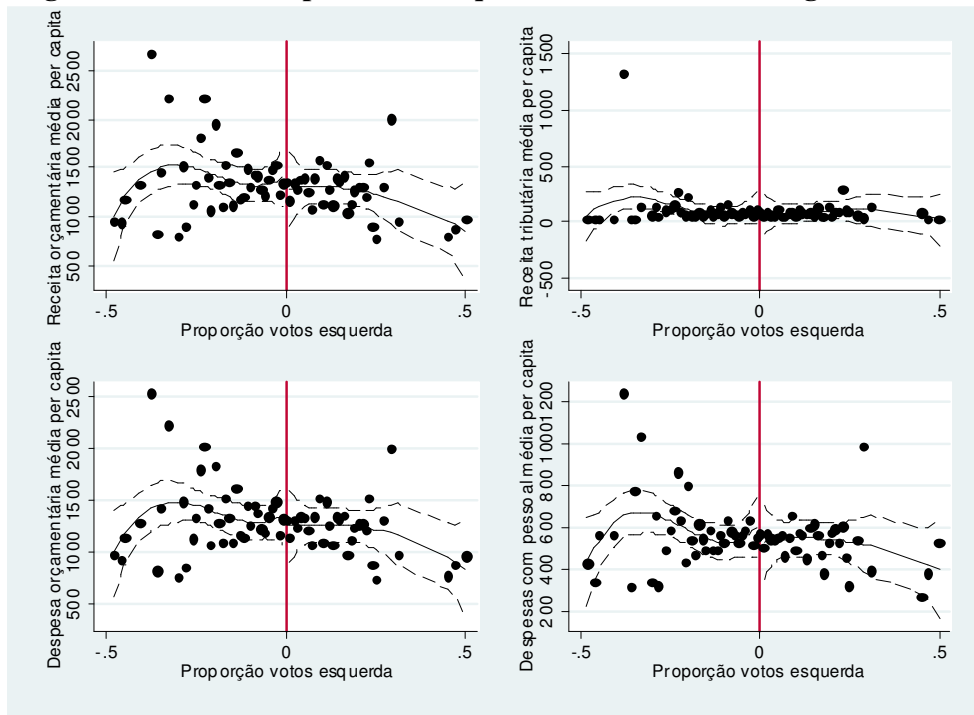
Fonte: Elaboração própria a partir de dados do TSE, FINBRA/STN, IPEADData e Munic/IBGE.

Erros-padrão em parênteses - *** p<0,01, ** p<0,05, * p<0,1.

O resultado apresentado pelas tabelas 3.4 a 3.7 pode ser mais bem visualizado com o auxílio de gráficos. Esses plotam os valores previstos e os intervalos de confiança, a 95%, gerados a partir da regressão do valor médio da variável de tamanho de governo, por ponto percentual de margem de vitória da esquerda, no polinômio cúbico desta variável, com as interações para a variável indicadora se o governo eleito é de esquerda⁶⁷. Os gráficos para as variáveis de receitas orçamentária e tributária, e despesas orçamentária e com pessoal são apresentados na figura 3.2. A análise visual desses gráficos reforça que não existe uma descontinuidade significativa na vizinhança do valor nulo de margem de vitória que caracteriza a alteração no grupo de partidos (esquerda ou direita) no controle do governo municipal.

⁶⁷ Novamente, a descrição desses gráficos será útil para todos os demais gráficos apresentados neste artigo.

Figura 3.2 – Efeito de partido de esquerda sobre tamanho de governo local



Fonte: Elaboração própria a partir de dados do TSE, FINBRA/STN, IPEAData e Munic/IBGE.

Uma vez verificada a inexistência de impacto da identificação partidária sobre o tamanho do governo associado, ou seja, sobre o valor absoluto (*per capita*) das variáveis *proxies* usualmente utilizadas. Resta-nos saber se o controle do governo municipal por um determinado grupo de partidos tem influência sobre a maneira pela qual se dá a composição de gastos e receitas?

Inicialmente, procuramos investigar a composição da origem das receitas tributárias dos governos municipais. Uma vez que as receitas tributárias dependem mais diretamente da ação do município para a arrecadação de tributos, impostos e taxas, e não se dá de maneira relativamente passiva tal como as transferências de fundos federais etc. As tabelas 3.8 e 3.9 mostram que, em geral, os municípios governados pela esquerda têm uma menor fração das suas receitas tributárias oriundas de impostos e uma maior fração derivadas de taxas. Essas são as evidências derivadas da especificação estimada por MQO e que podem nos levar a resultados equivocados. Entretanto, as demais especificações relativas às colunas 2 a 9 das tabelas 3.8 e 3.9 nos permitem afirmar que não existe uma diferença significativa da maneira pela qual os partidos de esquerda optam por compor suas

receitas tributárias entre taxas e impostos. Os dois gráficos superiores na figura 3.3 ilustram essa evidência.

Tabela 3.8 – Efeito de governos de esquerda sobre impostos/receitas tributárias

	Variável dependente: Impostos / receitas tributárias							
	OLS	OLS	RD	RD	RD	RD	RD	RD
	(1)	controles (2)	Linear (3)	Quadrática (4)	Cúbica (5)	Linear (6)	Quadrática (8)	Cúbica (9)
Partido esquerda	-0.0118	-0.00761	-0.00333	-0.00457	-0.00776	-0.00400	-0.00680	0.00496
	(0.00637)*	(0.00586)	(0.00800)	(0.00811)	(0.00944)	(0.00805)	(0.0103)	(0.0132)
Polinomio em margem vitória:								
Linear			-0.0260 (0.0322)	-0.0158 (0.0343)	0.0138 (0.0546)	-0.0395 (0.0395)	0.0611 (0.0997)	-0.0680 (0.231)
Quadrático				0.0948 (0.0795)	0.0796 (0.0779)		0.306 (0.216)	-0.597 (1.456)
Cúbico					-0.287 (0.263)			-1.465 (2.224)
Interação polinomio margem vitória com dummy para partido esquerda:								
Linear						0.0379 (0.0687)	-0.0980 (0.162)	-0.315 (0.377)
Quadrático							-0.185 (0.383)	3.239 (2.389)
Cúbico								-2.603 (3.662)
Controle:								
População 2004	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% menos de 4 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% mais de 11 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
IDH 2000	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Escolaridade prefeito	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Constante	0.894 (0.00437)***	0.882 (0.0984)***	0.880 (0.0984)***	0.881 (0.0985)***	0.884 (0.0985)***	0.879 (0.0984)***	0.888 (0.0985)***	0.881 (0.0985)***
N. observações	964	964	964	964	964	964	964	964
R-quadrado	0.004	0.171	0.172	0.172	0.173	0.172	0.173	0.174

Fonte: Elaboração própria a partir de dados do TSE, FINBRA/STN, IPEAData e Munic/IBGE.

Erros-padrão em parênteses - *** p<0,01, ** p<0,05, * p<0,1.

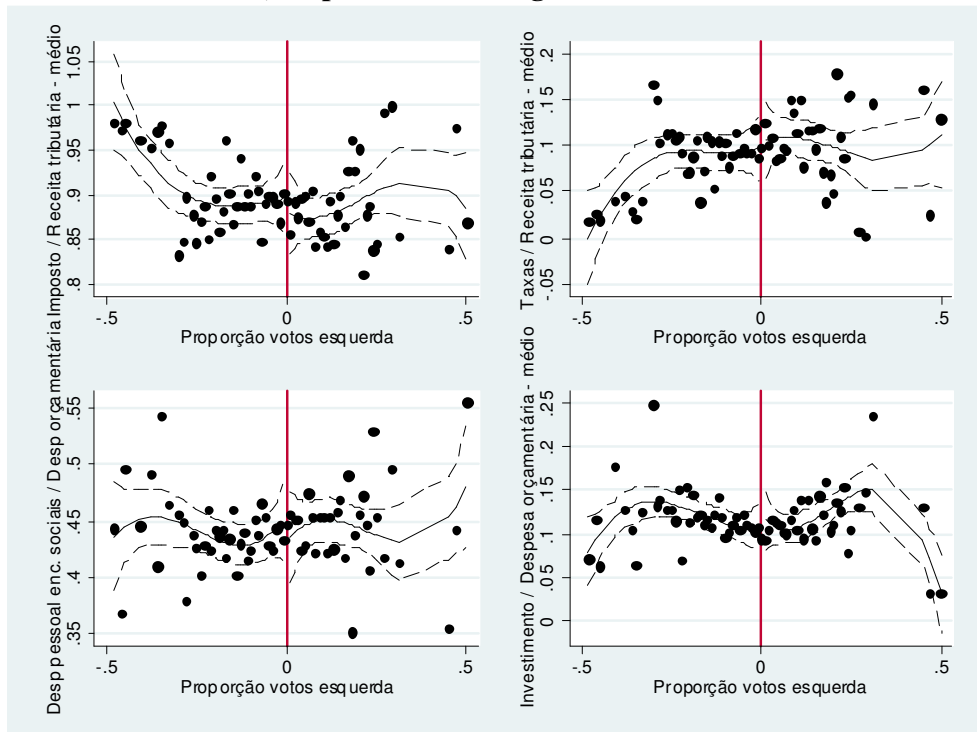
Tabela 3.9 – Efeito de governos de esquerda sobre taxas/receitas tributárias

	Variável dependente: Taxas / receitas tributárias							
	OLS	OLS	RD	RD	RD	RD	RD	RD
	(1)	controles	Linear	Quadrática	Cúbica	Linear	Quadrática	Cúbica
Partido esquerda	0.0115	0.00781	0.000582	0.00144	0.00284	0.00125	0.00219	-0.00624
	(0.00570)**	(0.00525)	(0.00730)	(0.00741)	(0.00858)	(0.00734)	(0.00941)	(0.0121)
Polinomio em margem vitória:								
Linear			0.0440	0.0368	0.0239	0.0575	0.0214	0.104
			(0.0283)	(0.0309)	(0.0480)	(0.0333)*	(0.0850)	(0.207)
Quadrático				-0.0661	-0.0595		-0.110	0.470
				(0.0723)	(0.0725)		(0.188)	(1.274)
Cúbico					0.126			0.940
					(0.238)			(1.915)
Interação polinomio margem vitória com dummy para partido esquerda:								
Linear						-0.0380	0.0125	0.187
						(0.0617)	(0.145)	(0.338)
Quadrático							0.0603	-2.396
							(0.348)	(2.156)
Cúbico								2.088
								(3.258)
Controle:								
População 2004	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% menos de 4 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% mais de 11 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
IDH 2000	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Escolaridade prefeito	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Constante	0.0935	0.0264	0.0304	0.0295	0.0285	0.0317	0.0284	0.0332
	(0.00390)***	(0.0877)	(0.0877)	(0.0879)	(0.0879)	(0.0877)	(0.0882)	(0.0883)
N. observações	964	964	964	964	964	964	964	964
R-quadrado	0.004	0.175	0.176	0.177	0.177	0.176	0.177	0.178

Fonte: Elaboração própria a partir de dados do TSE, FINBRA/STN, IPEADData e Munic/IBGE.

Erros-padrão em parênteses - *** p<0,01, ** p<0,05, * p<0,1.

Figura 3.3 - Efeito de partido de esquerda sobre fontes de receitas tributárias, despesas com encargos sociais e investimentos



Fonte: Elaboração própria a partir de dados do TSE, FINBRA/STN, IPEADData e Munic/IBGE.

Um resultado interessante e que caminha no sentido contrário é observado na tabela 3.10. Nessa, o interesse se dá pela parcela de receitas tributárias derivadas da arrecadação com o IPTU – o valor médio na amostra é de 15% – e inicialmente não se observa diferenças entre os municípios governados pela esquerda ou pela direita. Entretanto, dependendo da especificação adotada para o arcabouço de regressão descontínua, obtém-se evidência de um impacto positivo, importante e significativo do governo ser de esquerda sobre a importância relativa do IPTU para a composição das receitas tributárias do governo local. Esse impacto, apresentado nas colunas 5 e 8 da tabela 3.10, seria aproximadamente de 12% a 15% em relação à média da amostra. Apesar disso, o resultado deve ser interpretado com cautela, pois depende fortemente, tanto em significância estatística quanto em magnitude, da especificação escolhida para a regressão descontínua. Outra fonte importante de recursos que compõe as receitas tributárias dos municípios – imposto sobre serviços – também foi testada e as evidências apontam que não existe um impacto do governo local ser de esquerda sobre a

proporção de receita tributária proveniente desse imposto. Este resultado é apresentado na tabela A14 do anexo.

Tabela 3.10 – Efeito de governos de esquerda sobre IPTU/receitas tributárias

	Variável dependente: IPTU / receitas tributárias							
	OLS	OLS	RD	RD	RD	RD	RD	RD
		controles	Linear	Quadrática	Cúbica	Linear	Quadrática	Cúbica
	(1)	(2)	(3)	(4)	(5)	(6)	(8)	(9)
Partido esquerda	0.00437	-0.00319	0.00665	0.00815	0.0189	0.00825	0.0236	0.0151
	(0.00842)	(0.00576)	(0.00810)	(0.00842)	(0.00951)**	(0.00832)	(0.0106)**	(0.0144)
Polinômio em margem vitória:								
Linear			-0.0598	-0.0723	-0.171	-0.0276	-0.176	-0.0221
			(0.0347)*	(0.0388)*	(0.0555)***	(0.0425)	(0.109)	(0.279)
Quadrático				-0.115	-0.0640		-0.449	0.626
				(0.104)	(0.0807)		(0.257)*	(1.742)
Cúbico					0.962			1.745
					(0.283)***			(2.624)
Interação polinômio margem vitória com dummy para partido esquerda:								
Linear						-0.0907	-0.171	-0.134
						(0.0747)	(0.162)	(0.406)
Quadrático							1.229	-1.233
							(0.382)***	(2.503)
Cúbico								0.492
								(3.856)
Controle:								
População 2004	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% menos de 4 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% mais de 11 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
IDH 2000	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Escolaridade prefeito	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Constante	0.139	0.102	0.0962	0.0945	0.0874	0.0991	0.0935	0.0984
	(0.00578)***	(0.0883)	(0.0885)	(0.0885)	(0.0887)	(0.0884)	(0.0895)	(0.0894)
N. observações	964	964	964	964	964	964	964	964
R-quadrado	0.000	0.534	0.535	0.535	0.538	0.536	0.538	0.539

Fonte: Elaboração própria a partir de dados do TSE, FINBRA/STN, IPEAData e Munic/IBGE.

Erros-padrão em parênteses - *** p<0,01, ** p<0,05, * p<0,1.

Em seguida, estendemos a análise para diferenças na composição dos gastos públicos locais entre municípios governados por partidos de esquerda ou de direita. As principais classes de gastos públicos examinados foram despesas em saúde, em educação, com urbanismo e com administração. Para a amostra de municipalidades brasileiras em questão, os valores médios da parcela desses gastos em relação às despesas orçamentárias, são respectivamente de 21,67%, 28,72%, 8,57% e 17,0%. Ou seja, perfazem juntas aproximadamente 75% das despesas orçamentárias dos municípios pertencentes à amostra.

