



**UNIVERSIDADE FEDERAL DO CEARÁ**  
**FACULDADE DE ECONOMIA, ADMINISTRAÇÃO, ATUÁRIA E**  
**CONTABILIDADE (FEAAC)**  
**PÓS-GRADUAÇÃO EM ECONOMIA (CAEN)**

**DANIEL BARBOZA GUIMARÃES**

**AVALIAÇÕES DE ABORDAGENS METODOLÓGICAS DA EQUAÇÃO DE**  
**RENDIMENTOS APLICADAS EM CENÁRIOS SELECIONADOS DO MERCADO**  
**DE TRABALHO BRASILEIRO**

**FORTALEZA**

**2013**

**DANIEL BARBOZA GUIMARÃES**

**AVALIAÇÕES DE ABORDAGENS METODOLÓGICAS DA EQUAÇÃO DE  
RENDIMENTOS APLICADAS EM CENÁRIOS SELECIONADOS DO MERCADO DE  
TRABALHO BRASILEIRO**

Tese de Doutorado submetida à coordenação do Curso de Pós-Graduação em Economia - CAEN, da Universidade Federal do Ceará como requisito parcial para a obtenção do grau de Doutor em Economia.

Orientador: Prof. Ronaldo de Albuquerque e Arraes

FORTALEZA

2013

Dados Internacionais de Catalogação na Publicação  
Universidade Federal do Ceará  
Biblioteca de Pós Graduação em Economia - CAEN

- 
- G977a      Guimarães, Daniel Barboza  
            Avaliações de abordagens metodológicas da equação de rendimentos aplicadas em cenários selecionados no mercado de trabalho / Daniel Barboza Guimarães. – 2013.  
            117f.f. il. color., enc. ; 30 cm.
- Tese (doutorado) – Programa de Pós Graduação em Economia, CAEN, Universidade Federal do Ceará, Fortaleza, 2013.  
            Orientação: Prof. Dr. Ronaldo de Albuquerque e Arraes
1. Renda – Distribuição. 2. Rendimentos dos Trabalhadores. 3. Mercado de trabalho. 4. Equações de Rendimentos. Título.

---

CDD 339.2

DANIEL BARBOZA GUIMARÃES

AVALIAÇÕES DE ABORDAGENS METODOLÓGICAS DA EQUAÇÃO DE  
RENDIMENTOS APLICADAS EM CENÁRIOS SELECIONADOS DO MERCADO DE  
TRABALHO BRASILEIRO

Tese submetida à Coordenação do Curso de Pós-Graduação em Economia, da Universidade Federal do Ceará, como requisito parcial para a obtenção do grau de Doutor em Economia.

Aprovada em 08 de Agosto de 2013.

BANCA EXAMINADORA

---

Prof. Ronaldo de Albuquerque e Arraes (Orientador)  
Universidade Federal do Ceará – CAEN/UFC

---

Prof. Marcelo Bentez Diniz  
Universidade Federal do Pará – UFPA

---

Dr. Jimmy Lima de Oliveira  
Instituto de Pesquisa Econômica do Ceará - IPECE

---

Prof. Guilherme Diniz Irffi  
Universidade Federal do Ceará – Departamento de Economia Aplicada

---

Prof. Vitor Borges Monteiro  
Universidade Federal do Ceará – Campus de Sobral

## **AGRADECIMENTOS**

Agradeço primeiramente a Deus por ter me capacitado com saúde e por ter me dado uma família que sempre me incentivou e me apoiou em tudo o que busquei.

Aos meus pais, Francisco Teodoro Guimarães e Maria Cecília Guimarães, por terem me incentivado e proporcionado condições de me dedicar apenas aos meus estudos. À minha esposa Caroline e às minhas irmãs Danielly e Daianny por terem sempre ficado ao meu lado e me apoiado e dado força nos momentos de maiores dificuldade durante esta longa caminhada.

Ao Professor Ronaldo Arraes por ter aceitado esta orientação e pelo suporte teórico para a elaboração desta Tese.

Aos Professores Guilherme Irffi e Zilania Mariano e aos colegas de doutorado Cristiano Santos e Daniel Suliano pela parceria e dedicação nas estimativas realizadas neste trabalho.

Aos Profs. Vitor Borges Monteiro, Marcelo Bentez e Jimmy Lima de Oliveira por aceitarem o convite para participar da banca e pelas valiosas sugestões para o trabalho.

Aos colegas de doutorado, em especial aos amigos Elano e Glauber, que desde a graduação se mostraram amigos de verdade.

## RESUMO

Este estudo compreende duas análises para investigar alguns cenários específicos sobre os rendimentos de trabalhadores brasileiros. A primeira trata de investigar os principais determinantes dos rendimentos dos filhos, dentre os quais a ocorrência de transmissão das rendas dos pais. Estimacões de equações de rendimentos com distintos procedimentos metodológicos permitem captar o efeito transmissão (herança familiar) e o efeito dinâmico (análise de gerações superpostas), cujas verificações empíricas são baseadas nos dados da Pesquisa Nacional por Amostra de Domicílios (PNAD) para os anos de 1989 e 2009. São utilizados os métodos de Heckman em dois estágios e o modelo de variáveis instrumentais, haja vista a necessidade de correção do viés de seleção amostral e a endogeneidade presente em equações de rendimentos. Além destes, aplica-se o método Klein–Vella, por ser um procedimento corretivo para se obter o verdadeiro efeito médio, e não apenas um efeito para os indivíduos que são realmente afetados por variações nos instrumentos. Em seguida, devido à impossibilidade de se obter *contrafactuais* perfeitos nos dados da PNAD, estimam-se regressões quantílicas no sentido de aumentar o grau de confiabilidade nas estimativas através da homogeneização dos trabalhadores na amostra por faixas de rendimento. Das alternativas metodológicas, constatou-se que o método de variáveis instrumentais apresentou o pior ajuste preditivo. Quanto às demais abordagens, suas estimativas não permitiram concluir haver superioridade estatística de uma sobre a outra para fins de inferência. Portanto, pode-se concluir a partir destas que as rendas dos pais, principalmente a renda da mãe, exercem um impacto positivo e significativo de transmissão sobre as rendas dos seus filhos. Apesar da comprovação do efeito de transmissão de renda, este se mostrou com tendência declinante ao longo do tempo. A segunda análise visa testar possíveis diferenças de rendimentos por gênero em três níveis regionais (Brasil, Nordeste e Sul-Sudeste) e quatro níveis ocupacionais (todas as ocupações, dirigentes, serviços e técnicos). Para tanto, foram avaliados três enfoques metodológicos para estimacões da equação *minceriana* de rendimentos. Primeiramente avaliam-se os métodos de Mínimos Quadrados Ordinários e de Heckman em dois estágios para corrigir o viés de seleção. Para testar a existência de discriminação por gênero no mercado de trabalho emprega-se o método de *Propensity Score Matching*. Para inferir sobre tal discriminação aplica-se o método Oaxaca–Blinder, o qual permite decompor as diferenças salariais devidas às características individuais (explicada) e à discriminação (não explicada).

Independentemente do método utilizado nas estimações realizadas, comprova-se que existe diferença salarial a favor dos homens, cuja superioridade varia de 36,32% (Dirigentes no Brasil) a 118,72% (Técnicos no Nordeste). Além disso, conclui-se que tal diferença é resultado da presença de discriminação por gênero, uma vez que, em todas as estimações realizadas, o efeito da discriminação, o qual varia de 100,5% (Serviços no Sul-Sudeste) a 153,8% (Técnicos no Brasil), é maior que a diferença dos rendimentos, pois as mulheres têm, em média, atributos adquiridos (escolaridade, por exemplo) superiores aos dos homens. Portanto, caso não houvesse discriminação, há evidências na amostra utilizada de que os rendimentos das mulheres tenderiam a ser, em média, superiores aos dos homens.

Palavras-Chave: Transmissão Intergeracional de Renda, Discriminação por Gênero, Avaliações de Abordagens Metodológicas, Equações de Rendimentos.

## ABSTRACT

This study comprises two analyses to investigate some specific scenarios related to Brazilian workers' earnings. The first analysis focuses on the main determinants of the sons' and daughter's earnings, among which the incidence of transmission of parental income. Estimations of earning equations with different methodological procedures allow to capture the transmission effect and the dynamic effects (analysis of overlapping generations), whose empirical findings are based on data from the National Household Sample Survey (PNAD) for the years 1989 and 2009. Methods of Heckman's two-stage and instrumental variables models are used, in order to correct for sample selection bias and endogeneity in the model. In addition, it is applied Klein-Vella method as a corrective procedure to obtain the true average effect, not only an effect for individuals who are truly affected by changes in the instruments. Then, due to the impossibility of obtaining perfect counterfactual from PNAD data, quantile regressions are estimated to increase the reliability of estimates through homogenization of workers in the sample by income ranges. It was found that instrumental variables method presented the worst predictive fitting amongst the methodological procedures. Regarding the other approaches, the estimates did not show any statistical superiority of one over the others for inference purposes. Therefore, based upon anyone of these procedures it could be concluded that the parents' incomes, especially the mother's ones, play a positive and significant transmission impact on their sons' and daughter's earnings, although it has presented a declining tendency over time. The second analysis aims to test earning differences by gender at three regional levels (Brazil, Northeast and South-Southeast) and four occupational categories (all occupations, managers, technicians, general services). Three methodological approaches were evaluated in the estimation of Mincerian earnings equation. Firstly, the OLS and two-stage Heckman methods were performed to correct for selection bias. In order to test the existence of gender discrimination in the labor market it was initially employed the propensity score matching method, and then applied the Oaxaca-Blinder method to measure the magnitude of such discrimination, since it allows decomposing the earnings differences due to individual characteristics (explained) and discrimination (unexplained). Regardless of the method used in the estimations carried out, there is strong evidence toward an earning gap in favor of men, whose superiority ranges from 36.32% (managers in Brazil) to 118.72% (technicians in the Northeast). These differences are due to the presence of discrimination by gender since the effect of discrimination, which varies from

100.5% (services in South-Southeast to 153.8% (technicians in Brazil), is greater than the difference in earnings because women have, on average, acquired attributes (education, for example) higher than those of men. Therefore, if there were no discrimination in the labor market, there is evidence in the sample that men's earnings tend to be lower, on average, than those of women.

**Keywords:** Intergenerational Transmission of Income, Gender Discrimination, Evaluations of Methodological Approaches, Earning Equations.

## **LISTA DE FIGURAS**

Gráfico 1.1 - Síntese dos resultados para diferença salarial entre homens e mulheres, por regiões e ocupações.....	107
--	-----

## LISTA DE TABELAS

Tabela 1.1 - Estatísticas Descritivas das Variáveis, por Regiões e Ocupações, 1989 e 2009..	44
Tabela 1.2 – Estimação da Equação de Rendimentos por Alternativos Métodos em Todas as Ocupações. ....	46
Tabela 1.3 – Estimação da Equação de Rendimentos por Alternativos Métodos no Setor de Serviços. ....	47
Tabela 1.4 – Estimação da Equação de Rendimentos por Alternativos Métodos no Setor dos Profissionais das Indústrias de Produção e Transformação. ....	48
Tabela 1.5 – Estimação da Equação de Rendimentos por Alternativos Métodos com <i>Dummies</i> Regionais em Todas as Ocupações. ....	52
Tabela 1.8 – Estimação da Equação de Rendimentos através do Método de Regressão Quantílica em Todas as Ocupações. ....	57
Tabela 1.9 – Estimação da Equação de Rendimentos através do Método de Regressão Quantílica no Setor de Serviços.....	58
Tabela 1.10 – Estimação da Equação de Rendimentos através do Método de Regressão Quantílica no Setor dos Profissionais das Indústrias de Produção e Transformação. ....	59
Tabela 2.1 - Estatísticas Descritivas entre Homens e Mulheres, por Regiões e Ocupações, 2009.....	82
Tabela 2.2 - Estimações da Equação <i>Minceriana</i> por MQO por Regiões e Ocupações.....	94
Tabela 2.3 - Estimações da Equação <i>Minceriana</i> de Salários pelo Método de Heckman em Dois Estágios, por Regiões e Ocupações .....	98
Tabela 2.4 - Diferenciais de Renda entre Tratados e Controles com e sem Pareamento.....	100
Tabela 2.6 - Síntese dos Resultados das Diferenças Salariais Por Gênero. ....	107

## LISTA DE QUADROS

Quadro 1.1: Descrição das Variáveis retiradas da PNAD.....	40
Quadro 2.1 - Descrição das Variáveis retiradas da PNAD. ....	80

## SUMÁRIO

DIFERENCIAIS DE RENDIMENTOS INDIVIDUAIS E OS EFEITOS DE GERAÇÕES SUPERPOSTAS .....	13
1 INTRODUÇÃO .....	13
2 REVISÃO DA LITERATURA.....	17
2.1 Teorias sobre Mobilidade Intergeracional.....	18
2.2 Estudos Empíricos sobre Mobilidade Intergeracional .....	21
3 METODOLOGIA .....	29
3.1 Equações <i>Mincerianas</i> .....	29
3.2 Problemas Associados à Estimação de uma Equação de Rendimentos.....	33
3.2.1 <i>Método de Heckman em Dois Estágios</i> .....	34
3.2.2 <i>Método de Variáveis Instrumentais</i> .....	36
3.2.3 <i>Método de Klein e Vella</i> .....	37
3.2.4 <i>Regressão Quantílica</i> .....	38
3.3 Base de Dados.....	40
3.3.1 <i>Análise Descritiva dos Dados</i> .....	42
4 RESULTADOS DAS ESTIMAÇÕES DA EQUAÇÃO <i>MINCERIANA</i> .....	45
5 CONSIDERAÇÕES FINAIS.....	62
DIFERENCIAL DE SALÁRIO POR GÊNERO NAS OCUPAÇÕES E REGIÕES BRASILEIRAS.....	65
1 INTRODUÇÃO.....	65
2 REVISÃO DA LITERATURA.....	69
2.1 Estudos Empíricos sobre Diferenciais Salariais .....	71
3 METODOLOGIA.....	78
3.1 Modelo Empírico .....	78
3.2 Base de Dados.....	79
3.2.1 <i>Análise Descritiva dos Dados</i> .....	81
3.4 Avaliação do Impacto.....	83
3.4.1 <i>Propensity Score Matching</i> .....	86
3.4.1.1 <i>Um para Um</i> .....	87
3.4.1.2 <i>Pareamento por Vizinho mais Próximo</i> .....	87
3.4.1.3 <i>Distância (Raio)</i> .....	88

3.4.1.4 <i>Kernel Matching</i> .....	88
3.5 Decomposição dos Diferenciais Salariais .....	89
4 ANÁLISE E DISCUSSÃO DOS RESULTADOS .....	91
4.1 Estratégia 1: Estimções por MQO.....	92
4.2 Estratégia 2: Método de Heckman em dois estgios, razão inversa de Mills.....	95
4.3 Estratégia 3: Propensity Score Matching .....	99
4.4 Decomposição de Oaxaca-Blinder.....	102
4.5 Sínteses dos Resultados.....	105
5 CONSIDERAÇÕES FINAIS.....	108

## DIFERENCIAIS DE RENDIMENTOS INDIVIDUAIS E OS EFEITOS DE GERAÇÕES SUPERPOSTAS

### 1 INTRODUÇÃO

Apesar do declínio observado nos níveis de desigualdade de renda no Brasil em anos recentes, esta permanece extremamente elevada. Nesse sentido, Barros et al (2007) afirmam que, caso fosse mantida essa taxa de desaceleração da desigualdade no futuro, seriam necessários mais de 20 anos para o país atingir níveis similares aos da média dos países com maior grau de desenvolvimento. De acordo com o Relatório Territorial Brasil (2013), publicado pela OCDE, mesmo com o forte investimento do governo em programas de redução da pobreza nos últimos anos, o Brasil ainda é um dos países com maior desigualdade social do mundo apresentando a segunda pior distribuição de renda em *ranking* da OCDE.

Pero e Szerman (2008) argumentam que uma fonte potencial de manutenção da desigualdade é a transmissão intergeracional de renda, pois, se filhos de pais com maiores/menores níveis de renda tendem a possuir níveis equivalentes de renda, então, a desigualdade seria transmitida através das gerações. Os autores argumentam ainda que a transmissão intergeracional de *status* econômico, além de estar relacionada com a desigualdade de resultados, é identificada como a parte da desigualdade relativa às oportunidades, na qual se concentram, aproximadamente, 2/3 da desigualdade observada no Brasil, sendo, portanto, pior, sob a perspectiva da justiça social, do que a de resultados. Assim, políticas redistributivas apropriadas são justificadas por evidências de baixa mobilidade intergeracional.

Dunn (2007) argumenta que a influência gerada pela transmissão de salários sobre a desigualdade ao longo do tempo pode afetar significativamente o impacto das políticas públicas destinadas a reduzir a desigualdade transversal, isto é, entre indivíduos. Uma vez que o nível de transmissão intergeracional de salários também tem implicações para a manutenção da pobreza e arrefecimento do desenvolvimento econômico, conclui-se que a continuidade da transmissão de *status* econômico entre gerações é diretamente relacionada com o grau de persistência da desigualdade ao longo do tempo. Não obstante a importância desempenhada pelos efeitos da mobilidade intergeracional de renda para a determinação dos diferentes

componentes de aferição da desigualdade, ainda há escassez de estudos que têm se dedicado a analisar o tema no Brasil.

Diante do exposto, esta análise tem como principal objetivo verificar os principais determinantes dos diferenciais de rendimento entre pais e filhos no Brasil. No entanto, Menezes Filho (2002) argumenta que existe uma série de problemas associados a qualquer tentativa de estimar uma equação de rendimentos. Segundo o autor, o principal objetivo associado a qualquer procedimento que visa estimar tal equação, especialmente se o objetivo ulterior é simulação, diz respeito à questão causal.

Além disso, Menezes Filho (2002) afirma que, mesmo tendo uma interpretação causal, a estimativa gerada, a partir de uma equação *minceriana* de salários, pode ser válida apenas para a amostra em questão ou pode variar para diferentes valores das variáveis explicativas na própria amostra. Segundo o autor, a maneira mais simples de lidar com relações possivelmente não-lineares é através do uso de variáveis interativas, de termos quadráticos, como por exemplo, idade ao quadrado, ou através de variáveis *dummies* correspondendo a cada grupo em particular, como por exemplo, *dummies* regionais. Isso tem sido corrigido por Mariano e Arraes e Mariano (2011) ao implementarem coeficientes aleatórios para estimar a taxa de retorno da educação. Koenker e Basset (1978) *apud* Menezes Filho (2002) afirmam que os modelos de *matching* e regressões quantílicas são procedimentos mais sofisticados para tratar o problema de heterogeneidade presente na relação entre salários e características pessoais.

