1713
TEXTO PARA DISCUSSÃO

ipea

DIFERENCIAL DE SALÁRIOS ENTRE OS SETORES PÚBLICO E PRIVADO NO BRASIL: UM MODELO DE ESCOLHA ENDÓGENA

Ana Luiza Neves de Holanda Barbosa
Fernando de Holanda Barbosa Filho

ipea 47 anos
Por um Brasil desenvolvido

1713

TEXTO PARA DISCUSSÃO

Rio de Janeiro, fevereiro de 2012

DIFERENCIAL DE SALÁRIOS ENTRE OS SETORES PÚBLICO E PRIVADO NO BRASIL: UM MODELO DE ESCOLHA ENDÓGENA*

Ana Luiza Neves de Holanda Barbosa**
Fernando de Holanda Barbosa Filho ***

* Os autores agradecem os comentários de diversos participantes de seminários na Escola de Pós-Graduação em Economia (EPGE) da Fundação Getulio Vargas (FGV), na Universidade Federal Fluminense (UFF), no Programa de Pós-Graduação em Economia da Universidade Federal de Pernambuco (Pimes) e no Instituto de Pesquisa Econômica Aplicada (Ipea). Eventuais erros remanescentes são de nossa exclusiva responsabilidade.
** Técnica de Pesquisa e Planejamento da Diretoria de Estudos e Políticas Sociais (Disoc) do Ipea.
*** Pesquisador do Instituto Brasileiro de Economia (Ibre) da FGV.

Governo Federal

Secretaria de Assuntos Estratégicos da Presidência da República
Ministro Wellington Moreira Franco

ipea **Instituto de Pesquisa Econômica Aplicada**

Fundação pública vinculada à Secretaria de Assuntos Estratégicos da Presidência da República, o Ipea fornece suporte técnico e institucional às ações governamentais – possibilitando a formulação de inúmeras políticas públicas e programas de desenvolvimento brasileiro – e disponibiliza, para a sociedade, pesquisas e estudos realizados por seus técnicos.

Presidente
Marcio Pochmann

Diretor de Desenvolvimento Institucional
Geová Parente Farias

Diretor de Estudos e Relações Econômicas e Políticas Internacionais, Substituto
Marcos Antonio Macedo Cintra

Diretor de Estudos e Políticas do Estado, das Instituições e da Democracia
Alexandre de Ávila Gomide

Diretora de Estudos e Políticas Macroeconômicas
Vanessa Petrelli Corrêa

Diretor de Estudos e Políticas Regionais, Urbanas e Ambientais
Francisco de Assis Costa

Diretor de Estudos e Políticas Setoriais de Inovação, Regulação e Infraestrutura
Carlos Eduardo Fernandez da Silveira

Diretor de Estudos e Políticas Sociais
Jorge Abrahão de Castro

Chefe de Gabinete
Fabio de Sá e Silva

Assessor-chefe de Imprensa e Comunicação
Daniel Castro

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

Texto para Discussão

Publicação cujo objetivo é divulgar resultados de estudos direta ou indiretamente desenvolvidos pelo Ipea, os quais, por sua relevância, levam informações para profissionais especializados e estabelecem um espaço para sugestões.

As opiniões emitidas nesta publicação são de exclusiva e inteira responsabilidade do(s) autor(es), não exprimindo, necessariamente, o ponto de vista do Instituto de Pesquisa Econômica Aplicada ou da Secretaria de Assuntos Estratégicos da Presidência da República.

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

ISSN 1415-4765

JEL: J31, J45, C24

SUMÁRIO

SINOPSE

ABSTRACT

1 INTRODUÇÃO 7

2 MODELO ECONOMÉTRICO 11

3 BASE DE DADOS 15

4 RESULTADOS 23

5 CONCLUSÕES 31

REFERÊNCIAS 32

APÊNDICE 35

SINOPSE

Este artigo investiga o diferencial de salários entre os trabalhadores dos setores público e privado no Brasil. A análise tem como base os microdados da Pesquisa Nacional por Amostra de Domicílios (PNAD), de 2009, conduzida pelo Instituto Brasileiro de Geografia e Estatística (IBGE). A principal contribuição deste estudo é a estimação de um modelo de regressão com mudança endógena (*endogenous switching regression model*), que corrige o viés de seleção no processo de escolha setorial realizada pelos trabalhadores e permite a identificação de fatores determinantes na entrada do trabalhador no mercado de trabalho do setor público. O diferencial salarial público-privado ainda é calculado por gênero, uma vez que a oferta de trabalho é bastante diferenciada entre homens e mulheres no Brasil. Os resultados deste trabalho demonstram que os salários do setor público são mais altos do que aqueles do setor privado no Brasil. Em particular, o diferencial público-privado entre mulheres é maior do que entre homens, quando o modelo não é controlado por características não observáveis dos trabalhadores.

ABSTRACT[i]

This study investigates the wage gap between the public and private sectors in Brazil. The analysis is carried out with 2009 microdata from the Pesquisa Nacional por Amostra de Domicílios (PNAD) and it takes account its complex sample design. The main contribution of this study is the estimation of a switching endogenous regression model that corrects the selection bias in the choice of employment sector. This model allows for the identification of some factors that determine the entrance of the individual in the public sector labor market. The public-private wage gap is calculated by gender as labor supply varies significantly between women and men. The results show that public sector wages are higher than those in the private sector. In particular, the public-private wage gap for women is higher than for men.

---

i. *The versions in English of the abstracts of this series have not been edited by Ipea's editorial department.*
As versões em língua inglesa das sinopses (*abstracts*) desta coleção não são objeto de revisão pelo Editorial do Ipea.

1 INTRODUÇÃO

A literatura sobre determinação de salários apresenta diversas razões teóricas para explicar a existência de hiatos salariais entre diferentes grupos de trabalhadores. Falta também consenso a respeito de qual seria o melhor arcabouço para explicar a diferença de rendimentos entre os trabalhadores do setor público e do setor privado.[1] A comparação de salários entre estes grupos de trabalhadores torna-se mais complexa na medida em que os mercados de trabalho de ambos os setores são regidos por regras e fatores institucionais distintos (VAZ; HOFFMAN, 2007). Neste contexto, uma possível explicação para a discrepância salarial entre os setores público e privado trata das diferentes restrições defrontadas por cada setor. Enquanto o setor público está sujeito a restrições políticas, o setor privado é caracterizado, em geral, por restrições associadas ao problema de maximização de lucro.

A análise do diferencial de salários entre os setores público e privado é também importante sob o ponto de vista da eficiência e da equidade econômica. Considerações de equidade reforçam a questão da igualdade de bem-estar entre os trabalhadores de ambos os setores, enquanto aspectos de eficiência (minimização de custo) reforçam o fato de que um empregador (o governo, por exemplo) não deveria pagar salários maiores do que o necessário para atrair trabalhadores qualificados (HARTOG; OOSTERBEEK, 1993, p. 99).

Heitmueller (2004) argumenta que teoricamente não está claro por que empregados do setor público devem receber maiores salários do que aqueles do setor privado, em especial, quando todos os benefícios usufruídos pelos primeiros são considerados. O setor público na maior parte dos países oferece uma grande variedade de benefícios para seus empregados. Este também é o caso do Brasil. Uma característica bastante conhecida do setor público brasileiro é a de que seus empregados usufruem tanto da estabilidade no emprego quanto de aposentadoria integral.[2] Benefícios esses que seriam um tipo de seguro que os empregados pagariam ao entrar no mercado de trabalho do setor público. Portanto, a existência de tais benefícios implicaria uma

1. Para uma revisão abrangente sobre a literatura de hiatos salariais entre trabalhadores do setor público e do privado, a evidência empírica em diversos países e as principais características dos mercados de trabalho do setor público em geral, ver Ehrenberg e Schwartz (1986), Bender (1998) e Gregory e Borland (1999).
2. A Constituição brasileira garante o direito à estabilidade ao servidor que completar três anos de exercício efetivo em cargo obtido por concurso público.

menor compensação no setor público em comparação com o setor privado (GREGORY, 1990). Nesta mesma linha, Bender e Fernandes (2006) discorrem sobre a discussão das diferenças salariais entre estes setores no debate da reforma da previdência do funcionalismo público no Brasil.[3]

Ao contrário desta intuição econômica, todos os estudos empíricos referentes ao caso brasileiro demonstram a existência de um prêmio salarial positivo para os trabalhadores do setor público.[4] Ou seja, além da estabilidade no emprego e da aposentadoria integral, trabalhadores no setor público brasileiro recebem, na média, maiores salários do que os trabalhadores do setor privado. Tais evidências apontam para a necessidade de uma análise mais detalhada para o caso do Brasil, onde despesas com pessoal do governo consolidado representaram quase 13% do Produto Interno Bruto (PIB) em 2007.[5] Um estudo pormenorizado sobre os padrões de salários dos setores público e privado no Brasil tem relevância no debate sobre a eficiência do setor público e o papel do Estado na economia brasileira, além de ser fundamental para o conhecimento das características do mercado de trabalho do país.

A maior parte dos estudos aplicados ao caso brasileiro baseia-se na estimação de modelos de regressão com base em equações mincerianas de salários (MINCER, 1974).[6] As diferenças entre as características dos trabalhadores do setor público e as do

3. Os autores refutam o argumento de que "a manutenção da integralidade e paridade do salário do servidor público inativo com o da ativa é um justo direito porque, pressupostamente, ao longo da carreira (ativa) do servidor público, o mesmo (em média) recebe menos do que seu "igual" no setor privado." Neste caso, a integralidade e paridade seriam compensações pelo recebimento de salário menor durante a vida ativa do servidor público. Como bem documentado em Bender e Fernandes (2006) e Braga (2008), esse argumento não procede, pois o servidor público ativo em média recebe sistematicamente (e crescentemente) mais do que seu "igual" no setor privado.

4. Uma resenha detalhada da literatura empírica sobre o tema no Brasil pode ser encontrada em Barbosa (2009). Entre alguns dos principais estudos nesta literatura, ver: Foguel *et al.* (2000), Belluzo, Pazello e Anuatti-Neto. (2005), Bender e Fernandes (2006), Vaz e Hoffman (2007), Braga (2008).

5. Consolidação das Contas Públicas [Secretaria do Tesouro Nacional (STN, 2007)]. Vale salientar que esta estatística está subestimada já que os dados disponibilizados pela STN não contabilizam as despesas com pessoal dos trabalhadores de empresas estatais.