A tabela 3.11 apresenta as estimativas do impacto do governo local ser de esquerda sobre a fração de gasto público total alocado para saúde. Para quase todas as especificações adotadas no arcabouço de regressão descontínua o efeito é

significante e negativo. Este corresponde a uma redução de aproximadamente um ponto percentual (4,5% sobre a média da amostra) para as especificações lineares e quadráticas. Já para a especificação com termos cúbicos e interação com a variável indicadora se o governo é de esquerda, a magnitude deste impacto é aproximadamente o dobro. Assim governos de esquerda gastam proporcionalmente menos em saúde do que os de direita.

Tabela 3.11 – Efeito de governos de esquerda sobre gastos com saúde/despesas orçamentárias

	Variável dependente: Despesas com saúde / despesas orçamentárias							
	OLS	OLS	RD	RD	RD	RD	RD	RD
	(1)	controles (2)	Linear (3)	Quadrática (4)	Cúbica (5)	Linear (6)	Quadrática (8)	Cúbica (9)
Partido esquerda	-0.00360	-0.00357	-0.00888	-0.00763	-0.00778	-0.00766	-0.00935	-0.0208
	(0.00289)	(0.00285)	(0.00432)**	(0.00418)*	(0.00492)	(0.00416)*	(0.00529)*	(0.00685)***
Polinomio em margem vitória:								
Linear			0.0323	0.0219	0.0233	0.0568	0.0634	0.310
			(0.0186)*	(0.0176)	(0.0281)	(0.0251)**	(0.0641)	(0.135)**
Quadrático				-0.0959	-0.0966		0.0200	1.744
				(0.0506)*	(0.0476)**		(0.182)	(0.821)**
Cúbico					-0.0137			2.799
					(0.179)			(1.314)**
Interação polinomio margem vitória com dummy para partido esquerda:								
Linear						-0.0690	-0.0402	-0.0668
						(0.0366)*	(0.0836)	(0.188)
Quadrático							-0.141	-3.460
							(0.221)	(1.239)***
Cúbico								-0.228
								(1.924)
Controle:								
População 2004	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% menos de 4 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% mais de 11 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
IDH 2000	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Escolaridade prefeito	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Constante	0.218	0.153	0.156	0.155	0.155	0.159	0.158	0.165
	(0.00218)***	(0.0447)***	(0.0445)***	(0.0444)***	(0.0445)***	(0.0445)***	(0.0451)***	(0.0451)***
N. observações	964	964	964	964	964	964	964	964
R-quadrado	0.002	0.042	0.046	0.049	0.049	0.049	0.049	0.057

Fonte: Elaboração própria a partir de dados do TSE, FINBRA/STN, IPEAData e Munic/IBGE.

Erros-padrão em parênteses - *** p<0,01, ** p<0,05, * p<0,1.

Já para a parcela dos gastos públicos locais destinados para gastos em educação, a estimativa obtida por meio de regressão descontínua utilizando polinômio em margem de vitória com termos cúbicos e interação com a variável indicadora de governo de esquerda – coluna 9 da tabela 3.12 – aponta um efeito positivo desses governos. Assim, de acordo com essa estimativa, municipalidades governadas por partidos considerados de esquerda apresentam um incremento de

4,7% na fração média dos gastos totais alocados para educação. Uma ressalva a ser feita é que esse resultado não é robusto a diferentes especificações da regressão com descontinuidade e mesmo na única especificação para a qual observamos o resultado, seu nível de significância é de apenas 10%.

Tabela 3.12 – Efeito governos esquerda sobre gastos com educação/despesas orçamentárias

	Variável dependente: Despesas com educação / despesas orçamentárias							
	OLS	OLS	RD	RD	RD	RD	RD	RD
	(1)	controles (2)	Linear (3)	Quadrática (4)	Cúbica (5)	Linear (6)	Quadrática (8)	Cúbica (9)
Partido esquerda	-0.000435	0.00347	0.00800	0.00664	0.00658	0.00714	0.00829	0.0137
	(0.00441)	(0.00366)	(0.00514)	(0.00515)	(0.00580)	(0.00519)	(0.00626)	(0.00791)*
Polinômio em margem vitória:								
Linear			-0.0275	-0.0163	-0.0157	-0.0448	0.00832	-0.279
			(0.0242)	(0.0246)	(0.0360)	(0.0296)	(0.0722)	(0.162)*
Quadrático				0.104	0.103		0.161	-1.852
				(0.0678)	(0.0704)		(0.187)	(1.041)*
Cúbico					-0.00537			-3.269
					(0.238)			(1.593)**
Interação polinômio margem vitória com dummy para partido esquerda:								
Linear						0.0487	-0.0888	0.262
						(0.0493)	(0.104)	(0.243)
Quadrático							0.127	1.670
							(0.285)	(1.631)
Cúbico								4.032
								(2.538)
Controle:								
População 2004	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% menos de 4 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% mais de 11 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
IDH 2000	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Escolaridade prefeito	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Constante	0.287	0.461	0.458	0.460	0.460	0.457	0.463	0.460
	(0.00316)***	(0.0636)***	(0.0635)***	(0.0634)***	(0.0633)***	(0.0634)***	(0.0636)***	(0.0636)***
N. observações	964	964	964	964	964	964	964	964
R-quadrado	0.000	0.322	0.323	0.324	0.324	0.323	0.325	0.328

Fonte: Elaboração própria a partir de dados do TSE, FINBRA/STN, IPEADData e Munic/IBGE.

Erros-padrão em parênteses - *** p<0,01, ** p<0,05, * p<0,1.

Com relação às despesas com urbanismo, as evidências, derivadas da metodologia de regressão descontínua que permite diferentes inclinações abaixo e acima da margem de vitória nula, sugerem que os governos de esquerda têm um impacto negativo e significativo na fração dos gastos públicos destinados a essa classe de despesas. De acordo com a tabela 3.13, a magnitude do impacto oscila entre 1,0 a 1,5 pontos percentuais, ou seja, governos locais de esquerda gastam de 11% a 17% menos do que a média da proporção de gastos públicos destinados a urbanismo para as cidades que constituem a amostra.

As frações das despesas orçamentárias municipais com transporte e com assistência social também foram analisadas e de acordo com as tabelas A15 e A16, no apêndice desta tese, não existe um impacto significativo da identidade partidária sobre essas medidas de composição de gastos públicos.

Tabela 3.13 – Efeito governos esquerda sobre gastos com urbanismo/despesas orçamentárias

	Variável dependente: Despesas com urbanismo / despesas orçamentárias							
	OLS	OLS	RD	RD	RD	RD	RD	RD
	(1)	controles (2)	Linear (3)	Quadrática (4)	Cúbica (5)	Linear (6)	Quadrática (8)	Cúbica (9)
Partido esquerda	-0.000744	-0.00115	-0.00544	-0.00671	-0.00792	-0.00735	-0.0111	-0.0146
	(0.00324)	(0.00306)	(0.00428)	(0.00448)	(0.00499)	(0.00442)*	(0.00543)**	(0.00692)**
Polinômio em margem vitória:								
Linear			0.0261	0.0366	0.0478	-0.0124	-0.0446	0.121
			(0.0194)	(0.0222)*	(0.0311)	(0.0240)	(0.0600)	(0.139)
Quadrático				0.0974	0.0917		-0.0979	1.060
				(0.0648)	(0.0669)		(0.153)	(0.881)
Cúbico					-0.108			1.880
					(0.238)			(1.359)
Interação polinômio margem vitória com dummy para partido esquerda:								
Linear						0.108	0.267	0.0829
						(0.0434)**	(0.0912)***	(0.199)
Quadrático							-0.335	-1.348
							(0.254)	(1.319)
Cúbico								-2.115
								(2.037)
Controle:								
População 2004	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% menos de 4 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% mais de 11 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
IDH 2000	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Escolaridade prefeito	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Constante	0.0861	-0.0921	-0.0897	-0.0883	-0.0875	-0.0932	-0.0993	-0.0972
	(0.00219)***	(0.0547)*	(0.0550)	(0.0548)	(0.0549)	(0.0546)*	(0.0550)*	(0.0549)*
N. observações	964	964	964	964	964	964	964	964
R-quadrado	0.000	0.102	0.104	0.106	0.107	0.110	0.114	0.116

Fonte: Elaboração própria a partir de dados do TSE, FINBRA/STN, IPEADData e Munic/IBGE.

Erros-padrão em parênteses - *** p<0,01, ** p<0,05, * p<0,1.

Um último aspecto relacionado à composição dos gastos públicos locais refere-se à parcela desses recursos destinados a despesas de administração e possíveis diferenças nas importâncias relativas desses gastos para partidos considerados de esquerda vis-à-vis partidos de direita. A tabela 3.14 apresenta evidências bastante robustas de que governos de esquerda têm um impacto positivo e estatisticamente significativo sobre a fração dos gastos públicos destinados a financiar atividades administrativas dos governos locais. Além disso, é interessante notar como este efeito aumenta em magnitude conforme se opta por uma especificação mais flexível na abordagem de regressão descontínua. Além da

robustez e significância estatística, o impacto de governos de partidos de esquerda sobre a fração dos gastos públicos destinados à administração é de importante relevância econômica, pois se comparado ao valor médio desta medida de composição de despesas – 17% – na presente amostra, este impacto é de aproximadamente um acréscimo entre 11% a 22% na fração dos gastos públicos destinados a este fim.

Tabela 3.14 – Efeito governo esquerda sobre despesas com administração/despesas orçamentárias

	Variável dependente: Despesas com administração / despesas orçamentárias							
	OLS	OLS	RD	RD	RD	RD	RD	RD
	(1)	controles (2)	Linear (3)	Quadrática (4)	Cúbica (5)	Linear (6)	Quadrática (8)	Cúbica (9)
Partido esquerda	0.0119	0.0112	0.0196	0.0194	0.0197	0.0194	0.0226	0.0386
	(0.00435)***	(0.00430)***	(0.00605)***	(0.00615)***	(0.00688)***	(0.00619)***	(0.00779)***	(0.00983)***
Polinômio em margem vitória:								
Linear			-0.0507 (0.0262)*	-0.0488 (0.0274)*	-0.0517 (0.0411)	-0.0541 (0.0324)*	-0.0771 (0.0855)	-0.470 (0.166)***
Quadrático				0.0174 (0.0825)	0.0189 (0.0823)		-0.0698 (0.257)	-2.825 (1.014)***
Cúbico					0.0284 (0.310)			-4.473 (1.625)***
Interação polinômio margem vitória com dummy para partido esquerda:								
Linear						0.00955 (0.0557)	-0.0217 (0.136)	0.114 (0.275)
Quadrático							0.255 (0.408)	4.881 (1.752)***
Cúbico								1.458 (2.723)
Controle:								
População 2004	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% menos de 4 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% mais de 11 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
IDH 2000	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Escolaridade prefeito	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Constante	0.164 (0.00278)***	0.385 (0.0736)***	0.380 (0.0737)***	0.380 (0.0738)***	0.380 (0.0740)***	0.380 (0.0738)***	0.379 (0.0747)***	0.370 (0.0740)***
N. observações	964	964	964	964	964	964	964	964
R-quadrado	0.008	0.036	0.040	0.040	0.040	0.040	0.040	0.047

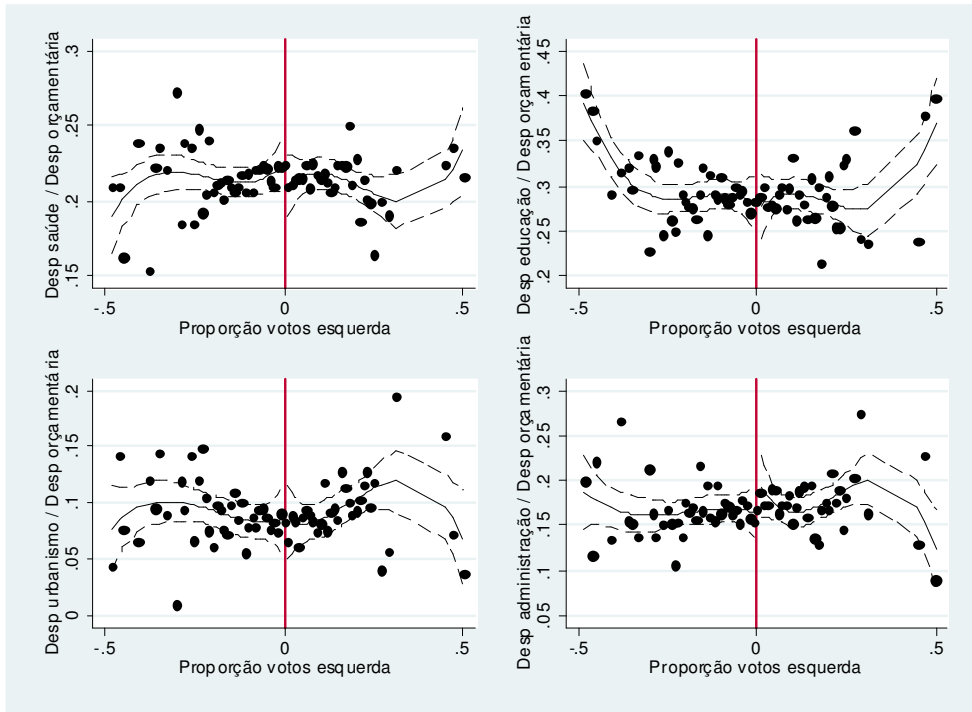
Fonte: Elaboração própria a partir de dados do TSE, FINBRA/STN, IPEADData e Munic/IBGE.

Erros-padrão em parênteses - *** p<0,01, ** p<0,05, * p<0,1.

As evidências de descontinuidades nas principais variáveis de composição de gastos públicos locais no presente estudo estão ilustradas nos gráficos da figura 3.4. Apesar da grande amplitude dos intervalos de confiança tornam-se mais evidentes, sobretudo, as descontinuidades referentes à fração das despesas com saúde (gráfico esquerdo da parte superior) e com administração (gráfico direito da parte inferior), em relação as despesas orçamentárias, de acordo com a margem de

vitória (0) que implica mudança na identidade partidária dos governos das municipalidades brasileiras.

Figura 3.4 – Efeito de partido de esquerda sobre a composição de gastos públicos



Fonte: Elaboração própria a partir de dados do TSE, FINBRA/STN, IPEADData e Munic/IBGE.

A partir dessa evidência, robusta e relevante do ponto de vista econômico, de que os governos de esquerda, em municípios nos quais os partidos de direita foram derrotados por pequena margem, gastam proporcionalmente mais em despesas relacionadas à administração. O próximo aspecto da investigação seria entender qual é o fator que leva a esse incremento relativo nessa classe de gasto para os partidos de esquerda. Uma primeira variável, relacionada ao tamanho de governo, que logo surge seria um maior contingente de funcionários em governos de esquerda, o que aumenta as necessidades de financiamento da máquina pública. Com o intuito de verificar se essa possível explicação seria crível do ponto de vista empírico, nós testamos se a fração de despesas com pessoal e encargos sociais em relação ao total de gastos públicos locais e o número de funcionários públicos per capita também apresentam uma descontinuidade significativa na proximidade do valor de margem de vitória que transfere o governo entre os

partidos de esquerda e direita. As tabelas 3.15 e 3.16 apresentam evidências de que esse suposto mecanismo não é estatisticamente distinto de zero, ou seja, não parece ser a questão de um estado maior – aproximado por um maior número de servidores públicos – que racionalizaria este resultado.