Dessa forma, para corrigir a questão do viés de seleção amostral e da endogeneidade presente em uma equação de salários, este artigo, assim como também aplicaram Mariano e Arraes (2011) e Suliano e Siqueira (2010), fará uso do método de Heckman em dois estágios e do modelo de variáveis instrumentais.

No entanto, uma vez que diferentes instrumentos produzem resultados distintos, os coeficientes estimados pelo método de variáveis instrumentais representam efeitos causais apenas para a subamostra de *compliers*, ou seja, indivíduos que são realmente afetados por variações nos instrumentos. Portanto, há a necessidade de se implementar uma metodologia

adicional que trate de tal correção. A metodologia desenvolvida por Klein e Vella (2010) é um procedimento corretivo para se obter o verdadeiro efeito médio, e não apenas um efeito local, conforme tem sido aplicado por Saniter (2012) para mensurar os retornos da educação sobre os rendimentos dos trabalhadores alemães. Portanto, além das duas técnicas já difundidas na literatura, Heckman em dois estágios e variáveis instrumentais, este artigo contribui com a literatura brasileira voltada à estimação da equação de rendimentos ao fazer uma aplicação dessa metodologia, permitindo, assim, confrontar com outros resultados obtidos no Brasil.

Diante do exposto, inicialmente, serão realizadas estimações da equação *minceriana* pelas três técnicas supracitadas, no intuito de verificar qual procedimento, ou quais dos procedimentos, apresentará o melhor ajuste, para estimar algumas variações da equação de rendimentos. Devido à indisponibilidade de uma base de dados apropriada, ou seja, um *survey*, este trabalho tratará da questão de transmissão de renda em uma única geração. Portanto, os resultados que serão apresentados serão interpretados como um possível efeito de transmissão de renda dos pais para os seus filhos. Como única alternativa disponível, a pesquisa utilizará dados da Pesquisa Nacional por Amostra de Domicílios (PNAD) referentes aos anos de 1989 e 2009, visto que, além de verificar o nível de transmissão de renda dos pais para os seus filhos em uma única geração, pretende-se captar possíveis variações de transmissão nas últimas décadas.

Além disso, Angrist e Krueger (1999) *apud* Menezes Filho (2002) afirmam que, a maneira ideal de garantir que a relação entre duas variáveis seja causal seria trabalhar com *contrafactuais*, ou seja, observar o mesmo trabalhador em duas situações diferentes, controlando por todos os demais condicionantes salariais (inclusive o tempo). Portanto, devido à incapacidade de se obter *contrafactuais* perfeitos a partir dos dados da PNAD, este artigo utilizar-se-á da metodologia de regressões quantílicas na tentativa de homogeneizar o máximo possível a amostra de valores.

Convém mencionar que, diferentemente dos demais trabalhos realizados para o Brasil, tais como os de Suliano e Siqueira (2010) e Mariano e Arraes (2011), este artigo visa, principalmente, captar um possível efeito transmissão de renda entre pais e filhos abordando

duas gerações distintas, através da inclusão dos salários do pai e da mãe e de uma *dummy* de tempo. Portanto, pode-se dizer que, através da estimação de uma equação *minceriana* de salários, este trabalho busca captar dois impactos, ou seja, um efeito transmissão (herança familiar) e um dinâmico (análise de gerações superpostas). Para isso, serão realizadas extensões da equação *minceriana*, como por exemplo, a inclusão de interações entre as variáveis de renda do pai e da mãe com *dummies* de gênero, ano e região, objetivando captar possíveis divergências de impacto das rendas paternas e maternas sobre os salários dos seus filhos e filhas levando em consideração o ano e as regiões brasileiras.

Este artigo está organizado de modo a contemplar, além desta introdução, mais quatro seções, iniciando-se com uma exposição dos trabalhos julgados mais relevantes que tratam sobre diferenciais de rendimento entre indivíduos. Em seguida estão descritos os dados e os procedimentos metodológicos aplicados. As duas seções finais são compostas de descrições dos resultados que permitirão extrair as principais conclusões sobre os objetivos e hipóteses propostos.

## 2 REVISÃO DA LITERATURA

Historicamente, as ciências humanas, notadamente Economia e Sociologia, tem contribuído com investigações em diferentes vertentes metodológicas acerca da interpretação sobre a mobilidade dos indivíduos na sociedade, referenciada na forma como tal mobilidade pode alterar a distribuição de riqueza e de poder, para a geração de bem-estar e igualdade de oportunidades. A análise sociológica de mobilidade dá-se por meio do movimento entre classes sociais ou *status* socioeconômico dos indivíduos, enquanto a econômica trata a mobilidade de renda e alocação nos estratos ocupacionais pelos indivíduos.

Antigo (2010) indica que a mobilidade de rendimentos (renda, salários) não decorre de um problema de escolha individual, portanto, a literatura, que estuda este fenômeno, é voltada para aspectos metodológicos de sua mensuração e à aplicação empírica. De acordo com a autora, a mobilidade de rendimentos é um reflexo de decisões tomadas *ex-ant* pelos indivíduos em investimento educacional, além de fatores independentes como condições macroeconômicas, fatores do mercado de trabalho e de cunho institucional que podem contribuir para a trajetória desse fenômeno.

A análise sobre a mobilidade pode ainda ser realizada de duas formas diferentes. No estudo da mobilidade intrageracional, compara-se o desempenho social de uma determinada pessoa ou grupo de pessoas com características comuns, durante um determinado período de tempo, desde a entrada no mercado de trabalho, até a posição ocupacional presente, por exemplo. Já a mobilidade intergeracional realiza uma comparação da posição social entre as gerações passadas e a geração atual, como, por exemplo, a condição de um filho em relação à condição de seu pai.

Diante disso, esta seção será contemplada em duas subseções. Primeiramente, será realizada uma discussão em torno dos trabalhos teóricos, que buscaram enfatizar os modelos e os determinantes da mobilidade de renda. Em seguida serão relacionados alguns trabalhos empíricos que buscaram estimar medidas de mobilidade, tanto na literatura internacional como na nacional.

## 2.1 Teorias sobre Mobilidade Intergeracional

A literatura teórica que trata a questão da mobilidade intergeracional tem se desenvolvida em duas abordagens principais (PETERS, 1992). A primeira, desenvolvida por Blau e Duncan (1967) foca diretamente a mobilidade intergeracional no sentido de avaliar o grau pelo qual o *status* socioeconômico dos pais afeta o do filho, ou seja, a posição na distribuição de renda é que está sendo estudada, e não o nível de renda em si. Portanto, os níveis de mobilidade social são medidos por correlações intergeracionais de *status* socioeconômico<sup>1</sup> e variáveis como educação e renda são utilizadas como *proxy* para o *background* familiar.

Conforme mencionado anteriormente, a análise de mobilidade intergeracional de *status* econômico e social tem uma longa tradição na Sociologia e por meio dela torna-se possível mensurar se esta reflete uma possível igualdade de oportunidades em uma sociedade. A igualdade de oportunidades pode ser ilustrada pela ascensão social das pessoas que não contam com *background* familiar favorável a isso.

Além da análise por *status* socioeconômico, Antigo (2010) indica que autores de formação sociológica também buscam responder à questão da mobilidade através do movimento entre classes sociais. Nessa abordagem, a mobilidade intergeracional é analisada por meio de classificações *cross-section* de ocupações de pais e filhos. Na literatura, essa análise é desenvolvida por meio do cálculo de índices alternativos que permitem comparações de classificações para diferentes sociedades ou, ainda, dentro de subgrupos de indivíduos pertencentes a uma mesma sociedade. As diferenças entre os grupos ocupacionais são determinadas a partir da relação com os meios de produção ou de acordo com a posição no mercado de trabalho, sem necessariamente expressar uma hierarquia social.

A segunda abordagem, desenvolvida por Jenks et al. (1979), Becker (1981) e Becker e Tomes (1986), é intimamente relacionada à literatura de capital humano e foca nos determinantes dos níveis de renda (PETERS, 1992). Esses autores foram os primeiros a discutir a questão da mobilidade de renda entre gerações e a considerar o papel da educação sugerindo, assim, um modelo teórico de transmissão intergeracional de *status* familiar.

---

<sup>1</sup> Para maiores detalhes ver Jencks (1990).

Em uma versão simplificada do modelo, Becker e Tomes (1979, *apud* Antigo, 2010) consideram a família composta por um único indivíduo em cada geração, e assumem que o capital humano do filho é escolhido pelo pai, como decisão da alocação ótima de sua renda permanente. Dessa forma, quando os rendimentos do pai aumentam, maior será o investimento em capital humano do filho, levando-o, portanto, a maiores rendimentos. Assim, a renda permanente do pai tem uma influência positiva nos rendimentos do filho.

Além disso, os autores ressaltam uma segunda fonte de correlação entre os rendimentos, resultante das habilidades entre pai e filho, em que fenômenos genéticos e sociais como transmissão de QI, rede social e preferências são considerados. Desse modo, a transmissão intergeracional de rendimentos de pai para filho é independente das decisões de investimento em capital humano dos filhos por parte dos pais e de sua restrição orçamentária.

A partir da abordagem desenvolvida por Becker e Tomes (1979), os estudos de mobilidade têm início na teoria econômica, onde um crescente grupo de pesquisadores passaram a estimar uma medida de transmissão de *status* socioeconômico. No entanto, os estudos sobre mobilidade só vieram a se desenvolver, dentro da literatura econômica, nos anos de 1990, a partir da disponibilidade de dados em painéis internacionais. Autores como Peters (1992), Fields e Ok (1996), Fortin e Lefebvre (1998), Solon (1999), Björklund e Jäntti (2000), Behrman et al (2001), Erikson e Goldthorpe (2002), Corak (2004), Ermisch e Nicoletti (2005) e Aydemir et al (2005) são alguns exemplos que abordam essa questão.

Devido à limitação de dados organizados em painel, alguns autores como, por exemplo, Ermisch e Francesconi (2004) estimam a elasticidade intergeracional por meio de um escore de prestígio ocupacional (*Hope-Goldthorpe score*)<sup>2</sup> de pais e filhos, partindo da constatação de que estes são fortemente relacionados aos rendimentos dos indivíduos. No entanto, conforme indicam Ermisch e Francesconi (2002) essa medida é comumente utilizada em Sociologia, mas raramente é usada nos trabalhos da área de Economia.

Tendo em vista a inexistência de uma base de dados que contemple pais e filhos entre gerações, Arellano e Meghir (1992) e Angrist e Krueger (1992) propuseram um procedimento alternativo para contornar esse problema através da combinação de duas

---

<sup>2</sup> Este índice foi desenvolvido por Goldthorpe e Hope (1974). Para maiores detalhes sobre essa metodologia, ver Ermisch e Francesconi (2002).

amostras distintas, cujo método denominou-se *Two-Sample Instrumental Variables* (TSIV). No primeiro estágio utilizam-se informações de uma amostra de pais para estimar uma equação do tipo *minceriana*, cujos coeficientes dos regressores obtidos são aplicados, no segundo estágio, em outra amostra de filhos para se obter a renda predita dos pais. Finalmente, estima-se por Mínimos Quadrados Ordinários (MQO), uma equação tendo a renda do filho como função da renda do pai. Com relação a este procedimento, Dunn (2007) indica que a inovação de duas amostras é importante, visto que a amostra utilizada no primeiro estágio não precisa incluir os verdadeiros pais dos filhos observados na segunda amostra. O único requerimento<sup>3</sup> para consistência é que a relação entre salário e educação dos pais no primeiro estágio, seja a mesma para os pais dos filhos observados na amostra do segundo estágio.

Após a metodologia proposta por Arellano e Meghir (1992) e Angrist e Krueger (1992), autores como Dearden, Machin, e Reed (1997), Björklund e Jäntti (1997), e Grawe (2001, 2004), dentre outros, aplicaram uma estrutura similar em estudos sobre mobilidade intergeracional, e que passou a ser denominada de *Two-Sample Two-Stage Least Squares* (TS2SLS). De acordo com Inoue e Solon (2006) o TS2SLS é uma técnica computacionalmente mais conveniente e produz estimadores assintoticamente mais eficientes que os de TSIV. Além disso, esses autores indicam que o estimador de TS2SLS difere do de TSIV por inserir uma matriz, que pode ser vista como um termo de correção para as diferenças entre as duas amostras<sup>4</sup>.

Além das metodologias econométricas supracitadas, trabalhos como os de Ferreira e Veloso (2006), Figueiredo, Netto Junior e Porto Junior (2007), Figueiredo (2010) e Antigo (2011) também utilizam as matrizes de transição markoviana nos estudos sobre mobilidade. Tais matrizes apresentam as probabilidades de mobilidade intergeracional nas frações de distribuição de renda (quantis, decis, dentre outros). Portanto, a partir do uso dessas matrizes pode-se verificar padrões de não-linearidade nos processos de mobilidade.

---

<sup>3</sup> Embora este requerimento não possa ser testado, ele é indicado para selecionar pais na primeira amostra com características observáveis similares às dos pais dos filhos observados na segunda amostra.

<sup>4</sup> Para um melhor entendimento das diferenças entre os estimadores de TSIV e TS2SLS, ver Inoue e Solon (2006).

A seguir serão apresentados alguns trabalhos empíricos abrangendo a literatura de mobilidade de renda intergeracional, os quais abordam principais fatos estilizados relacionados ao tema, bem como as bases metodológicas dessas investigações.

## 2.2 Estudos Empíricos sobre Mobilidade Intergeracional

Os estudos empíricos sobre mobilidade intergeracional, tanto de renda como de educação, são limitados devido à escassez de conjuntos de dados existentes que contenham informações sobre o rendimento ou renda de pessoas em idade adulta de duas gerações diferentes.

As evidências sugeridas pelos estudos realizados, abrangendo essa literatura, podem diferir entre classes sociais, grupos de renda, grau de desenvolvimento de um país ou região e pelas inúmeras características dos indivíduos e de seus pais.

Uma hipótese sugere que a mobilidade pode ser maior entre as famílias pertencentes às classes inferiores e intermediárias da distribuição de renda, onde os salários são o componente mais importante do nível de renda e os gastos com capital humano são maiores quando comparados com famílias mais ricas. Também tem surgido certa preocupação de que o sistema de bem-estar cria uma classe de bem-estar permanente que persiste ao longo de gerações e que resulta em pouca mobilidade ascendente para os indivíduos pertencentes às famílias mais pobres.

Utilizando o arcabouço teórico desenvolvido por Becker e Tomes<sup>5</sup> juntamente com dados do *National Longitudinal Surveys of Labor Market Experience* (NLS), Peters (1992) estima empiricamente a importância de características de *background* familiar em influenciar a relação sistemática entre a renda de um filho na idade adulta e a renda dos seus pais. Os resultados indicam que existe uma diferença no padrão de mobilidade entre filhos e filhas e entre renda e salário. A autora ainda verifica que a característica de *background* familiar mais importante no processo de transmissão de renda entre gerações é a renda, no entanto, a adição de outras características familiares e interações entre renda e estas variáveis de *background* familiar melhoraram o ajustamento do modelo.

---

<sup>5</sup> Para maiores detalhes, ver Becker e Tomes (1979, 1986) e Becker (1981).

Visando determinar o grau de mobilidade de renda entre gerações no Canadá durante as décadas de 1980 e 1990, bem como, investigar se este grau de mobilidade está mudando ao longo do tempo, Fortin e Lefebvre (1998) utilizam dados do Censo Decenal Canadense para o período de 1951 a 1991 e informações do *General Social Survey* (GSS) de 1986 e 1994 para estimar um modelo log-linear padrão com o uso de variáveis instrumentais. Os autores verificam que o grau de mobilidade de renda intergeracional no Canadá é maior que os encontrados nos trabalhos realizados para os Estados Unidos e Reino Unido e que existe um maior grau de transmissão de *status* entre pais e filhas do que entre pais e filhos. Os resultados da análise de corte indicam que o grau de mobilidade de renda está crescendo com o passar do tempo e que esta mobilidade é maior para as pessoas mais jovens.

Com o intuito de gerar estimativas mais confiáveis das recentes variações na mobilidade de renda entre as gerações Lee e Solon (2006) utilizam informações do *Panel Study of Income Dynamics* (PSID) para os anos de 1977 a 2000. Os autores utilizam o logaritmo da renda familiar anual dos filhos e filhas com idades entre 25 e 48 anos como variável dependente. Como os dados que são utilizados não trazem informações a cerca da renda de longo prazo dos indivíduos, uma medida multi-ano<sup>6</sup> de renda dos pais é utilizada como *proxy* para a renda de longo prazo. Os resultados indicam que a mobilidade de renda entre gerações nos Estados Unidos não mudou drasticamente ao longo das décadas de 1980 e 1990.

Utilizando dados do *Swedish Military Enlistment*, Hirvonen (2010) investiga os fatores que contribuem para a mobilidade de renda entre gerações na Suécia. A partir da estimação de um modelo em que a renda do pai impacta direta e indiretamente na renda dos filhos, constata que o impacto do nível educacional, no processo de transmissão de renda, decresce para os filhos com 12 ou mais anos de estudo. A autora também verifica que o nível educacional e as habilidades cognitivas e não-cognitivas apresentam impactos mais acentuados nos quantis superiores de distribuição de renda dos pais e que o salário da filha é um pouco menos dependente do salário do seu pai, no entanto, a educação tem impacto um pouco mais pronunciado para as filhas.

---

<sup>6</sup> Esta medida é o logaritmo da média da renda familiar anual durante o período em que os filhos tinham entre 15 e 17 anos de idade. Por exemplo, para filhos e filhas nascidos em 1952, a variável renda dos pais será o logaritmo da média da renda dos pais para os anos de 1967 a 1969.

Com base em um painel de famílias chilenas, Celhay, Sanhueza e Zubizarreta (2009) analisam a mobilidade de renda intergeracional no período 1996 – 2006 através de matrizes de mobilidade, de onde verificam que a mobilidade é maior nos decis intermediários de renda, porém, decresce com o aumento dos estratos de renda. Os resultados indicam baixa mobilidade de renda quando comparados com países desenvolvidos e que a elasticidade renda é maior para os homens. Além disso, a análise de corte sugere que a mobilidade de renda decresce com o tempo.