6. Mincer (1974) desenvolve o modelo teórico que justifica a adoção da forma funcional semilogarítimica de rendimentos. Dois importantes conceitos econômicos estão implícitos neste modelo de rendimentos: *i*) uma equação de preços ou uma função hedônica de rendimentos que revela como o mercado de trabalho recompensa o investimento em capital humano do trabalhador, seja por educação seja por experiência no trabalho; *ii*) a taxa de retorno à educação (HECKMAN; LOCHNER; TODD, 2003). Esta última pode ser comparada com a taxa de juros que determina a otimalidade de investimentos em capital humano, ou seja, agentes decidem investir em capital humano tendo em vista o custo e o benefício de cada ano a mais de escolaridade. Este arcabouço permeia a teoria do capital humano (de investimento ou de acumulação de capital humano) de Becker (1962, 1975).

setor privado são explicitamente levadas em conta quando duas regressões separadas são estimadas, uma para cada setor.[7] O universo de análise nesta comparação entre os salários observados (ou previstos) destes setores abrange apenas os trabalhadores que já estão empregados em um dos setores e não se considera o fato de que os trabalhadores escolheram (ou foram escolhidos) para trabalhar em cada um dos mesmos. Há, portanto, um viés de autosseleção na medida em que os trabalhadores se alocam entre os setores de forma não aleatória. Assim, as diferenças observadas entre os momentos das distribuições de salários não têm uma interpretação causal imediata. Diferenciais de salários com base em estimações que não contemplem a correção do viés de autosseleção dos trabalhadores geram, portanto, estimativas viesadas e inconsistentes. De fato, uma comparação mais apropriada de diferenciais salariais é aquela em que se utilizam os salários potenciais (ou *wage offers*) nos dois setores corrigidos pelo viés de autosseleção.

Estudos empíricos sobre a relação entre o viés de seleção e o diferencial salarial público e privado são escassos no Brasil.[8] Com exceção de Barbosa e Barbosa Filho (2008) e de Ponczek, Botelho e Daulins (2008), nenhum estudo publicado investigou o diferencial de salários entre trabalhadores do setor público e do setor privado no Brasil levando-se em conta o fato de que a seleção dos trabalhadores entre estes setores não ocorre de forma aleatória.[9] O objetivo deste estudo é preencher essa lacuna ao analisar o viés de seleção no cálculo do hiato salarial público-privado no Brasil. Em particular, ao especificar o processo de alocação entre setores dos trabalhadores, está se levando em conta suas preferências, associadas a características demográficas e produtivas que possam influenciar em sua escolha na entrada no mercado de trabalho do setor público. Importante ressaltar que o diferencial público-privado é calculado por gênero, na medida em que a participação no mercado de trabalho é bastante diferenciada

7. É importante notar, contudo, que a estimação de duas equações de salários é equivalente à estimação de uma única equação, em que cada regressor interage com a variável *dummy* do setor. Neste caso, o coeficiente da variável dummy para o setor público reflete somente o hiato salarial não explicado pelas características individuais. As interações das variáveis de emprego no setor público com todas as covariadas também apresentam problemas na medida em que, neste caso, a análise do hiato salarial global (ou marginal) depende das distribuições covariadas. Além disso, nestes modelos, os hiatos salariais são os mesmos para todos os indivíduos com as mesmas características observáveis. Esta última limitação é resolvida com o emprego do método de regressões quantílicas (BELLUZO; PAZELLO; ANUATTI-NETO, 2005).

8. No que tange ao diferencial de salários entre os trabalhadores do setor público e privado, são inúmeros os estudos da literatura internacional realizados que tratam do problema de autosseleção. De forma geral, estes estudos indicam que há uma diferença significativa entre os modelos com e sem a correção do viés de autosseleção (BARBOSA, 2009).

9. Com base de dados em uma estrutura em painel da Pesquisa Mensal de Emprego (PME), Ponczek, Botelho e Daulins (2008) levam em conta o viés potencial causado pelas características não observáveis (e invariantes no tempo) dos trabalhadores no cálculo do diferencial de salários público-privado. Os autores sugerem que o prêmio salarial favorável ao setor público tem magnitudes muito menores (em torno de 5%) do que as encontradas em estudos anteriores.

entre homens e mulheres. A análise desagregada dos salários entre homens e mulheres também busca pontuar questões associadas à discriminação por gênero no mercado de trabalho no Brasil.[10]

A estimação do diferencial de salários controlado pelo viés de autosseleção tem como base um modelo de regressão com mudança endógena (*endogenous switching regression model*), arcabouço que nos permite testar empiricamente como um indivíduo se autosseleciona entre esses dois setores de acordo com suas vantagens comparativas. A ideia de que indivíduos se autosselecionam no setor (ou ocupação) em que eles têm maior produtividade data do modelo clássico de Roy (1951).[11] Uma importante vantagem do modelo adotado neste trabalho é o tratamento endógeno da escolha de setor e a possibilidade de se avaliar o efeito simultâneo entre a seleção de setor e os salários dos trabalhadores.

A base de dados deste trabalho provém da PNAD, conduzida pelo IBGE. A PNAD é uma pesquisa anual e a utilizada aqui é a de 2009. Um procedimento estatístico muito relevante em análises empíricas, associado com o desenho amostral complexo da PNAD é adotado nas estimações do modelo.

Os resultados deste estudo demonstram que os salários do setor público são mais altos do que aqueles do setor privado no Brasil. A educação (em especial, o nível superior de ensino) é um fator que apresenta efeitos significativos na probabilidade de entrada de homens e mulheres no mercado de trabalho do setor público. Por outro lado, trabalhadores residentes nas regiões Sudeste e Sul têm menor probabilidade de entrar no setor público. Um ponto importante a ser destacado é que, com exceção das mulheres, a aplicação de um modelo endógeno de escolha de se trabalhar no setor público, ou seja, a introdução do termo de correção do viés de seleção nas estimações das equações de rendimentos subestima, ainda que de forma pouco significativa, o diferencial salarial público-privado não controlado pelo viés de seleção do trabalhador.

10. Estudos empíricos sugerem que discriminação salarial por gênero é alta no Brasil. Com base na Pesquisa Nacional por Amostra de Domicílios (PNAD) do Instituto Brasileiro de Geografia e Estatística (IBGE), de 2006, o estudo de Santos, Bastos e Rocha (2008), por exemplo, mostra que a renda média dos homens é 1,5 mais alta do que a renda média das mulheres. Ver também Kassouf (1998).

11. Roy (1951) foi um dos primeiros a discutir problemas de autosseleção ao considerar um exemplo de duas ocupações: caça e pesca. Neste caso, um indivíduo escolheria uma destas ocupações de acordo com a sua produtividade e, portanto, os mesmos se autosselecionariam a partir de suas vantagens comparativas. A distribuição observada de rendas dos caçadores e pescadores seria determinada por tais escolhas.

O diferencial público-privado com a correção do viés de seleção entre mulheres é extremamente baixo, se comparado com o diferencial do modelo sem a correção do viés de seleção. Tal resultado pode indicar má especificação das variáveis que tentam captar características não observáveis das trabalhadoras, tais como a presença de filhos ou de idosos na família.

Além desta introdução, este trabalho tem mais quatro seções. A próxima seção descreve o modelo econométrico usado para identificar os fatores determinantes no problema da escolha entre setores sob a perspectiva do trabalhador. A terceira seção descreve a base de dados, e também apresenta algumas estatísticas descritivas no que diz respeito ao mercado de trabalho brasileiro. A quarta seção apresenta os principais resultados deste estudo. Por fim, a última seção é dedicada a algumas conclusões. O apêndice expõe uma breve descrição do desenho amostral complexo da PNAD adotado em 2006 e seus diferentes métodos de amostragem.

2 MODELO ECONOMÉTRICO

Modelos de regressão com mudança endógena (*endogenous swithing regression*) aplicados à análise de viés de autosseleção envolvem um mecanismo que seleciona os agentes entre dois estados possíveis da economia.[12] Lee (1978) foi um dos primeiros a utilizar tal arcabouço empiricamente para analisar os diferenciais de salários entre trabalhadores filiados a sindicatos e não membros de sindicatos nos Estados Unidos. Desde então, tal metodologia tem sido aplicada em uma variedade de temas subjacentes a mercado de trabalho (HECKMAN; HONORE, 1990).[13]

O modelo de regressão com mudança endógena em questão trata de um processo endógeno que seleciona o indivíduo em um dos dois estados: setor público ou setor privado. O estado em que um indivíduo é selecionado determina o regime de salário ao qual o trabalhador se defronta. A modelagem simultânea dos rendimentos dos

12. Modelos econométricos relacionados com desequilíbrio e modelos que analisam viés de autosseleção têm bastante semelhança e podem ser considerados como sistemas com mudanças estruturais. Este trabalho está centrado no segundo tipo de modelo. Uma excelente análise de modelos com mudanças estruturais (*switching regression models*) pode ser encontrada em Maddala (1986).
13. Além de mercado de trabalho, estes modelos têm sido aplicados em inúmeros outros problemas, tais como demanda por habitação e educação (MADDALA, 1986, p. 1.637).

trabalhadores dos setores (público ou privado) e da decisão de participação no mercado de trabalho em um destes setores pode ser descrita como:[14]

$$\ln (w_1) = X\beta_1 + u_1 \quad (1)$$

$$\ln (w_2) = X\beta_2 + u_2 \quad (2)$$

e

$$I = 1 \text{ (setor público) se } I^* = Z\theta + \varepsilon \geq 0 \quad (3)$$

$$I = 0 \text{ (setor privado) se } I^* = Z\theta + \varepsilon < 0 \quad (4)$$

em que $\ln (w_j)$ é o logaritmo natural dos salários para trabalhadores do setor público (j = 1) e do setor privado (j = 2), respectivamente; X é uma matriz de variáveis demográficas; β_j é o vetor de coeficientes associados com X e u_j é o erro normalmente distribuído. As equações (1) e (2) incluem controles em X para características demográficas (gênero e cor), produtivas (nível educacional e experiência) e a localização geográfica do trabalhador. O nível educacional do trabalhador é representado pelas variáveis FUNDAMENTAL, MÉDIO e SUPERIOR, correspondentes aos anos de escolaridade completos dos respectivos níveis de ensino. A variável experiência (EXPERIENCIA) é definida como a idade do trabalhador menos os anos de escolaridade menos seis. Esta medida embute a hipótese de que todos os trabalhadores iniciam a vida escolar aos seis anos de idade e nenhum tempo é gasto fora da força de trabalho ou da escola. O capital humano pode ser depreciado, seguindo uma curva parabólica. Assim, a experiência é também incluída na sua forma quadrática (EXPERIENCIA2), a fim de se levar em conta a possível depreciação do capital humano.

Uma hipótese no método de estimação das regressões das equações (1) e (2) em separado, uma para trabalhadores do setor público e outra para os trabalhadores do setor privado, é a de que $Eu_1 = 0$ e $Eu_2 = 0$, o que implica que trabalhadores são distribuídos de forma aleatória entre o setor público e o setor privado. Esta hipótese é questionável se as características não observáveis dos trabalhadores que determinam

14. Muito da discussão desta seção pode ser encontrado em Maddala (1983, p. 223-224) e Lokshin e Sajaia (2004).

os salários são correlacionadas com características não observáveis que determinam a escolha do setor de emprego. Se um diferencial salarial público-privado existe, trabalhadores preferem um setor ao outro e um processo de seleção endógena determina a seleção dos trabalhadores entre os dois setores.[15] Além disso, quando os trabalhadores são heterogêneos, há a possibilidade de que se autosselecionem no setor em que são mais produtivos, como advoga o modelo de Roy (1951).