Tabela 3.15 – Efeito governo esquerda sobre despesas com pessoal/despesas orçamentárias

	Variável dependente: despesas com pessoal e encargos sociais / despesas orçamentárias							
	OLS	OLS	RD	RD	RD	RD	RD	RD
	(1)	controles (2)	Linear (3)	Quadrática (4)	Cúbica (5)	Linear (6)	Quadrática (8)	Cúbica (9)
Partido esquerda	0.00884	0.00964	0.00812	0.00822	0.00726	0.00890	0.00883	0.0120
	(0.00407)**	(0.00402)**	(0.00590)	(0.00607)	(0.00670)	(0.00606)	(0.00753)	(0.00976)
Polinômio em margem vitória:								
Linear			0.00926	0.00842	0.0173	0.0249	0.105	0.0274
			(0.0259)	(0.0280)	(0.0387)	(0.0323)	(0.0772)	(0.179)
Quadrático				-0.00774	-0.0123		0.242	-0.299
				(0.0851)	(0.0919)		(0.217)	(1.144)
Cúbico					-0.0862			-0.878
					(0.323)			(1.790)
Interação polinômio margem vitória com dummy para partido esquerda:								
Linear						-0.0441	-0.206	-0.180
						(0.0556)	(0.125)*	(0.277)
Quadrático							0.0375	0.952
							(0.393)	(1.789)
Cúbico								0.275
								(2.889)
Controle:								
População 2004	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% menos de 4 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% mais de 11 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
IDH 2000	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Escolaridade prefeito	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Constante	0.437	0.592	0.592	0.592	0.593	0.594	0.602	0.601
	(0.00273)***	(0.0746)***	(0.0745)***	(0.0745)***	(0.0744)***	(0.0746)***	(0.0744)***	(0.0747)***
N. observações	964	964	964	964	964	964	964	964
R-quadrado	0.005	0.032	0.032	0.032	0.032	0.033	0.035	0.035

Fonte: Elaboração própria a partir de dados do TSE, FINBRA/STN, IPEADData e Munic/IBGE.

Erros-padrão em parênteses - *** p<0,01, ** p<0,05, * p<0,1.

Tabela 3.16 – Efeito governo esquerda sobre número funcionários públicos per capita

	Variável dependente: Número de funcionários públicos per capita no município em 2005							
	OLS	OLS	RD	RD	RD	RD	RD	RD
	(1)	controles	Linear	Quadrática	Cúbica	Linear	Quadrática	Cúbica
Partido esquerda	-0.00238 (0.00114)**	-0.00165 (0.00105)	-0.00144 (0.00155)	-0.00169 (0.00157)	-0.00272 (0.00177)	-0.00167 (0.00157)	-0.00300 (0.00196)	-0.00109 (0.00268)
Polinomio em margem vitória:								
Linear			-0.00129 (0.00711)	0.000759 (0.00754)	0.0103 (0.0105)	-0.00591 (0.00882)	0.0106 (0.0207)	-0.00507 (0.0486)
Quadrático				0.0189 (0.0216)	0.0140 (0.0196)		0.0502 (0.0583)	-0.0594 (0.317)
Cúbico					-0.0926 (0.0642)			-0.178 (0.500)
Interação polinomio margem vitória com dummy para partido esquerda:								
Linear						0.0130 (0.0148)	0.0125 (0.0313)	-0.0332 (0.0798)
Quadrático							-0.105 (0.0846)	0.453 (0.521)
Cúbico								-0.545 (0.805)
Controle:								
População 2004	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% menos de 4 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% mais de 11 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
IDH 2000	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Escolaridade prefeito	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Constante	0.0395 (0.000816)***	0.0330 (0.0199)*	0.0328 (0.0199)*	0.0331 (0.0199)*	0.0338 (0.0199)*	0.0324 (0.0199)	0.0333 (0.0201)*	0.0322 (0.0201)
N. observações	964	964	964	964	964	964	964	964
R-quadrado	0.004	0.166	0.166	0.166	0.168	0.166	0.168	0.169

Fonte: Elaboração própria a partir de dados do TSE, FINBRA/STN, IPEAData e Munic/IBGE.

Erros-padrão em parênteses - *** $p < 0,01$, ** $p < 0,05$, * $p < 0,1$.

Outras medidas relacionadas tanto com despesas associadas a funcionários públicos quanto ao seu número ou composição também foram investigadas e corroboram a evidência acima discutida de inexistência de impacto da identidade partidária sobre essas decisões políticas. Esses resultados são apresentados nas tabelas A17 e A18, além dos gráficos da figura A4, no apêndice desta tese.

3.6 Conclusão

Ao contrário dos resultados obtidos por Ferreira & Gyourko (2009), de inexistência de impactos partidários para o contexto das cidades dos E.U.A., e em consonância, pelo menos parcial, com as evidências obtidas para a realidade das unidades de governo locais da Suécia por Pettersson-Lindbom (2008). Os nossos

resultados mostram que a identidade partidária, ao nível das municipalidades brasileiras, tem impacto significativo sobre as decisões políticas de gastos e receitas.

Apesar de, inicialmente, verificarmos a inexistência de impacto da ideologia partidária sobre o tamanho do governo, isto é, sobre o valor absoluto (per capita) das medidas de despesas e receitas. Ao analisarmos a relação entre o modo como se dá a composição de gastos e receitas nos municípios e o controle das prefeituras por grupos políticos mais orientados a esquerda ou direita, obtemos indícios de efeitos da identidade partidária ao nível local. Em relação à composição das receitas, encontramos evidência de um impacto positivo e significativo do governo ser de esquerda sobre a importância relativa do IPTU para a composição das receitas tributárias do governo local. No entanto, esse resultado deve ser interpretado com cautela, pois depende da especificação escolhida para a regressão com descontinuidade.

Não obstante, as principais evidências que encontramos referem-se ao perfil de gastos públicos em diferentes áreas de responsabilidade dos governos locais. Os governos de esquerda gastam proporcionalmente menos do que os de direita em saúde e urbanismo. O fato de o governo ser de esquerda implica uma redução de 4,5% em relação à participação dos gastos em saúde sobre as despesas orçamentárias. Já para a proporção de gastos destinados a urbanismo, esse impacto está entre 11% a 17% para as cidades que constituem a amostra. Por outro lado, nós encontramos evidências bastante robustas de que partidos de esquerda têm um impacto positivo relevante sobre a fração dos gastos públicos destinados a financiar atividades administrativas dos governos locais. No entanto, essa maior fração de gastos em atividades administrativas não está relacionada a um maior contingente de funcionários em governos de esquerda, o que poderia aumentar a necessidade de financiamento da máquina pública.

4

Referências Bibliográficas

AARONSON, D.; BARROW, L.; SANDER, W. Teachers and student achievement in the Chicago public high schools. **Journal of labor economics**, v. 25, n. 1, p. 95-135, 2007.

ALTONJI, J. G.; BLANK, R. M. Race and gender in the labor market. In: ASHENFELTER, O (Ed.); CARD, D (Ed.). **Handbook of Labor Economics**, Amsterdam: Elsevier, v. 3C, p. 3143-259, 1999.

ARCAS, P. H. **Implicações da progressão continuada e do SARESP na avaliação escolar: tensões, dilemas e tendências**. São Paulo, 2009. 178p. Tese (Doutorado em Educação) – Faculdade de Educação, Universidade de São Paulo.

ARROW, K. J. The theory of discrimination. In: Ashenfelter, O. (Ed.); Rees, A (Ed.). **Discrimination in labor markets**. Princeton University Press, p. 3-33, 1974.

BECKER, G. S. **The economics of discrimination**. 2. ed. London: The University of Chicago Press, 178 p., 1971.

BERTRAND, M.; GOLDIN, C.; KATZ, L. F. Dynamics of the gender gap for young professionals in the financial and corporate sectors. **American Economic Journal: Applied Economics**, v. 2, p. 228-55, 2010.

BESLEY, T.; CASE, A. Political Institutions and Policy Choices: Evidence from the United States. **Journal of Economic Literature**, v. 41, n. 1, p. 7-73, 2003.

BLACK, D. On the rationale of group decision making. **Journal of Political Economy**, v. 56, p. 23-34, 1948.

BLAU, F. D.; KAHN, L. M. The feasibility and importance of adding measures of actual experience to cross-sectional data collection. **NBER Working Paper**, n. 17.241, National Bureau of Economic Research, Cambridge, p. 48, 2011.

BOARDMAN, A. E.; MURNANE, R. J. Using panel data to improve estimates of the determinants of educational achievement. **Sociology of Education**, v. 52, p. 113-21, 1979.

CAHUC, P.; ZYLBERBERG, A. **Labor Economics**, Cambridge: The Massachusetts Institute of Technology Press, 880 p., 2004.

CAMARGO, J. M.; SERRANO, F. Os Dois Mercados: homens e mulheres na indústria brasileira. **Revista Brasileira de Economia**, Rio de Janeiro, v. 37, n. 4, p. 435-48, 1983.

CAVALIERI, C. H.; FERNANDES, R. Diferenciais de Salário por Gênero e Cor: uma comparação entre as regiões metropolitanas brasileiras. **Revista de Economia Política**, v. 18, n. 1, p. 158-175, 1998.

CHATTOPADHYAY, R.; DUFLO, E. Women as Policy Makers: Evidence from a Randomized Policy Experiment in India. **Econometrica**, v. 72, n. 5, p. 1409-1443, 2004.

CHETTY, R.; FRIEDMAN, J. N.; ROCKOFF, J. E. The long term impacts of teachers: teacher value-added and student outcomes in adulthood. **NBER Working Paper**, n. 17.699, National Bureau of Economic Research, Cambridge, 94 p., 2011.

CLARK, D.; MARTORELL, P; ROCKOFF, J. School principals and school performance. **National Center for Analysis of Longitudinal Data in Education Research Working Paper**, n. 38, 56 p., 2009.

CUNHA, F.; HECKMAN, J. The technology of skill formation. **The American Economic Review**, v. 97, n. 2, p. 31-47, 2007.

_____, F.; _____, J. J.; SCHENNACH, S. M. Estimating the technology of cognitive and noncognitive skill formation. **Econometrica**, v. 78, n. 3, p. 883-931, 2010.

DEE, T. S. A teacher like me: does race, ethnicity or gender matter? **The American Economic Review**, v. 95, n. 2, p. 158-165, 2005.

DEE, T. S. Teachers and the gender gaps in student achievement. **The Journal of Human Resources**, v. 42, n. 3, p. 528-54, 2007.

DOWNS, A. **An economic theory of democracy**. 1. ed. New York: Harper & Row, 1957.

FABIAN, W. Quality matters: the expulsion of professors and the consequences for PhD student outcomes in nazi Germany. **Journal of Political Economy**, v. 118, n. 4, p. 787-831.

FERREIRA, F.; GYOURK, J. (2009). Do Political Parties Matter? Evidence from U.S. Cities. **The Quarterly Journal of Economics**, v. 124, n. 1, p. 349-397, 2009.

FIGUEIREDO, A. C.; LIMONGI, F. **Executivo e Legislativo na nova ordem constitucional**. Rio de Janeiro: FGV, 1999.

FOLKE, O. (2009). Shades of Brown and Green: Party Effects in Proportional Election Systems. **Working Paper**, 56 p., 2011.

FORTIN, N; LEMIEUX, T.; FIRPO, S. Decomposition methods in economics. In: CARD, D (Ed.); ASHENFELTER, O (Ed.). **Handbook of Labor Economics**, Amsterdam: North Holland, v. 4A, p. 1-102, 2011.

FRYER JR., R. G. Financial incentives and student achievement: evidence from a randomized trials. **The Quarterly Journal of Economics**, v. 126, n. 4, p. 1755-98, 2011.

GIUBERTI, A. C.; MENEZES-FILHO, N. Discriminação de rendimentos por gênero: uma comparação entre o Brasil e os Estados Unidos. **Economia Aplicada**, v. 93, n. 3, p. 369-83, 2005.

GLEWWE, P.; KREMER, M. Schools, teachers and education outcomes in developing countries. In: HANUSHEK, E. A.; WELCH, F. **Handbook of the Economics of Education**, v. 2, North-Holland, Amsterdam, p. 945-1017, 2006.

GOLDHABER, D. Everyone's doing it, but what does teacher testing tell us about teacher effectiveness? **Journal of Human Resources**, v. 42, p. 765-94, 2007.

GOLDIN, C; KATZ, L. F. Transitions: career and family life cycles of the educational elite. **American Economic Review: Papers & Proceedings**, v. 98, n. 2, p. 363-69, 2008.

GUIMARÃES, R.; SITARAM, A.; JARDON, L.; TAGUCHI, S.; ROBINSON, L. The effect of teacher content knowledge on student achievement: a quantitative case analysis of six brazilian states. In: **Population Association of America Annual Meeting, 2013**, New Orleans, Population Association of America 2013 Annual Meeting Program, 17p., 2013.

HANUSHEK, E. A.; RIVKIN, S. G. Generalizations about using value-added measures of teacher quality. **American Economic Review**, v. 100, n. 2, p. 267-71, 2010.

IMBEAU, L.; PÉTRY, F.; LAMARI, M. Left-right party ideology and government policies: A meta-analysis. **European Journal of Political Research**, v. 40, p. 1–29, 2001.

JACOB, B. A.; LEFGREN, L. Can principals identify effective teachers? Evidence on subjective performance evaluation in education. **Journal of Labor Economics**, v. 26, n. 1, p. 101-36, 2008.

KANE, T.; STAIGER, D. O. Estimating teacher impacts on student achievement: an experimental evaluation. **NBER Working Paper**, n. 14.607, National Bureau of Economic Research, Cambridge, 51 p., 2008.

KANE, T.; ROCKOFF, J. E.; STAIGER, D. O. What does certification tell us about teacher effectiveness? Evidence from New York city. **Economics of Education Review**, v. 27, p. 615-31, 2008.

KASSOUF, A. L. Wage Gender Discrimination and Segmentation in the Brazilian Labor Market. **Economia Aplicada**, São Paulo, v.2, n.2, 1998.

LANGONI, C. G. **Distribuição da Renda e Desenvolvimento Econômico do Brasil**. Rio de Janeiro: Editora Expressão e Cultura, 315p., 1973.

LAVY, V. What makes an effective teacher? Quasi-experimental evidence. **NBER Working Paper**, n. 16.885, National Bureau of Economic Research, Cambridge, 43 p., 2011.

LEE, D.; MORETTI, E.; BUTLER, D. Do Voters Affect or Elect Policies? Evidence from the U.S. House. **The Quarterly Journal of Economics**, v. 119, n. 3, 2004.

LEME, C.; WAJNMAN, S. Tendências de Coorte nos Diferenciais de Rendimento por Sexo. In: HENRIQUES, R. M. (org.). **Desigualdade e Pobreza no Brasil**. Rio de Janeiro: IPEA, p. 251-70, 2000.