Para resolver o problema da ausência de uma base de dados apropriada<sup>7</sup>, Mocetti (2007) utiliza a metodologia de TS2SLS, citada anteriormente, para analisar a mobilidade de renda intergeracional na Itália. O autor ainda se utiliza da regressão quantílica<sup>8</sup> para explorar aspectos de não-linearidade na relação entre a renda de um pai e de seus filhos. Os resultados da estimação por TS2SLS mostram que a mobilidade intergeracional de renda é menor na Itália que em outros países desenvolvidos. Além disso, pela estimação por regressão quantílica, o autor encontra evidências de uma forte persistência entre os quantis superiores de renda dos filhos.

Não obstante os estudos sobre mobilidade de salários terem uma vasta quantidade de distintos países, poucos foram realizados para países em desenvolvimento (DUNN, 2007). De acordo com Pero e Szerman (2008) é surpreendente o número reduzido de estudos que têm se dedicado a analisar o tema no Brasil, embora realcem que parte da literatura que estuda as fontes de desigualdade no Brasil se divide em diferentes focos centrais, como o papel do *background* familiar na determinação de salários (Lam; Schoeni, 1993), a mobilidade intergeracional de educação (Ferreira; Veloso, 2003b; Barros et al. 2001) e de ocupação (do Valle e Silva; Pastore (2000).

Os estudos empíricos sobre mobilidade intergeracional de renda no Brasil ainda enfrentam outro grave problema, visto que, além de restrições nas informações dos dados existentes, o país não possui um conjunto de dados estruturado em forma de painel, essencial

---

<sup>7</sup> O SHIW é um modelo representativo da população italiana conduzindo pelo Banco da Itália e que contém informações de renda e características demográficas e socioeconômicas dos indivíduos.

<sup>8</sup> Esta metodologia foi concebida por Koenker e Basset (1978) e que passou a ser utilizada cada vez mais em vários campos. Tal metodologia é capaz de quantificar o efeito de uma variável explicativa (renda do pai) ao longo da distribuição de renda do filho, e não somente na média, como na estimação por Mínimos Quadrados Ordinários.

para análises sistemáticas de mobilidade. Por esse motivo os diversos trabalhos realizados para analisar a questão da mobilidade de renda no Brasil utilizam dados da PNAD<sup>9</sup>. No entanto, conforme argumentam Pero e Szerman (2008), apenas as PNADs de 1973, 1976, 1982, 1988 e 1996 podem ser utilizadas em estudos sobre mobilidade, visto que, estas pesquisas contêm suplementos sobre mobilidade social, ou seja, carregam informações sobre o nível de escolaridade do pai do chefe de família. Diante de todos esses problemas, a conclusão desses trabalhos é que, de uma maneira geral, há significativa transmissão de *status* entre as gerações, o que contribui para a persistência das desigualdades.

Mesmo com todos os problemas supracitados, diversos autores utilizaram dados da PNAD aplicados no procedimento de TSIV para mensurar a mobilidade intergeracional de renda no Brasil. Ferreira e Veloso (2006) utilizam tais dados e verificam que o grau de mobilidade intergeracional de salários no Brasil é menor do que os observados em países desenvolvidos, porém, internamente, sua ocorrência é maior para indivíduos residentes em regiões mais desenvolvidas (Sudeste), bem como para os de cor branca. Os autores também verificam evidências de não-linearidade no padrão de mobilidade, visto que, o grau de persistência cai, em média, com o salário do pai.

Na mesma linha de Ferreira e Veloso (2006), Dunn (2007) utiliza dados das PNADs de 1982, 1988 e 1996 para verificar o grau de transmissão de ganhos entre gerações no Brasil e examinar os efeitos destes ganhos nas medidas de mobilidade intergeracional. Os resultados apontam uma elasticidade de salário intertemporal na ordem de 0.688. Já Pero e Szerman (2008) utilizam dados da PNAD de 1977 e 1996 para estimar o grau de persistência de *status* econômico para diferentes conceitos de renda e de amostras, e verificam que, as estimativas da persistência intergeracional de renda para o Brasil variam entre 0.67 e 0.80, dependendo do conceito de renda utilizado, sendo mais altas quando a renda familiar per capita é considerada.

No entanto, mesmo com todos os problemas mencionados anteriormente, observa-se um substancial crescimento na utilização de metodologias alternativas nos trabalhos que se

---

<sup>9</sup> Ferreira e Veloso (2006) indicam que a PNAD é um conjunto de dados adequado para o estudo de mobilidade intergeracional por duas razões. Primeiramente, ela permite um estudo da persistência de *status* econômico entre gerações de um grande país em desenvolvimento que apresenta um persistente e elevado grau de desigualdade de renda nas últimas décadas. E também pelo elevado número de observações, permitindo assim, estimativas de não-linearidade no padrão de mobilidade intergeracional.

propõe estudar o problema de mobilidade de renda no Brasil. Dunn (2007) ainda utiliza informações de pais e filhos morando na mesma residência, em 1996, para estimar uma medida de elasticidade utilizando as metodologias de MQO e VI. O autor verifica que, independentemente do método utilizado, o Brasil apresenta um dos maiores níveis de transmissão de ganhos, quanto comparado com estudos realizados para outros países e que a educação pode ser o caminho mais significativo pelo qual os ganhos são transmitidos entre as gerações.

Já Pero e Szerman (2008) aplicam as metodologias propostas por Atkinson (1983) para avaliação da mobilidade segundo seus efeitos sobre o bem-estar social; de Benabou e Ok (2001), para avaliar a progressividade da estrutura de mobilidade no Brasil; e de Bigard, Guillotin e Lucifora (1998), para a identificação das barreiras à mobilidade de renda, e constatam que, a estrutura de mobilidade do Brasil reduz o Gini em 20 pontos percentuais, o que equivale a dizer que a desigualdade de oportunidades é responsável por algo em torno de 67% da desigualdade observada. Portanto, os índices de mobilidade reforçaram as evidências das estimações por regressões, apontando o Brasil como o país com menor mobilidade intergeracional de renda, quando comparado com países desenvolvidos.

Visando identificar os fatores inerentes às mobilidades de renda e de educação no período de 1987 a 2003, Figueiredo, Netto Junior e Porto Junior (2007) fazem uso da medida de polarização de renda proposta por Duclos et al. (2004), do cálculo das matrizes de transição de probabilidade proposto por Aebi et al. (2001), das funções de densidade kernel, índices de desigualdade, modelo de pseudo painel e as matrizes de transição markoviana. Em linhas gerais, os autores constatam que, houve uma diminuição do índice de gini para todas as regiões brasileiras tornando as distribuições mais equitativas e um aumento no índice de polarização para todas as regiões, onde a polarização no Nordeste beneficiou o menor estrato de renda, ao passo que, na região Sudeste há um aumento da polarização nos estratos superiores de renda e do ponto de vista da mobilidade de renda, observa-se um esvaziamento das classes intermediárias.

Com o intuito de mensurar o grau de mobilidade intrageracional de renda no Brasil no período de 1995 a 2005, Figueiredo (2010) utiliza os dados da PNAD juntamente com a abordagem axiomática da mobilidade e da construção da matriz de transição

markoviana por meio do método de entropias relativas. Após a captação da dinâmica de renda, via matrizes de transição, o autor calcula os índices de mobilidade contidos em Prais (1955) e Shorrocks (1978). Os resultados indicam que o Brasil apresenta uma baixa mobilidade de renda intrageracional, indicando que, a sua estrutura social é relativamente rígida e que os negros apresentam um maior índice de mobilidade, no entanto, tendem a transitar para os estratos inferiores da distribuição de renda e que o movimento dos brancos ocorre de maneira análoga. Além disso, quanto mais jovem maior é a probabilidade do indivíduo migrar para o topo da distribuição de rendimentos. O autor ainda verifica que quanto maior o nível educacional, maior é a mobilidade de renda. Finalmente, os resultados mostram que as regiões Sudeste e Nordeste apresentam os maiores índices de mobilidade, no entanto, não há uma diferença expressiva entre as regiões brasileiras.

Alguns autores passaram a utilizar outras pesquisas para analisar os determinantes da mobilidade intergeracional de renda no Brasil. Em um desses trabalhos, Netto Junior, Ramalho e Aragon (2010) utilizam informações dos Censos Demográficos de 1991 e 2000. Os autores estimam um modelo *logit* ordenado para analisar os rebatimentos do histórico (*background*) familiar sobre a mobilidade de renda/ocupação dos jovens adultos no Brasil<sup>10</sup> e um *probit* bivariado para investigar os determinantes da mobilidade conjunta de renda e de educação<sup>11</sup>, considerando a interdependência presente nestes processos. Os autores concluem que os indivíduos brancos, do sexo masculino, filhos de pais letrados, em lares biparentais e residentes no Sudeste apresentam uma maior probabilidade de atingirem níveis de renda mais elevados. No caso de famílias monoparentais, os filhos de pessoas de referência do sexo feminino apresentam uma maior probabilidade de se mover para estratos de renda superiores se comparado a lares monoparentais chefiados por homens.

---

<sup>10</sup> Como o modelo *logit* ordenado requer uma variável dependente discreta e ordenada, os autores realizaram um ranqueamento a partir da distribuição de renda condicionada às diferenças de remuneração por ocupação/profissão, ou seja, utilizaram-se as médias de remuneração do trabalho principal de cada categoria profissional e fez-se uma ordenação da distribuição de renda por cinco intervalos de percentis: Estrato I - renda menor ou igual ao percentil 20% (R\$ 331,08); Estrato II - remuneração entre os percentis 20% e 40% (R\$ 485,34) da distribuição; Estrato III - renda entre os percentis 40% e 60% (R\$ 681,88); Estrato IV - remuneração entre os percentis 60% e 80% (R\$ 1.187,13) e Estrato V - renda superior ao percentil 80% da distribuição.

<sup>11</sup> Também foram criados estratos ordenados de educação para filhos, chefes e cônjuges dos domicílios: Estrato I - indivíduos com menos de 1 ano de estudo completo; Estrato II - pessoas com histórico de 1 a 4 anos de estudo; Estrato III - de 5 a 8 anos; Estrato IV - de 9 a 11 anos e Estrato V - pessoas com mais de 11 anos de escolaridade.

Buscando analisar a mobilidade de rendimentos no Brasil e verificar se esta contribui ou não para a queda da desigualdade, considerando o período anterior e posterior à queda da desigualdade de renda do país, Antigo (2010) estima um pseudo-painel dinâmico e um modelo de idade-período-coorte, que capta os efeitos de idade, período e coorte sobre a mobilidade, utilizando dados da PNAD entre 1993 e 2007 e da Pesquisa Mensal de Emprego (PME) abrangendo os períodos de 1992 a 2001 e 2002 a 2009. Como os grupos por renda agem diferentemente, recorre-se ao método das regressões quantílicas, com o uso de modelos autoregressivos de primeira ordem para mensurar a mobilidade ao longo dos quantis da distribuição. Os resultados mostram que a queda da desigualdade foi acompanhada por uma maior mobilidade ascendente de rendimentos na base da distribuição, sobretudo do trabalho principal, a partir de 2001. As variáveis de educação e de salário mínimo têm um papel chave para explicar a mobilidade dos mais pobres.

Apesar do conhecimento de diferenças acentuadas entre as regiões brasileiras, poucos são os trabalhos que buscam investigar a mobilidade intergeracional de renda em determinada região. Vieira e Bangolin (2008) realizaram uma análise descritiva dos dados da PNAD para os anos de 1984, 1989, 1995, 1999 e 2004 com o intuito de analisar a trajetória da mobilidade intergeracional da renda média relativa, de uma amostra da população da Região Sul. A mobilidade intergeracional, foi realizada em dois cenários. O primeiro considerava a mesma idade para ambas as gerações, ou seja, quando tinham 20 anos de idade, observou-se que a geração atual, em 2004, ingressou no mercado de trabalho com uma condição mais favorável, em termos de renda, educação e ocupação, do que a geração anterior em 1984, quando estava em torno de 20 anos de idade. Na segunda análise intergeracional, considerou-se o mesmo ano para a análise de ambas as gerações, isto é, o ano de 2004, em que a geração anterior estava ao redor dos 40 anos de idade, enquanto que a geração atual, para esse mesmo período, estava com 20 anos de idade. Em virtude da geração anterior, ter sofrido maior retração da renda no período pós-estabilização, resultou em uma situação menos favorável do que apresentava em 1984, e, portanto, a geração atual, para esta análise também se encontra em uma situação com menor desigualdade de renda relativa.

Portanto, pode-se concluir, de uma maneira geral, que existe uma significativa transmissão de *status* entre as gerações no Brasil, ou seja, pelas evidências verificadas nos

trabalhos supracitados, o Brasil, comparado com países desenvolvidos, apresenta um baixo nível de mobilidade intergeracional de renda.

Diante disso, este trabalho visa contribuir com a literatura voltada aos determinantes de rendimento no Brasil. Conforme mencionado anteriormente, devido à indisponibilidade de uma base de dados apropriada, este artigo não tratará da questão da mobilidade intergeracional de renda, no entanto, espera-se que, a partir da estimação da equação mineriana, seja possível esboçar um direcionamento do impacto das rendas dos pais na renda dos filhos a partir da aferição de um efeito transmissão (herança familiar) e um dinâmico (análise de gerações superpostas).

### 3 METODOLOGIA

Neste capítulo são discutidos os fundamentos metodológicos utilizados nesta análise. Primeiramente, é realizada uma discussão acerca do arcabouço teórico desenvolvido por Mincer (1974) que deu origem à equação *minceriana* de salários. Em seguida são abordados os possíveis problemas que podem emergir ao se estimar uma equação de rendimentos, bem como, as questões metodológicas associadas à resolução desses problemas. E finalmente, é realizada uma breve explanação sobre a base de dados utilizada para estimar a equação de salários.

#### 3.1 Equações *Mincerianas*

As discussões sobre o retorno do nível de escolaridade no salário dos indivíduos foram iniciadas por Schultz (1961) e Ben-Porath (1967). No entanto, tais discussões só foram intensificadas a partir de 1974, quando Mincer formalizou teoricamente a derivação da equação de rendimentos, onde a renda de um indivíduo seria explicada pelos seus níveis de educação e experiência.

Heckman et al (2006) argumenta que o modelo utilizado por Mincer (1974) gera uma especificação algebricamente similar da equação de salários desenvolvida por Mincer (1958), apesar de possuir uma motivação diferente. É muito menos ligado a um modelo de otimização subjacente, embora alguns dos pressupostos sejam motivados pelo modelo de investimento de capital humano dinâmico de Ben-Porath (1967).

O Modelo de Mincer (1974) enfatiza a dinâmica dos ganhos de ciclo de vida e as relações entre os ganhos observados, os ganhos potenciais, e o investimento em capital humano, tanto para a escolarização formal como para o investimento em *on-the-job*. Considerando  $P_t$  como sendo os ganhos potenciais na idade  $t$ , então, uma fração destes ( $k_t$ ) representará o custo total de investimentos em treinamento ( $C_t$ ), ou seja,  $C_t = k_t P_t$ . Sendo  $\rho_t$  o retorno médio dos investimentos em formação feitos na idade  $t$ , então os ganhos potenciais na idade  $t$  são:

$$P_t \equiv P_{t-1}(1 + k_{t-1}\rho_{t-1}) \equiv \prod_{j=0}^{t-1} (1 + k_j\rho_j) P_0 \quad (1)$$

Definindo educação formal como os anos investidos em tempo integral ( $k_t = 1$ ), desde o início da vida, para se obter uma taxa de retorno ( $\rho_s$ ), constante em todos os anos de escolaridade, e supondo que a taxa de retorno para o investimento pós-escolar é constante ao longo do tempo e é igual a  $\rho_0$ , então:

$$\ln P_t \equiv \ln P_0 + s \ln(1 + \rho_s) + \sum_{j=s}^{t-1} \ln(1 + k_j\rho_0) \approx \ln P_0 + s\rho_s + \rho_0 \sum_{j=s}^{t-1} k_j \quad (2)$$

De acordo com Heckman et al (2006), o modelo de Mincer aproxima-se do modelo de Bem-Porath (1967) por assumir uma taxa de investimento pós-escolar linearmente declinante, ou seja,  $k_{s+x} = \kappa(1 - (x/T))$  onde  $x = t - s \geq 0$  representa o nível de experiência no trabalho na idade  $t$ . A duração de tempo no trabalho,  $T$ , é independente dos anos de escolaridade. Portanto, a relação entre salários potenciais, escolaridade e experiência será definida da seguinte maneira.

$$\ln P_{x+s} \approx \ln P_0 + s\rho_s + \left(\rho_s\kappa + \frac{\rho_0\kappa}{2T}\right)x - \frac{\rho_0\kappa}{2T}x^2 \quad (3)$$

Como salários observados são iguais aos salários potenciais menos os custos de investimento, então:

$$\begin{aligned} \ln w(s, x) &\approx \ln P_{x+s} - \kappa \left(1 - \frac{x}{T}\right) \\ &= (\ln P_0 - \kappa) + \rho_s s + \left(\rho_0 s + \frac{\rho_0\kappa}{2T} + \frac{\kappa}{T}\right)x - \frac{\rho_0\kappa}{2T}x^2 \\ &= \alpha_0 + \rho_s s + \beta_0 x + \beta_1 x^2 \end{aligned} \quad (4)$$

Onde  $s$  é o nível de escolaridade,  $\rho_s$  é a taxa média de retorno da escolaridade (assumida ser a mesma para todos os níveis de escolaridade) entre os indivíduos, e  $x$  é a experiência adquirida pós-educação. A equação (4) é, portanto, denominada como “função salário do capital humano”, ou popularmente conhecida como “equação *minceriana*”, que impulsionou os estudos sobre os determinantes salariais na literatura econômica.

Mincer (1974) concebeu, portanto, uma equação para salários que seria dependente de fatores explicativos, tais como: escolaridade, experiência e diversos outros atributos individuais como, por exemplo, o gênero. Portanto, a formulação teórica desenvolvida por Mincer (1974) constitui-se em um poderoso arcabouço utilizado para estimar retornos à educação, à qualidade da educação, à experiência, entre outros.

Carvalho, Neri e Silva (2006) realçam ainda mais a teoria desenvolvida por Mincer (1974), indicando que este integrou a teoria do investimento em capital humano dentro de um contexto empírico, compatível com a teoria econômica. Desde então, essa formulação passou a ser amplamente utilizada em estudos do mercado de trabalho devido ao fato de o coeficiente estimado da variável educacional poder ser interpretado como a taxa marginal de retorno da educação ou retorno esperado da produtividade do trabalhador (HECKMAN, 2005; FLEISCHHAUER, 2007).