As equações (3) e (4) determinam a escolha de se trabalhar em um determinado setor – são usualmente chamadas de *switching equations*.[16] A variável indicadora latente I^* descreve o processo de seleção e, portanto, determina o setor em que o indivíduo trabalha. Embora esta última variável não seja observada de forma completa, sua realização dicotômica I, contudo, o é. Se $I^* > 0$, então $I = 1$, caso contrário $I = 0$. O vetor $Z = [X|Y]$ inclui todas as variáveis relacionadas com características produtivas e demográficas que estão nas equações de salários originais (X). O vetor Z também inclui as restrições de exclusão que podem ajudar no problema de identificação da escolha de setor (Y). O número de filhos entre 0 e 15 anos do trabalhador (FILHO0_15), o número de idosos na família (IDOSO80_99), a condição de ser chefe de família (CHEFE) e a renda não proveniente do trabalho (RENDA NTRAB) são as variáveis usadas na identificação do modelo.

Como Heitmueller (2004) aponta, a identificação na equação de escolha de setor, se público ou privado, é muito mais difícil do que aquela da decisão de participação do indivíduo na entrada no mercado de trabalho de forma geral. A literatura sugere que as características de *background* social e familiar não têm impacto nos salários, mas sim na decisão entre qual o setor (se público ou privado) o trabalhador irá escolher trabalhar. Inúmeras variáveis, tais como nível educacional, escolhas ocupacionais ou setoriais dos pais são comumente usadas para identificação das escolhas setoriais feitas pelos trabalhadores. Como a PNAD de 2006 não fornece tais informações, estas variáveis não foram utilizadas nas regressões.

15. Além do problema da endogeneidade, a simples estimação das equações (1) e (2) pode produzir estimativas inconsistentes com base em um viés de seleção amostral relacionado com a exclusão de não participantes no mercado de trabalho (HECKMAN, 1979).

16. O modelo apresentado nesta seção é semelhante aos comumente usados para análise dos hiatos salariais entre trabalhadores filiados a sindicatos e não membros de sindicatos (LEE, 1978; ROBINSON; TOMES, 1984). Em geral, não há um nome aceito para este modelo. Cameron e Triverdi (2005, p. 555) o chamam de Modelo de Roy. Maddala (1983, p. 225) analisa em detalhes tal modelo e o nomeia de *switching regression* com *switching* endógeno (*switching regression model with endogenous switching*). Amemiya (1985, p. 399) o descreve como modelo *tobit* tipo 5.

Uma estimação das equações (1) e (2) produz estimativas de salários não viesadas somente se ε não for correlacionado com u_1 e u_2, respectivamente. Contudo, se preferências não observáveis influenciam o processo de seleção e os determinantes de salários, esta hipótese é violada e comparações de salários entre os dois setores geram estimativas inconsistentes. As equações (3) e (4) resumem o processo de seleção em que um trabalhador decide se trabalha no setor público ou no setor privado.

A estratégia econométrica em um modelo de regressão com mudança endógena é realizada em dois estágios. No primeiro estágio, um modelo *probit* é estimado a fim de se determinar as variáveis que têm efeito na probabilidade de se entrar no setor público (equações (3) e (4)). No segundo estágio, os termos de seleção λ_i (razão inversa de Mills), respectivos para cada setor, são adicionados em cada equação de salários. A importância da razão inversa de Mills é tornar possível a correção do viés de seleção quando uma equação de rendimentos é estimada. A inclusão da razão inversa de Mills (LAMBDA) nas equações de rendimentos leva em conta explicitamente a decisão do trabalhador na escolha entre o setor público e o setor privado, de forma que este termo elimina o viés de seletividade e consequentemente nos dá estimativas consistentes dos parâmetros das equações de salários. Este modelo, portanto, tenta controlar qualquer vantagem ou desvantagem comparativa que os dois grupos tenham em seus respectivos setores, e corrigir o viés que a seleção amostral pode gerar em cada determinante dos rendimentos. A equação de salários é então estimada de uma forma consistente pelo método de Mínimos Quadrados Ordinários (MQO):

$$\ln(w_1) = X\beta_1 + \lambda_1(Z\theta)\gamma_1 u_1 \quad (5)$$

$$\ln(w_2) = X\beta_2 + \lambda_2(-Z\theta)\gamma_2 u_2 \quad (6)$$

As equações (5) e (6) correspondem aos setores público e privado, respectivamente. A robustez dos resultados é testada pela estimação de equações de regressão para ambos os setores, com e sem termos de seleção.

3 BASE DE DADOS

A base de dados deste trabalho provém da PNAD de 2009. Como uma pesquisa de desenho amostral complexo, a PNAD utiliza técnicas de amostragem estratificada e de conglomeração e com probabilidades desiguais de seleção que, se não levadas em conta, podem gerar resultados estatísticos pouco precisos. Portanto, o tratamento incorreto do desenho amostral complexo de uma pesquisa compromete a confiabilidade das estimativas pontuais e seus respectivos erros-padrão, o que afeta o nível de significância e um eventual teste de hipóteses. O apêndice descreve o desenho amostral complexo da PNAD e apresenta os seus métodos de amostragem.

Em estudos que se baseiam em pesquisas que utilizam métodos complexos amostrais, é útil calcular o chamado *design effect* (*deff*), ou a razão entre variância estimada, que leva em conta o desenho amostral efetivo e a variância estimada, supondo que os dados sejam de uma amostra aleatória simples do mesmo tamanho.[17] O *deff* é usado para identificar qualquer perda de uma precisão da estimativa. Altos valores de *deff* (> 1) indicam que as estimativas da variância obtida ao não se levar em conta o plano amostral complexo subestimam a verdadeira variância das estimativas do modelo. O impacto do desenho amostral complexo em nossas estimativas (*deff*) será apresentado juntamente com os resultados de nosso modelo na seção 4. Em seguida, serão descritas algumas estatísticas sobre o mercado de trabalho dos setores público e privado no Brasil.

3.1 ESTATÍSTICAS DESCRITIVAS E DESCRIÇÃO DE VARIÁVEIS

A composição da força de trabalho do setor público e privado no Brasil é apresentada na tabela 1. O emprego do setor público como um todo representa mais de 14% da força de trabalho no país em 2009. A tabela mostra uma desagregação por tipo de contrato de trabalho em cada setor. Do setor público no Brasil, quatro categorias de trabalhadores são apresentadas: *i*) funcionários públicos estatutários (FPE); *ii*) Funcionários públicos celetistas; *iii*) funcionários públicos sem carteira (FPS); e *iv*) militares. A categoria dos funcionários públicos estatutários consiste dos funcionários que passaram em concurso público, enquanto a dos funcionários públicos celetistas é composta por trabalhadores cujo vínculo empregatício é regido pela Consolidação

17. Ver Kish (1965).

das Leis do Trabalho (CLT).[18] Os funcionários públicos sem carteira (os trabalhadores "informais" do setor público) não têm um contrato oficial de trabalho e, em geral, são temporários.[19] A inclusão de cada um destes grupos cobre todas as atividades da administração pública nos três níveis de governo (federal, estadual e local) e nas empresas públicas. Quando o mercado de trabalho do setor público é analisado por gênero, observa-se o mesmo padrão na desagregação por contrato de trabalho. Tanto para homens quanto para mulheres, a maior parte do funcionalismo público é representada por funcionários públicos estatutários.

Com relação ao setor privado, os trabalhadores são classificados entre aqueles que possuem e não possuem um contrato formal de trabalho. A maior parte da força de trabalho do setor privado é composta por trabalhadores que têm carteira de trabalho e que, portanto, fazem parte do setor formal da economia (tabela 1). Nota-se que as mulheres têm uma maior participação no mercado informal de trabalho do que os homens. Em torno de 23% da força de trabalho feminina empregada não têm carteira assinada. Esta estatística se deve muito provavelmente ao grande número de mulheres que são trabalhadoras domésticas sem carteira de trabalho. Por sua vez, os trabalhadores por conta própria têm uma maior representatividade em relação às mulheres.

TABELA 1
Emprego total por setor no Brasil – 2009

	Emprego total		Mulheres		Homens	
	Milhões	%	Milhões	%	Milhões	%
Setor público	9,4	14,2	5,3	18,2	4,1	11,2
Funcionário público estatutário (FPE)	5,8	8,7	3,4	11,6	2,4	6,4
Funcionário público com carteira (Celetista)	1,8	2,7	0,9	3,2	0,8	2,2
Funcionário público sem carteira (FPS)	1,7	2,5	1,0	3,4	0,7	1,8
Militares	0,3	0,4	0,0	0,0	0,3	0,7
Setor privado	56,9	85,8	23,9	81,8	33,0	88,8
Trabalhadores com carteira assinada	30,3	45,6	12,4	42,5	17,8	48,1
Trabalhadores sem carteira assinada	12,4	18,7	6,7	23,1	5,6	15,2
Conta-própria	11,2	16,8	3,9	13,2	7,3	19,6
Empregadores	3,1	4,7	0,9	3,0	2,2	5,9
Total	66,3	100,0	29,2	100,0	37,1	100,0

Fonte: PNAD/IBGE (2009).

18. Os funcionários públicos estatutários são regidos pelo Regime Jurídico Único (RJU).
19. A Constituição Federal (CF) de 1988 determina formalmente todas as categorias de funcionários públicos e a emenda constitucional (EC) nº 19 de 1998 regulamentou a contratação de funcionários públicos não estatutários.

O universo de análise deste trabalho abrange todos os indivíduos de 18 a 65 anos de idade, moradores de área urbana e que reportaram uma jornada semanal entre 20 e 72 horas e uma remuneração positiva em sua ocupação principal na semana de referência da pesquisa. Trabalhadores que exercem atividades agrícolas foram excluídos da análise. Uma série de filtros foi feita na subpopulação original com o objetivo de refinar a comparação de rendimentos entre trabalhadores do setor público e os equivalentes do setor privado. Do setor público, a amostra compreende apenas os funcionários públicos estatutários e os funcionários públicos com carteira (celetistas). Com relação ao setor privado, apenas os trabalhadores com e sem carteira assinada foram considerados. Assim, a amostra usada para o cálculo do diferencial público-privado ficou restrita a 85.526 indivíduos (cerca de 66% da população original reportada pela PNAD).[20]

As tabelas 2 e 3 apresentam, respectivamente, as definições das variáveis e estatísticas descritivas referentes às características dos trabalhadores do setor público e do setor privado.

A maior parte dos trabalhadores do setor público é composta de mulheres, representando em torno de 57% do total de trabalhadores deste setor, enquanto o valor correspondente para o setor privado é aproximadamente de 37% (tabela 3). Algumas características, válidas tanto para homens quanto para mulheres, são interessantes de destacar. Os funcionários públicos têm mais idade e mais anos de escolaridade em comparação com os trabalhadores do setor privado. Com relação à distribuição espacial da força de trabalho, o emprego do setor público e do setor privado está sobrerrepresentado na região Sudeste do país. Em torno de 43% (44%) dos(as) funcionários(as) do setor público e 52% (53%) dos funcionários(as) do setor privado estão localizados nesta região (tabela 3). Com relação às características de trabalho, vale ressaltar que 38% dos(as) trabalhadores(as) do setor público são membros de sindicatos, enquanto para o setor privado esta estatística corresponde a 21%, para homens e 19% para mulheres. Os trabalhadores do setor público permanecem no emprego mais do que o dobro do tempo dos trabalhadores do setor privado.