MACHADO, C. Selection, heterogeneity and the gender wage gap. **Working Paper**, 62 p., 2012.

MADALOZZO, R. Occupational segregation and the gender wage gap in Brazil: an empirical analysis. **Economia Aplicada**, v. 14, n. 2, p. 147-68, 2010.

MANNING, A; SWAFFIELD, J. The gender gap in early-career wage growth. **The Economic Journal**, v. 118, p. 983-1024, 2008.

METZLER, J.; WOESMANN, L. The impact of teacher subject knowledge on student achievement: evidence from within-teacher within-student variation. **Journal of Development Economics**, v. 99, p. 486-96, 2012.

NORDMAN, C. J.; ROUBAUD, F. Reassessing the gender gap in Madagascar: Does labor force attachment really matter? **Economic Development and Cultural Change**, v. 57, n.4, p. 785-808, 2009.

PANDE, R. Can Mandated Political Representation Provide Disadvantaged Minorities Policy Influence? Theory and Evidence from India. **American Economic Review**, v. 93, n. 4, p. 1132-51, 2003.

PERSSON, T.; TABELLINI, G. E. **Political Economics: Explaining Economic Policy** (Zeuthen Lectures). Cambridge: The Massachusetts Institute of Technology Press, p. 468, 2000.

PETTERSSON-LIDBOM, P. Do parties matter for economic outcomes: a regression-discontinuity approach. **Journal of the European Economic Association**, v. 6 n. 5, p. 1037–56, 2008.

PHELPS, E. The statistical theory of racism and sexism. **American Economic Review**, v. 62, p. 569-661, 1972.

POWER, T. **The Political Right in Postauthoritarian Brazil: Elites, Institutions, and Democratization**. University Park, PA: Penn State Press, 2000.

RAMOS, L; FERREIRA, V. Padrões espacial e setorial da evolução da informalidade no período 1991-2005. **Pesquisa e Planejamento Econômico**, Rio de Janeiro, v. 36, n.3, p. 471-88, 2006.

REGAN, T. L.; OAXACA, R. L. Work experience as a source of specification error in earnings models: implications for gender wage decompositions. **Journal of Population Economics**, v. 22, p. 463-99, 2009.

RIVKIN, S. G.; HANUSHEK, E. A.; KAIN, J. F. Teachers, schools and academic achievement. **Econometrica**, v. 73, n. 2, p. 417-58, 2005.

ROCKOFF, J. The impact of individual teachers on student achievement: evidence from panel data. **The American Economic Review**, v. 94, n. 2, p. 247-252, 2004.

ROCKOFF, J; SPERONI, C. Subjective and objective evaluations of teacher effectiveness. **The American Economic Review: Papers & Proceedings**, v. 100, p. 261-66, 2010.

ROTHSTEIN, J. Teacher quality in educational production: tracking, decay and student achievement. **The Quarterly Journal of Economics**, v. 125, n. 1, p. 175-214, 2010.

SANTIBAÑEZ, L. Why we should care if teachers get A's: teachers test scores and student achievement in Mexico. **Economics of Education Review**, v. 25, p. 510-20, 2006.

SCORZAFAVE, L. G.; MENEZES-FILHO, N. A. Participação feminina no mercado de trabalho brasileiro: evolução e determinantes. **Pesquisa e Planejamento Econômico**, Rio de Janeiro, v. 31, n. 3, p. 441-78, 2001.

SCORZAFAVE, L. G.; PAZELLO, E. T. Using Normalized Equations to Solve the Indetermination Problem in the Oaxaca-Blinder Decomposition: An Application to the Gender Wage Gap in Brazil. **Revista Brasileira de Economia**, Rio de Janeiro, v. 61, n. 4, p. 535-48, 2007.

TODD, P. E.; WOLPIN, K. I. On the specification and estimation of the production function for cognitive achievement. **Economic Journal**, v. 113, p. F3-F33, 2003.

TODD, P. E.; WOLPIN, K. I. The production of cognitive achievement in children: home, school and racial test score gaps. **Journal of Human Capital**, v. 1, n. 1, p. 91-136, 2007.

WAYNE, A. J.; YOUNGS, P. Teacher characteristics and student achievement gains: a review. **Review of Educational Research**, v. 73, n. 1, p. 89-122, 2003.

ZUCCO JUNIOR, C.; POWER, T. Estimating Ideology of Brazilian Legislative Parties, 1990–2005. **Latin American Research Review**, v. 44, n.1, p. 218–46, 2009.

Apêndices:

Apêndice 1 – Procedimentos para verificar as inconsistências na amostra inicial da RAIS/MTE.

Aqui apresentamos uma descrição dos algoritmos utilizados nos testes para inconsistências em relação aos dados dos trabalhadores presentes na RAIS/MTE e os procedimentos de correção dessas informações, quando possível. Além de outros ajustes que se fizeram necessários devido às particularidades da RAIS/MTE.

Primeiramente, identificamos os casos de duplicidade de informações para a mesma combinação entre trabalhador e ano. Eram 137 observações repetidas para as quais todas as demais variáveis relevantes eram idênticas. Neste caso optamos por eliminar as observações repetidas. Adicionalmente para 352 observações na base de dados havia apenas a informação sobre o número identificador dos trabalhadores (código PIS). Estas observações também foram excluídas. Após esses primeiros procedimentos a amostra inicial tinha 1.364.646 trabalhadores e 10.308.428 observações ou vínculos empregatícios.

No segundo procedimento, nós analisamos a conformidade das informações relativas ao sexo dos indivíduos. Do total de 1.364.646 trabalhadores, para 55.276 havia alguma inconsistência quanto ao preenchimento da informação sobre gênero por parte das firmas em pelo menos um ano entre os dezesseis compreendidos na análise. Assim para 1.309.370 indivíduos a informação sobre gênero era consistente ao longo de todo o painel – isso representa 96% da amostra em questão. Para os demais indivíduos com alguma inconsistência na informação sobre gênero, nós optamos por aplicar a seguinte regra: se ao longo dos dezesseis anos mais do que 75% das observações afirmavam que o indivíduo era do sexo feminino, então nós consideramos esse trabalhador como uma mulher. A mesma regra foi aplicada para os homens. No final restaram 17.452 indivíduos para os quais não foi possível determinar de maneira inequívoca – de acordo com a nossa regra – o respectivo sexo. Estes foram eliminados da amostra, o que representa 134.894 observações. Após esse procedimento referente a análise do gênero dos indivíduos, a base de dados permaneceu com 1.347.194 indivíduos, o que

representa 98,72% do total de indivíduos anterior ao exame das informações sobre o gênero, ou 10.173.534 observações.

O terceiro procedimento refere-se ao teste e correção de possíveis inconsistências – quando possível – para a variável de idade dos trabalhadores. Para 733 indivíduos não havia informação alguma sobre a idade dos indivíduos ao longo dos anos em que estes estavam presentes na base de dados. Por conseguinte, esses trabalhadores foram eliminados da amostra. Do restante de trabalhadores na amostra (1.346.461), para 96,29%, ou seja, 1.296.572 indivíduos, as informações sobre suas idades ao longo do intervalo de tempo avaliado eram consistentes entre si e não apresentavam valores *missing*. Para os demais 49.889 trabalhadores havia pelo menos um de dois possíveis problemas com os dados: (i) valores *missing* para pelo menos uma observação por indivíduo; e/ou (ii) inconsistência ao longo do tempo da informação fornecida pelas firmas sobre as idades. Com o auxílio das etapas que serão descritas abaixo, nós pudemos recuperar as informações de maneira consistente inter-temporalmente para 43.048 empregados – o que representa 86,3% dos indivíduos que apresentavam problemas na informação sobre idade.

O primeiro passo foi o preenchimento dos valores *missing* para o grupo de indivíduos para os quais as demais observações continham informações consistentes acerca da idade destes para os diferentes anos observados. No segundo passo, analisamos o conjunto de trabalhadores que não apresentavam valores *missing* para a idade, mas tinham alguma informação inconsistente sobre as suas idades ao longo do painel. Para tratar essas informações, nós confrontamos os valores de idade do trabalhador no início e no final do painel com a variação de números de anos entre a primeira e a última observação de cada indivíduo na amostra. Se esse conjunto de informações estava consistente entre si, nós simplesmente interpolamos a idade do trabalhador para as demais observações nos períodos intermediários que apresentavam problemas. O passo três é muito semelhante ao anterior, com a alteração que nós permitíamos a existência de inconformidades da idade ao longo das observações intermediárias e alguns valores *missing*. E corrigimos as informações intermediárias sobre idades caso houvesse consistências entre os valores inicial e final da idade dos indivíduos e a variação no número de anos entre essas observações na extremidade do painel.

Após os três passos acima, para os demais indivíduos presentes na amostra que permaneceram com inconsistência nas informações sobre sua idade ao longo do tempo e não tinham valores *missing* para essa variável, nós procedemos da seguinte maneira. O quarto passo comparava a idade do trabalhador no último período com a informação presente na segunda observação do painel. Se essa diferença fosse igual a variação no número de anos entre essas duas observações então nós corrigíamos a idade na primeira observação desde que houvesse no mínimo quatro observações na base de dados para o respectivo indivíduo. O quinto passo é muito semelhante ao quarto com a diferença de que nós comparamos a informação sobre idades na última e na terceira observação. Se estas fossem compatíveis com a variação no número de anos entre as observações associadas e o indivíduo estivesse presente em pelo menos seis períodos, nós corrigimos a idade nas duas primeiras observações do painel. Os demais procedimentos são a extensão desses passos quatro e cinco para o restante de trabalhadores que ainda apresentavam problemas na informação sobre idade. A cada novo passo um menor número de indivíduos se encaixava nos requerimentos para correção da idade. Por último, para 6.841 trabalhadores não foi possível reconstruir de maneira harmoniosa a série de idade com nossa regra e estes foram eliminados da amostra. Assim, após os procedimentos referentes a verificação das informações sobre idade a amostra permaneceu com 1.339.620 trabalhadores e 10.126.894 observações.

Por último, o quarto procedimento refere-se a investigação da consistência das informações sobre escolaridade dos trabalhadores ao longo do painel prestadas pelas firmas. Para 251.842 trabalhadores não foi possível determinar reconstruir de maneira adequada à série de faixas de escolaridade ao longo dos anos observados na amostra. Esses trabalhadores estavam associados a 2.538.762 observações que foram excluídas da base de dados. Depois desse procedimento de compatibilização das informações sobre escolaridade dos empregados restaram na amostra 1.083.762 indivíduos ou 7.489.693 observações.

Apêndice 2 – Evidências sobre a validade externa dos impactos dos professores estimados a partir da amostra de alunos com informações sobre a atribuição e notas dos seus docentes.

A percepção derivada da figura 2.1 pode ser corroborada por algumas estatísticas dessas distribuições que são apresentadas na tabela A1. O painel A apresenta as evidências para a disciplina de matemática e o painel B para língua portuguesa. De maneira geral, os valores dos desvios-padrão são muito semelhantes entre as duas populações. Em relação às medidas de posição central dessas distribuições, nós podemos observar pequenas diferenças favoráveis ao conjunto de estudantes que tem seus docentes identificados.

Uma alternativa mais formal para testar empiricamente a igualdade entre as duas distribuições dos resíduos – da regressão da nota no SARESP 2009 sobre a proficiência passada em 2007 – para cada disciplina separadamente é o teste de Kolmogorov-Smirnov. Os resultados também são apresentados na tabela A1 e apontam que as distribuições são estatisticamente distintas. Seguindo o padrão descrito na comparação das estatísticas dessas distribuições, o teste indica que a aquisição de proficiência líquida do conjunto de alunos para os quais os professores não são identificados tem valores inferiores do que para o grupo de estudantes – com professores identificados – que utilizamos nos nossos exercícios empíricos. Para matemática, a maior diferença entre as funções distribuição é de 0,0318 e o p-valor igual a zero. Já para língua portuguesa, esses valores são iguais a 0,0206 e zero, respectivamente. O teste combinado apresenta p-valor igual a zero para ambas as disciplinas o que rejeita a igualdade entre as duas distribuições.

Portanto, a partir da análise em conjunto das evidências apresentadas na figura 2.1 e tabela A1, apesar das distribuições serem distintas, essas diferenças não parecem ser de magnitude elevada. Assim as evidências sugerem a existência de um efeito da qualidade do professor sobre a produção de proficiência mesmo para os alunos para os quais não há a disponibilidade de dados sobre a atribuição dos professores.

Tabela A1 – Comparação das distribuições de ganho de desempenho entre os SARESP 2007 e 2009: grupo de estudantes para os quais foi possível identificar os professores vs. os demais alunos

Painel A: Matemática - Resíduo da regressão do desempenho no SARESP 2009 na nota no SARESP 2007

Teste Kolmogorov-Smirnov para igualdade das funções distribuições entre duas amostras

Menor grupo	Distância máxima	P-valor	P-valor corrigido
Grupo de alunos com docentes não identificados	0,0318	0,000	
Grupo de alunos com docentes identificados	-0,0001	0,998	
Teste Kolmogorov-Smirnov combinado	0,0318	0,000	0,000

Estadísticas das amostras

	Total	Docentes identificados	Docentes não identificados
Média	0,000	1,328	-1,128
Mediana	0,404	1,740	-0,731
Desvio-padrão	33,058	33,128	32,956
Número total de alunos	338.104	155.272	182.832

Painel B: Língua Portuguesa - Resíduo da regressão do desempenho no SARESP 2009 na nota no SARESP 2007

Teste Kolmogorov-Smirnov para igualdade das funções distribuições entre duas amostras

Menor grupo	Distância máxima	P-valor	P-valor corrigido
Grupo de alunos com docentes não identificados	0,0206	0,000	
Grupo de alunos com docentes identificados	-0,0001	0,999	
Teste Kolmogorov-Smirnov combinado	0,0206	0,000	0,000

Estadísticas das amostras

	Total	Docentes identificados	Docentes não identificados
Média	0,000	0,767	-0,718
Mediana	0,569	1,307	-0,092
Desvio-padrão	31,088	31,054	31,103
Número total de alunos	336.898	162.892	174.006

Fonte: Elaboração própria, dados do SARESP 2009, SARESP 2007 e dados administrativos da atribuição de professores da SEE/SP.