Portanto, diante da grande importância da formulação teórica desenvolvida por Mincer (1974), ratificada pela sua grande utilização para estimar equações de salários, verificado em Langoni (1973), Lam e Levinson (1992), Menezes Filho (2001), Duarte (2003), Heckman (2005), Carvalho, Neri e Silva (2006), Suliano e Siqueira (2010), Mariano e Arraes (2011), este trabalho utilizará o modelo teórico desenvolvido por Mincer (1974) com a finalidade de mensurar um possível grau de transmissão de salário entre pais e filhos no Brasil em uma única geração. Diante disso, a equação aqui proposta apresentará a seguinte forma.

$$\ln Y_i = \beta_0 + \beta_1 E_i + \beta_2 I_i + \beta_3 I_i^2 + \beta_4 E_i * A_i + \beta_5 E_i * G_i + \beta_6 RP_i + \beta_7 RM_i + \sum_{j=1}^3 \theta_{ji} X_{ji} + \varepsilon_i \quad (5)$$

As variáveis  $Y_i$ ,  $E_i$ ,  $I_i$ ,  $RP_i$  e  $RM_i$  representam a renda do trabalho principal, anos de estudo, idade, renda do pai e renda da mãe do filho  $i$ , respectivamente. O vetor  $X$  é composto pelas *dummies* de gênero ( $G$ ), raça ( $R$ ) e urbana ( $U$ ). Além disso, serão incluídas interações entre a variável anos de estudo do filho e as *dummies* de gênero ( $E_i * G_i$ ) e de ano ( $E_i * A_i$ ) visando captar possíveis divergências da taxa de retorno da educação entre filhas e filhas e entre os anos de 1989 e 2009, respectivamente.

O presente trabalho consiste, inicialmente, em aplicativos empíricos sobre a comparação teórica entre três métodos de estimação para a equação (5), onde dois destes, o método de Heckman em dois estágios e o procedimento de variáveis instrumentais (VI), já foram utilizados em outros trabalhos, tais como Suliano e Siqueira (2010) e Mariano e Arraes (2011), que visavam determinar os condicionantes dos rendimentos dos indivíduos, enquanto que o método desenvolvido por Klein e Vella (2010), doravante KV (2010), ainda foi pouco explorado a nível internacional e desconhecida sua aplicação no âmbito nacional. Portanto, este trabalho irá contribuir com a literatura voltada à estimação de uma equação de rendimentos, pois incorporará no debate uma nova metodologia que terá grande utilidade para futuros trabalhos a serem realizados nesta área. Os três métodos acima mencionados serão descritos mais adiante.

Realizadas as estimações da equação (5), pelos três métodos supracitados, escolher-se á o modelo que apresentar o melhor ajuste para se estimar variações na equação de salários, a fim de permitir retornos médios variáveis. Serão, portanto, acrescentadas interações das variáveis renda do pai e renda da mãe com *dummies* de gênero e de ano, na tentativa de se verificar se o impacto das rendas, do pai e da mãe, na renda dos filhos apresentam diferenças de acordo com o gênero do filho e entre os dois anos aqui utilizados, respectivamente. Além disso, serão acrescentadas, à equação (5), interações das variáveis renda do pai e renda da mãe com uma *dummy* regional, referente às regiões Centro-Oeste, Sudeste e Sul, na tentativa de verificar se o impacto das rendas, paterna e materna, divergem entre as regiões brasileiras.

$$\begin{aligned} \ln Y_i = & \beta_0 + \beta_1 E_i + \beta_2 Exp_i + \beta_3 Exp_i^2 + \beta_4 E_i * A_i + \beta_5 E_i * G_i + \beta_6 RP_i \\ & + \beta_7 RM_i + \beta_8 RP_i * A_i + \beta_9 RM_i * A_i + \beta_{10} RP_i * G_i + \beta_{11} RM_i * G_i \\ & + \beta_{12} RP_i * COSESUL_i + \beta_{13} RM_i * COSESUL_i + \sum_{j=1}^3 \theta_{ji} X_{ji} + \varepsilon_i \end{aligned} \quad (6)$$

Após a estimação da equação (6) espera-se que os resultados obtidos sejam suficientes para indicar um direcionamento na questão da determinação da renda dos jovens no Brasil. Mais especificamente, espera-se um impacto positivo das rendas, paternas e maternas, sobre a renda dos filhos, induzindo, portanto, que existe certo grau de transmissão de rendimentos entre as gerações e que tais impactos ocorram de maneira diferenciada entre as regiões brasileiras, ou seja, espera-se que nas regiões Centro-Oeste, Sudeste e Sul as rendas

dos pais exerçam um maior impacto na renda dos seus descendentes quando comparadas as regiões Norte e Nordeste.

No entanto, conforme argumentado anteriormente, existe uma série de problemas associados a qualquer tentativa de estimar uma equação de rendimentos, os quais serão descritos em seguida.

### **3.2 Problemas Associados à Estimação de uma Equação de Rendimentos**

O primeiro problema que surge ao se estimar uma equação *minceriana* de salários é o viés de seleção amostral. Segundo Heckman (1979) o viés de seleção amostral pode surgir praticamente por dois motivos: pela auto-seleção de indivíduos ou unidades amostrais investigados e pelas decisões de seleção da amostra pelos analistas. O autor ainda argumenta que a utilização de amostras selecionadas de forma não aleatória para estimar relações comportamentais acarretam em um viés ocasionado pelo problema de omissão de dados.

De acordo com Mariano e Arraes (2011) ao se utilizar dados extraídos da PNAD, deve-se considerar que as informações coletadas são fornecidas pelas pessoas que trabalhavam no período da pesquisa, ou seja, os salários observados nas PNADs estão relacionados com a decisão de um indivíduo trabalhar ou não e esta pode estar correlacionada a fatores não observáveis que afetam seu rendimento. Portanto, além de o salário depender da oferta de emprego, pode também ser uma função da estratégia de “*job search*” do indivíduo, o que remete ao fato de o mesmo ter implícito um salário de reserva abaixo do qual não aceitaria participar do mercado de trabalho, resultando, assim, em um possível viés de seleção amostral (SULIANO; SIQUEIRA, 2010).

Além disso, Cameron e Trivedi (2005) indicam que a omissão e/ou a característica endógena de uma variável em uma equação, especificamente de rendimentos, poderá gerar estimadores viesados e inconsistentes quando se aplica o usual método de mínimos quadrados ordinários (MQO). Portanto, um grande problema que as equações *mincerianas* de salários enfrentam é o possível problema de endogeneidade entre educação e salários, visto que, a causalidade que geralmente vai da educação para salários também vai dos salários para a educação (SULIANO; SIQUEIRA, 2010).

Diversos autores passaram, portanto, a utilizar o método de VI para solucionar o problema da endogeneidade entre salários e educação. No entanto, embora teoricamente atraente, o método de variáveis instrumentais nem sempre pode ser facilmente implementado na prática, uma vez que, depende da disponibilidade e da validade dos instrumentos. Além disso, a interpretação dos coeficientes de VI é muitas vezes limitada ao *local average treatment effect* (LATE), visto que, tais coeficientes representam os efeitos causais apenas para a subamostra de *compliers* (SANITER, 2012).

Diante do exposto, pode-se concluir que a estimação de equações *mincerianas* pelo Método de Mínimos Quadrados Ordinários gera estimativas viesadas e inconsistentes, tanto pelo viés de seleção amostral, quanto pela endogeneidade entre educação e salários. Portanto, as próximas subseções se reservam a apresentar as metodologias utilizadas neste artigo visando solucionar o viés de seleção amostral, a endogeneidade entre educação e salário e a heterogeneidade presente na amostra de valores.

Além disso, a estimativa gerada, a partir de uma equação *minceriana* de salários, pode ser válida apenas para a amostra em questão ou pode variar para diferentes valores das variáveis explicativas na própria amostra. Portanto, a equação *minceriana* também será estimada pelo método de regressões quantílicas, visto que, este é um procedimento mais sofisticado para tratar o problema de heterogeneidade presente na relação entre salários e características pessoais, permitindo, portanto, a obtenção de estimativas diferenciadas para os diferentes níveis de rendimento dos filhos.

### 3.2.1 Método de Heckman em Dois Estágios

Para solucionar o viés de seletividade amostral, será utilizado o procedimento de Heckman (1979) em dois estágios. O primeiro estágio consiste em estimar a seguinte equação de participação no mercado de trabalho:

$$L_i^* = \beta_i' X_i + \mu_i \quad (7)$$

Onde  $L_i^*$  é a variável que designa a participação no mercado de trabalho e  $X_i$  representa um conjunto de variáveis que explicam tal participação. A variável latente  $L_i^*$  é não observada, no entanto, o seu sinal determinará o valor que a variável binária observada  $L_i$  assumirá.

$$\begin{aligned} L_i &= 1 \text{ se } L_i^* > 0 \\ L_i &= 0 \text{ se } L_i^* \leq 0 \end{aligned} \quad (8)$$

Será utilizado o modelo *Probit* para computar a probabilidade de que  $L_i$  assumo o valor 1, ou seja, para calcular a probabilidade do individuo trabalhar, dadas algumas variáveis explicativas, visto que, este modelo pode ser desenvolvido através do uso de uma variável não observada, comumente chamada de variável latente assumida possuir uma certa distribuição de probabilidade (DAVIDSON; MACKINNON, 2004). Essa probabilidade será representada da seguinte maneira.

$$\begin{aligned} \Pr(L_i = 1) &= \Pr(L_i^* > 0) = \Pr(\beta_i' X_i + \mu_i > 0) \\ &= \Pr(\mu_i > -\beta_i' X_i) = \Pr(\mu_i < \beta_i' X_i) = \Phi(\beta_i' X_i) \end{aligned} \quad (9)$$

Onde  $\Phi(\beta_i' X_i)$  é a função de probabilidade cumulativa da distribuição normal padrão. O segundo estágio do método de Heckman (1979) consiste em estimar a equação de rendimentos levando em consideração o viés de seleção amostral.

$$\ln(W_i) = \gamma_i' Z_i + \varepsilon_i \quad (10)$$

Onde  $W$  é o salário de um indivíduo,  $Z$  é o vetor de características individuais e  $\varepsilon$  é o vetor de termos de erros aleatórios. Dessa forma, o viés de seleção amostral pode ser observado da seguinte maneira.

$$\begin{aligned} E(W_i / L_i^* > 0) &= E(W_i / \mu_i > -\beta_i' X_i) = \gamma_i' Z_i + E(\varepsilon_i / \mu_i > -\beta_i' X_i) \\ &= \gamma_i' Z_i + \frac{\text{cov}(\mu_i, \varepsilon_i)}{\sigma_\mu} \frac{\phi(\beta_i' X_i)}{\Phi(\beta_i' X_i)} = \gamma_i' Z_i + \frac{\text{cov}(\mu_i, \varepsilon_i)}{\sigma_\mu} \varphi_i \end{aligned} \quad (11)$$

Onde  $\phi$  é a função densidade de probabilidade da normal padrão e  $\varphi_i$  é denominada razão inversa de Mills. De acordo com Heckman (1979), a razão inversa de Mills é utilizada como variável explicativa na equação (10) com o objetivo de eliminar o viés de seleção amostral. Portanto, a significância dessa variável indicará a presença do viés de seletividade amostral. A equação *minceriana* utilizada para solucionar o problema do possível viés de seletividade amostral será, portanto, representada da seguinte maneira.

$$\ln(W_i) = \gamma_i' Z_i + \lambda_i \varphi_i + \varepsilon_i \quad (12)$$

### 3.2.2 Método de Variáveis Instrumentais

Conforme explicitado anteriormente, outro tipo de viés encontrado em uma equação *minceriana* de salários decorre da endogeneidade entre educação e salário, ou seja, ambas as variáveis passam a ser determinadas dentro do modelo. Portanto, conforme argumentam Wooldridge (2002) e Stock e Watson (2004), para corrigir os efeitos maléficos ocasionados pela possível endogeneidade presente na estimação de uma equação de rendimentos, deve-se utilizar o método de mínimos quadrados em dois estágios (MQ2E) com o uso de variáveis instrumentais (VI). Além disso, para que o instrumento seja válido, ele precisa ser correlacionado com a variável endógena (educação) e não apresentar correlação com o termo de erro.

De acordo com Mariano e Arraes (2011), existem diversos trabalhos que buscam corrigir o viés da endogeneidade da educação através de VI, no entanto, eles acabam divergindo quando se trata sobre as escolhas dos instrumentos, seja por inadequação voluntária ou indisponibilidade de dados. Visando corrigir a possível endogeneidade entre educação e salário presente em uma equação de rendimentos, Mariano e Arraes (2011) utilizaram o número de componentes da família e acesso a internet como instrumentos para a variável anos de estudo, a qual é considerada endógena em uma equação de rendimentos.

Já Suliano e Siqueira (2010), para estimar as taxas de retornos da educação para estados da região Nordeste do Brasil a partir da PNAD, utilizam como instrumento o número de pessoas na família do indivíduo, com o argumento de que dadas as limitações orçamentárias, indivíduos que decidem ter mais filhos passam a investir menos em educação. Salvato e Silva (2008) utilizaram a renda de aluguel como instrumento na estimação para a região metropolitana de Belo Horizonte, justificando que este é um dos rendimentos de não trabalho e que pode ser entendida como uma *proxy* para o nível de renda da família de origem, mostrando assim as condições da família em que o indivíduo foi educado.

No entanto, conforme argumentam Mariano e Arraes (2011), um problema que persiste e que a maioria dos trabalhos desconsidera refere-se à qualidade destes instrumentos. Heckman e Urzua (2009) relatam uma série de problemas que podem emergir a partir de uma

estimação por VI, como por exemplo, deficientes instrumentos provêm estimativas viesadas; estimativas de VI advêm de fortes restrições tomadas a priori sobre os dados; em modelos heterogêneos, diferentes instrumentos fornecem estimativas distintas; estimativas de VI, dependendo dos instrumentos utilizados e hipóteses feitas, geram diferentes estimativas do retorno da educação, as quais podem levar a incorretas interpretações.

### 3.2.3 Método de Klein e Vella

De acordo com Saniter (2012), uma das alternativas para se realizar inferência sobre o *average treatment effect* (EMT) de toda a população é através do método desenvolvido por Klein e Vella (2010), o qual é tratado a seguir. Considere o seguinte sistema envolvendo a equação *minceriana* padrão (13) e outra que permite a possibilidade de educação ser endógena (14):

$$\ln(W) = X\beta + \delta S + \mu \quad (13)$$

$$S = Y\varphi + v \quad (14)$$

Onde  $W$  é o salário,  $S$  representa os anos de estudo e  $X$  e  $Y$  são vetores de variáveis explicativas, consideradas exógenas, que podem ser idênticos. Endogeneidade está presente se, e somente se,  $Cov(\mu, v) \equiv \sigma_{\mu, v} \neq 0$ . Para evitar que isso ocorra, a seguinte hipótese é formulada:  $\mu = \lambda v + \varepsilon$ , onde  $v$  e  $\varepsilon$  são não correlacionadas ( $\sigma_{\varepsilon, v} = 0$ ). Então, a equação (13) pode ser reescrita como:

$$\ln(W) = X\beta + \delta S + \lambda v + \varepsilon \quad (15)$$

Onde  $v$  é denominado termo de controle e  $\lambda$  representa o grau de endogeneidade no sistema. Em assim procedendo, a equação controlada (15) estaria livre de endogeneidade ( $\sigma_{\varepsilon, v} = 0$ ), mas, como  $\sigma_{S, v} \neq 0$ , a estimação de (15) por MQO geraria estimadores viesados e inconsistentes. Uma alternativa para remediar esse problema seria acrescentar novos regressores e utilizar o método de variáveis instrumentais, muito embora, conforme explicitado anteriormente, apresente algumas deficiências.

Nesse sentido, a abordagem de Klein e Vella (2010) utiliza as restrições de exclusão de segundo momento na forma de heterocedasticidade para medir identificação. A

idéia central dessa abordagem se baseia na possibilidade de identificar os coeficientes de interesse se o impacto de  $v$  nas variáveis explicativas não for constante, e que tal impacto possa ser estimado. Assim, para a implementação dessa técnica, inicialmente substitui-se  $v$  em (15) pela sua estimativa  $\hat{v}$  obtida a partir da estimação da equação (14) por MQO. Em seguida, deriva-se o estimador MQO de  $\lambda$  da equação,  $\mu = \lambda v + \varepsilon$ , o qual é dado por:

$$\hat{\lambda} = \frac{\hat{\sigma}_{\mu,v}}{\hat{\sigma}_v^2} \Rightarrow \text{plim}\hat{\lambda} = \lambda = \frac{\sigma_{\mu,v}}{\sigma_v^2} = \frac{\sigma_{\mu,v}}{\sigma_v\sigma_\mu} \cdot \frac{\sigma_\mu}{\sigma_v} = \rho_{\mu,v} \cdot \frac{\sigma_\mu}{\sigma_v} \quad (16)$$

Onde  $\rho_{\mu,v}$ ,  $\sigma_v$  e  $\sigma_u$  representam o coeficiente de correlação entre  $u$  e  $v$  e os desvios padrão de  $v$  e  $u$ , respectivamente. Considere que os erros  $v$  e  $\mu$  sejam heterocedásticos, e que os valores estimados dos desvios padrão de suas distribuições são, por hipótese plausível, dadas por,  $H_v(X_v)$  e  $H_u(X_u)$ , as quais variam de acordo com as variáveis contidas no vetor  $X$  e/ou  $Y$ . Substituindo essa hipótese na equação (16) para obtenção de  $\hat{\lambda}$  e sua inclusão na equação (15), obtém-se a equação final a ser estimada:

$$\ln(W) = X\beta + \delta S + \rho_{\mu,v} \frac{H_u(X_u)}{H_v(X_v)} \hat{v} + \varepsilon \quad (17)$$

Os vetores  $X_v$  e  $X_u$  podem ser diferentes ou idênticos<sup>12</sup>. Portanto, assumindo que o termo de controle aprimorado  $(H_u(X_u)/H_v(X_v))\hat{v} \neq \text{constante}$ , então, os regressores não são mais combinações lineares e os parâmetros podem ser estimados por MQO. Para estimar  $H_v(X_v)$  e  $H_u(X_u)$ , Klein e Vella (2010) propõem um procedimento que envolve estimação não paramétrica. Diante disso, para obter o termo de controle aprimorado, este artigo se baseará em Saniter (2012) e utilizará o procedimento de três passos de Ichimura (1993)<sup>13</sup>.

### 3.2.4 Regressão Quantílica

Conforme mencionado anteriormente, as estimativas geradas em uma específica equação econométrica variam entre amostras, bem como com a quantidade de regressores nela incluída. Nesse sentido, a taxa de retorno da educação em uma equação de salário deve variar de acordo com o nível educacional que se meça o efeito marginal, ou mesmo com o

<sup>12</sup> Neste artigo,  $H_u = H_v$ , o qual é composto pelas seguintes variáveis: anos de estudo do filho, renda do pai, renda da mãe, idade, idade ao quadrado, gênero, raça e urbana.