20. Desse total, 15.047 representam os trabalhadores do setor público restrito (só com funcionários públicos estatutários e funcionários públicos com carteira); 65.439 trabalhadores fazem parte do setor privado (somente os trabalhadores com e sem carteira assinada de trabalho).

TABELA 2
Variáveis do modelo

Variáveis	Descrição de variáveis
SALÁRIO	salário mensal
SALÁRIO PADRÃO	salário mensal padronizado
LN SALÁRIO	logaritmo do salário padronizado
HORAS	número de horas trabalhadas por semana
GÊNERO	1 = Mulher, 0 caso contrário
RAÇA	1 = Branco, 0 caso contrário
IDADE	idade do trabalhador
EXPERIENCIA	anos de experiência
EXPERIENCIA2	anos de experiência ao quadrado
EDUCAÇÃO	número de anos de escolaridade
BASICO	1 = ensino básico completo, 0 caso contrário (c.c.)
FUNDAMENTAL	1 = ensino fundamental completo, 0 c.c.
MEDIO	1 = ensino médio completo, 0 c.c.
SUPERIOR	1 = ensino superior completo, 0 c.c.
ANOS TRAB	número de anos no trabalho
SINDICALIZADO	1 = Membro de sindicato, 0 c.c
NORTE	1 = Indivíduo residente na região Norte
NORDESTE	1 = Indivíduo residente no Nordeste
CENTRO-OESTE	1 = Indivíduo no Centro-Oeste
SUDESTE	1 = Indivíduo residente no Sudeste
SUL	1 = Indivíduo residente no Sul
RENDA NTRAB	renda não proveniente do trabalho
FILHO0_15	no de filhos de 0 a 15 anos
IDOSO80_99	no de idosos
CHEFE	1 = chefe de família, 0 c.c.
CONJUGE	1 = cônjuge, 0 c.c.
FILHO	1 = filho, 0 c.c.

Fonte: Elaboração própria.

TABELA 3
Estaísticas descritivas do setor público e do setor privado no Brasil – 2009

Variáveis	Setor público[1]		Setor privado		Homens				Mulheres			
					Setor público[1]		Setor privado		Setor público[1]		Setor privado	
	Média	Desvio-padrão[2]	Média	Desvio-padrão[2]	Média	Desvio-padrão[2]	Média	Desvio-padrão[2]	Média	Desvio-padrão[2]	Média	Desvio-padrão[2]
EQUAÇÃO DE SALÁRIOS												
SALÁRIO	R$ 1.869,16	26,753	R$ 1.018,20	13,358	R$ 2.270,33	42,324	R$ 1.106,77	19,100	R$ 1.569,70	23,337	R$ 860,87	8,826
SALÁRIO PADRÃO	R$ 2.008,38	27,567	R$ 962,52	11,825	R$ 2.324,92	43,279	R$ 1.023,09	16,257	R$ 1.771,59	24,384	R$ 853,20	9,184
LN SALÁRIO PADRÃO	R$ 7,26	0,010	R$ 6,61	0,005	R$ 7,36	0,014	R$ 6,65	0,006	R$ 7,18	0,011	R$ 6,52	0,006
HORAS	37,81	0,081	43,49	0,045	40,25	0,111	44,34	0,052	35,98	0,100	42,02	0,062
GÊNERO	0,57	0,004	0,37	0,002	-	-	-	-	-	-	-	-
RAÇA	0,57	0,005	0,53	0,004	0,53	0,007	0,50	0,004	0,59	0,006	0,58	0,005
IDADE	41,58	0,099	33,56	0,049	41,79	0,149	34,12	0,060	41,43	0,123	32,56	0,074
EXPERIENCIA	17,61	0,958	15,18	0,369	18,27	1,280	16,45	0,485	17,14	1,260	13,33	0,549
EXPERIENCIA2	421,13	43,314	333,17	14,812	438,89	59,552	376,37	2,048	408,61	57,146	274,73	20,968
EDUCAÇÃO	13,04	0,274	9,71	0,131	12,71	0,392	9,21	0,177	13,33	0,322	10,48	0,171
BASICO	0,03	-	0,15	-	0,05	-	0,19	-	0,02	-	0,09	-
FUNDAMENTAL	0,09	-	0,23	-	0,12	-	0,26	-	0,06	-	0,18	-
MEDIO	0,30	-	0,46	-	0,29	-	0,41	-	0,31	-	0,54	-
SUPERIOR	0,58	-	0,16	-	0,55	-	0,14	-	0,62	-	0,19	-
ANOS TRAB	11,89	0,097	4,40	0,030	12,43	0,138	4,74	0,038	11,49	0,119	3,80	0,040
SINDICALIZADO	0,38	0,006	0,20	0,003	0,38	0,007	0,21	0,003	0,38	0,007	0,19	0,004
NORTE	0,08	0,003	0,05	0,001	0,09	0,003	0,06	0,002	0,08	0,003	0,05	0,002
NORDESTE	0,24	0,006	0,18	0,004	0,24	0,007	0,18	0,004	0,24	0,006	0,16	0,004
CENTRO-OESTE	0,09	0,003	0,07	0,002	0,10	0,004	0,07	0,002	0,09	0,003	0,07	0,002
SUDESTE	0,44	0,007	0,53	0,005	0,43	0,008	0,52	0,005	0,44	0,008	0,53	0,006
SUL	0,15	0,004	0,17	0,004	0,14	0,005	0,16	0,004	0,15	0,006	0,18	0,005

(continua)

(continuação)

Variáveis	Setor público[1]		Setor privado		Homens				Mulheres			
					Setor público[1]		Setor privado		Setor público[1]		Setor privado	
	Média	Desvio-padrão[2]	Média	Desvio-padrão[2]	Média	Desvio-padrão[2]	Média	Desvio-padrão[2]	Média	Desvio-padrão[2]	Média	Desvio-padrão[2]
EQUAÇÃO DE PARTICIPAÇÃO												
RENDA NTRAB	R$ 222,39	8,317	R$ 62,84	2,085	R$ 235,48	13,881	R$ 59,60	2,356	R$ 213,16	8,735	R$ 67,69	3,311
FILHO0_15	0,69	0,009	0,78	0,006	0,71	0,013	0,81	0,007	0,67	0,0104	0,73	0,007
IDOSO80_99	0,03	0,002	0,02	0,001	0,02	0,002	0,02	0,001	0,04	0,0024	0,02	0,001
CHEFE	0,50	0,004	0,43	0,002	0,73	0,006	0,54	0,003	0,33	0,0057	0,23	0,003
CONJUGE	0,31	0,004	0,20	0,002	0,09	0,004	0,10	0,002	0,47	0,0059	0,37	0,004
FILHO	0,15	0,003	0,29	0,002	0,14	0,005	0,28	0,003	0,15	0,0042	0,32	0,004

Fonte: PNAD/IBGE (2009).

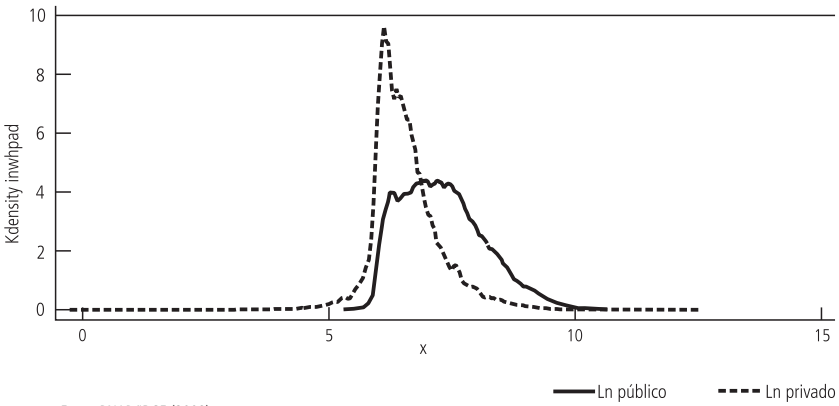
Notas: [1] O conceito de setor público adotado neste trabalho compreende apenas os funcionários públicos estatutários e os funcionários públicos com carteira assinada.

[2] Desvio-padrão linearizado (estatísticas baseadas no desenho amostral complexo da PNAD).

A variável correspondente aos salários é a renda mensal, estritamente positiva, recebida pelos trabalhadores de seu trabalho principal. Nossa medida de renda de trabalho, tanto do setor público quanto do setor privado, corresponde ao rendimento do trabalho padronizado pelo número de horas trabalhadas por semana.[21]

Tanto a média quanto o desvio-padrão do salário padronizado do setor público é maior do que estas mesmas estatísticas do setor privado. A diferença bruta entre o salário médio padronizado do setor público e o do setor privado é extremamente alta (cerca de 108,66%). Quando o cálculo do hiato salarial é realizado entre as rendas nominais dos dois setores, o salário do setor público é superior cerca de 84%, o que evidencia que parte do diferencial do rendimento padronizado é explicado pela diferença de jornada de trabalho entre o setor público e o setor privado. De fato, importante ressaltar que a média de horas semanais trabalhadas no setor público em questão (37,81) é quase 13% menor do que a média de horas semanais trabalhadas no setor privado (43,49). O gráfico 1 apresenta as distribuições de salários dos trabalhadores do setor público e do setor privado.

GRÁFICO 1
Ln do rendimento padronizado – setor público e setor privado



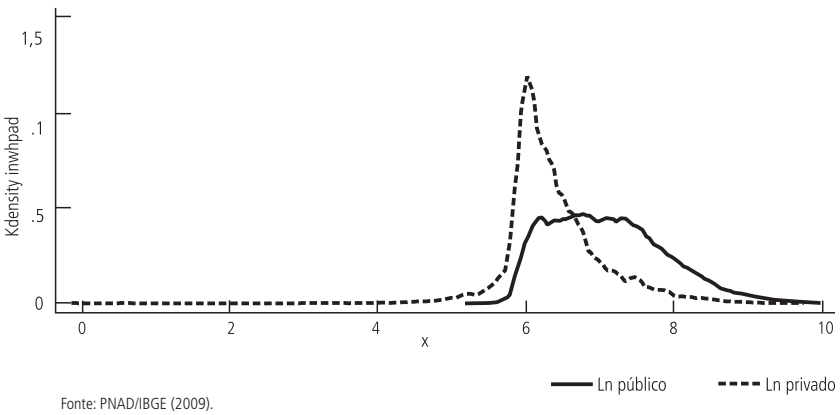
Fonte: PNAD/IBGE (2009).