Tabela A2 – Equação de determinação de salários para os trabalhadores entre 18 e 55 anos por faixa de escolaridade com a inclusão de controles para a ocupação no ano 2009

Ln salário médio real	Amostra Geral - trabalhadores entre 18 e 55 anos em 2009			
	Ensino Superior I	Ensino Médio II	Ensino Fundamental III	Semi-Analfabeto IV
Mulher	-0.256	-0.173	-0.164	-0.228
	(0.00601)***	(0.00207)***	(0.00247)***	(0.0125)***
Num. anos fora mercado 2004-08	-0.0451	-0.0228	-0.0177	-0.0229
	(0.00349)***	(0.00100)***	(0.00107)***	(0.00456)***
Num. anos fora mercado 1999-03	-0.0239	-0.00249	-0.000160	0.00646
	(0.00340)***	(0.00111)**	(0.00114)	(0.00472)
Num. anos fora mercado 1994-98	0.0108	0.0140	0.00792	-0.00455
	(0.00531)**	(0.00190)***	(0.00170)***	(0.00729)
Ano inteiro em 2009	0.101	0.0163	-0.0162	-0.0816
	(0.00619)***	(0.00187)***	(0.00225)***	(0.0103)***
Num. anos inteiros 2004-08	0.0631	0.0550	0.0357	0.0220
	(0.00228)***	(0.000765)***	(0.000844)***	(0.00370)***
Num. anos inteiros 1999-03	0.00955	0.0319	0.0240	0.0243
	(0.00246)***	(0.000981)***	(0.00105)***	(0.00451)***
Num. anos inteiros 1994-98	0.0326	0.0488	0.0392	0.0260
	(0.00257)***	(0.00118)***	(0.00118)***	(0.00450)***
Num anos setor público	-0.00179	0.000549	0.00860	0.0188
	(0.00161)	(0.00133)	(0.00239)***	(0.00766)**
Dummies Idade	Sim	Sim	Sim	Sim
Estados	Sim	Sim	Sim	Sim
Jornada de trabalho	Sim	Sim	Sim	Sim
Tempo parcial	Sim	Sim	Sim	Sim
Ocupações	Sim	Sim	Sim	Sim
Constante	7.602	7.509	6.927	7.332
	(0.111)***	(0.0204)***	(0.0362)***	(0.188)***
Observações	65,326	287,738	149,085	8,981
R-quadrado	0.448	0.416	0.430	0.396

Fonte: Elaboração própria, dados da RAIS/MTE 1994 a 2009.

Erros padrão entre parênteses - *** p<0,01; ** p<0,05; * p<0,1.

Tabela A3 – Equação de determinação de salários para seis coortes de trabalhadores com inclusão de controles para ocupação no ano de 2009

Ln salário médio horário	Amostra Geral - trabalhadores entre 18 e 55 anos em 2009					
	Coorte 1987-91 18 a 22 anos	Coorte 1982-86 23 a 27 anos	Coorte 1977-81 28 a 32 anos	Coorte 1972-76 33 a 37 anos	Coorte 1967-71 38 a 42 anos	Coorte 1962-66 43 a 47 anos
	I	II	III	IV	V	VI
Mulher	-0.0732	-0.145	-0.183	-0.226	-0.247	-0.261
	(0.00287)***	(0.00302)***	(0.00379)***	(0.00506)***	(0.00600)***	(0.00722)***
Num. anos fora mercado 2004-08	-0.00614	-0.0202	-0.0253	-0.0274	-0.0236	-0.0214
	(0.00512)	(0.00157)***	(0.00146)***	(0.00186)***	(0.00226)***	(0.00286)***
Num. anos fora mercado 1999-03		0.0166	0.000357	-0.00200	-0.00335	-0.00629
		(0.00415)***	(0.00154)	(0.00170)	(0.00202)*	(0.00244)***
Num. anos fora mercado 1994-98			0.00245	-5.92e-05	0.00775	0.0110
			(0.00489)	(0.00236)	(0.00270)***	(0.00333)***
Ano inteiro em 2009	0.00651	0.0269	0.0312	0.0237	0.0188	0.0129
	(0.00272)**	(0.00262)***	(0.00323)***	(0.00425)***	(0.00524)***	(0.00619)**
Num. anos inteiros 2004-08	0.0573	0.0408	0.0477	0.0478	0.0508	0.0558
	(0.00227)***	(0.00105)***	(0.00116)***	(0.00150)***	(0.00180)***	(0.00213)***
Num. anos inteiros 1999-03		0.0176	0.0156	0.0179	0.0243	0.0275
		(0.00252)***	(0.00133)***	(0.00158)***	(0.00189)***	(0.00217)***
Num. anos inteiros 1994-98			0.00888	0.0157	0.0257	0.0383
			(0.00309)***	(0.00170)***	(0.00174)***	(0.00195)***
Num anos setor público	-0.00846	-0.0163	-0.00732	-0.00966	-0.00172	0.00183
	(0.0124)	(0.00363)***	(0.00243)***	(0.00209)***	(0.00243)	(0.00232)
Dummies Idade	Sim	Sim	Sim	Sim	Sim	Sim
Estados	Sim	Sim	Sim	Sim	Sim	Sim
Jornada de trabalho	Sim	Sim	Sim	Sim	Sim	Sim
Tempo parcial	Sim	Sim	Sim	Sim	Sim	Sim
Faixas escolaridade	Sim	Sim	Sim	Sim	Sim	Sim
Ocupações	Sim	Sim	Sim	Sim	Sim	Sim
Observações	80,157	117,715	103,702	71,630	54,278	42,779
R-quadrado	0.288	0.420	0.516	0.563	0.593	0.622

Fonte: Elaboração própria, dados da RAIS/MTE 1994 a 2009.

Erros padrão entre parênteses - *** p<0,01; ** p<0,05; * p<0,1.

Tabela A4 – Estimativas do diferencial de salários entre gêneros por ano para a coorte 1972 - 1976

Coorte 1972 - 1976								
Anos	1994	1995	1996	1997	1998	1999	2000	2001
	18 a 22 anos	19 a 23 anos	20 a 24 anos	21 a 25 anos	22 a 26 anos	23 a 27 anos	24 a 28 anos	25 a 29 anos
Diferencial de salários entre gêneros - modelo básico / equação (2)								
Amostra geral	-0.0834 (0.00680)***	-0.108 (0.00561)***	-0.104 (0.00553)***	-0.105 (0.00550)***	-0.104 (0.00567)***	-0.102 (0.00579)***	-0.107 (0.00592)***	-0.125 (0.00590)***
Continuamente empregados	-0.0548 (0.0154)***	-0.0716 (0.0149)***	-0.0699 (0.0148)***	-0.0726 (0.0150)***	-0.0830 (0.0152)***	-0.0705 (0.0156)***	-0.0723 (0.0161)***	-0.0700 (0.0166)***
Diferencial de salários entre gêneros - modelo completo / equação (3)								
Amostra geral	-0.160 (0.00663)***	-0.214 (0.00514)***	-0.228 (0.00499)***	-0.247 (0.00487)***	-0.263 (0.00494)***	-0.275 (0.00496)***	-0.292 (0.00495)***	-0.312 (0.00482)***
Continuamente empregados	-0.170 (0.0147)***	-0.208 (0.0134)***	-0.228 (0.0132)***	-0.244 (0.0137)***	-0.263 (0.0133)***	-0.267 (0.0134)***	-0.279 (0.0138)***	-0.292 (0.0140)***
N. Observações (Amostra geral)	42,829	52,724	55,984	59,488	60,751	60,875	62,970	65,632
Anos	2002	2003	2004	2005	2006	2007	2008	2009
	26 a 30 anos	27 a 31 anos	28 a 32 anos	29 a 33 anos	30 a 34 anos	31 a 35 anos	32 a 36 anos	33 a 37 anos
Diferencial de salários entre gêneros - modelo básico / equação (2)								
Amostra geral	-0.131 (0.00589)***	-0.147 (0.00585)***	-0.171 (0.00588)***	-0.184 (0.00584)***	-0.192 (0.00575)***	-0.205 (0.00568)***	-0.229 (0.00570)***	-0.233 (0.00573)***
Continuamente empregados	-0.0693 (0.0167)***	-0.0827 (0.0167)***	-0.101 (0.0170)***	-0.0871 (0.0172)***	-0.0874 (0.0174)***	-0.0981 (0.0174)***	-0.111 (0.0176)***	-0.0969 (0.0180)***
Diferencial de salários entre gêneros - modelo completo / equação (3)								
Amostra geral	-0.320 (0.00473)***	-0.330 (0.00464)***	-0.351 (0.00464)***	-0.359 (0.00456)***	-0.365 (0.00447)***	-0.375 (0.00441)***	-0.393 (0.00439)***	-0.393 (0.00443)***
Continuamente empregados	-0.293 (0.0140)***	-0.317 (0.0137)***	-0.329 (0.0142)***	-0.317 (0.0141)***	-0.320 (0.0144)***	-0.336 (0.0142)***	-0.350 (0.0142)***	-0.344 (0.0146)***
N. Observações (Amostra geral)	66,369	65,936	67,222	68,279	69,388	70,782	72,423	71,635
N. Observações (Continuamente empregados)	8,416	8,416	8,416	8,416	8,416	8,416	8,416	8,416

Fonte: Elaboração própria, dados da RAIS/MTE 1994 a 2009.

Erros padrão entre parênteses - *** p<0,01; ** p<0,05; * p<0,1.

Tabela A5 – Estimativas do diferencial de salários entre gêneros por ano para a coorte 1967 - 1971

Coorte 1967 - 1971								
Anos	1994	1995	1996	1997	1998	1999	2000	2001
	23 a 27 anos	24 a 28 anos	25 a 29 anos	26 a 30 anos	27 a 31 anos	28 a 32 anos	29 a 33 anos	30 a 34 anos
Diferencial de salários entre gêneros - modelo básico / equação (2)								
Amostra geral	-0.172 (0.00711)***	-0.191 (0.00673)***	-0.181 (0.00685)***	-0.187 (0.00694)***	-0.188 (0.00716)***	-0.189 (0.00730)***	-0.205 (0.00736)***	-0.221 (0.00740)***
Continuamente empregados	-0.188 (0.0156)***	-0.208 (0.0153)***	-0.173 (0.0156)***	-0.169 (0.0156)***	-0.166 (0.0159)***	-0.141 (0.0160)***	-0.143 (0.0163)***	-0.135 (0.0167)***
Diferencial de salários entre gêneros - modelo completo / equação (3)								
Amostra geral	-0.324 (0.00644)***	-0.361 (0.00568)***	-0.362 (0.00564)***	-0.380 (0.00564)***	-0.383 (0.00578)***	-0.382 (0.00588)***	-0.398 (0.00583)***	-0.415 (0.00579)***
Continuamente empregados	-0.378 (0.0133)***	-0.417 (0.0125)***	-0.398 (0.0125)***	-0.398 (0.0125)***	-0.404 (0.0127)***	-0.385 (0.0125)***	-0.395 (0.0129)***	-0.396 (0.0128)***
N. Observações (Amostra geral)	50,958	54,652	53,646	53,498	52,565	51,254	51,197	52,344
Anos	2002	2003	2004	2005	2006	2007	2008	2009
	31 a 35 anos	32 a 36 anos	33 a 37 anos	34 a 38 anos	35 a 39 anos	36 a 40 anos	37 a 41 anos	38 a 42 anos
Diferencial de salários entre gêneros - modelo básico / equação (2)								
Amostra geral	-0.235 (0.00734)***	-0.244 (0.00732)***	-0.262 (0.00727)***	-0.266 (0.00718)***	-0.274 (0.00700)***	-0.281 (0.00694)***	-0.295 (0.00693)***	-0.294 (0.00695)***
Continuamente empregados	-0.143 (0.0165)***	-0.143 (0.0165)***	-0.154 (0.0167)***	-0.151 (0.0169)***	-0.151 (0.0169)***	-0.149 (0.0169)***	-0.160 (0.0171)***	-0.144 (0.0173)***
Diferencial de salários entre gêneros - modelo completo / equação (3)								
Amostra geral	-0.421 (0.00568)***	-0.422 (0.00557)***	-0.430 (0.00554)***	-0.428 (0.00547)***	-0.433 (0.00532)***	-0.434 (0.00526)***	-0.443 (0.00526)***	-0.434 (0.00528)***
Continuamente empregados	-0.406 (0.0127)***	-0.406 (0.0127)***	-0.415 (0.0128)***	-0.412 (0.0130)***	-0.412 (0.0128)***	-0.401 (0.0130)***	-0.412 (0.0131)***	-0.387 (0.0133)***
N. Observações (Amostra geral)	52,123	51,376	52,020	52,427	53,088	53,729	54,737	54,285
N. Observações (Continuamente empregados)	11,494	11,494	11,494	11,494	11,494	11,494	11,494	11,494

Fonte: Elaboração própria, dados da RAIS/MTE 1994 a 2009.

Erros padrão entre parênteses - *** p<0,01; ** p<0,05; * p<0,1.

Tabela A6 – Estimativas do diferencial de salários entre gêneros por ano para a coorte 1962 - 1966

Coorte 1962 - 1966								
Anos	1994	1995	1996	1997	1998	1999	2000	2001
	28 a 32 anos	29 a 33 anos	30 a 34 anos	31 a 35 anos	32 a 36 anos	33 a 37 anos	34 a 38 anos	35 a 39 anos
Diferencial de salários entre gêneros - modelo básico / equação (2)								
Amostra geral	-0.245 (0.00858)***	-0.267 (0.00828)***	-0.258 (0.00844)***	-0.266 (0.00853)***	-0.262 (0.00873)***	-0.284 (0.00878)***	-0.281 (0.00883)***	-0.309 (0.00893)***
Continuamente empregados	-0.217 (0.0187)***	-0.229 (0.0184)***	-0.199 (0.0184)***	-0.198 (0.0183)***	-0.187 (0.0187)***	-0.194 (0.0186)***	-0.177 (0.0189)***	-0.180 (0.0193)***
Diferencial de salários entre gêneros - modelo completo / equação (3)								
Amostra geral	-0.441 (0.00721)***	-0.469 (0.00654)***	-0.457 (0.00660)***	-0.475 (0.00655)***	-0.471 (0.00665)***	-0.476 (0.00669)***	-0.471 (0.00675)***	-0.489 (0.00670)***
Continuamente empregados	-0.461 (0.0149)***	-0.480 (0.0143)***	-0.464 (0.0140)***	-0.465 (0.0138)***	-0.461 (0.0138)***	-0.468 (0.0137)***	-0.455 (0.0141)***	-0.457 (0.0142)***
N. Observações (Amostra geral)	45,144	47,508	45,940	45,310	44,228	42,723	42,512	42,854
Anos	2002	2003	2004	2005	2006	2007	2008	2009
	36 a 40 anos	37 a 41 anos	38 a 42 anos	39 a 43 anos	40 a 44 anos	41 a 45 anos	42 a 46 anos	43 a 47 anos
Diferencial de salários entre gêneros - modelo básico / equação (2)								
Amostra geral	-0.318 (0.00887)***	-0.331 (0.00882)***	-0.348 (0.00876)***	-0.342 (0.00869)***	-0.329 (0.00851)***	-0.345 (0.00839)***	-0.343 (0.00839)***	-0.328 (0.00844)***
Continuamente empregados	-0.187 (0.0191)***	-0.182 (0.0190)***	-0.191 (0.0191)***	-0.185 (0.0192)***	-0.190 (0.0191)***	-0.184 (0.0191)***	-0.184 (0.0192)***	-0.166 (0.0194)***
Diferencial de salários entre gêneros - modelo completo / equação (3)								
Amostra geral	-0.488 (0.00664)***	-0.488 (0.00660)***	-0.501 (0.00655)***	-0.497 (0.00648)***	-0.484 (0.00642)***	-0.482 (0.00630)***	-0.487 (0.00625)***	-0.470 (0.00627)***
Continuamente empregados	-0.459 (0.0140)***	-0.457 (0.0140)***	-0.465 (0.0141)***	-0.455 (0.0141)***	-0.455 (0.0140)***	-0.437 (0.0140)***	-0.437 (0.0142)***	-0.415 (0.0145)***
N. Observações (Amostra geral)	42,508	41,751	41,725	41,984	42,427	42,681	43,346	42,788
N. Observações (Continuamente empregados)	10,766	10,766	10,766	10,766	10,766	10,766	10,766	10,766

Fonte: Elaboração própria, dados da RAIS/MTE 1994 a 2009.