<sup>13</sup> Para um melhor detalhamento do procedimento de Ichimura (1993) ver Saniter (2012).

ambiente econômico em que o indivíduo está inserido, além de problemas de heterogeneidade observada na amostra. Menezes-Filho (2002) indica que a inclusão de interações entre variáveis explicativas e *dummies*, como, por exemplo, de gênero e regiões, e os modelos de *matching* e regressões quantílicas são alternativas para solucionar a possível heterogeneidade presente nos resultados.

Os modelos de *matching* explicitam o processo de agregação de efeitos causais para diferentes valores das co-variadas, que está implícito no procedimento de MQO, levando em conta a distribuição amostral dessas covariadas [ver Angrist e Krueger (1999)].

O método de regressão quantílica, introduzido na literatura especializada por Koenker e Basset (1978), permite a estimação de efeitos que variam ao longo da distribuição condicional da variável dependente [ver Menezes-Filho, Fernandes e Picchetti (1999)]. Considere, portanto, que  $\{x_t: t = 1, \dots, T\}$  seja uma sequência de  $K$  vetores de uma matriz conhecida e que  $\{y_t: t = 1, \dots, T\}$  seja uma amostra aleatória da variável aleatória  $Y$  no processo de regressão  $u_t = y_t - x_t\beta$ , então o  $\theta$  – ésimo quantil amostral,  $0 < \theta < 1$ , pode ser definido como qualquer solução do seguinte problema de minimização.

$$\min_{\beta} \left[ \sum_{t \in \{t: y_t \geq x_t\beta\}} \theta |y_t - x_t\beta| + \sum_{t \in \{t: y_t < x_t\beta\}} (1 - \theta) |y_t - x_t\beta| \right] \quad (18)$$

A grande vantagem é que a estimação pelo método de regressão quantílica permite identificar as variações inter e intraquantis, o que o método de regressão por MQO não comporta, pois é baseado na média da distribuição condicional.

Portanto, devido à natureza dos dados que serão utilizados, impossibilitando a formação dos *matchings*, este trabalho fará uso de interações entre algumas variáveis explicativas e *dummies* de gênero, visando captar impactos diferenciados entre filhos e filhas nos anos de 1989 e 2009 e nas regiões brasileiras. A aplicação do método de regressão quantílica visa homogeneizar as estimativas geradas, fazendo com que estas possam ser consideradas para todos os indivíduos presentes em cada um dos estratos ou grupos considerados.

### 3.3 Base de Dados

Para a realização deste exercício empírico utilizar-se-á dados da Pesquisa Nacional por Amostra de Domicílios (PNAD) do Instituto Brasileiro de Geografia e Estatística (IBGE) referente aos anos de 1989 e 2009. O quadro 1.1 sintetiza as variáveis que foram selecionadas e que serão utilizadas nas estimações dos modelos que serão descritos na próxima seção.

Quadro 1.1: Descrição das Variáveis retiradas da PNAD.

Variáveis	Descrição	Valores Assumidos
Y	Renda do Filho do Trabalho Principal <sup>14</sup>	Contínuo
G	Gênero do Filho	1 – Homem 0 – Mulher
I	Idade do Filho	15 a 35 anos
R	Raça do Filho	1 – Branco 0 - Não Branco
E	Anos de Estudo do Filho	0 - Sem Instrução ou Menos de 1 Ano 1 - 1 Ano de Estudo 2 - 2 Anos de Estudo 3 - 3 Anos de Estudo 4 - 4 Anos de Estudo 5 - 5 Anos de Estudo 6 - 6 Anos de Estudo 7 - 7 Anos de Estudo 8 - 8 Anos de Estudo 9 - 9 a 11 Anos de Estudo 10 - 12 ou Mais Anos de Estudo
E*G	Interação entre a variável Anos de Estudo do Filho e a <i>dummy</i> de gênero	Contínuo
E*A	Interação entre a variável Anos de Estudo do Filho e a <i>dummy</i> de ano	Contínuo
U	Urbana	1 - Região Urbana 0 - Região Não Urbana
A	Ano	1 – 2009 0 – 1989
Exp	Experiência	Idade Filho-Anos de Estudo Filho-6
COSSE	<i>Dummy</i> referente às Regiões Centro-Oeste, Sul-Sudeste	Assume valor 1 para os indivíduos residentes em uma dessas 3 Regiões
RP	Renda do Pai do Trabalho	Contínuo

<sup>14</sup>Valores monetários deflacionados a preços de 2009 com base no Deflator para rendimentos da PNAD(INPC – IPEA) – GAMMA\_DEFLPNAD).

Quadro 1.1: Descrição das Variáveis retiradas da PNAD.

Variáveis	Descrição	Valores Assumidos
	Principal	
RM	Renda da Mãe do Trabalho Principal	Contínuo
RP*G	Interação entre a variável Renda do Pai e a <i>dummy</i> de gênero	Contínuo
RM*G	Interação entre a variável Renda do Mãe e a <i>dummy</i> de gênero	Contínuo
RP*A	Interação entre a variável Renda do Pai e a <i>dummy</i> de ano	Contínuo
RM*A	Interação entre a variável Renda da Mãe e a <i>dummy</i> de ano	Contínuo
RP*COSSE	Interação entre a variável Renda do Pai e a <i>dummy</i> COSSE	Contínuo
RM*COSSE	Interação entre a variável Renda da Mãe e a <i>dummy</i> COSSE	Contínuo

Fonte: Elaboração Própria a partir dos Dados da PNAD

Mesmo com o conhecimento de que apenas as PNADs de 1973, 1976, 1982, 1988 e 1996, conforme argumentado por Pero e Szerman (2008), podem ser utilizadas em estudos sobre mobilidade, optou-se por utilizar as PNADs referentes aos anos de 1989 e 2009, visto que, o objetivo deste trabalho é verificar se as rendas dos pais influenciam o rendimento dos seus filhos em uma mesma geração, além de captar possíveis variações nas últimas décadas. Além disso, a utilização dos dados referentes ao ano de 2009 permite aferir resultados mais recentes, visto que, os demais trabalhos no âmbito nacional reportam resultados com base no ano de 1996.

Portanto, a partir das variáveis relatadas no quadro acima, serão realizadas estimações da equação *minceriana* de salários, pelas técnicas citadas anteriormente, considerando toda a amostra e duas subamostras, uma considerando os indivíduos que trabalham no setor de serviços, e a outra, aqueles que trabalham nas indústrias de Produção e Transformação.

Convém mencionar que foram impostas algumas restrições à amostra original através dos seguintes filtros: 1) Exclusão de militares, pois a rigidez de suas atividades, e a determinação de seus rendimentos não são ditadas pelas “forças” de mercado; 2) Exclusão de trabalhadores rurais, haja vista as particularidades inerentes em suas atividades serem incomparáveis às dos trabalhadores urbanos; 3) Inclusão apenas dos filhos com idade entre 15 e 35 anos residentes no domicílio dos pais; 4) A diferença de idade entre pais e filhos seja no mínimo 15 anos; 5) Exclusão de filhos que estudam e trabalham, atendendo hipótese teórica do modelo de Mincer, no qual há um *trade off* entre educação e experiência.

### 3.3.1 Análise Descritiva dos Dados

Após apresentar a fonte e a descrição dos dados, tanto por nível ocupacional quanto regional, se faz pertinente realizar uma análise descritiva dos dados para se ter, inicialmente, algumas evidências sobre transmissão de rendimentos. A Tabela 1.1 contempla as estatísticas descritivas considerando a amostra completa para o Brasil e as macrorregiões Centro-Oeste, Sul-Sudeste e Norte-Nordeste<sup>15</sup>, bem como por ocupação nos setores de serviços e indústrias de produção e transformação.

Observe que, independentemente do ano e do nível ocupacional considerado, os filhos residentes nas regiões Centro-Oeste e Sul-Sudeste recebem, em média, salários superiores aos dos jovens residentes na região Norte-Nordeste. Além disso, percebe-se que, em geral, os jovens brasileiros apresentam, em média, salários maiores no ano de 2009 quando comparados aos de 1989<sup>16</sup>, exceto para os profissionais da área de serviços no Norte-Nordeste.

No tocante a idade, verifica-se que tanto em 1989 como em 2009, os filhos da região Norte-Nordeste são, em média, mais velhos que os jovens do Centro-Oeste, Sul-Sudeste, no entanto, esta diferença é quase imperceptível. Não obstante, observa-se que em 2009, a idade média dos trabalhadores jovens brasileiros, apresenta uma elevação quando

---

<sup>15</sup> Optou-se por realizar estimações para o Norte e Nordeste e Centro-Oeste, Sudeste e Sul para tecer um comparativo entre as regiões. Vale ressaltar que no Sul-Sudeste encontram-se as pessoas que recebem, em média, os maiores salários, ao passo que, no Nordeste encontram-se os trabalhadores com remunerações inferiores. Portanto, esta distinção é para apenas verificar se há divergências na transmissão de renda entre pais e filhos em regiões consideradas ricas e pobres no Brasil.

<sup>16</sup> Salários convertidos a preços de 2009.

comparada à do ano de 1989, independentemente dos níveis regionais e ocupacionais considerados.

Apesar de ter ocorrido uma elevação nos níveis de escolaridade dos filhos em 2009 para todos os níveis regionais e ocupacionais, verifica-se que, com exceção dos técnicos na região Norte-Nordeste no ano de 1989, os jovens brasileiros localizados nas macrorregiões Centro-Oeste e Sul-Sudeste apresentam, em média, um maior nível de escolaridade quando comparados aos filhos do Norte-Nordeste.

Em relação às rendas paternas e maternas, percebe-se que, em média, o salário do pai é superior ao da mãe, independentemente do ano, ocupação e região considerados. Além disso, observa-se que, em média, os pais residentes nas regiões Centro-Oeste, Sul e Sudeste, em qualquer período e ocupação, recebem em média rendimentos superiores aos dos pais do Norte e Nordeste. Quanto à análise temporal, não se verifica, de maneira absoluta, uma trajetória monótona das rendas paternas e maternas, visto que, em algumas ocupações e regiões, tais rendas apresentam um crescimento, ao passo que, em outras, estas variáveis apresentam uma trajetória decrescente.

No tocante ao gênero, nota-se que, exceto para os trabalhadores da área de serviços no Centro-Oeste, Sul e Sudeste em 1989, a maior parte da amostra é composta por homens. Além disso, a maioria dos jovens considerados reside em áreas urbanas e são considerados não brancos.

De uma maneira geral, se observa que os rendimentos dos jovens, estão relacionados, à idade, ao nível educacional, à região dos profissionais, renda dos pais, visto que, em média, os filhos que apresentam uma idade mais elevada, um maior nível de escolaridade, residem nas regiões Sul e Sudeste e tem pais com maiores salários, apresentam, em média, os maiores rendimentos.

Portanto, a partir dos números apresentados na tabela 1.1 e devidamente comentados nos parágrafos supracitados, pode-se inferir sobre a existência de certo nível de transmissão de rendimentos dos pais para seus filhos, onde este decorre principalmente, via renda da mãe.

Tabela 1.1 - Estatísticas Descritivas das Variáveis, por Regiões e Ocupações, 1989 e 2009.

Variáveis	Norte-Nordeste				Centro-Oeste, Sul-Sudeste			
	1989		2009		1989		2009	
	Média	DP	Média	DP	Média	DP	Média	DP
<b>Amostra Toda</b>								
REN FILHO	549.61	1010.79	583.35	754.76	751.62	1075.88	854.59	849.21
E	5.82	3.20	8.16	2.11	6.44	2.62	8.68	1.59
I	22.40	4.47	23.95	4.76	22.01	4.60	23.93	4.64
G	0.65	0.48	0.63	0.48	0.62	0.49	0.63	0.48
R	0.29	0.45	0.29	0.46	0.65	0.48	0.62	0.49
U	0.86	0.35	0.88	0.32	0.88	0.33	0.94	0.24
RENPAI	1138.64	3496.51	993.61	1975.16	1322.62	2457.76	1415.52	1885.00
RENMAE	602.88	1124.25	655.38	1075.41	794.24	1251.06	904.72	1165.76
<b>Serviços</b>								
REN FILHO	446.70	586.48	425.69	366.64	588.66	882.52	652.68	495.60
E	5.53	3.04	8.21	1.95	5.95	2.50	8.70	1.42
I	21.82	4.60	23.27	4.72	21.43	4.55	23.24	4.56
G	0.58	0.49	0.53	0.50	0.47	0.50	0.51	0.50
R	0.29	0.46	0.29	0.45	0.65	0.48	0.59	0.49
U	0.85	0.35	0.90	0.30	0.85	0.35	0.95	0.22
RENPAI	998.86	2021.28	832.71	1150.29	1134.00	2166.33	1209.48	1308.97
RENMAE	528.73	842.14	518.85	644.52	622.63	900.18	741.05	827.77
<b>Profissionais das Indústrias de Produção e Transformação</b>								
REN FILHO	475.64	747.97	501.14	472.00	724.12	877.99	790.94	623.87
E	5.42	3.15	7.64	2.33	6.14	2.64	8.28	1.80
I	22.29	4.44	23.94	4.66	21.96	4.62	23.76	4.64
G	0.67	0.47	0.83	0.38	0.72	0.45	0.83	0.38
R	0.26	0.44	0.25	0.43	0.64	0.48	0.59	0.49
U	0.82	0.39	0.84	0.37	0.88	0.33	0.92	0.27
RENPAI	857.33	1712.20	739.77	1107.26	1193.68	1993.89	1136.70	1212.28
RENMAE	489.64	769.78	500.01	650.20	740.58	984.01	767.66	871.65

Fonte: Elaboração Própria a partir dos Dados da PNAD/1989-2009.

Nota: Valores monetários medidos a preços constantes de 2009.

#### 4 RESULTADOS DAS ESTIMAÇÕES DA EQUAÇÃO *MINCERIANA*

Para a estimação do modelo pelo procedimento de variável instrumental (VI) utilizou-se, na ausência de outras mais adequadas, o número de componentes da família como instrumento para a variável anos de estudo do filho, assim como Suliano e Siqueira (2010) e Mariano e Arraes (2011)<sup>17</sup>. Cálculos próprios demonstram que a correlação entre essas duas variáveis é negativa com magnitude 0,44 (amostra toda), 0,46 (serviços) e 0,41 (técnicos e profissionais das indústrias de transformação e produção) e significância de 1%, indicando, portanto, a validade de sua utilização como instrumento.

As tabelas 1.2, 1.3 e 1.4 fornecem as estimações da equação *minceriana* pelos métodos de MQO, de Heckman em dois estágios, de Klein e Vella (2010), por uma combinação dos métodos de Heckman e Klein e Vella (2010) e VI. As estimações foram realizadas considerando todas as ocupações (tabela 1.2), e duas subamostras de trabalhadores: uma para o setor de serviços (tabela 1.3) e outra para as indústrias de produção e transformação (tabela 1.4). A razão para tal subdivisão é uma tentativa de homogeneizar a amostra por categorias próprias das atividades de trabalho na tentativa de gerar resultados médios mais precisos à realidade da atividade econômica.

Observa-se que o modelo estimado pelo método de VI diverge dos demais. O Erro Quadrado Médio (EQM) gerado por sua estimativa é superior ao dos outros modelos, indicando que, este método apresenta um pior ajuste para fins preditivos. Outro ponto que pode ser destacado é que quando se utiliza a estimação pelo método de VI a maioria dos coeficientes estimados aumentam em magnitude, inclusive o retorno da educação, o qual apresentou um aumento de, no mínimo, 100%, em todas as estimações realizadas, superestimando assim, as previsões dos retornos sobre os rendimentos. Além disso, em alguns casos, os coeficientes estimados utilizando-se o método de VI, apresentam variações de sinal e algumas variáveis, que nos demais modelos são estatisticamente significantes, passam a ser insignificantes.

---

<sup>17</sup> Estes autores também utilizaram a variável acesso à internet como instrumento, no entanto, como a PNAD de 1989 não apresenta esta variável, optou-se por utilizar apenas o número de componentes da família.

Tabela 1.2 – Estimação da Equação de Rendimentos por Alternativos Métodos em Todas as Ocupações.

Variáveis	Métodos				
	MQO	Heckman	KV	Heckman e KV	VI
E	0.139 (0.000)	0.145 (0.000)	0.139 (0.000)	0.145 (0.000)	0.799 (0.000)
I	0.119 (0.000)	0.0813 (0.000)	0.119 (0.000)	0.0812 (0.000)	-0.885 (0.000)
I <sup>2</sup>	-0.00170 (0.000)	-0.00102 (0.000)	-0.00170 (0.000)	-0.00102 (0.000)	0.0166 (0.000)
G	0.872 (0.000)	0.879 (0.000)	0.872 (0.000)	0.879 (0.000)	4.552 (0.000)
R	0.188 (0.000)	0.179 (0.000)	0.188 (0.000)	0.179 (0.000)	-0.0962 (0.052)
U	0.155 (0.000)	0.148 (0.000)	0.155 (0.000)	0.148 (0.000)	-0.333 (0.000)
E*G	-0.0701 (0.000)	-0.0706 (0.000)	-0.0700 (0.000)	-0.0706 (0.000)	-0.515 (0.000)
E*A	0.00606 (0.000)	0.00617 (0.000)	0.00607 (0.000)	0.00618 (0.000)	-0.0540 (0.000)
RP	0.0000172 (0.008)	0.0000193 (0.003)	0.0000172 (0.008)	0.0000192 (0.003)	0.0000416 (0.000)
RM	0.000255 (0.000)	0.000251 (0.000)	0.000255 (0.000)	0.000251 (0.000)	0.000122 (0.000)
$\lambda$		-1.966 (0.000)		-1.965 (0.000)	-39.88 (0.000)
$\rho$			-0.0000435 (0.000)	-0.0000428 (0.000)	
Const.	2.706 (0.000)	3.745 (0.000)	2.707 (0.000)	3.745 (0.000)	23.38 (0.000)
Obs	21838	21838	21838	21838	21838
R <sup>2</sup>	0.486	0.486	0.486	0.486	
EQM	0.603	0.603	0.603	0.603	0.950
F	1203.3 (0.000)	1102.7 (0.000)	1130.6 (0.000)	1045.1 (0.000)	

Fonte: Elaboração Própria a partir dos Dados da PNAD.

Nota: p-valor entre parênteses.