No que diz respeito à análise específica por gênero, nota-se que, entre as mulheres, o diferencial de salários padronizados público-privado chega a 108% enquanto para os homens é de 127%. Os gráficos 2 e 3 apresentam as distribuições de salários do setor público e do setor privado para mulheres e homens, respectivamente. Nota-se

21. O rendimento padronizado é a renda mensal dividida pelo número de horas trabalhadas na semana multiplicada por 40 horas.

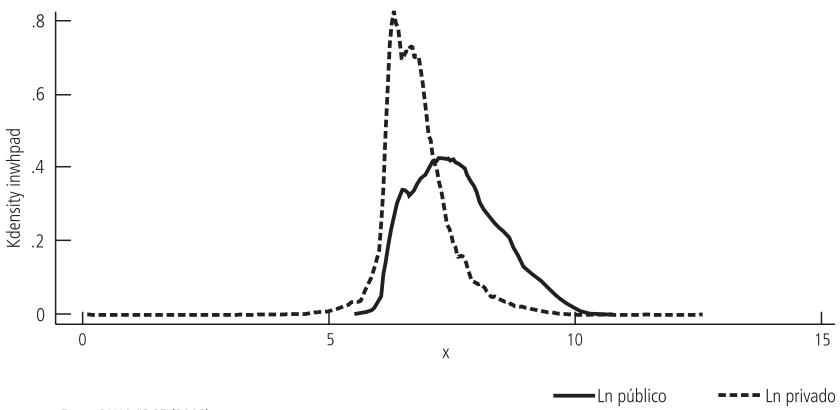
uma forte concentração de trabalhadores(as) que ganham em torno do salário mínimo no setor privado.[22] Um dado interessante a se observar é que, para os trabalhadores homens, o setor privado tem limites inferiores e superiores maiores do que o setor público no suporte da distribuição dos salários. No caso das mulheres, o setor público apresenta um limite superior maior do que o do setor privado.

GRÁFICO 2
Ln do rendimento padronizado-Mulheres – setor público e setor privado



Fonte: PNAD/IBGE (2009).

GRÁFICO 3
Ln do rendimento padronizado-Homens – setor público e setor privado



Fonte: PNAD/IBGE (2009).

22. Braga (2008) apresenta estas mesmas evidências para os trabalhadores do setor público e do setor privado de forma geral.

4 RESULTADOS

4.1 EQUAÇÃO DE ESCOLHA DO SETOR

Os coeficientes estimados do modelo probit para a participação no mercado de trabalho do setor público são apresentados na tabela 4. Os efeitos marginais correspondentes aos fatores que influenciam a probabilidade de um indivíduo trabalhar no setor público são calculados nos valores médios de cada variável. Um coeficiente positivo sugere que a variável contribui para o aumento da probabilidade de se obter emprego no setor público.

A educação afeta de forma positiva a escolha do indivíduo (homens e mulheres) em trabalhar no setor público. Trabalhadores com um nível maior de escolaridade têm maior probabilidade de buscar emprego no setor público (em especial, aqueles que têm nível superior de ensino). Quanto à experiência, os resultados também apontam uma relação positiva nas oportunidades de trabalho no setor público. Tal fato é indicado pelo coeficiente positivo para variável experiência (EXPERIENCIA); o coeficiente negativo para experiência ao quadrado (EXPERIENCIA2) está associado à depreciação do capital humano. Com respeito à variável referente cor do trabalhador, o coeficiente da mesma não é estatisticamente significativo para as amostras compostas por homens e mulheres.

Outro resultado interessante diz respeito à localização geográfica do trabalhador. Nas regiões mais desenvolvidas do país, como Sul, Sudeste e Centro-Oeste, as oportunidades do emprego no setor público não são tão atraentes para os trabalhadores quanto nas regiões Norte e Nordeste. Tal fato se verifica muito provavelmente porque os mercados de trabalho no setor privado são mais dinâmicos do que os do setor público naquelas três regiões. No que diz respeito à região Norte, apenas os trabalhadores homens aí residentes têm maiores chances de entrar no setor público do que os mesmos residentes em outras regiões do país.[23]

23. A variável associada à região Nordeste foi omitida das estimações para se evitar perfeita colinearidade.

TABELA 4
Equação de Escolha de Setor - Modelo Probit (2009)

Setor Público	*Probit*	*Deff*	*Probit* mulheres	*Deff*	*Probit* homens	*Deff*
EXPERIENCIA	0.0604	1.0323	0.0779	0.9814	0.0415	1.1246
	(0.0201)***		(0.0267)***		(0.0294)	
EXPERIENCIA2	–0.0006	1.0563	–0.0009	0.9187	–0.0002	1.1475
	(0.0005)		(0.0006)		(0.0007)	
FUNDAMENTAL	0.5561	1.0841	0.4851	1.2294	0.6242	1.0586
	(0.2988)*		(0.6089)		(0.3401)*	
MÉDIO	0.9902	1.1808	1.0682	1.2629	0.9765	1.0778
	(0.3029)***		(0.5979)*		(0.3362)***	
SUPERIOR	2.0486	1.2215	2.1420	1.2550	2.0294	1.0705
	(0.3238)***		(0.6290)***		(0.3457)***	
RAÇA	0.0202	1.1173	–0.0099	1.0972	0.0377	1.0531
	(0.1312)		(0.1893)		(0.1744)	
GÊNERO	0.2993	1.0695	–	–	–	–
	(0.1221)**					
NORTE	0.2812	.7198	0.1807	0.7842	0.2787	0.6682
	(0.1827)		(0.3163)		(0.2125)	
CENTRO-OESTE	–0.2763	.7561	–0.4991	0.7458	–0.2019	0.7659
	(0.2072)		(0.3511)		(0.2610)	
SUDESTE	–0.4315	1.1125	–0.2591	1.1188	–0.5818	0.9488
	(0.1544)***		(0.2197)		(0.1976)***	
SUL	–0.4564	1.1318	–0.3052	0.9188	–0.6165	0.9220
	(0.1799)**		(0.2294)		(0.2313)***	
CHEFE DE FAMÍLIA	–0.0117	1.0226	–0.0686	1.1316	0.0476	1.1106
	(0.1210)		(0.1851)		(0.1770)	
RENDA NTRAb	0.0002	1.3644	0.0004	1.1896	0.0001	1.3007
	(0.0001)		(0.0003)		(0.0001)	
FILHO0_15	–0.1155	1.2406	–0.0911	1.1896	–0.1343	1.0377
	(0.0678)*		(0.0950)		(0.0875)	
IDOSO80_99	–1.0632	0.6482	–1.1018	0.6981	0.0000	–
	(0.4506)**		(0.5081)**		(0.0000)	
CONSTANTE	–2.7584	1.1180	–2.7835	1.2010	–2.5204	1.0850
	(0.5168)***		(0.6403)***		(0.4616)***	
OBSERVAÇÕES	85.526		34.601		50.925	

Fonte: PNAD/IBGE (2009).

Nota: Erro-padrão linearizado entre parênteses; * nível de significância de 10%, ** 5% e *** 1%.

Alguns resultados apresentam diferenças significativas entre homens e mulheres, no entanto. Enquanto para as mulheres a variável que identifica a presença de idosos na família sugere um efeito negativo (com nível de significância de 5%) na probabilidade de se entrar no setor público, para os homens esta variável não é estatisticamente significativa. A questão de identificação pelo controle da presença ou pelo número de filhos ou de pessoas idosas na família já foi tratada de forma abrangente na literatura sobre oferta de salários e participação da força de trabalho (KILLINGSWORTH; HECKMAN, 1986). Em análise sobre o viés de seleção presente na escolha de se trabalhar entre o setor público e o setor privado, Tansel (2005) e Heitmueller (2004) encontram evidências de que o número de filhos tem um efeito positivo para os homens e um efeito negativo para mulheres, com relação à entrada no setor público, resultados contrários aos encontrados aqui. A tabela 3 aponta que, para a amostra como um todo (ou seja, quando não se distingue os trabalhadores por gênero), a presença de filhos entre 0 e 15 anos desestimula a entrada do trabalhador no setor público (ao nível de significância de 10%). Com relação às variáveis que representam a condição do trabalhador na família – ser chefe de família (CHEFE) – e a renda não proveniente do trabalho, as mesmas não são estatisticamente significativas para as amostras.[24]

O *deff* nos diz quanto de informação ganhamos ou perdemos ao se levar em conta o desenho amostral complexo de uma pesquisa em vez de assumir o método por amostra aleatória simples. Como colocado por Silva, Pessoa e Lila (2002), os valores do *deff* longe de 1 indicam que ao não se levar em conta desenho amostral complexo nas estimações do modelo, as estimativas ficam viesadas e incorretas. Altos valores (> 1) para *deff* indicam que o estimador de variância obtido sem levar em conta o desenho amostral complexo subestima o verdadeiro estimador variância. Os resultados claramente revelam que ignorar o desenho amostral complexo não gera efeitos significativos na análise de regressão no caso da PNAD de 2009.[25] Os valores do

24. A renda não proveniente do trabalho (RENDA NTRAB) representa todos os tipos de renda que não provêm dos rendimentos do trabalho (aluguel, pensões, entre outras).

25. Os resultados de Silva, Pessoa e Lila (2002) e Carvalho (2005) são muito mais críticos do que os resultados encontrados neste estudo. Enquanto que Silva, Pessoa e Lila. (2002) apresentam algumas estatísticas descritivas para a PNAD de 1998, Carvalho (2005) analisa o modelo de salários com ênfase em discriminação por gênero e raça com base na PNAD de 2003. Ambos os estudos recomendam fortemente o uso de modelagem estatística adequada para pesquisas com desenho amostral complexo. Silva, Pessoa e Lila (2002) encontram valores de 13,7, por exemplo. Já Santos *et al.* (2008) analisam fatores determinantes da entrada de um indivíduo no mercado de trabalho e geram uma equação de rendimentos para o Brasil, com base de dados proveniente da PNAD para o ano de 2006. Os autores também encontram resultados altos para o deff (também em torno de 13) e sugerem, portanto, que o plano amostral complexo é extremamente importante para a obtenção de estatísticas das variâncias dos regressores robustas e não viesadas.

deff variam de 0,65 a 1,36 (tabela 3).[26] Todos os resultados das estimações das equações de salários apresentados na próxima subseção (4.2) também são baseados no desenho amostral complexo da PNAD. Os valores do *deff* das estimativas também não são muito diferentes daqueles apresentados na tabela 4.

4.2 EQUAÇÕES DE SALÁRIOS

A estimação de duas regressões de salários separadas para os trabalhadores do setor público e do setor privado permite identificar algumas diferenças potenciais entre os mercados de trabalho destes setores, mas não leva em conta o viés de seleção presente na escolha setorial. Apesar de as estimativas por mínimos quadrados poder oferecer informações interessantes acerca dos determinantes dos salários dos indivíduos, existe uma possibilidade de que os coeficientes das covariadas destas equações de rendimentos sejam viesados. A aplicação do modelo de regressão com mudança endógena busca, portanto, corrigir este viés de seleção. Ao modelar as preferências dos trabalhadores através do modelo *probit* apresentado na seção anterior, o diferencial de salários público-privado pode ser estimado com a correção do viés de seleção presente no problema da escolha do trabalhador em entrar no setor público. São dois os modelos estimados: *i*) modelo de regressão pelo método por MQO, ou seja, sem a correção do viés de seleção; *ii*) modelo de regressão com mudança endógena (*endogenous switching regression model*). As tabelas 5 e 6 reportam os resultados das estimações de ambos os modelos para mulheres e homens, respectivamente.