Erros padrão entre parênteses - *** p<0,01; ** p<0,05; * p<0,1.

Tabela A7 – Estimativas do diferencial de salários entre gêneros por ano para a coorte 1957 - 1961

Coorte 1957 - 1961								
Anos	1994	1995	1996	1997	1998	1999	2000	2001
	33 a 37 anos	34 a 38 anos	35 a 39 anos	36 a 40 anos	37 a 41 anos	38 a 42 anos	39 a 43 anos	40 a 44 anos
Diferencial de salários entre gêneros - modelo básico / equação (2)								
Amostra geral	-0.373 (0.0107)***	-0.376 (0.0102)***	-0.375 (0.0103)***	-0.378 (0.0105)***	-0.360 (0.0107)***	-0.370 (0.0109)***	-0.369 (0.0108)***	-0.386 (0.0110)***
Continuamente empregados	-0.368 (0.0248)***	-0.362 (0.0241)***	-0.339 (0.0239)***	-0.321 (0.0238)***	-0.315 (0.0240)***	-0.304 (0.0239)***	-0.301 (0.0241)***	-0.291 (0.0244)***
Diferencial de salários entre gêneros - modelo completo / equação (3)								
Amostra geral	-0.555 (0.00831)***	-0.562 (0.00766)***	-0.555 (0.00764)***	-0.559 (0.00770)***	-0.549 (0.00775)***	-0.550 (0.00794)***	-0.550 (0.00805)***	-0.554 (0.00797)***
Continuamente empregados	-0.600 (0.0181)***	-0.590 (0.0172)***	-0.570 (0.0170)***	-0.557 (0.0167)***	-0.558 (0.0167)***	-0.545 (0.0166)***	-0.543 (0.0170)***	-0.539 (0.0169)***
N. Observações (Amostra geral)	35,234	37,005	35,540	34,907	33,645	32,127	31,824	31,867
Anos	2002	2003	2004	2005	2006	2007	2008	2009
	41 a 45 anos	42 a 46 anos	43 a 47 anos	44 a 48 anos	45 a 49 anos	46 a 50 anos	47 a 51 anos	48 a 52 anos
Diferencial de salários entre gêneros - modelo básico / equação (2)								
Amostra geral	-0.385 (0.0110)***	-0.395 (0.0110)***	-0.402 (0.0111)***	-0.395 (0.0111)***	-0.366 (0.0109)***	-0.365 (0.0109)***	-0.380 (0.0109)***	-0.353 (0.0111)***
Continuamente empregados	-0.301 (0.0239)***	-0.300 (0.0240)***	-0.295 (0.0241)***	-0.298 (0.0240)***	-0.296 (0.0238)***	-0.292 (0.0238)***	-0.301 (0.0239)***	-0.278 (0.0241)***
Diferencial de salários entre gêneros - modelo completo / equação (3)								
Amostra geral	-0.550 (0.00792)***	-0.562 (0.00793)***	-0.563 (0.00807)***	-0.553 (0.00802)***	-0.532 (0.00805)***	-0.531 (0.00804)***	-0.539 (0.00805)***	-0.519 (0.00820)***
Continuamente empregados	-0.545 (0.0164)***	-0.540 (0.0164)***	-0.536 (0.0168)***	-0.543 (0.0165)***	-0.531 (0.0166)***	-0.532 (0.0167)***	-0.541 (0.0171)***	-0.514 (0.0176)***
N. Observações (Amostra geral)	31,318	30,335	29,978	29,707	30,003	29,813	29,887	29,374
N. Observações (Continuamente empregados)	7,974	7,974	7,974	7,974	7,974	7,974	7,974	7,974

Fonte: Elaboração própria, dados da RAIS/MTE 1994 a 2009.

Erros padrão entre parênteses - *** p<0,01; ** p<0,05; * p<0,1.

Tabela A8 – Estimativas do diferencial de salários entre os gêneros para a amostra geral e coortes jovens: nascidos entre 1974-76, 1977-79 e 1980-82

Diferencial de salários entre gêneros - modelo completo / equação (3)								
Coorte 1974 a 1976								
Anos	1994	1995	1996	1997	1998	1999	2000	2001
	18 a 20 anos	19 a 21 anos	20 a 22 anos	21 a 23 anos	22 a 24 anos	23 a 25 anos	24 a 26 anos	25 a 27 anos
Mulher	-0,0999 (0,00906)***	-0,176 (0,00652)***	-0,198 (0,00628)***	-0,224 (0,00607)***	-0,246 (0,00615)***	-0,256 (0,00617)***	-0,276 (0,00620)***	-0,292 (0,00599)***
N. Observações (Amostra geral)	22,629	29,991	33,046	35,843	37,027	37,574	39,220	41,196
Anos	2002	2003	2004	2005	2006	2007	2008	2009
	26 a 28 anos	27 a 29 anos	28 a 30 anos	29 a 31 anos	30 a 32 anos	31 a 33 anos	32 a 34 anos	33 a 35 anos
Mulher	-0,299 (0,00588)***	-0,313 (0,00577)***	-0,338 (0,00578)***	-0,347 (0,00566)***	-0,351 (0,00554)***	-0,364 (0,00548)***	-0,385 (0,00549)***	-0,389 (0,00555)***
N. Observações (Amostra geral)	41,942	41,864	42,726	43,522	44,302	45,235	46,341	45,704
Coorte 1977 a 1979								
Anos	1997	1998	1999	2000	2001	2002	2003	2004
	18 a 20 anos	19 a 21 anos	20 a 22 anos	21 a 23 anos	22 a 24 anos	23 a 25 anos	24 a 26 anos	25 a 27 anos
Mulher	-0,0907 (0,00623)***	-0,131 (0,00573)***	-0,154 (0,00555)***	-0,180 (0,00537)***	-0,211 (0,00514)***	-0,227 (0,00502)***	-0,240 (0,00484)***	-0,265 (0,00481)***
N. Observações (Amostra geral)	27,263	33,746	37,966	42,420	46,880	48,783	49,503	51,557
Anos	2005	2006	2007	2008	2009			
	26 a 28 anos	27 a 29 anos	28 a 30 anos	29 a 31 anos	30 a 32 anos			
Mulher	-0,283 (0,00478)***	-0,288 (0,00468)***	-0,311 (0,00463)***	-0,332 (0,00460)***	-0,336 (0,00470)***			
N. Observações (Amostra geral)	53,591	55,284	56,709	58,527	57,874			
Coorte 1980 a 1982								
Anos	2000	2001	2002	2003	2004	2005	2006	2007
	18 a 20 anos	19 a 21 anos	20 a 22 anos	21 a 23 anos	22 a 24 anos	23 a 25 anos	24 a 26 anos	25 a 27 anos
Mulher	-0,0768 (0,00549)***	-0,118 (0,00471)***	-0,135 (0,00446)***	-0,164 (0,00418)***	-0,189 (0,00405)***	-0,217 (0,00399)***	-0,227 (0,00392)***	-0,257 (0,00392)***
N. Observações (Amostra geral)	31,449	42,510	49,185	52,666	57,590	61,895	65,232	68,035
Anos	2008	2009						
	26 a 28 anos	27 a 29 anos						
Mulher	-0,287 (0,00392)***	-0,288 (0,00401)***						
N. Observações (Amostra geral)	70,910	70,607						

Fonte: Elaboração própria, dados da RAIS/MTE 1994 a 2009.

Erros padrão entre parênteses - *** $p < 0,01$; ** $p < 0,05$; * $p < 0,1$.

Tabela A9 – Estimativas do diferencial de salários entre os gêneros para a amostra de trabalhadores continuamente empregados e coortes jovens: nascidos entre 1974-76, 1977-79 e 1980-82

Diferencial de salários entre gêneros - modelo completo / equação (3)								
Coorte 1974 a 1976								
Anos	1994	1995	1996	1997	1998	1999	2000	2001
	18 a 20 anos	19 a 21 anos	20 a 22 anos	21 a 23 anos	22 a 24 anos	23 a 25 anos	24 a 26 anos	25 a 27 anos
Mulher	-0.110 (0.0202)***	-0.179 (0.0175)***	-0.201 (0.0179)***	-0.212 (0.0189)***	-0.240 (0.0178)***	-0.237 (0.0183)***	-0.266 (0.0188)***	-0.280 (0.0187)***
Anos	2002	2003	2004	2005	2006	2007	2008	2009
	26 a 28 anos	27 a 29 anos	28 a 30 anos	29 a 31 anos	30 a 32 anos	31 a 33 anos	32 a 34 anos	33 a 35 anos
Mulher	-0.272 (0.0193)***	-0.305 (0.0191)***	-0.316 (0.0198)***	-0.306 (0.0196)***	-0.294 (0.0201)***	-0.322 (0.0195)***	-0.340 (0.0197)***	-0.325 (0.0205)***
N. Observações (Continuamente empregados)	4,308	4,308	4,308	4,308	4,308	4,308	4,308	4,308
Coorte 1977 a 1979								
Anos	1997	1998	1999	2000	2001	2002	2003	2004
	18 a 20 anos	19 a 21 anos	20 a 22 anos	21 a 23 anos	22 a 24 anos	23 a 25 anos	24 a 26 anos	25 a 27 anos
Mulher	-0.101 (0.0119)***	-0.147 (0.0122)***	-0.172 (0.0128)***	-0.195 (0.0130)***	-0.203 (0.0133)***	-0.216 (0.0135)***	-0.249 (0.0133)***	-0.294 (0.0134)***
Anos	2005	2006	2007	2008	2009			
	26 a 28 anos	27 a 29 anos	28 a 30 anos	29 a 31 anos	30 a 32 anos			
Mulher	-0.304 (0.0137)***	-0.318 (0.0136)***	-0.314 (0.0138)***	-0.324 (0.0140)***	-0.327 (0.0143)***			
N. Observações (Continuamente empregados)	7,020	7,020	7,020	7,020	7,020	7,020	7,020	7,020
Coorte 1980 a 1982								
Anos	2000	2001	2002	2003	2004	2005	2006	2007
	18 a 20 anos	19 a 21 anos	20 a 22 anos	21 a 23 anos	22 a 24 anos	23 a 25 anos	24 a 26 anos	25 a 27 anos
Mulher	-0.0796 (0.00878)***	-0.126 (0.00888)***	-0.138 (0.00912)***	-0.180 (0.00905)***	-0.207 (0.00919)***	-0.232 (0.00941)***	-0.248 (0.00964)***	-0.274 (0.00971)***
Anos	2008	2009						
	26 a 28 anos	27 a 29 anos						
Mulher	-0.297 (0.0100)***	-0.283 (0.0105)***						
N. Observações (Continuamente empregados)	11,671	11,671	11,671	11,671	11,671	11,671	11,671	11,671

Fonte: Elaboração própria, dados da RAIS/MTE 1994 a 2009.

Erros padrão entre parênteses - *** $p < 0,01$; ** $p < 0,05$; * $p < 0,1$.

Tabela A10 – Efeito governo esquerda sobre impostos reais per capita

Variável dependente: Impostos médios (2005 a 2008) per capita em R\$ 2008								
	OLS	OLS	RD	RD	RD	RD	RD	RD
		controles	Linear	Quadrática	Cúbica	Linear	Quadrática	Cúbica
	(1)	(2)	(3)	(4)	(5)	(6)	(8)	(9)
Partido esquerda	-7.592	-9.815	0.565	-1.792	-3.894	-1.403	-6.436	-14.30
	(6.567)	(5.428)*	(9.244)	(8.403)	(9.258)	(8.546)	(9.871)	(13.56)
Polinomio em margem vitória:								
Linear			-63.08	-43.54	-24.12	-102.8	4.344	350.4
			(54.38)	(45.69)	(59.85)	(76.47)	(143.7)	(343.1)
Quadrático				180.7	170.6		325.2	2,752
				(142.3)	(130.6)		(485.0)	(2,163)
Cúbico					-188.9			3,941
					(399.1)			(3,018)
Interação polinomio margem vitória com dummy para partido esquerda:								
Linear						111.6	18.37	-351.6
						(85.76)	(173.7)	(423.5)
Quadrático							-371.9	-2,616
							(503.1)	(2,787)
Cúbico								-4,241
								(4,096)
Controle:								
População 2004	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% menos de 4 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% mais de 11 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
IDH 2000	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Escolaridade prefeito	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Constante	79.05	-173.6	-179.4	-176.7	-175.4	-183.0	-174.8	-170.1
	(5.388)***	(79.95)**	(80.56)**	(80.61)**	(80.26)**	(80.99)**	(80.33)**	(81.00)**
N. observações	964	964	964	964	964	964	964	964
R-quadrado	0.001	0.291	0.293	0.295	0.295	0.295	0.296	0.297

Fonte: Elaboração própria a partir de dados do TSE, FINBRA/STN, IPEADData e Munic/IBGE.

Erros-padrão em parênteses - *** p<0,01, ** p<0,05, * p<0,1.

Tabela A11 – Efeito governo esquerda sobre investimentos reais per capita

	Variável dependente: Investimentos médios (2005 a 2008) per capita em R\$ 2008							
	OLS	OLS	RD	RD	RD	RD	RD	RD
		controles	Linear	Quadrática	Cúbica	Linear	Quadrática	Cúbica
	(1)	(2)	(3)	(4)	(5)	(6)	(8)	(9)
Partido esquerda	-1.102	-3.731	6.137	4.230	4.832	2.197	-1.709	-10.77
	(7.347)	(6.760)	(9.484)	(9.633)	(10.73)	(9.675)	(11.28)	(14.51)
Polinomio em margem vitória:								
Linear			-59.97	-44.16	-49.72	-139.4	-330.3	34.30
			(43.25)	(44.57)	(64.95)	(59.12)**	(123.4)***	(298.7)
Quadrático				146.2	149.1		-580.4	1,977
				(130.7)	(135.1)		(320.3)*	(1,985)
Cúbico					54.13			4,151
					(461.5)			(3,013)
Interação polinomio margem vitória com dummy para partido esquerda:								
Linear						223.5	712.2	353.9
						(91.70)**	(176.7)***	(450.1)
Quadrático							-436.7	-3,030
							(434.1)	(2,931)
Cúbico								-4,097
								(4,533)
Controle:								
População 2004	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% menos de 4 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% mais de 11 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
IDH 2000	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Escolaridade prefeito	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Constante	153.4	32.75	27.22	29.35	28.95	20.05	-3.531	1.866
	(4.727)***	(120.4)	(120.8)	(120.9)	(121.1)	(120.5)	(120.3)	(120.5)
N. observações	964	964	964	964	964	964	964	964
R-quadrado	0.000	0.118	0.120	0.121	0.121	0.125	0.132	0.133

Fonte: Elaboração própria a partir de dados do TSE, FINBRA/STN, IPEADa e Munic/IBGE.