Tabela 1.3 – Estimação da Equação de Rendimentos por Alternativos Métodos no Setor de Serviços.

Variáveis	Métodos				
	MQO	Heckman	KV	Heckman e KV	VI
E	0.129 (0.000)	0.134 (0.000)	0.129 (0.000)	0.134 (0.000)	0.720 (0.000)
I	0.133 (0.000)	0.0983 (0.000)	0.133 (0.000)	0.0984 (0.000)	-0.980 (0.006)
I <sup>2</sup>	-0.00209 (0.000)	-0.00145 (0.001)	-0.00209 (0.000)	-0.00145 (0.001)	0.0187 (0.005)
G	0.938 (0.000)	0.948 (0.000)	0.937 (0.000)	0.947 (0.000)	3.553 (0.000)
R	0.141 (0.000)	0.135 (0.000)	0.141 (0.000)	0.135 (0.000)	-0.117 (0.181)
U	0.127 (0.000)	0.125 (0.000)	0.127 (0.000)	0.125 (0.000)	-0.254 (0.056)
E*G	-0.0776 (0.000)	-0.0778 (0.000)	-0.0775 (0.000)	-0.0777 (0.000)	-0.387 (0.000)
E*A	0.0107 (0.000)	0.0107 (0.000)	0.0107 (0.000)	0.0107 (0.000)	-0.0943 (0.007)
RP	0.0000452 (0.000)	0.0000491 (0.000)	0.0000450 (0.000)	0.0000489 (0.000)	0.000106 (0.000)
RM	0.000296 (0.000)	0.000291 (0.000)	0.000296 (0.000)	0.000291 (0.000)	0.0000799 (0.271)
$\lambda$		-1.084 (0.055)		-1.080 (0.056)	-25.89 (0.002)
$\rho$			-0.0000353 (0.000)	-0.0000339 (0.000)	
Const.	2.550 (0.000)	3.287 (0.000)	2.552 (0.000)	3.286 (0.000)	21.14 (0.000)
Obs	8088	8088	8088	8088	8088
R <sup>2</sup>	0.493	0.493	0.493	0.493	
EQM	0.565	0.565	0.565	0.565	0.902
F	521.5	476.4	775.5	712.4	
Valor P	0.000	0.000	0.000	0.000	0.000

Fonte: Elaboração Própria a partir dos Dados da PNAD.

Nota: p-valor entre parênteses.

Tabela 1.4 – Estimação da Equação de Rendimentos por Alternativos Métodos no Setor dos Profissionais das Indústrias de Produção e Transformação.

Variáveis	Métodos				
	MQO	Heckman	KV	Heckman e KV	VI
E	0.103 (0.000)	0.113 (0.000)	0.102 (0.000)	0.112 (0.000)	1.091 (0.000)
I	0.0955 (0.000)	0.0202 (0.377)	0.0966 (0.000)	0.0214 (0.351)	-1.019 (0.002)
I <sup>2</sup>	-0.00138 (0.000)	0.00000677 (0.988)	-0.00140 (0.000)	-0.0000150 (0.973)	0.0191 (0.001)
G	0.672 (0.000)	0.687 (0.000)	0.672 (0.000)	0.686 (0.000)	6.991 (0.000)
R	0.187 (0.000)	0.171 (0.000)	0.187 (0.000)	0.172 (0.000)	-0.0223 (0.730)
U	0.175 (0.000)	0.142 (0.000)	0.176 (0.000)	0.143 (0.000)	-0.642 (0.008)
E*G	-0.0470 (0.000)	-0.0480 (0.000)	-0.0470 (0.000)	-0.0480 (0.000)	-0.820 (0.001)
E*A	0.00998 (0.000)	0.0102 (0.000)	0.0100 (0.000)	0.0102 (0.000)	-0.0533 (0.007)
RP	0.0000406 (0.000)	0.0000441 (0.000)	0.0000406 (0.000)	0.0000441 (0.000)	0.0000151 (0.219)
RM	0.000327 (0.000)	0.000314 (0.000)	0.000327 (0.000)	0.000314 (0.000)	0.0000940 (0.186)
$\lambda$		-5.108 (0.000)		-5.103 (0.000)	-59.83 (0.000)
$\rho$			0.000893 (0.160)	0.000860 (0.180)	
Const.	3.247 (0.000)	5.695 (0.000)	3.238 (0.000)	5.684 (0.000)	28.87 (0.000)
Obs	8573	8573	8573	8573	8573
R <sup>2</sup>	0.443	0.444	0.443	0.444	
EQM	0.583	0.582	0.583	0.582	1.123
F	370.0	345.6	336.9	317.3	
Valor P	0.000	0.000	0.000	0.000	0.000

Fonte: Elaboração Própria a partir dos Dados da PNAD.

Nota: p-valor entre parênteses.

Em vista desses resultados, cabe destacar que, correções de endogeneidade presente em uma equação de salários através de instrumentos inadequados podem estar gerando estimativas viesadas, portanto, com previsões distantes da realidade, em acordo com Mariano e Arraes (2011). Para esses autores, tal ocorrência deve-se à qualidade e limitação de informações na PNAD acerca dos instrumentos que podem, ao invés de corrigir, criar vieses nas estimativas. Os autores ainda sugerem que esses resultados reforçam a precariedade de

variáveis disponibilizadas para estimação de um modelo mais sofisticado e alertam para a necessidade de ampliar o conteúdo das amostras.

Os demais modelos, por outro lado, por apresentarem resultados bastante similares, a escolha de um deles para gerar previsões deve recair em questões apenas operacionais. Além disso, observa-se que as metodologias de Heckman em dois estágios e Klein e Vella (2010) são eficazes para solucionar os problemas de viés de seleção amostral, endogeneidade entre educação e salário e heterogeneidade presente na amostra de valores. Na estimação pelo método de Heckman em dois estágios, o coeficiente da variável razão inversa de Mills ( $\lambda$ ) é estatisticamente significativa, indicando que, devido à seletividade amostral, a não inclusão desta variável, em uma equação de salários, fornecerá estimativas viesadas e, no modelo estimado pelo método de Klein e Vella (2010), o coeficiente do termo de controle completo ( $\rho$ ), apesar de ter apresentado um valor muito pequeno, também é estatisticamente significativa, exceto na subamostra referente aos técnicos e trabalhadores das indústrias de produção e transformação, indicando que, existem retornos heterogêneos da escolaridade.

Diante dos resultados aqui verificados, este trabalho ratifica a utilização dos métodos de Heckman em dois estágios e de Klein e Vella (2010) em trabalhos voltados à estimação de uma equação de rendimentos. Desse modo, os comentários que serão realizados se aterão apenas a essas duas metodologias. Convém ressaltar, que não faz parte do propósito deste artigo realizar uma comparação entre as duas metodologias supracitadas, mas sim ratificar ainda mais a utilização do método de Heckman em dois estágios e, dos resultados obtidos nessa amostra ser referência desta metodologia alternativa para trabalhos futuros.

Em relação às variáveis explicativas, verifica-se que os coeficientes estimados apresentaram o mesmo sinal, significância estatística e valores bastante próximos nas duas metodologias citadas no parágrafo anterior, à exceção dos coeficientes das variáveis idade (I) e idade ao quadrado ( $I^2$ ), na metodologia de Heckman em dois estágios, aplicada a subamostra dos técnicos e trabalhadores das indústrias de produção e transformação, os quais se apresentam insignificantes.

O coeficiente da variável anos de estudo do filho (E) apresenta sinal positivo em todas as estimações realizadas, indicando que filhos mais escolarizados apresentam, em média, maiores salários. Observa-se também que, quando a estimação é realizada pelo método

de VI o coeficiente dessa variável aumenta em magnitude, superestimando assim o retorno da educação sobre os salários. Esse resultado corrobora com os encontrados por Mariano e Arraes (2011) e Suliano e Siqueira (2010).

Visando verificar se o nível de escolaridade exerce impactos diferenciados nos salários quanto ao gênero do indivíduo e o período que está sendo considerado, foram realizadas interações entre a variável anos de estudo do filho e as *dummies* de gênero (E\*G) e ano (E\*A), respectivamente. Verifica-se que o coeficiente da primeira variável apresenta sinal negativo, nas duas metodologias e nas três amostras utilizadas, indicando que o impacto de um ano a mais de estudo é maior, *ceteris paribus*, no rendimento das filhas. Tal resultado também foi verificado em um estudo realizado por Hirvonen (2010) na Suécia, onde a autora constatou que a educação tem impacto um pouco mais pronunciado no salário das filhas. Além disso, Leone e Baltar (2006) indicam que, apesar de as diferenças de rendimento entre homens e mulheres no Brasil ainda continuarem significativas, estas diminuíram na década de 90, uma vez que, as mulheres avançaram em termos de acesso aos cargos mais bem remunerados. Analogamente, o coeficiente da segunda variável apresenta sinal positivo, nas duas metodologias e nas três amostras utilizadas, o que significa que, o impacto de um ano a mais de estudo é maior, *ceteris paribus*, no salário dos indivíduos em 2009.

Os coeficientes das *dummies* de gênero (G), raça (R) e urbana (U) apresentam, em todas as estimações, sinais condizentes com os já verificados na literatura, indicando que, filhos brancos do gênero masculino e residentes em áreas urbanas, apresentam, em média, os maiores salários.

Com relação às principais variáveis explicativas que visam captar um possível efeito transmissão de renda, observa-se que os coeficientes das variáveis renda do pai (RP) e renda da mãe (RM) apresentam sinais positivos, portanto, atestando tal efeito. Esse resultado corrobora com Dunn (2007) e Pero e Szerman (2008), os quais afirmam que no Brasil existe um elevado nível de transmissão de salários de pais para filhos.

Devido à metodologia VI utilizaram-se as variáveis idade (I) e idade ao quadrado ( $I^2$ ) nas estimações apresentadas na tabela. No entanto, como se verificou, essa metodologia apresenta o pior ajuste, quando comparada às demais. Diante disso, as demais tabelas omitirão estimativas dessa metodologia. Além disso, as variáveis supracitadas serão substituídas pelas

variáveis experiência e experiência ao quadrado, respectivamente. A não inclusão dessas variáveis nas estimações da tabela 1 deve-se à presença da variável anos de estudo do filho nas suas composições, inviabilizando, portanto, as suas presenças nas estimações por VI.

Além do mais, devido as grandes disparidades entre as regiões brasileiras, utilizar-se-á interações entre as variáveis renda do pai e renda da mãe e uma *dummy* regional referente às regiões Centro-Oeste, Sudeste e Sul. Convém ressaltar que a inclusão dessa interação é de grande importância, visto que, o coeficiente da variável interagida indicará um resultado mais confiável, pois representará impactos associados aos rendimentos de filhos e filhas pertencentes a essas regiões, ou seja, indivíduos que habitam em localidades que apresentam características mais homogêneas. Além disso, serão acrescentadas interações das rendas paterna e materna com *dummies* de gênero e ano, na tentativa de captar impactos diferenciados destas rendas nos salários dos filhos e filhas e nos dois períodos utilizados, respectivamente. As tabelas 1.5, 1.6 e 1.7 apresentam os resultados dessas estimações para todas as ocupações, setor de serviços e profissionais das indústrias de produção e transformação respectivamente.

Tabela 1.5 – Estimação da Equação de Rendimentos por Alternativos Métodos com *Dummies* Regionais em Todas as Ocupações.

Variáveis	MQO	Heckman	KV	Heckman e KV
E	0.171 (0.000)	0.171 (0.000)	0.171 (0.000)	0.171 (0.000)
Exp	0.0683 (0.000)	0.0515 (0.000)	0.0682 (0.000)	0.0515 (0.000)
Exp <sup>2</sup>	-0.00142 (0.000)	-0.000868 (0.000)	-0.00142 (0.000)	-0.000868 (0.000)
G	0.882 (0.000)	0.882 (0.000)	0.882 (0.000)	0.881 (0.000)
R	0.169 (0.000)	0.160 (0.000)	0.169 (0.000)	0.160 (0.000)
U	0.159 (0.000)	0.151 (0.000)	0.159 (0.000)	0.151 (0.000)
E*G	-0.0676 (0.000)	-0.0672 (0.000)	-0.0676 (0.000)	-0.0672 (0.000)
E*A	0.0127 (0.000)	0.0125 (0.000)	0.0127 (0.000)	0.0125 (0.000)
RP	0.00000917 (0.249)	0.0000112 (0.157)	0.00000907 (0.254)	0.0000111 (0.160)
RM	0.000276 (0.000)	0.000273 (0.000)	0.000276 (0.000)	0.000273 (0.000)
RP*A	0.0000347 (0.000)	0.0000348 (0.000)	0.0000347 (0.000)	0.0000348 (0.000)
RM*A	-0.000112 (0.000)	-0.000112 (0.000)	-0.000112 (0.000)	-0.000112 (0.000)
RP*G	-0.0000218 (0.003)	-0.0000215 (0.003)	-0.0000217 (0.003)	-0.0000214 (0.003)
RM*G	-0.00000404 (0.847)	-0.00000486 (0.817)	-0.00000408 (0.846)	-0.00000490 (0.815)
RP*COSSE	0.0000235 (0.006)	0.0000228 (0.008)	0.0000235 (0.006)	0.0000228 (0.008)
RM*COSSE	0.0000385 (0.016)	0.0000396 (0.014)	0.0000386 (0.016)	0.0000396 (0.014)
$\lambda$		-1.939 (0.000)		-1.937 (0.000)
$\rho$			-0.0000328 (0.000)	-0.0000312 (0.000)
Const.	3.723 (0.000)	4.409 (0.000)	3.723 (0.000)	4.408 (0.000)
R <sup>2</sup>	0.495	0.495	0.495	0.495
F	800.7	755.6	822.0	766.0
Valor P	0.000	0.000	0.000	0.000

Fonte: Elaboração Própria a partir dos Dados da PNAD. Nota: p-valor entre parênteses.

Tabela 1.6 – Estimação da Equação de Rendimentos por Alternativos Métodos com *Dummies* Regionais no Setor de Serviços.

Variáveis	MQO	Heckman	KV	Heckman e KV
E	0.161 (0.000)	0.162 (0.000)	0.161 (0.000)	0.162 (0.000)
Exp	0.0529 (0.000)	0.0156 (0.065)	0.0529 (0.000)	0.0157 (0.065)
Exp <sup>2</sup>	-0.000934 (0.000)	0.000319 (0.337)	-0.000933 (0.000)	0.000318 (0.339)
G	0.961 (0.000)	0.959 (0.000)	0.961 (0.000)	0.958 (0.000)
R	0.115 (0.000)	0.0973 (0.000)	0.115 (0.000)	0.0973 (0.000)
U	0.131 (0.000)	0.124 (0.000)	0.130 (0.000)	0.124 (0.000)
E*G	-0.0815 (0.000)	-0.0783 (0.000)	-0.0815 (0.000)	-0.0783 (0.000)
E*A	0.0225 (0.000)	0.0222 (0.000)	0.0225 (0.000)	0.0222 (0.000)
RP	0.0000318 (0.136)	0.0000413 (0.050)	0.0000314 (0.143)	0.0000410 (0.053)
RM	0.000258 (0.000)	0.000248 (0.000)	0.000258 (0.000)	0.000248 (0.000)
RP*A	0.0000176 (0.163)	0.0000143 (0.245)	0.0000178 (0.159)	0.0000145 (0.241)
RM*A	-0.000157 (0.000)	-0.000156 (0.000)	-0.000157 (0.000)	-0.000156 (0.000)
RP*G	-0.00000635 (0.667)	-0.00000635 (0.656)	-0.00000590 (0.691)	-0.00000601 (0.675)
RM*G	0.0000292 (0.444)	0.0000246 (0.516)	0.0000290 (0.447)	0.0000245 (0.518)
RP*COSSE	0.0000154 (0.301)	0.0000177 (0.231)	0.0000153 (0.305)	0.0000176 (0.233)
RM*COSSE	0.000137 (0.000)	0.000137 (0.000)	0.000137 (0.000)	0.000137 (0.000)
$\lambda$		-2.840 (0.000)		-2.836 (0.000)
$\rho$			-0.0000286 (0.005)	-0.0000216 (0.027)
Const.	3.785 (0.000)	4.871 (0.000)	3.786 (0.000)	4.869 (0.000)
R <sup>2</sup>	0.502	0.505	0.502	0.505
F	334.3	318.4	985.5	671.5
Valor P	0.000	0.000	0.000	0.000

Fonte: Elaboração Própria a partir dos Dados da PNAD. Nota: p-valor entre parênteses.

Tabela 1.7 – Estimação da Equação de Rendimentos por Alternativos Métodos com *Dummies* Regionais no Setor de Profissionais das Indústrias de Transformação e Produção.

Variáveis	MQO	Heckman	KV	Heckman e KV
E	0.129 (0.000)	0.129 (0.000)	0.129 (0.000)	0.129 (0.000)
Exp	0.0506 (0.000)	0.0203 (0.008)	0.0506 (0.000)	0.0204 (0.007)
Exp <sup>2</sup>	-0.000944 (0.000)	0.00000178 (0.995)	-0.000945 (0.000)	-0.000000750 (0.998)
G	0.666 (0.000)	0.669 (0.000)	0.666 (0.000)	0.669 (0.000)
R	0.162 (0.000)	0.146 (0.000)	0.162 (0.000)	0.146 (0.000)
U	0.183 (0.000)	0.150 (0.000)	0.183 (0.000)	0.151 (0.000)
E*G	-0.0555 (0.000)	-0.0548 (0.000)	-0.0555 (0.000)	-0.0548 (0.000)
E*A	0.0271 (0.000)	0.0266 (0.000)	0.0271 (0.000)	0.0266 (0.000)
RP	-0.0000335 (0.038)	-0.0000292 (0.068)	-0.0000335 (0.038)	-0.0000292 (0.068)
RM	0.000359 (0.000)	0.000349 (0.000)	0.000360 (0.000)	0.000349 (0.000)
RP*A	0.0000461 (0.010)	0.0000449 (0.013)	0.0000461 (0.010)	0.0000449 (0.013)
RM*A	-0.000237 (0.000)	-0.000236 (0.000)	-0.000237 (0.000)	-0.000236 (0.000)
RP*G	0.0000146 (0.322)	0.0000149 (0.314)	0.0000146 (0.322)	0.0000149 (0.312)
RM*G	0.0000693 (0.107)	0.0000643 (0.127)	0.0000693 (0.108)	0.0000640 (0.130)
RP*COSSE	0.0000649 (0.000)	0.0000635 (0.000)	0.0000649 (0.000)	0.0000635 (0.000)
RM*COSSE	0.0000313 (0.329)	0.0000341 (0.279)	0.0000313 (0.329)	0.0000341 (0.279)
$\lambda$		-5.178 (0.000)		-5.190 (0.000)
$\rho$			0.0000753 (0.931)	0.000357 (0.567)
Const.	4.099 (0.000)	5.860 (0.000)	4.099 (0.000)	5.865 (0.000)
R <sup>2</sup>	0.459	0.462	0.459	0.462
F	247.7	238.4	233.2	225.2
Valor P	0.000	0.000	0.000	0.000

Fonte: Elaboração Própria a partir dos Dados da PNAD. Nota: p-valor entre parênteses.