Com relação ao modelo de regressão por mínimos quadrados, nota-se que, com algumas poucas exceções, todas as variáveis têm seus coeficientes estatisticamente significativos, com sinais adequados, nas equações de rendimentos dos trabalhadores do setor privado. Este fato se verifica para as estimações dos salários da amostra que abrangem tanto as mulheres quanto os homens (tabela 5). Um maior nível de escolaridade, por exemplo, aumenta os retornos à educação do trabalhador, tanto no setor público quanto no setor privado. Há um efeito igualmente positivo para a variável raça no setor privado, o que sinaliza que os salários potenciais dos trabalhadores(as) brancos(as) são mais altos do que os dos não brancos.

26. Diferentes pacotes computacionais, tais como Statistical Package for the Social Science (SPSS), STATA e Statistical Analysis System (SAS), têm diferentes métodos para análise estatística e econométrica com base em pesquisas amostrais complexas.

TABELA 5
Equações log-salários para os setores público e privado no Brasil – 2009

LN salários	Modelo de regressão MQO				Modelo *switching regression* – regressão com seletividade corrigida			
	Setor público	*Deff*	Setor privado	*Deff*	Setor público	*Deff*	Setor privado	*Deff*
EXPERIENCIA	0.0749	1.1211	0.0408	1.1519	0.0727	1.0245	0.0400	1.2227
	(0.0154)***		(0.0064)***		(0.0199)***		(0.0077)***	
EXPERIENCIA2	–0.0012	1.1206	–0.0007	1.1434	–0.0012	1.0706	–0.0007	1.1644
	(0.0003)***		(0.0002)***		(0.0003)***		(0.0002)***	
FUNDAMENTAL	0.1279	1.0243	0.1895	1.1345	0.1033	1.0421	0.1827	1.2550
	(0.1836)		(0.0608)***		(0.2311)		(0.0720)**	
MÉDIO	0.9410	0.9287	0.5186	1.2027	0.8988	1.0165	0.5061	1.3523
	(0.1578)***		(0.0703)***		(0.2850)***		(0.1002)***	
SUPERIOR	1.6860	0.8937	1.2507	1.2102	16.025	1.0082	12176	1.2349
	(0.1484)***		(0.0933)***		(0.4999)***		(0.2152)***	
RAÇA	–0.1292	0.8999	0.0700	1.1115	–0.1291	0.8783	0.0700	1.1133
	(0.1116)		(0.0411)*		(0.1117)		(0.0412)*	
MULHER	–0.3903	0.9897	–0.3205	1.1455	–0.4016	1.0218	–0.3245	1.1272
	(0.0981)***		(0.0409)***		(0.1172)***		(0.0445)***	
NORTE	–0.0816	0.6563	0.1418	0.6276	–0.0894	0.7313	0.1385	0.6509
	(0.1956)		(0.0806)*		(0.2052)		(0.0846)	
CENTRO-OESTE	0.1201	0.7520	0.1574	0.9025	0.1296	0.7595	0.1615	0.9189
	(0.2264)		(0.0671)**		(0.2296)		(0.0693)**	
SUDESTE	0.1160	1.0618	0.2957	1.2135	0.1312	1.0517	0.3023	1.1917
	(0.1275)		(0.0573)***		(0.1519)		(0.0647)***	
SUL	0.2590	0.8580	0.3606	1.0703	0.2759	0.8946	0.3667	1.1228
	(0.1384)*		(0.0603)***		(0.1633)*		(0.0675)***	
CONSTANTE	5.5793	1.0765	5.5739	1.1335	5.7316	1.0375	5.5824	1.1749
	(0.2054)***		(0.0934)***		(0.8890)***		(0.1050)***	
LAMBDA	-		-		–0.0518	1.0672	0.0507	1.2463
					(0.2858)		(0.3163)	
Observações	16.175		69.351		16.175		69.351	
R^2	0.49		0.40		0.49		0.40	

Fonte PNAD/IBGE (2009).

Nota: Erro-padrão linearizado entre parênteses; * nível de significância de 10%, ** 5% e *** 1%.

No que diz respeito às regiões geográficas, as mulheres e os homens residentes nas regiões Sul e Sudeste e que trabalham no setor privado são os que apresentam os coeficientes mais altos (em ambos os modelos) e, portanto, maiores retornos em relação aos salários, se comparado com outras regiões (tabelas 6 e 7). Quanto ao termo de seleção (variável LAMBDA), construído a partir da equação de escolha do setor público, as estimativas dos coeficientes desta variável não se mostraram estatisticamente significativas.[27]

27. Heitmueller (2004) encontra este mesmo resultado para a amostra que compõe trabalhadoras.

TABELA 6
Equações log-salários para setores público e privado no Brasil (mulheres) – 2009

LN salários	Modelo de regressão MQO				Modelo *switching regression* – regressão com seletividade corrigida			
	Setor público	*Deff*	Setor privado	*Deff*	Setor Público	*Deff*	Setor Privado	*Deff*
EXPERIENCIA	0.0234	1,1231	0.0403	1,3018	0.0004	0,8292	0.0350	1,3617
	(0.0227)		(0.0126)***		(0.0259)		(0.0165)**	
EXPERIENCIA2	–0.0004	1,0898	–0.0011	1,2808	–0.0002	0,8757	–0.0010	1,3165
	(0.0004)		(0.0003)***		(0.0004)		(0.0004)***	
FUNDAMENTAL	–0.2597	0,9788	0.0188	1,1783	–0.0240	0,0000	–0.0133	1,2338
	(0.1908)		(0.1116)		(0.2803)		(0.1076)	
MÉDIO	0.0000	1,1200	0.4110	1,1889	0.0000	0,8410	0.3381	1,2098
	(0.0000)		(0.1503)***		(0.0000)		(0.1407)**	
SUPERIOR	0.7087	1,3142	0.9917	1,1183	0.3059	0,8381	0.8284	1,0981
	(0.1595)***		(0.1736)***		(0.4528)		(0.2863)***	
RAÇA	–0.1024	1,0145	0.0454	1,0963	–0.1441	0,8647	0.0470	1,0293
	(0.2026)		(0.0821)		(0.2152)		(0.0811)	
NORTE	–0.1028	0,7203	0.2715	0,6933	–0.1719	0,6197	0.2547	0,7274
	(0.2398)		(0.1094)**		(0.2502)		(0.1215)**	
CENTRO-OESTE	0.1290	0,7556	0.0514	0,8216	0.3442	0,8003	0.0825	0,9502
	(0.2269)		(0.1258)		(0.3355)		(0.1189)	
SUDESTE	0.0946	1,0416	0.4024	1,0065	0.1792	1,0156	0.4227	1,0325
	(0.1901)		(0.1063)***		(0.1991)		(0.1022)***	
SUL	0.1920	0,9623	0.3415	0,8624	0.3067	0,8049	0.3575	0,8526
	(0.2407)		(0.0951)***		(0.2665)		(0.0896)***	
CONSTANTE	6,7463	1,1550	5.4919	1,0593	7.7786	0,8665	5.5532	1,1216
	(0.2669)***		(0.1582)***		(0,9887)***		(0,1780)***	
LAMBDA	-		-		–0.5094	0,8873	0.2024	1,2760
					(0.4933)		(3571)	
Observações	9.150		25.541		9.150		25.541	
R^2	0,32		0,36		0,34		0,36	

Fonte: PNAD/IBGE (2009).

Nota: Erro-padrão linearizado entre parênteses; * nível de significância de 10%, ** 5% e *** 1%.

Em geral, as magnitudes dos coeficientes diferem entre os dois modelos: o modelo de regressão por MQO e o modelo com seletividade corrigida (modelo de s*witching regression*). Para os trabalhadores do sexo feminino, os coeficientes das variáveis EXPERIENCIA, EXPERIENCIA2, as associadas à educação (MÉDIO, SUPERIOR) são menores em valores absolutos, no setor privado, do que no modelo de regressão por MQO (tabela 6). Os mesmos resultados se aplicam para trabalhadores homens do setor privado (tabela 7).

TABELA 7
Equações log-salários para setores público e privado no Brasil (homens) – 2009

LN salários	Modelo de regressão MQO				Modelo *switching regression* – regressão com seletividade corrigida			
	Setor público	*Deff*	Setor privado	*Deff*	Setor público	*Deff*	Setor privado	*Deff*
EXPERIENCIA	0.0683	0.8270	0.0415	0.9689	0.0782	0.9911	0.0365	1,0639
	(0.0785)		(0.0095)***		(0.0777)		(0.0124)***	
EXPERIENCIA2	–0.0009	0.8351	–0.0006	0.9609	–0.0010	0.9907	–0.0006	0.9927
	(0.0021)		(0.0002)***		(0.0021)		(0.0002)**	
FUNDAMENTAL	0.1486	0.9286	0.2218	0.8796	0.3108	0.7940	0.1798	1.0280
	(0.2606)		(0.0770)***		(0.3737)		(0.0961)*	
MÉDIO	0.9276	0.8147	0.5473	1.0648	10926	0.9036	0.5230	1.1225
	(0.4168)**		(0.0933)***		(0.4928)**		(0.1168)***	
SUPERIOR	1.3618	0.7760	1.2553	1.0979	17.678	0.8694	11627	1.1772
	(0.3767)***		(0.1441)***		(0.8432)**		(0.2512)***	
RAÇA	–0.1995	0.9460	0.1794	1.0634	–0.2249	0.9728	0.1814	1.1408
	(0.2949)		(0.0642)*		(0.2776)		(0.0644)***	
NORTE	0.2566	0.6104	0.0177	0.5504	0.2768	0.8974	0.0136	0.6456
	(0.3023)		(0.1049)		(0.3215)		(0.1055)	
CENTRO-OESTE	0.2015	0.7301	0.0913	0.8252	0.1917	0.7954	0.0963	0.8269
	(0.3318)		(0.1070)		(0.3327)		(0.1077)	
SUDESTE	0.0135	0.9296	0.2305	1.1311	–0.0728	1.0209	0.2441	0.9953
	(0.3005)		(0.0786)***		(0.3592)		(0.0860)***	
SUL	0.5120	0.8825	0.3146	1.0580	0.3497	0.9243	0.3319	1,0432
	(0.4463)		(0.0942)***		(0.5773)		(0.1058)***	
CONSTANTE	5.8194	0.8686	5.5280	0.9673	5.1772	0,9172	5.5789	1.1519
	(0.6539)***		(0.1347)***		(1.3976)***		(0.1533)***	
LAMBDA	-				–0.2121	0.8968	0.1564	1.0314
					(0,4232)***		(0,4238)	
Observações	7.025		43.900		7.025		43.900	
R^2	0.31		0.41		0,52		0,44	

Fonte: PNAD/IBGE (2009).