Erros-padrão em parênteses - *** p<0,01, ** p<0,05, * p<0,1.

Tabela A12 – Efeito governo esquerda sobre superávit per capita

	Variável dependente: Superávit médio (2005 a 2008) per capita em R\$ 2008							
	OLS	OLS	RD	RD	RD	RD	RD	RD
	(1)	controles (2)	Linear (3)	Quadrática (4)	Cúbica (5)	Linear (6)	Quadrática (8)	Cúbica (9)
Partido esquerda	3.246	0.661	4.129	4.364	2.203	4.758	2.364	3.973
	(5.219)	(4.910)	(6.545)	(6.568)	(7.757)	(6.563)	(8.325)	(10.09)
Polinomio em margem vitória:								
Linear			-21.08 (27.86)	-23.03 (27.33)	-3.058 (47.22)	-8.383 (37.10)	60.54 (94.96)	8.328 (187.3)
Quadrático				-18.04 (56.45)	-28.37 (58.91)		209.3 (216.2)	-156.7 (1,221)
Cúbico					-194.2 (242.0)			-594.3 (1,932)
Interação polinomio margem vitória com dummy para partido esquerda:								
Linear						-35.71 (50.91)	-116.9 (127.0)	-78.09 (283.6)
Quadrático							-167.2 (295.0)	295.3 (1,812)
Cúbico								439.1 (2,731)
Controle:								
População 2004	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% menos de 4 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% mais de 11 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
IDH 2000	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Escolaridade prefeito	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Constante	38.73 (3.640)***	112.6 (84.65)	110.7 (84.42)	110.4 (84.59)	111.8 (84.48)	111.8 (84.26)	117.7 (84.39)	116.7 (84.21)
N. observações	964	964	964	964	964	964	964	964
R-quadrado	0.000	0.141	0.141	0.141	0.142	0.142	0.142	0.142

Fonte: Elaboração própria a partir de dados do TSE, FINBRA/STN, IPEAData e Munic/IBGE.

Erros-padrão em parênteses - *** p<0,01, ** p<0,05, * p<0,1.

Tabela A13 – Efeito governo esquerda sobre imposto sobre serviços/receitas tributárias

	Variável dependente: Imposto sobre serviços / receitas tributárias							
	OLS	OLS	RD	RD	RD	RD	RD	RD
		controles	Linear	Quadrática	Cúbica	Linear	Quadrática	Cúbica
	(1)	(2)	(3)	(4)	(5)	(6)	(8)	(9)
Partido esquerda	0.00212	0.00166	0.00429	0.00233	-0.00297	0.00202	-0.00765	-0.00773
	(0.0106)	(0.0102)	(0.0139)	(0.0142)	(0.0163)	(0.0141)	(0.0178)	(0.0232)
Polinomio em margem vitória:								
Linear			-0.0159	0.000270	0.0492	-0.0618	0.00658	0.475
			(0.0599)	(0.0634)	(0.0979)	(0.0760)	(0.189)	(0.440)
Quadrático				0.150	0.125		0.207	3.494
				(0.157)	(0.150)		(0.445)	(2.792)
Cúbico					-0.476			5.337
					(0.508)			(4.213)
Interação polinomio margem vitória com dummy para partido esquerda:								
Linear						0.129	0.231	-0.694
						(0.125)	(0.283)	(0.676)
Quadrático							-0.789	-0.731
							(0.658)	(4.340)
Cúbico								-10.74
								(6.531)
Controle:								
População 2004	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% menos de 4 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% mais de 11 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
IDH 2000	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Escolaridade prefeito	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Constante	0.355	0.315	0.314	0.316	0.319	0.310	0.310	0.311
	(0.00727)***	(0.189)*	(0.189)*	(0.189)*	(0.190)*	(0.189)	(0.190)	(0.191)
N. observações	964	964	964	964	964	964	964	964
R-quadrado	0.000	0.082	0.082	0.082	0.083	0.083	0.083	0.086

Fonte: Elaboração própria a partir de dados do TSE, FINBRA/STN, IPEAData e Munic/IBGE.

Erros-padrão em parênteses - *** p<0,01, ** p<0,05, * p<0,1.

Tabela A14 – Efeito governo esquerda despesas com transporte/despesas orçamentárias

	Variável dependente: despesas com transporte / despesas orçamentárias							
	OLS	OLS	RD	RD	RD	RD	RD	RD
		controles	Linear	Quadrática	Cúbica	Linear	Quadrática	Cúbica
	(1)	(2)	(3)	(4)	(5)	(6)	(8)	(9)
Partido esquerda	-0.00411	-0.00520	-0.00665	-0.00583	-0.00372	-0.00586	-0.00316	-0.00725
	(0.00311)	(0.00271)*	(0.00392)*	(0.00399)	(0.00462)	(0.00396)	(0.00513)	(0.00665)
Polinomio em margem vitória:								
Linear			0.00879	0.00201	-0.0175	0.0247	-0.00960	0.112
			(0.0155)	(0.0164)	(0.0255)	(0.0193)	(0.0507)	(0.119)
Quadrático				-0.0627	-0.0526		-0.104	0.749
				(0.0393)	(0.0384)		(0.114)	(0.735)
Cúbico					0.190			1.385
					(0.141)			(1.104)
Interação polinomio margem vitória com dummy para partido esquerda:								
Linear						-0.0447	-0.0420	-0.119
						(0.0311)	(0.0770)	(0.177)
Quadrático							0.212	-0.964
							(0.194)	(1.092)
Cúbico								-0.866
								(1.674)
Controle:								
População 2004	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% menos de 4 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% mais de 11 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
IDH 2000	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Escolaridade prefeito	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Constante	0.0473	-0.0526	-0.0518	-0.0527	-0.0541	-0.0504	-0.0523	-0.0499
	(0.00215)***	(0.0474)	(0.0474)	(0.0474)	(0.0475)	(0.0473)	(0.0476)	(0.0477)
N. observações	964	964	964	964	964	964	964	964
R-quadrado	0.002	0.247	0.247	0.248	0.249	0.248	0.249	0.250

Fonte: Elaboração própria a partir de dados do TSE, FINBRA/STN, IPEAData e Munic/IBGE.

Erros-padrão em parênteses - *** p<0,01, ** p<0,05, * p<0,1.

Tabela A15 – Efeito governo esquerda despesas com assistência social/despesas orçamentárias

	Variável dependente: Despesas com assistência social / despesas orçamentárias							
	OLS	OLS	RD	RD	RD	RD	RD	RD
	(1)	controles (2)	Linear (3)	Quadrática (4)	Cúbica (5)	Linear (6)	Quadrática (8)	Cúbica (9)
Partido esquerda	-0.00148	-0.000889	-0.00235	-0.00217	-0.000855	-0.00228	-0.000242	0.000585
	(0.00132)	(0.00127)	(0.00183)	(0.00190)	(0.00212)	(0.00186)	(0.00239)	(0.00309)
Polinomio em margem vitória:								
Linear			0.00888	0.00736	-0.00476	0.0103	-0.0237	-0.0455
			(0.00778)	(0.00876)	(0.0124)	(0.00942)	(0.0257)	(0.0566)
Quadrático				-0.0140	-0.00778		-0.103	-0.256
				(0.0246)	(0.0263)		(0.0629)	(0.341)
Cúbico					0.118			-0.248
					(0.0954)			(0.519)
Interação polinomio margem vitória com dummy para partido esquerda:								
Linear						-0.00413	0.0146	0.0245
						(0.0165)	(0.0409)	(0.0858)
Quadrático							0.155	0.394
							(0.127)	(0.541)
Cúbico								0.109
								(0.881)
Controle:								
População 2004	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% menos de 4 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% mais de 11 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
IDH 2000	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Escolaridade prefeito	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Constante	0.0396	0.0166	0.0174	0.0172	0.0163	0.0175	0.0152	0.0147
	(0.000950)***	(0.0266)	(0.0266)	(0.0266)	(0.0266)	(0.0267)	(0.0268)	(0.0269)
N. observações	964	964	964	964	964	964	964	964
R-quadrado	0.001	0.061	0.062	0.063	0.064	0.062	0.065	0.065

Fonte: Elaboração própria a partir de dados do TSE, FINBRA/STN, IPEAData e Munic/IBGE.

Erros-padrão em parênteses - *** p<0,01, ** p<0,05, * p<0,1.

Tabela A16 – Efeito governo esquerda despesas diretas com pessoal /despesas orçamentárias

Variável dependente: Despesas com aplicações diretas de recursos com pessoal / despesas orçamentárias								
	OLS	OLS	RD	RD	RD	RD	RD	RD
	controles		Linear	Quadrática	Cúbica	Linear	Quadrática	Cúbica
	(1)	(2)	(3)	(4)	(5)	(6)	(8)	(9)
Partido esquerda	0.00670	0.00788	0.00535	0.00510	0.00379	0.00565	0.00479	0.00815
	(0.00416)	(0.00408)*	(0.00601)	(0.00619)	(0.00679)	(0.00618)	(0.00766)	(0.00982)
Polinomio em margem vitória:								
Linear			0.0154	0.0174	0.0295	0.0214	0.0964	0.0421
			(0.0264)	(0.0285)	(0.0389)	(0.0325)	(0.0773)	(0.179)
Quadrático				0.0194	0.0131		0.228	-0.153
				(0.0868)	(0.0938)		(0.218)	(1.146)
Cúbico					-0.118			-0.617
					(0.326)			(1.792)
Interação polinomio margem vitória com dummy para partido esquerda:								
Linear						-0.0171	-0.149	-0.176
						(0.0564)	(0.127)	(0.279)
Quadrático							-0.0333	0.942
							(0.408)	(1.805)
Cúbico								-0.342
								(2.939)
Controle:								
População 2004	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% menos de 4 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% mais de 11 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
IDH 2000	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Escolaridade prefeito	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Constante	0.433	0.592	0.594	0.594	0.595	0.594	0.602	0.600
	(0.00274)***	(0.0748)***	(0.0746)***	(0.0746)***	(0.0745)***	(0.0747)***	(0.0744)***	(0.0748)***
N. observações	964	964	964	964	964	964	964	964
R-quadrado	0.003	0.043	0.043	0.043	0.043	0.043	0.045	0.045

Fonte: Elaboração própria a partir de dados do TSE, FINBRA/STN, IPEADData e Munic/IBGE.

Erros-padrão em parênteses - *** p<0,01, ** p<0,05, * p<0,1.

Tabela A17 – Efeito governo esquerda sobre número de funcionários estatutários per capita

Variável dependente: Número de funcionários públicos estatutários per capita no município em 2005								
	OLS	OLS	RD	RD	RD	RD	RD	RD
	controles		Linear	Quadrática	Cúbica	Linear	Quadrática	Cúbica
	(1)	(2)	(3)	(4)	(5)	(6)	(8)	(9)
Partido esquerda	-0.00198	-0.00163	-0.000664	-0.000753	-0.000638	-0.000712	-1.13e-05	0.00290
	(0.00102)*	(0.000976)*	(0.00144)	(0.00143)	(0.00170)	(0.00144)	(0.00189)	(0.00252)
Polinomio em margem vitória:								
Linear			-0.00586	-0.00511	-0.00617	-0.00683	-0.00843	-0.0993
			(0.00612)	(0.00606)	(0.00985)	(0.00798)	(0.0205)	(0.0445)**
Quadrático				0.00689	0.00744		-0.00481	-0.641
				(0.0154)	(0.0150)		(0.0540)	(0.276)**
Cúbico					0.0103			-1.033
					(0.0580)			(0.430)**
Interação polinomio margem vitória com dummy para partido esquerda:								
Linear						0.00273	-0.0115	0.0517
						(0.0120)	(0.0294)	(0.0675)
Quadrático							0.0589	0.895
							(0.0747)	(0.418)**
Cúbico								0.712
								(0.634)
Controle:								
População 2004	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% menos de 4 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% mais de 11 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
IDH 2000	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Escolaridade prefeito	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Constante	0.0251	0.0248	0.0243	0.0244	0.0243	0.0242	0.0245	0.0227
	(0.000758)***	(0.0174)	(0.0173)	(0.0174)	(0.0173)	(0.0173)	(0.0174)	(0.0174)
N. observações	963	963	963	963	963	963	963	963
R-quadrado	0.004	0.102	0.103	0.103	0.103	0.103	0.104	0.109

Fonte: Elaboração própria a partir de dados do TSE, FINBRA/STN, IPEADData e Munic/IBGE.

Erros-padrão em parênteses - *** p<0,01, ** p<0,05, * p<0,1.

Tabela A18 – Efeito governo esquerda sobre funcionários temporários/total funcionários

	Variável dependente: Proporção de funcionários comissionados e sem vínculo permanente / total funcionários							
	OLS	OLS	RD	RD	RD	RD	RD	RD
	(1)	controles (2)	Linear (3)	Quadrática (4)	Cúbica (5)	Linear (6)	Quadrática (8)	Cúbica (9)
Partido esquerda	0.0155	0.0176	-0.00191	-0.00177	-0.0233	-0.000948	-0.0326	-0.0333
	(0.0122)	(0.0117)	(0.0165)	(0.0172)	(0.0188)	(0.0169)	(0.0205)	(0.0264)
Polinomio em margem vitória:								
Linear			0.119	0.117	0.317	0.138	0.549	0.769
			(0.0749)	(0.0848)	(0.112)***	(0.0876)	(0.203)***	(0.503)
Quadrático				-0.0108	-0.114		1.245	2.788
				(0.246)	(0.209)		(0.511)**	(3.246)
Cúbico					-1.936			2.504
					(0.642)***			(4.966)
Interação polinomio margem vitória com dummy para partido esquerda:								
Linear						-0.0550	-0.104	-0.511
						(0.166)	(0.337)	(0.792)
Quadrático							-2.477	-2.652
							(0.903)***	(5.267)
Cúbico								-4.714
								(8.284)
Controle:								
População 2004	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% menos de 4 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
% mais de 11 anos de estudo	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
IDH 2000	Não	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Escolaridade prefeito	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Constante	0.247	0.393	0.405	0.404	0.419	0.406	0.430	0.431
	(0.00816)***	(0.224)*	(0.224)*	(0.224)*	(0.224)*	(0.224)*	(0.225)*	(0.226)*
N. observações	963	963	963	963	963	963	963	963
R-quadrado	0.002	0.095	0.097	0.097	0.103	0.097	0.103	0.103

Fonte: Elaboração própria a partir de dados do TSE, FINBRA/STN, IPEADData e Munic/IBGE.

Erros-padrão em parênteses - *** p<0,01, ** p<0,05, * p<0,1.