As variáveis utilizadas nas estimações apresentadas nas tabelas 1.2, 1.3 e 1.4 (E, G, R, U, RP e RM), apresentam o mesmo comportamento nas tabelas 1.5, 1.6 e 1.7, salvo algumas exceções, visto que, em alguns casos o coeficiente da variável RP mostra-se estatisticamente insignificante. Portanto, os comentários que serão realizados sobre os resultados verificados nas tabelas 1.5, 1.6 e 1.7 não tratarão dessas variáveis.

Verifica-se que o coeficiente da variável experiência (Exp) apresenta sinal positivo, em todas as estimações realizadas, indicando que, independentemente da idade, indivíduos mais experientes, ou seja, com mais tempo de trabalho, recebem em média, maiores salários. Já o coeficiente da variável experiência ao quadrado (Exp<sup>2</sup>) mostrou-se estatisticamente insignificante nas subamostras dos profissionais da área de serviços e entre os técnicos e trabalhadores das indústrias de transformação e produção, onde tal resultado pode ser explicado pela baixa idade dos indivíduos amostrados. Já nas estimativas realizadas para a amostra toda, tal coeficiente mostra-se estatisticamente significativo e apresenta sinal negativo, corroborando, portanto, com os resultados já difundidos na literatura.

Quanto às interações entre as variáveis de renda do pai e da mãe e as *dummies* de gênero (RP\*G e RM\*G) e ano (RP\*A e RM\*A), verifica-se que, apenas o coeficiente da variável RM\*A mostra-se estatisticamente significativo, apresentando sinal negativo, em todas as estimações realizadas, indicando que o impacto da renda da mãe sobre a renda dos filhos era maior em 1989 do que em 2009. Portanto, esse resultado, gera certo indicativo de que, através da renda da mãe, o Brasil esteja apresentando uma redução no grau de transmissão de renda, contribuindo, portanto, para uma melhoria nos níveis de mobilidade. Já o coeficiente da variável RM\*G mostra-se estatisticamente insignificante em todas as estimações realizadas, logo, pode-se concluir que não há impactos diferenciados da renda materna nos salários dos filhos e das filhas. Os coeficientes das demais variáveis (RP\*A e RP\*G) alternam entre significância e insignificância dependendo da amostra considerada, portanto, não se pode chegar a uma conclusão robusta acerca dos seus impactos.

Em relação às interações das variáveis de renda do pai e da mãe com a *dummy* referente às regiões Centro-Oeste, Sudeste e Sul (RP\*COSSE e RM\*COSSE), verifica-se que na subamostra de serviços o coeficiente da variável RP\*COSSE e na subamostra dos técnicos e profissionais das indústrias de transformação e produção o coeficiente da variável

RM\**COSSE* mostram-se estatisticamente insignificantes. Nos demais casos, tais coeficientes são significantes e apresentam sinais positivos. Diante disso, pode-se afirmar que há fortes indicativos para a ocorrência de maiores impactos de transmissão de renda de pais para filhos nas regiões mais desenvolvidas.

Esse resultado corrobora com os demais trabalhos revistos na literatura brasileira sobre o assunto, visto que, estes também verificaram que nas regiões mais ricas existe um maior grau de transmissão de renda de pai para filho, indicando, portanto, que no Brasil o grau de imobilidade socioeconômico é elevado, visto que, filhos de pais mais ricos receberão, em média, os maiores salários quando comparados com os jovens advindos das famílias pertencentes às classes menos favorecidas.

Conforme explicitado anteriormente, este trabalho também utiliza a metodologia de regressão quantílica na tentativa de obter impactos mais precisos, das variáveis explicativas, ao considerar os níveis de renda dos filhos. As tabelas 1.8, 1.9 e 1.10 apresentam, portanto, os resultados das estimações via regressões quantílicas para todas as ocupações, setor de serviços e profissionais das indústrias de produção e transformação respectivamente.

Tabela 1.8 – Estimação da Equação de Rendimentos através do Método de Regressão Quantílica em Todas as Ocupações.

Variáveis	q10		q25		q50		q75		q90	
E	0.24	(0.199)	0.188	(0.000)	0.142	(0.000)	0.0639	(0.000)	0.118	(0.000)
Exp	0.0315	(0.950)	0.0112	(0.035)	0.0333	(0.000)	0.164	(0.000)	0.0601	(0.745)
Exp <sup>2</sup>	-0.00038	(0.985)	0.000315	(0.122)	-0.00032	(0.081)	-0.00466	(0.000)	-0.00131	(0.747)
G	1.527	(0.424)	1.183	(0.000)	0.733	(0.000)	0.224	(0.000)	0.547	(0.589)
R	0.185	(0.000)	0.107	(0.000)	0.104	(0.000)	0.125	(0.000)	0.0819	(0.037)
U	0.245	(0.713)	0.119	(0.000)	0.0576	(0.000)	0.416	(0.000)	0.0346	(0.950)
E*G	-0.138	(0.534)	-0.113	(0.000)	-0.0644	(0.000)	-0.0114	(0.004)	-0.0347	(0.716)
E*A	0.0338	(0.000)	0.0374	(0.000)	0.0256	(0.000)	0.0384	(0.000)	0.00606	(0.451)
RP	-1.1E-05	(0.693)	-6.3E-06	(0.493)	1.08E-05	(0.088)	-5.4E-05	(0.000)	0.000029	(0.047)
RM	0.000174	(0.023)	0.000251	(0.000)	0.000394	(0.000)	0.000579	(0.000)	0.00054	(0.000)
RP*A	-2.8E-06	(0.917)	1.64E-05	(0.031)	2.82E-05	(0.000)	-7.5E-05	(0.000)	5.49E-05	(0.680)
RM*A	-7.8E-05	(0.475)	-0.00018	(0.000)	-0.00021	(0.000)	-0.0003	(0.000)	-0.0002	(0.470)
RP*G	-8.9E-06	(0.841)	-1.2E-05	(0.080)	-8.3E-06	(0.103)	5.44E-05	(0.000)	9.69E-06	(0.535)
RM*G	1.21E-05	(0.907)	4.33E-05	(0.130)	3.63E-05	(0.305)	-5.3E-05	(0.004)	-6.2E-05	(0.469)
RP*COSSE	5.57E-05	(0.114)	4.28E-05	(0.000)	0.000015	(0.006)	8.48E-05	(0.000)	1.73E-05	(0.225)
RM*COSSE	3.41E-05	(0.374)	0.000102	(0.000)	0.000135	(0.000)	-0.00013	(0.000)	8.89E-06	(0.954)
$\lambda$	-5.195	(0.924)	-3.978	(0.000)	-1.471	(0.000)	4.139	(0.000)	-0.826	(0.923)
$\rho$	-9.1E-05	(0.923)	-5.8E-05	(0.946)	-2.2E-05	(0.992)	-5.7E-06	(0.997)	4.35E-05	(0.986)
Const.	4.122	(0.819)	4.805	(0.000)	4.656	(0.000)	2.902	(0.000)	5.121	(0.136)
Pseudo R <sup>2</sup>	0.2808		0.2879		0.3069		0.3655		0.399	

Fonte: Elaboração própria a partir dos dados da PNAD.

Nota: p-valor entre parênteses.

Tabela 1.9 – Estimação da Equação de Rendimentos através do Método de Regressão Quantílica no Setor de Serviços.

Variáveis	q10		q25		q50		q75		q90	
E	0.202	(0.000)	0.184	(0.000)	0.147	(0.000)	0.104	(0.000)	0.0982	(0.000)
Exp	0.0208	(0.388)	-0.0102	(0.438)	-0.00829	(0.196)	0.0166	(0.008)	0.0289	(0.003)
Exp <sup>2</sup>	0.00013	(0.889)	0.00103	(0.046)	0.00109	(0.000)	9.59E-05	(0.708)	-0.00037	(0.339)
G	1.186	(0.000)	1.237	(0.000)	0.997	(0.000)	0.766	(0.000)	0.783	(0.000)
R	0.119	(0.000)	0.0856	(0.000)	0.0688	(0.000)	0.0482	(0.000)	0.0638	(0.000)
U	0.195	(0.000)	0.0841	(0.062)	0.042	(0.050)	0.0552	(0.006)	0.0807	(0.000)
E*G	-0.085	(0.000)	-0.115	(0.000)	-0.0941	(0.000)	-0.0599	(0.000)	-0.0534	(0.000)
E*A	0.0389	(0.000)	0.0481	(0.000)	0.0409	(0.000)	0.0284	(0.000)	0.02	(0.004)
RP	4.89E-06	(0.898)	2.48E-05	(0.334)	4.04E-05	(0.025)	5.36E-05	(0.008)	5.64E-05	(0.016)
RM	0.000146	(0.047)	0.000244	(0.000)	0.000418	(0.000)	0.000654	(0.000)	0.000713	(0.000)
RP*A	-1.3E-06	(0.945)	-1.5E-05	(0.267)	1.72E-05	(0.226)	4.26E-05	(0.069)	0.000016	(0.603)
RM*A	-9.7E-05	(0.023)	-0.00019	(0.000)	-0.00033	(0.000)	-0.00035	(0.000)	-0.0003	(0.000)
RP*G	-1.7E-05	(0.303)	-3.3E-06	(0.823)	-7.5E-06	(0.389)	2.55E-06	(0.886)	0.000011	(0.633)
RM*G	2.38E-05	(0.517)	2.58E-05	(0.474)	4.45E-05	(0.191)	-9.6E-05	(0.079)	-0.00018	(0.000)
RP*COSSE	6.62E-05	(0.017)	3.96E-05	(0.004)	-1.3E-06	(0.929)	-2.2E-05	(0.256)	1.36E-05	(0.560)
RM*COSSE	0.000112	(0.003)	0.000125	(0.000)	0.000206	(0.000)	0.000163	(0.000)	0.000056	(0.222)
Λ	-4.148	(0.000)	-4.184	(0.000)	-2.534	(0.000)	-1.687	(0.000)	-1.193	(0.011)
P	-8.7E-05	(0.995)	-4.7E-05	(0.995)	-1.4E-05	(0.997)	2.05E-05	(0.994)	6.63E-05	(0.977)
Const.	4.112	(0.000)	4.963	(0.000)	5.051	(0.000)	5.223	(0.000)	5.258	(0.000)
Pseudo R <sup>2</sup>	0.3403		0.3321		0.2975		0.3284		0.3582	

Fonte: Elaboração própria a partir dos dados da PNAD.

Nota: p-valor entre parênteses.

Tabela 1.10 – Estimação da Equação de Rendimentos através do Método de Regressão Quantílica no Setor dos Profissionais das Indústrias de Produção e Transformação.

Variáveis	q10		q25		q50		q75		q90	
E	0.234	(0.000)	0.127	(0.000)	0.0845	(0.000)	0.0611	(0.000)	0.0769	(0.000)
Exp	0.00315	(0.882)	-0.0014	(0.873)	0.0206	(0.020)	0.0221	(0.000)	0.0529	(0.000)
Exp <sup>2</sup>	0.000468	(0.533)	0.000638	(0.055)	-1.6E-05	(0.959)	-0.00023	(0.269)	-0.00115	(0.003)
G	1.756	(0.000)	0.764	(0.000)	0.347	(0.000)	0.265	(0.000)	0.174	(0.014)
R	0.202	(0.000)	0.0874	(0.000)	0.0934	(0.000)	0.0679	(0.000)	0.0618	(0.004)
U	0.287	(0.000)	0.147	(0.000)	0.0458	(0.007)	0.0126	(0.397)	-0.0144	(0.591)
E*G	-0.166	(0.000)	-0.0763	(0.000)	-0.0281	(0.000)	-0.0127	(0.046)	9.59E-05	(0.993)
E*A	0.0475	(0.000)	0.0479	(0.000)	0.0322	(0.000)	0.023	(0.000)	0.00659	(0.298)
RP	-9.9E-05	(0.009)	-3.6E-05	(0.097)	-1.1E-05	(0.512)	-1.4E-05	(0.243)	-1.9E-05	(0.224)
RM	0.000263	(0.000)	0.000243	(0.000)	0.000461	(0.000)	0.00068	(0.000)	0.000579	(0.000)
RP*A	-1.2E-05	(0.592)	9.35E-06	(0.661)	5.98E-05	(0.001)	8.61E-05	(0.000)	0.000102	(0.001)
RM*A	-0.00022	(0.001)	-0.00025	(0.000)	-0.00029	(0.000)	-0.0003	(0.000)	-0.00022	(0.000)
RP*G	5.38E-05	(0.020)	0.00002	(0.287)	7.37E-06	(0.518)	2.21E-05	(0.143)	5.89E-05	(0.003)
RM*G	5.23E-05	(0.330)	0.00013	(0.001)	7.36E-05	(0.083)	-3.7E-05	(0.200)	-7.5E-05	(0.159)
RP*COSSE	0.000115	(0.001)	7.11E-05	(0.000)	3.51E-05	(0.023)	2.84E-05	(0.033)	2.22E-05	(0.209)
RM*COSSE	4.47E-05	(0.379)	0.000138	(0.000)	0.000146	(0.002)	4.42E-05	(0.092)	3.46E-05	(0.163)
Λ	-10.56	(0.001)	-5.836	(0.000)	-2.69	(0.006)	-3.244	(0.001)	-1.05	(0.492)
P	-0.00064	(0.935)	0.000218	(0.702)	0.00172	(0.196)	0.00232	(0.110)	-0.00321	(0.510)
Const.	5.777	(0.000)	5.933	(0.000)	5.572	(0.000)	6.201	(0.000)	5.638	(0.000)
Pseudo R <sup>2</sup>	0.238		0.262		0.2878		0.3412		0.3815	

Fonte: Elaboração própria a partir dos dados da PNAD.

Nota: p-valor entre parênteses.

Conforme verificado nas tabelas 1.2 a 1.7, os coeficientes estimados pelos métodos de MQO, Heckman em dois estágios, Klein e Vella, e um modelo que combina estes dois últimos apresentaram valores bastante aproximados, cabendo realçar, portanto, que estimativas obtidas através de métodos com extensões corretivas nesta amostra não se diferenciam daquelas via MQO.

Os resultados das regressões quantílicas expostos nas tabelas 1.8, 1.9 e 1.10 se baseiam no modelo que utiliza as duas medidas corretivas em conjunto para medir possíveis variações de impacto entre os diferentes quantis de renda. Constata-se que existem realmente impactos diferenciados entre os quantis de renda, como por exemplo, pode-se citar o retorno da educação, que chega a apresentar variações de mais de 10 pontos percentuais, nas três amostras consideradas, entre os quantis de renda. Portanto, a partir das estimativas obtidas, pode-se constatar que o adicional salarial correspondente a um ano a mais de estudo não é o mesmo para os que estão nos quantis inferiores e superiores da distribuição de rendimentos.

Quanto às principais variáveis de interesse, renda do pai e da mãe, verifica-se que a primeira, na maior parte dos quantis e nas três amostras consideradas, mostra-se estatisticamente insignificante, indicando que a rejeição de um possível impacto da renda paterna sobre a dos filhos, contrário a aceitação do efeito transmissão positivo e significativo da renda materna. Deste resultado, constata-se que tal efeito é mais forte nos quantis de renda mais elevada.

O sinal negativo do coeficiente da variável  $RM^*A$ , nas três amostras e nos quantis nos quais ele é significativo, indicam uma redução neste grau de transmissão de renda ao longo do tempo, visto que, *ceteris paribus*, o impacto de um aumento na renda materna era maior no salário dos filhos em 1989 do que em 2009. Entretanto, verifica-se que os coeficientes dessa variável apresentam-se praticamente constantes entre os quantis, indicando que a redução do grau de herança familiar está ocorrendo de maneira uniforme entre os quantis de rendimento. Portanto, a partir desses resultados pode-se corroborar com a tese de que, mesmo apresentando uma trajetória decrescente com o passar dos anos, existe um considerável grau de transmissão de renda no Brasil, principalmente nos quantis superiores de renda.

Os coeficientes das variáveis  $RP^*A$ ,  $RM^*G$  e  $RM^*G$  alternam entre significância e insignificância dependendo da amostra e dos quantis considerados, portanto, não se pode

chegar a uma conclusão robusta acerca de que a renda do pai esteja contribuindo para mudanças nos níveis de transmissão ao longo do tempo e de que as rendas paternas e maternas impactem de maneira diferenciada nos rendimentos de filhos e filhas, respectivamente.

Em relação às interações das variáveis de renda do pai e da mãe com a *dummy* referente às regiões Centro-Oeste, Sudeste e Sul (RP\*COSSE e RM\*COSSE), verifica-se que no quantil referente aos indivíduos de maior renda, os coeficientes destas variáveis mostram-se estatisticamente insignificantes nas três amostras consideradas. Já no quantil referente aos indivíduos de renda mais baixa, o coeficiente da variável RP\*COSSE é significativo nas amostras dos profissionais da área de serviços e entre os técnicos e trabalhadores das indústrias de transformação e produção, já o coeficiente da variável RM\*COSSE é significativo apenas na amostra de serviços.

Nos demais quantis, o coeficiente da variável RM\*COSSE é significativo, em todas as amostras, e verifica-se uma trajetória crescente em seus valores ao passar-se para quantis referentes a indivíduos com maiores rendimentos. Já para o coeficiente da variável RP\*COSSE, além de se mostrar insignificante em alguns outros quantis, não se pode inferir uma tendência na sua trajetória. Portanto, a partir dos resultados obtidos pelo método de regressões quantílicas pode-se verificar que existe transmissão de renda das mães para os filhos, e que este não ocorre de maneira linear, pois é mais acentuado, nos maiores estratos de rendimento. No entanto, percebe-se que está havendo uma redução neste grau de transmissão de rendimento com o passar dos anos.

## 5 CONSIDERAÇÕES FINAIS

Além de prover alguns pontos contributivos em aplicações empíricas que abordam o problema da endogeneidade da educação em uma equação de rendimentos individuais ao comparar distintas metodologias, a análise inova no sentido de incluir a essa questão um efeito dinâmico e um efeito transmissão de renda entre gerações, embora, por limitação de dados, tenham sido mensurados para apenas uma geração.