Nota: Erro-padrão linearizado entre parênteses; * nível de significância de 10%, ** 5% e *** 1%.

4.3 DIFERENCIAIS DE SALÁRIOS

A tabela 8 mostra que o diferencial salarial público-privado é extremamente alto no Brasil. Nota-se uma diferença significativa entre os diferenciais salariais não condicionais (não controlados) e os diferenciais controlados por características observáveis e não observáveis dos trabalhadores. Com base no modelo por mínimos quadrados (modelo de regressão MQO), portanto, sem o condicionamento na escolha de seleção, este diferencial representa 12,8% para os trabalhadores homens. Para as mulheres, este diferencial é mais alto, cerca de 18%. No que diz respeito ao diferencial salarial público-privado que tem como base o modelo endógeno de *switching regression*, um ponto

merece destaque. O diferencial público-privado entre mulheres é extremamente baixo e bem inferior ao diferencial encontrado no modelo sem a correção da endogeneidade. Uma possível explicação para tal resultado é que o termo que trata do viés de seleção não é estatisticamente significativo e, portanto, as características não observáveis não estariam sendo captadas de forma adequada neste modelo. Uma alternativa para a análise deste diferencial seria a inclusão de variáveis associadas às características de *background* social e familiar que melhor identificassem a escolha do setor (público ou privado) na qual o trabalhador irá escolher trabalhar.

TABELA 8
Teste de Média (teste t) para log-salários no setor público e no setor privado – 2009

	Salários não condicionais		Modelo de regressão MQO		Modelo *switching regression* regressão com seletividade corrigida	
	Setor público	Setor privado	Setor público	Setor privado	Setor público	Setor privado
Total						
Setor público – setor privado						
Log-salários padronizados	7,3059	6,5650	6,3773	6,1803	6,3643	6,188
Erro-padrão linearizado	(0.0066)**	(0.0025)**	(0.0121)**	(0.0082)**	(0.0123)**	(0.0081)**
Diferença nas médias	0.7410		0.1970		0.1763	
(test t)	(0,0061)**		(0,0146)**		(0,0147)**	
Hiato salarial (%) diferença no exp (salários padronizados)	109,78 %		21,77 %		19.28 %	
Mulheres						
Setor público – setor privado						
Log-salários padronizados	7.2238	6.4946	6,2404	6,0670	6,2522	6,2351
Erro-padrão linearizado	(0.0082)**	(0.0040)**	(0,0106)**	(0,0079)**	(0,0105)**	(0,0078)**
Diferença nas médias	0.7293		0,1731		0,0172	
(test t)	(0,0082)**		(0,0132)**		(0,0131)**	
Hiato salarial (%) diferença no exp (salários padronizados)	107,34%		18,93 %		1,74 %	
Homens						
Setor público – setor privado						
Log-salários padronizados	7,4133	6,6057	6,4325	6,3118	6,4301	6,3163
Erro padrão linearizado	(0,0107)**	(0,0032)**	(0,0151)**	(0,0088)**	(0,0155)**	(0,0089)**
Diferença nas médias	0,8076		0,1207		0,1138	
(test t)	(0,0092)**		(0,0175)**		(0,0035)**	
Hiato salarial (%) diferença no exp (salários padronizados)	124,25 %		12,83 %		12,05 %	

Fonte: PNAD/IBGE (2009).

Nota: Erro-padrão linearizado entre parênteses; ** nível de significância de 5%.

5 CONCLUSÕES

Este estudo analisa a possibilidade da existência de um viés de seleção no cálculo do hiato salarial entre os setores público e privado no Brasil. Com base em um modelo de regressão com mudança endógena (*endogenous switching regression model*), são analisados fatores que incentivam os indivíduos a trabalhar no setor público e seus impactos sobre o cálculo do diferencial salarial público-privado no Brasil são investigados.

Os resultados demonstram que os salários do setor público são mais altos do que aqueles do setor privado no Brasil. Em particular, o diferencial público-privado entre mulheres é mais alto do que entre homens, quando o modelo só trata das características observáveis dos trabalhadores. Fatores importantes que influenciam a decisão de se trabalhar no setor público estão relacionados ao nível de escolaridade do trabalhador, resultados estes amplamente consistentes com a maior parte da literatura existente sobre o tema (PANIZZA; QIANG, 2005; HEITMUELLER, 2004; STELCNER; VAN DER GAAG;VIJVERBERG, 1989), entre outros.

Por fim, vale ressaltar que o presente trabalho utilizou procedimentos de modelagem estatística relacionados com o desenho amostral complexo da PNAD. Os resultados indicam que o uso de tais procedimentos não gera discrepâncias quantitativas e qualitativas notáveis em relação às estimações sem levar em conta os métodos de amostragem adotados pela PNAD.

Diversas são as direções possíveis para trabalhos futuros sobre o tema. A primeira concerne ao cálculo dos diferenciais salariais para ocupações específicas dos trabalhadores do setor público e do setor privado. Os grupos dos trabalhadores do setor público e do setor privado diferem devido a fatores institucionais e há heterogeneidade significativa (intragrupos) em relação aos tipos de ocupação ou região, por exemplo.

Outra extensão interessante deste tema se relaciona à decomposição do diferencial de salários em um montante referido aos retornos das características produtivas do trabalhador, às dotações e aos termos relativos ao viés de seleção amostral do problema de escolha endógena tratado neste estudo. Uma decomposição (à Oaxaca-Blinder) permite identificar, assim, os principais fatores responsáveis pelas discrepâncias salariais entre os setores público e privado no Brasil.

A inclusão no modelo de escolha setorial de variáveis relacionadas com o *background* familiar do agente também se revela uma importante extensão deste trabalho. Como Bender (2003) e Hartog e Ooesrbeek (1993) apontam, variáveis como nível de educação dos pais e ocupação impactam de forma significativa a escolha individual do emprego no setor público e no setor privado. Outra variável de escolha de setor não observável está relacionada com a aversão ao risco do agente. Os funcionários públicos são mais avessos ao risco do que aqueles que trabalham no setor privado? Tais incertezas requerem análises futuras. Em um estudo recente, Pfeifer (2008), com base de dados da Alemanha sugere que os indivíduos mais avessos ao risco buscam, de fato, emprego no setor público e que a tomada de risco é recompensada com salários mais altos para trabalhadores no setor privado. Fica claro que estes e outros resultados devem ser aplicados para a evidência empírica existente.

REFERÊNCIAS

AMEMIYA, T. **Advanced econometrics**. Cambridge, MA: Harvard University Press, 1985.

BARBOSA, A. L. H. **Diferencial salarial público-privado**: uma resenha da literatura. 2009. Mimeografado.

_____; BARBOSA FILHO, F. H. **Public-private wage gap and endogenous sector choice in Brazil**. 13th Latin American and Caribean Economic Association (LACEA), Rio de Janeiro: nov. 2008.

BECKER, G. Investment in human capital: a theoretical analysis. **Journal of Political Economy**, v. 70, Oct. 1962 (Special supplement).

_____. **Human capital**: a theoretical and empirical analysis. Nova York: Columbia University Press, 1975.

BELLUZO, W.; PAZELLO, E.; ANUATTI-NETO, F. Distribuição de salários e diferencial público-privado no Brasil. **Revista Brasileira de Economia**, v. 59, n. 4. p. 511-533, out./dez. 2005.

BENDER, K. A. The central government private sector wage differential. J**ournal of Economic Surveys**, v. 12, n. 2, p. 177-220, 1998.

_____. Examining equality between public and private sector wage distributions. **Economic Inquiry**, v. 41, p. 62-79, 2003.

_____.; FERNANDES, R. Gastos públicos com pessoal: uma análise de emprego e salário no setor público brasileiro no período 1992-2004. In: ENCONTRO NACIONAL DE ECONOMIA, 24. **Anais** ... Salvador: ANPEC, 2006.

BRAGA, B. G. **Educação, experiência e o hiato salarial entre o setor público e privado no Brasil**. Tese (Mestrado) – Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2008.

CAMERON, A. C.; TRIVERDI, P. K. **Microeconometrics** – methods and applications. Cambridge University Press, 2005.

CARVALHO, A. P. **Decomposição do diferencial de salários no Brasil em 2003**: uma aplicação dos procedimentos de Oaxaca e Heckman em pesquisas amostrais complexas. 2005. Dissertação (Mestrado) - Escola Nacional de Ciências Estatísticas/IBGE, Rio de Janeiro, 2005.

COCHRAN, W. G. **Sampling techniques**. 3. ed. New York: J. Wiley, 1977.

EHRENBERG, R. G.; SCHWARTZ, J. L. Public sector markets. In: ASHENFELTER, O.; LAYARD, R. (Ed.). **Handbook of Labor Economics**. Amsterdam: North-Holland, 1986. v. 2, p. 1.219-1.268.

FOGUEL, M. N. *et al.* The public-private wage gap in Brazil. **Revista Brasileira de Economia**, v. 54, n. 4, p. 433-472, out./dez., 2000.

GREGORY, R. G. Public sector pay. In: GREGORY, M.; THOMPSON, A. (Ed.). **A portrait of pay, 1970-1982**: an analysis of the new earinings survey. Oxford: Clarendon Press, 1990. p. 172-205.

_____.; BORLAND, J. Recent developments in public sector labor markets. In: ASHENFELTER, O. C.; CARD, D. (Ed.). **Handbook of Labor Economics**. North-Holland, Amsterdam, 1999. v. 3 cap. 53, p. 3.573-3.630.

HARTOG, J.; OOSTERBEEK, H. Public and private sector wages in the Netherlands. **European Economic Review**, v. 37, p. 97-114, 1993.

HECKMAN, J. J. Sample selection bias as a specification error. **Econometrica**, v. 47, n. 1, p. 153-161, 1979.

HECKMAN, J. J.; HONORE, B. E. The empirical content of the Roy Model. **Econometrica**, v. 58, n. 5, p. 1.121-1.149, Sept. 1990.

HECKMAN, J. J.; LOCHNER, L. J.; TODD, P. E. **Fifity years of Mincer earnings regressions**. Bonn: Institute for the Study of Labor (IZA), 2003 (Discussion Paper, n. 775).

HEITMUELLER, A. **Public-private sector wage differentials in Scotland**: an endogenous switching model. Bonn: Institute for the Study of Labor (IZA), 2004 (Discussion Paper, n. 992).

IBGE. PNAD. **Notas Técnicas, 2006**. Rio de Janeiro: IBGE, 2006.

KASSOUF, A. L. Wage gender discrimination and segmentation in the Brazilian labor market. **Economia Aplicada**, v. 2, n. 2, 1998.

KILLINGSWORTH, M.; HECKMAN, J. J. Female labor supply: a survey. In: ASHENFELTER, O.; LAYARD, R. (Ed.). **Handbook of Labor Economics**. Amsterdam: Elsevier Science Publishers, 1986. v. 1.

KISH, L. **Survey Sampling**. Nova York: Willey, 1965.