Figura A1 – Distribuição das notas na parte objetiva da prova de promoção para os professores de matemática e língua portuguesa

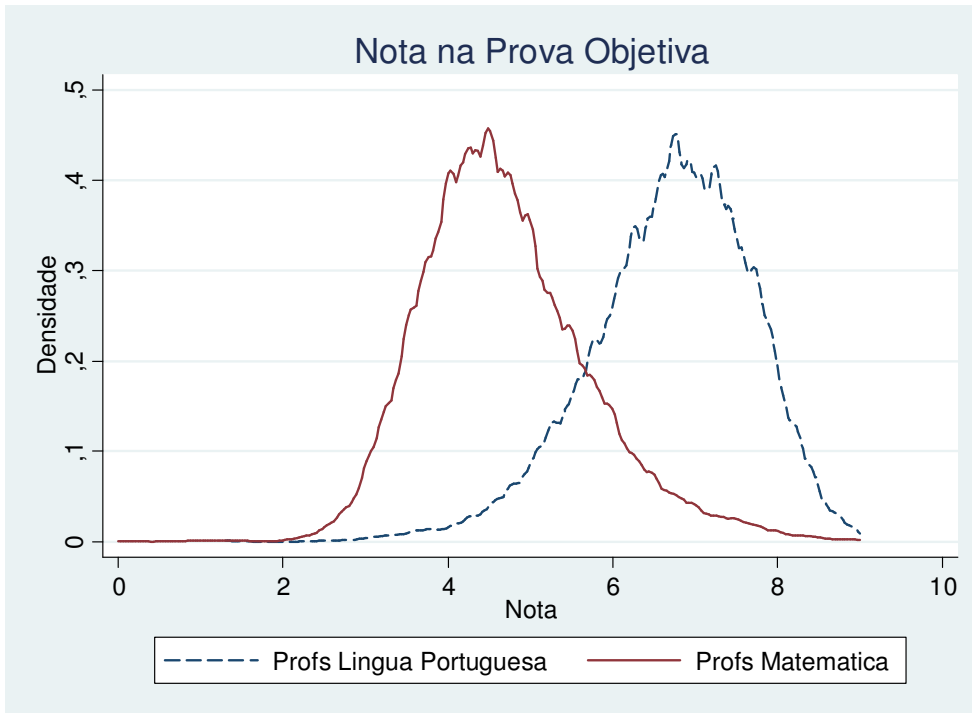


Figura A2 – Relação entre a nota média em matemática e a nota na parte objetiva da prova de promoção dos professores

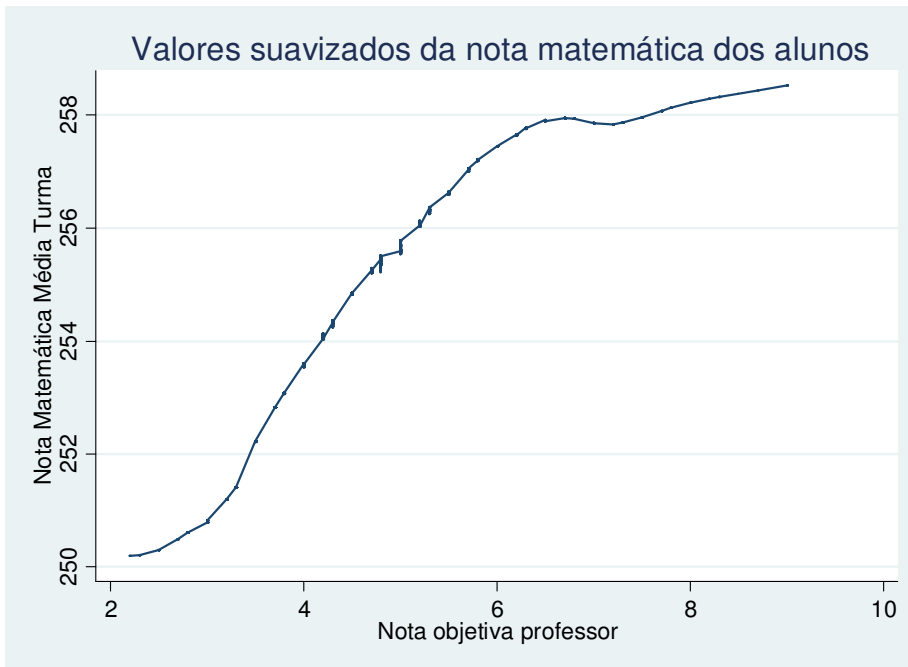


Gráfico A2.1 – Valores suavizados da regressão local da média por turma das notas em matemática no SARESP 2009 e a nota do respectivo professor.

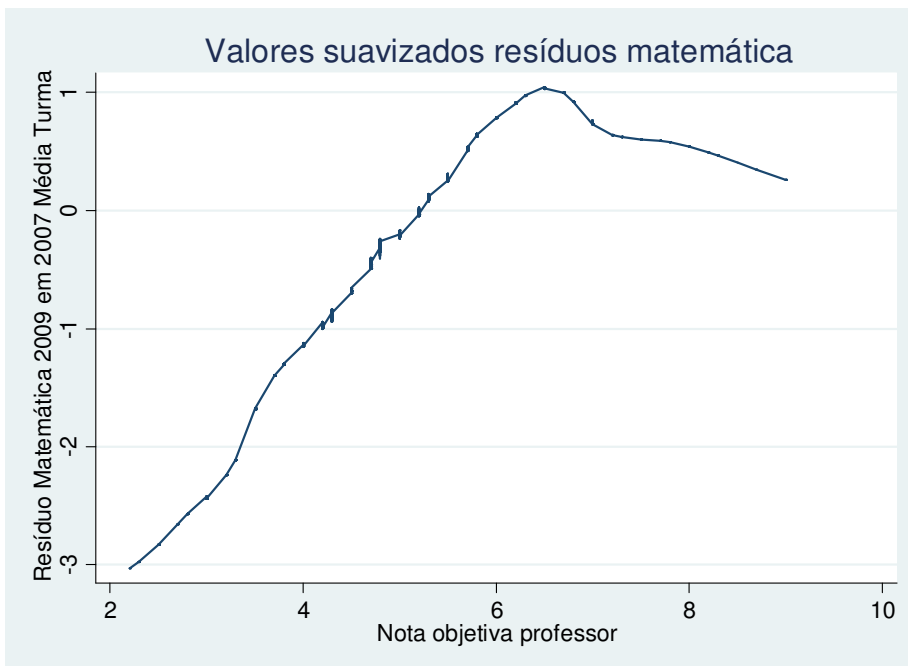


Gráfico A2.2 – Valores suavizados da regressão local da média por turma dos resíduos das notas em matemática no SARESP 2009 e a nota do respectivo professor.

Figura A3 – Relação entre a nota média em língua portuguesa e a nota na parte objetiva da prova de promoção dos professores

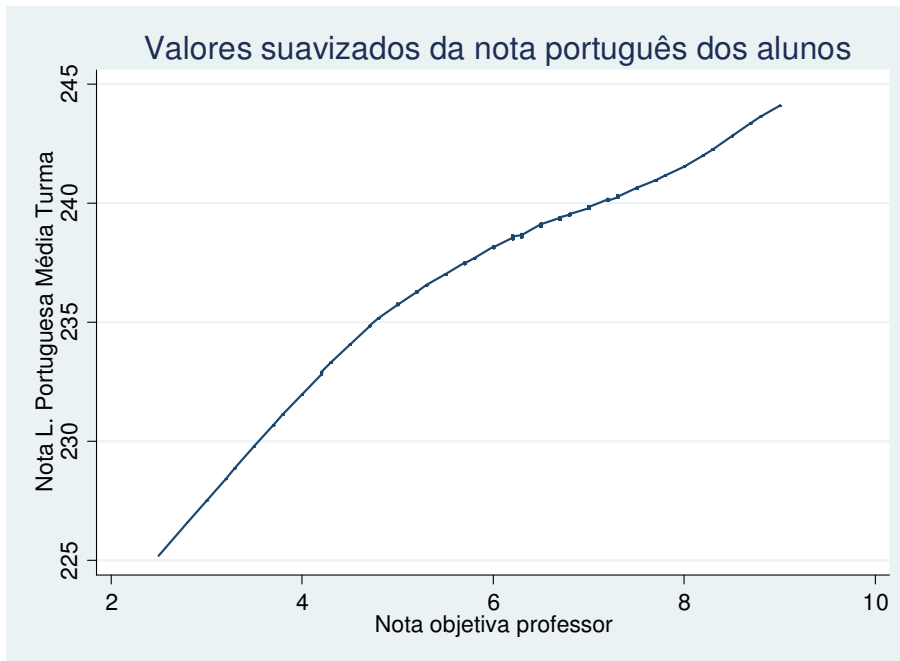


Gráfico A3.1 – Valores suavizados da regressão local da média por turma das notas em língua portuguesa no SARESP 2009 e a nota do respectivo professor.

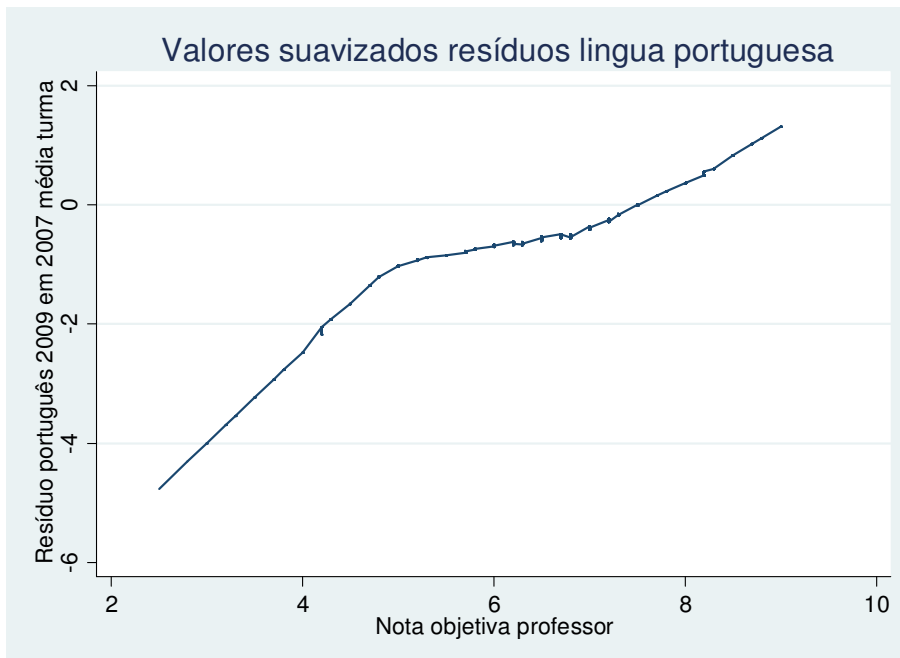
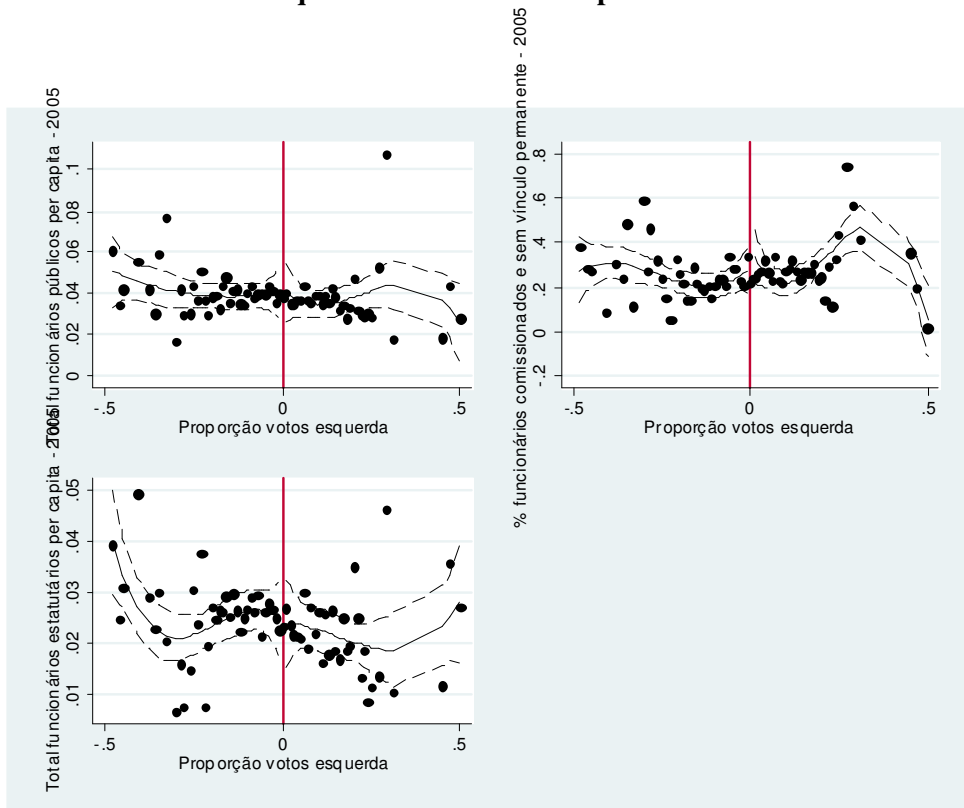


Gráfico A3.2 – Valores suavizados da regressão local da média por turma dos resíduos das notas em língua portuguesa no SARESP 2009 e a nota do respectivo professor.

Figura A4 – Efeito de partido de esquerda sobre o total e composição do quadro de funcionários públicos



Fonte: Elaboração própria a partir de dados do TSE, FINBRA/STN, IPEADData e Munic/IBGE.

Anexos:

Anexo 1 – Perguntas referentes aos professores matemática do questionário dos alunos do SARESP 2009.

SARESP 2009

Questionário do Aluno

6ª e 8ª Séries / 7º e 9º Anos Ensino Fundamental



40. Agora vamos falar das aulas de Matemática na sua escola. (Marque <u>apenas</u> uma resposta para cada situação) O seu professor de Matemática:		Em todas ou quase todas as aulas	Em algumas aulas	Nunca
40.1.	costuma aguardar muito tempo até que os alunos façam silêncio para iniciar a aula.	(A)	(B)	(C)
40.2.	é exigente com relação ao trabalho dos alunos.	(A)	(B)	(C)
40.3.	incentiva os alunos a melhorarem o seu desempenho.	(A)	(B)	(C)
40.4.	é atencioso e auxilia os alunos a realizarem suas tarefas.	(A)	(B)	(C)
40.5.	explica a matéria até que todos os alunos entendam.	(A)	(B)	(C)
40.6.	passa lições de casa.	(A)	(B)	(C)
40.7.	corrige as lições de casa dos alunos.	(A)	(B)	(C)
40.8.	mostra interesse no aprendizado de todos os alunos.	(A)	(B)	(C)
40.9.	dá aos alunos oportunidade de participarem e expressarem suas opiniões.	(A)	(B)	(C)
40.11.	propõe atividades de resolução de problemas variados.	(A)	(B)	(C)
40.12.	relaciona os conteúdos de Matemática às situações do cotidiano.	(A)	(B)	(C)
40.13.	indica pesquisa em livros.	(A)	(B)	(C)
40.14.	propõe trabalhos para serem realizados em grupos.	(A)	(B)	(C)
40.15.	utiliza jogos e brincadeiras nas aulas.	(A)	(B)	(C)
40.16.	costuma mostrar a aplicação dos conteúdos estudados na Matemática em outras disciplinas, como Geografia, Artes, Ciências e outras.	(A)	(B)	(C)