LEE, L. F. Unionism and wages rates: a simultaneous equations model with qualitative and limited dependent variables. **International Economic Review**, v. 19, p. 415-433, 1978.

LOKSHIN, M.; SAJAIA, Z. Maximum likelihood estimation of endogenous switching regression models. **Stata Journal**, v. 4, n. 2, p. 282-289, 2004.

MADDALA, G. S. **Limited-dependent and qualitative variables in economics**. Cambridge, UK: Cambridge University Press, 1983.

_____. Disiquilibrium, self-selection, and switching models. In: GRILICHES, Z.; INTRILIGATOR, M. D. (Ed.). **Handbook of Econometrics**. 8, Amsterdam: North-Holland, 1986. v. 3, cap. 28, p. 1.633-1.68

MINCER, J. **Schooling, experience and earnings**. Nova York: Columbia University Press, 1974 (National Bureau of Economics Research).

PANIZZA, U.; QIANG, C. Z. Public-private wage differential and gender gap in Latin America: spoiled bureaucrats and exploited women? **The Journal of Socio Economics**, v. 34, p. 810-833, 2005.

PONCZEK, V. P.; BOTELHO, F. B.; DAULINS, E. Evaluating the wage differential between public and private sectors in Brazil. In: ENCONTRO BRASILEIRO DE ECONOMETRIA (EBE), 30. **Anais** ... Salvador: Sociedade Brasileira de Econometria (SBE), 2008.

PFEIFER, C. **Risk aversion and sorting into public sector employment**. Bonn: Institute for the Study of Labor (IZA), 2008 (Discussion Paper, n. 3.503).

ROBINSON, C.; TOMES, N. Union wage differentials in the public and private sector: a simultaneous equations specification. **Journal of Labor Economics**, v. 2, p. 106-127, Jan. 1984.

ROY, A. Some thoughts on the distributional of earnings. **Oxford Economic Papers**, n. 3, p. 135-146, 1951.

SANTOS, G. C.; *et al*. E. Participation in the Brazilian labor market and income determinants: a heckman's procedure analysis. In: LATIN AMERICAN AND CARIBEAN ECONOMIC ASSOCIATION (LACEA), 13. **Anais**... Rio de Janeiro, nov. 2008.

SANTOS, G. C.; BASTOS, P. M. A.; ROCHA, L. E.V. Determinantes da renda do trabalho no Brasil no período de 2002 a 2006. In: CONGRESSO DA SOBER, 96. **Anais** ... Rio Branco, 2008.

STN. **Consolidação das Contas Públicas de 2007**. Brasília: STN, 2007.

SILVA, P. L. N.; PESSOA, D. C.; LILA, M. F. Análise estatística de dados da PNAD: incorporando a estrutura do plano amostral. **Ciênca & Saúde Coletiva**, v. 7, n. 4, p. 659-670, 2002.

STELCNER, M.; VAN DER GAAG, J.; VIJVERBERG, W. A Switching regression model of public-private sector wage differentials in Peru: 1985-86. **The Journal of Human Resources**, v. 24, n. 3, 545-559, 1989.

TANSEL, A. Public-private choice, wage differentials, and gender in Turkey. **Economic Development and Cultural Change**, University of Chicago Press, v. 53, n. 2, p. 453-477, 2005.

VAZ, D. V.; HOFFMAN, R. Remuneração nos serviços no Brasil: o contraste entre funcionários públicos e privados. **Economia e Sociedade**, Campinas, v. 16, n. 2, p. 30, 2007.

APÊNDICE

DESENHO AMOSTRAL DA PNAD

A PNAD, conduzida pelo IBGE, é de abrangência nacional e fornece informações sobre as características socioeconômicas e demográficas da população brasileira. Realizada pela primeira vez em 1967 e tornando-se anual a partir de 1976, o tamanho amostral da PNAD chega a aproximadamente 116 mil domicílios por pesquisa.[28]

A maior parte das fontes de informações socioeconômicas e demográficas brasileiras se baseia em pesquisas por amostragem complexa. Portanto, elas não utilizam, por questões de custos e de eficiência, amostras aleatórias simples com reposição, em que variáveis são tratadas de forma independente e estatisticamente distribuídas (i.i.d). Entre os diferentes métodos de amostragem probabilística utilizados pela PNAD para determinar sua amostra representativa estão os de amostragem por estratificação e por conglomerados.[29] Uma pesquisa com desenho amostral complexo (como é o caso da PNAD) apresenta métodos de amostragem por estratificação e conglomeração e de probabilidades desiguais que, se não forem levados em conta na estimação de modelos, podem gerar estimativas viesadas e invalidam testes de hipóteses.

28. A PNAD não é realizada em anos de censo demográfico e, por razões excepcionais, não foi realizada no ano de 1994.
29. Para uma descrição de todos os tipos de técnicas de amostragem, ver Cochran (1977).

A amostragem estratificada é caracterizada pela seleção de uma amostra de cada subgrupo da população considerada. Neste método, a população é agrupada em subgrupos relativamente homogêneos (estratos); cada elemento da população deve estar em um único estrato. Assim, os elementos da população são divididos em grupos distintos e fica mais fácil e eficiente escolher, independentemente, uma amostra aleatória simples ou sistemática dentro de cada um desses grupos (estratos). A unidade amostral na amostragem por estratificação, portanto, é o elemento de cada estrato. Este tipo de amostragem tem como principal vantagem aumentar a representatividade da amostra no universo pesquisado ao reduzir o erro amostral, o que resulta em estimativas mais precisas para cada um dos subgrupos. É uma das formas de amostragem mais utilizadas, pois a maior parte das populações é composta por estratos bem definidos, havendo maior homogeneidade entre as unidades amostrais dentro de cada estrato do que entre as unidades amostrais de estratos diferentes.

Na técnica de amostragem por conglomerados, a população também é dividida em grupos e a análise estatística é feita em uma população destes conglomerados. Diferentemente da amostragem por estratificação, a unidade amostral na conglomeração é representada por cada grupo (conglomerado) de elementos com características semelhantes. Assim, neste caso, cada grupo é visualizado como um grupo representativo da população. A utilização de tal método será tanto melhor quanto maior a heterogeneidade da população; como regra geral, o número de elementos em um conglomerado deve ser pequeno em relação ao tamanho da população e o número de conglomerados, razoavelmente grande.

De forma resumida, tanto na amostragem estratificada, como na amostragem por conglomerado, a população deve estar dividida em grupos. Na amostragem estratificada, entretanto, uma amostra aleatória (simples ou sistemática) é selecionada dentro de cada grupo (estrato), enquanto na amostragem por conglomerado selecionam-se amostras aleatórias, simples ou sistemáticas, de grupos, e todos os elementos dentro dos grupos (conglomerados) selecionados farão parte da amostra.

Em geral, estimativas de variância e de desvio-padrão e os parâmetros de algum tipo de modelo são influenciados não somente pelos pesos amostrais, mas também por amostras estratificadas e por conglomerados. Portanto, ignorar o desenho amostral complexo de uma pesquisa compromete a confiabilidade das estimativas pontuais e seus respectivos desvios-padrão, o que afeta o nível de significância e um eventual teste de hipóteses.

A PNAD é realizada por meio de uma amostra probabilística de domicílios estratificada e de conglomeração obtida em três estágios de seleção: unidades primárias – municípios; unidades secundárias –- setores censitários; e unidades terciárias – unidades domiciliares (domicílios particulares e unidades de habitação em domicílios coletivos).[30]

Quanto ao seu plano de estratificação, primeiro, há uma estratificação geográfica que divide o país em 36 estratos "naturais"; 27 correspondem a todos os 27 estados federais. Os 9 estratos restantes correspondem às áreas metropolitanas dos seguintes estados: Pará, Ceará, Pernambuco, Bahia, Minas Gerais, Rio de Janeiro, São Paulo, Paraná e Rio Grande do Sul.[31] Estes 9 estratos consistem de municípios da área metropolitana localizada na capital de cada estado, e há uma estratificação adicional dos municípios.

Com relação à conglomeração, na seleção das unidades primárias e secundárias (municípios e setores censitários) da PNAD da primeira década deste século, foram adotadas a divisão territorial e a malha setorial vigentes em 1º de agosto de 2000 e utilizadas para a realização do Censo Demográfico 2000.

No primeiro estágio, as unidades (municípios) foram classificadas em duas categorias: autorrepresentativas (probabilidade 1 de pertencer à amostra) e não autorrepresentativas. Os municípios pertencentes à segunda categoria passaram por um processo de estratificação e, em cada estrato, foram selecionados com reposição e com probabilidade proporcional à população residente obtida no Censo Demográfico 2000. No segundo estágio, as unidades (setores censitários) foram selecionadas, em cada município da amostra, também com probabilidade proporcional e com reposição, sendo utilizado o número de unidades domiciliares existentes por ocasião do Censo Demográfico 2000 como medida de tamanho. No último estágio foram selecionados, com equiprobabilidade, em cada setor censitário da amostra, os domicílios particulares e as unidades de habitação em domicílios coletivos para investigação das características dos moradores e da habitação.

30. Uma descrição cuidadosa do plano de amostragem da PNAD pode ser encontrada em Silva, Pessoa e Lila (2002). Ver também Notas Técnicas da PNAD de 2006.
31. Estes 9 estratos produzem 18 estratos na medida em que incluem 2 estratos "naturais": um que contém todos os municípios de áreas metropolitanas localizadas na capital do estado, e outro que contém os municípios restantes (nos estados em que as áreas metropolitanas foram selecionadas, um estrato contém uma área metropolitana e outro contém a área estadual restante, excluindo a área metropolitana).

© Instituto de Pesquisa Econômica Aplicada – **ipea** 2012

EDITORIAL

Coordenação
Cláudio Passos de Oliveira

Supervisão
Andrea Bossle de Abreu

Revisão
Eliezer Moreira
Elisabete de Carvalho Soares
Fabiana da Silva Matos
Lucia Duarte Moreira
Luciana Nogueira Duarte
Míriam Nunes da Fonseca

Editoração
Roberto das Chagas Campos
Aeromilson Mesquita
Aline Cristine Torres da Silva Martins
Carlos Henrique Santos Vianna
Maria Hosana Carneiro Cunha

Capa
Luís Cláudio Cardoso da Silva

Projeto Gráfico
Renato Rodrigues Bueno

Livraria do Ipea

SBS – Quadra 1 – Bloco J – Ed. BNDES, Térreo.
70076-900 – Brasília – DF
Fone: (61) 3315-5336

Correio eletrônico: livraria@ipea.gov.br
Tiragem: 500 exemplares

Logotipo/slogan de governo removido em observancia ao disposto no art. 40 da Lei n 9.504/1997

Missão do Ipea
Produzir, articular e disseminar conhecimento para aperfeiçoar as políticas públicas e contribuir para o planejamento do desenvolvimento brasileiro.

ISSN 1415-4765

ipea Instituto de Pesquisa Econômica Aplicada

SECRETARIA DE ASSUNTOS ESTRATÉGICOS DA PRESIDÊNCIA DA REPÚBLICA