Utilizando dados da PNAD para os anos de 1989 e 2009, estimam-se equações *mincerianas* de rendimentos pelos métodos de MQO, Heckman em Dois Estágios e Mínimos Quadrados em Dois Estágios (MQ2E) com o uso de VI, onde, o método de Heckman leva em conta o problema de seletividade amostral, ao passo que, o de VI corrige uma possível endogeneidade presente em uma equação de rendimentos. Além disso, esta análise apresenta uma inovação à literatura voltada à estimação de uma equação de rendimentos, pois apresenta uma metodologia pouco explorada em trabalhos internacionais e ainda não utilizada em artigos nacionais, denominada de método de Klein e Vella (2010).

As estimações revelaram que o método de variáveis instrumentais apresentou pior performance para fins preditivos relativo aos demais. Esse resultado está em conformidade com Mariano e Arraes (2011) ao atestarem que a correção do problema da endogeneidade através da utilização de instrumentos conduz a estimativas menos confiáveis. Isso acontece devido a restrições que ocorrem quanto às fontes de dados disponíveis. Convém mencionar que para a amostra aqui utilizada, os coeficientes obtidos pelas metodologias de Heckman em dois estágios e de Klein e Vella (2010) apresentaram valores bastante aproximados, não se podendo, portanto, chegar a uma conclusão a cerca de qual dos dois modelos torna-se mais apropriado.

Em relação aos resultados, verificou-se que os coeficientes da grande maioria das variáveis apresentaram o mesmo comportamento já difundido nesta literatura, onde tal fato pode ser comprovado ao se inspecionar os trabalhos de Mariano e Arraes (2011), Suliano e Siqueira (2010), Salvato e Silva (2008), Carvalho, Neri e Silva (2006), dentre outros. Um resultado de certa forma não esperado, porém verificado por Hirvonen (2010), foi o da interação entre a variável anos de estudo do filho e a *dummy* de gênero, visto que, verificou-se

que o impacto de um ano a mais de estudo é maior no salário das filhas. A partir de alguns indícios, verificados em Leone e Baltar (2006), acredita-se que esse resultado possa ser explicado pelo acentuado crescimento das mulheres no mercado de trabalho nas últimas décadas, buscando maiores qualificações para competirem de forma mais igualitária com os homens.

No que se refere ao principal objetivo deste trabalho, pode-se inferir que existe transmissão de renda dos pais para seus filhos, uma vez que os coeficientes das variáveis renda do pai e mãe apresentaram sinais positivos. Em seguida foram estimadas variações da equação *minceriana*, na tentativa de se verificar possíveis efeitos diferenciados de transmissão de renda de pais para filhos no Brasil, através da utilização de interações das variáveis renda do pai e da mãe com *dummies* de gênero, ano e uma *dummy* referente às regiões Centro-Oeste, Sudeste e Sul.

No tocante às interações utilizadas para verificar o efeito transmissão, por gênero e regiões, conclui-se que, as rendas paternas e maternas exercem impactos positivos e semelhantes nos salários de filhos e filhas. Ratificaram-se, ainda, os resultados já encontrados na literatura de que em regiões mais ricas, as rendas dos pais, principalmente, neste estudo, através da renda da mãe, visto que, a renda paterna não apresentou certa consistência nos exercícios realizados, exerce um maior impacto sobre as rendas dos seus filhos, indicando, de certa forma, certo nível de imobilidade de renda no Brasil, visto que, *ceteris paribus*, filhos de pais mais ricos apresentam maiores salários.

Em seguida, devido à incapacidade de se obter *contrafactuais* perfeitos a partir dos dados da PNAD, estimou-se a equação *minceriana* pela metodologia de regressões quantílicas. A partir dos resultados obtidos pode-se concluir que a renda da mãe impacta positivamente a renda dos filhos, sendo este, maior nos quantis de renda mais elevada, indicando, portanto, a transmissão de renda das mães para os filhos não ocorre de maneira linear, sendo, portanto, mais acentuada, nos maiores estratos de renda. No entanto, percebe-se que está havendo uma redução neste grau de transmissão de renda, visto que, o impacto de um aumento na renda materna era maior no salário dos filhos em 1989 do que em 2009.

Diante do exposto, pode-se concluir que, os resultados encontrados neste exercício empírico apontam para a existência de transmissão de renda de pais para filhos no

Brasil, principalmente nos extratos mais elevados de renda, indicando, portanto, que filhos de pais mais ricos tendem a ser mais ricos, e que, esta transmissão, ocorre principalmente, através da renda da mãe. No entanto, se observou uma redução nos níveis de transmissão entre os anos de 1989 e 2009, o que indica que o Brasil vem apresentando uma melhoria nos níveis de mobilidade.

## **DIFERENCIAL DE SALÁRIO POR GÊNERO NAS OCUPAÇÕES E REGIÕES BRASILEIRAS**

### **1 INTRODUÇÃO**

A participação da mulher na atividade econômica no Brasil aumentou durante as últimas décadas, verificada a partir da década de 1970 (HOFFMANN; LEONE, 2004). Em relatório elaborado pelo Instituto Brasileiro de Geografia e Estatística (IBGE), a partir dos dados coletados pela Pesquisa Mensal de Emprego (PME), observou-se que a participação das mulheres na população ocupada continuou a apresentar trajetória crescente no início deste século, visto que o percentual de mulheres ocupadas passou de 43% em 2003 para 45,4% em 2011, ao passo que, o percentual de homens ocupados apresentou uma redução de 57% em 2003 para 54,6% em 2011.

Diversas causas passaram a ser apontadas como os principais determinantes da evolução feminina no mercado de trabalho brasileiro. Para Costa (1990) a produção crescente de bens e serviços antes produzidos no domicílio, desde produtos alimentícios até cuidar de crianças, doentes e velhos, a pronunciada queda da fecundidade, que encurtou o tempo gasto em gestações e na guarda das crianças pequenas, e a urbanização, colocaram a mulher em contato com o mercado de trabalho.

Já para Scorzafave e Menezes-Filho (2006) e Souza, Rios-Neto e Queiroz (2011) a melhoria do nível educacional das mulheres e a redução das taxas de fecundidade, respectivamente, são os principais responsáveis pelo aumento da taxa de participação das mulheres no mercado de trabalho no Brasil.

Além disso, a consolidação da participação da mulher no mercado de trabalho não se refletiu somente na aproximação, entre homens e mulheres, das taxas de participação, mas também na diminuição do hiato salarial entre os gêneros. De acordo com Hoffmann e Leone (2004), em 1981, o rendimento médio do trabalho da mulher equivalia a 55,7% do rendimento médio do trabalho do homem e essa relação passou a ser de 70,6% em 2002. No entanto, de acordo com o IBGE/PME (2012), pelo terceiro ano consecutivo, o rendimento feminino mantém basicamente a mesma proporção (72,3%) em relação ao rendimento dos homens, visto que, em 2003 as mulheres recebiam 70,8% do que recebia, em média, um homem.

Diante destes fatos, percebe-se que aumentou a participação feminina no mercado de trabalho brasileiro, entretanto, existe um diferencial de rendimentos negativos médios para as mulheres, em relação aos homens, e que, esta diferença que vinha seguindo uma trajetória decrescente, não vem apresentando redução nos últimos anos. Portanto, de acordo com Sachsida e Loureiro (1998), para fins de política econômica, é importante verificar a existência, ou não, de diferenciais de salários entre gêneros, a partir do aumento da demanda por mão-de-obra feminina, incentivado pelo crescimento da economia.

De acordo com Flabbi (2005), a diferença salarial entre homens e mulheres, condicionada às características de produtividade observadas, têm sido considerada uma possível indicação de preconceito contra as mulheres no mercado de trabalho. No entanto, segundo o autor, não há provas conclusivas sobre se esses diferenciais são devido à discriminação no mercado de trabalho ou às diferenças de produtividade não observadas.

Além disso, Sachsida e Loureiro (1998) argumentam que a simples verificação das diferenças salariais entre homens e mulheres não pode ser caracterizada como discriminação salarial, a não ser que essa diferença salarial não esteja associada ao capital humano. De acordo com Becker (1957) há indício de discriminação se homens e mulheres fossem substitutos perfeitos e ganhassem salários diferentes. Enquanto que Oaxaca (1973) defende a ideia de que a cultura, tradições e discriminação tendem a gerar restrições à participação da mulher no mercado de trabalho. Essas influências combinam para gerar uma distribuição ocupacional desfavorável para as mulheres e criar diferenças salariais entre homens e mulheres pertencentes à mesma ocupação.

De acordo com Richard (2007) se as diferenças salariais existentes no mercado de trabalho são atribuídas à discriminação, ao invés de diferenças de dotações de capital humano, pode-se limitar o desenvolvimento humano e econômico de milhões de pessoas. Pois, conforme argumentam Polachek e Kim (1994), se o diferencial de salários entre os gêneros surge em função da existência de oportunidades desiguais, originadas de discriminação, então a economia perde eficiência alocativa, visto que, postos de trabalho que exigem alta qualificação talvez não sejam ocupados pelos trabalhadores mais adequados.

Com esteio nestas argumentações, percebe-se que a caracterização de existência de discriminação salarial entre homens e mulheres não é tão óbvia quanto parece, assim como bem argumentou, Sachsida e Loureiro (1998).

Diante do exposto, este capítulo visa analisar a existência de uma possível diferença salarial entre homens e mulheres no mercado de trabalho no Brasil. Para alcançar tal objetivo, utilizam-se dados da Pesquisa Nacional por Amostra de Domicílios (PNAD) referente ao ano de 2009 para realizar estimações da equação *minceriana* de salários. Primeiramente, serão realizadas estimações pelo método de Mínimos Quadrados Ordinários (MQO) e, em seguida, pelo procedimento de Heckman em dois estágios, visto que, conforme argumentado na primeira análise deste estudo, este método corrige o viés de seleção, fato este corroborado pelos resultados encontrados em Carvalho, Neri e Silva (2006), Suliano e Siqueira (2010) e Mariano e Arraes (2011).

Além desses métodos, para verificar se o diferencial de salário é decorrente do gênero, emprega-se o método de *propensity score matching (PSM)*<sup>18</sup> para realizar uma comparação de salários entre homens e mulheres a partir de uma amostra adequada; isto é, controlando por características observáveis dos indivíduos como, por exemplo, educação, idade, ocupação e região censitária. Assim, com a utilização dessa metodologia, espera-se obter indivíduos que comporão grupo de controle (mulheres) com características bastante similares aos indivíduos do grupo de tratamento (homens). Dessa forma, a comparação seria realizada entre homens e mulheres considerados praticamente substitutos perfeitos, em relação às características observáveis, e, portanto, a diferença salarial deve-se exclusivamente ao gênero, indicando, portanto, a presença de discriminação.

Além disso, para decompor o diferencial de salários entre homens e mulheres será utilizada a metodologia desenvolvida por Oaxaca (1973) e Blinder (1973), a qual permite decompor as diferenças salariais em dois efeitos: da discriminação (não explicada) e das diferenças entre as características individuais (explicada). Portanto, nos mesmos moldes de Gravelle e Hole (2008), são realizadas decomposições do diferencial de salário médio entre homens e mulheres a partir da equação *minceriana* considerando a estimação pelo método de Heckman em dois estágios.

---

<sup>18</sup> Utiliza-se de quatro formas de pareamento, a saber: i) um para um; ii) vizinho mais próximo; iii) pela distância do score de propensão; e, iv) pela densidade de kernel.

Em virtude das diferenças entre as regiões e as ocupações<sup>19</sup> no Brasil, este capítulo também tem como objetivo verificar possíveis diferenciais salariais, entre homens e mulheres, dentro de uma mesma região e ocupação. Para isso, as metodologias supracitadas serão efetuadas em três níveis regionais (Brasil, Nordeste e Sudeste-Sul) e quatro níveis ocupacionais (todas as ocupações, dirigentes, serviços<sup>20</sup> e técnicos<sup>21</sup>).

Além desta introdução, este artigo será composto por mais quatro seções. Em seguida será realizada uma revisão dos trabalhos realizados que buscaram analisar o diferencial de salários e a discriminação salarial entre homens e mulheres, tanto em nível nacional como internacional. Na terceira seção será apresentada a metodologia utilizada para a consecução dos resultados. Em seguida serão apresentados e discutidos os principais resultados obtidos. E, finalmente, serão tecidas as principais conclusões deste estudo.

---

<sup>19</sup> A PNAD agrega os indivíduos nos seguintes níveis ocupacionais: Dirigentes em geral, Profissionais das ciências e das artes, Técnicos de nível médio, Trabalhadores de serviços administrativos, Trabalhadores dos serviços, Vendedores e prestadores de serviço do comércio, Trabalhadores agrícolas, Trabalhadores da produção de bens e serviços e de reparação e manutenção e Membros das forças armadas e auxiliares. Neste trabalho, optou-se por excluir os Profissionais das ciências e das artes, os Trabalhadores agrícolas e os Membros das forças armadas e auxiliares.

<sup>20</sup> A categoria de serviços contempla os Trabalhadores de serviços administrativos, Trabalhadores dos serviços e Vendedores e prestadores de serviço do comércio.

<sup>21</sup> A categoria dos técnicos engloba os Técnicos de nível médio e Trabalhadores da produção de bens e serviços e de reparação e manutenção.

## 2 REVISÃO DA LITERATURA

Visando verificar a existência de diferença salarial e discriminação em gênero no mercado de trabalho, as investigações na literatura empregam diferentes metodologias aplicadas a informações sobre os determinantes do salário dos trabalhadores, tanto a partir de suas características individuais quanto do mercado de trabalho, buscando identificar possíveis causas para tal ocorrência.

Oaxaca (1973) e Blinder (1973) contribuíram com essa questão ao apontarem que tais investigações eram tratadas com abordagens essencialmente descritivas, e em decorrência, elaboraram um procedimento metodológico de decomposições de efeitos a fim de identificar a existência de discriminação no mercado de trabalho. No entanto, conforme argumentam Carvalho, Neri e Silva (2006) as medidas descritivas, apesar de apresentarem os diferenciais de rendimentos entre os grupos de comparações como, por exemplo, cor e sexo, não são suficientes para avaliar o quanto deste diferencial é explicado pela discriminação ou pelas características individuais (escolaridade, experiência de trabalho, local de residência e idade, por exemplo).

Os estudos sobre os determinantes salariais na literatura econômica foram impulsionados a partir de Mincer (1974) que concebeu uma equação para salários, denominada como “função salário do capital humano”, conhecida como *equação minceriana*, cuja especificação original expressa os rendimentos como função de escolaridade e experiência. Extensões desse modelo incorporaram diversos outros atributos individuais como fatores explicativos: gênero, raça, habilidades, etc.

Para estimar os determinantes salariais a partir do arcabouço teórico desenvolvido por Mincer (1974) foram aplicadas inicialmente para comprovação empírica, diversos métodos estatísticos convencionais, predominando o método de MQO. No entanto, segundo Heckman (1979) ao estimar uma equação de rendimentos deve-se levar em consideração o viés de seletividade, que pode surgir tanto pela auto-seleção de indivíduos ou unidades amostrais investigados quanto pelas decisões de seleção da amostra pelos analistas.

Complementando o argumento de Heckman (1979), Neuman e Oaxaca (2001) indicam que o viés de seletividade pode ser encontrado em duas etapas do processo de

emprego: na fase em que o indivíduo resolve ingressar na força de trabalho empregada e quando uma ocupação específica ou um *status* profissional é escolhido. Para solucionar o problema do possível viés de seletividade amostral, utiliza-se o procedimento de Heckman (1979) em dois estágios, conforme tem sido amplamente aplicado na literatura<sup>22</sup>.

No entanto, a simples aferição de diferenças salariais, obtidas a partir da estimação de uma equação de rendimentos, não garante que exista discriminação no mercado de trabalho. As estimativas da presença de discriminação no mercado de trabalho tornaram-se rotina a partir da metodologia desenvolvida por Blinder (1973) e Oaxaca (1973) que permite realizar uma decomposição das diferenças salariais em dois efeitos, discriminação (não explicada) e diferenças entre as características individuais (explicada).

De acordo com Neuman e Oaxaca (2001) essas metodologias baseiam-se na utilização de equações de salários, estimadas separadamente para dois grupos de trabalhadores, visando a decomposição das diferenças salariais. Muitas vezes, selecionam-se a estrutura salarial para o grupo de trabalhadores que se acredita ser dominante no mercado de trabalho (pelo menos em relação ao grupo de comparação). As diferenças nas características médias dos dois grupos são ponderadas pelos coeficientes estimados e somadas para obter a parcela atribuída ao capital humano no diferencial salarial global. A porção atribuída à discriminação no diferencial salarial global é o resíduo deixado após a exclusão da parcela atribuída ao capital humano.

Após os trabalhos pioneiros de Oaxaca (1973) e Blinder (1973), Neumark (1988), Oaxaca e Ransom (1988), e Oaxaca e Ransom (1994) utilizam uma abordagem mais geral para decomposições salariais, onde a estrutura de salário não discriminatória é estimada a partir de uma amostra combinada dos dois grupos de comparação. De acordo com Neuman e Oaxaca (2001) essa abordagem permite que o componente de discriminação seja desagregado em *overpayment* (favoritismo) e *underpayment* (discriminação pura).

Outras técnicas, não derivadas da metodologia de Oaxaca (1973) e Blinder (1973), passaram a ser utilizadas para estimar as diferenças de rendimentos. Na tentativa de controlar os efeitos salariais individuais autores como Polachek e Kim (1994) e Rosholm e

---

<sup>22</sup> Ver, por exemplo: Carvalho, Neri e Silva (2006), Suliano e Siqueira (2010), Mariano e Arraes (2011), Neuman e Oaxaca (2001) e Richard (2007).

Smith (1996) utilizam a metodologia de dados em painel para estimar o diferencial de rendimento entre homens e mulheres, enquanto outros, como Hirsch e Macpherson (1994), Hirsch e Schumacher (1992), e Macpherson e Hirsch (1995), passaram a incorporar o gênero e composições étnicas de cada ocupação como determinantes de salários ocupacionais para medir a discriminação no mercado de trabalho.

Em outro conjunto de estudos, como os de Brown et al. (1980), Miller (1987), Reilly (1991), e Neuman e Silber (1996), a contribuição da segregação ocupacional ao diferencial de salário foi estimada separadamente, de modo que a diferença entre os salários dos grupos em questão foi agora decomposto em três componentes: dotação, discriminação salarial e segregação.

Além disso, Caliendo e Kopeinig (2008) argumentam que a metodologia de *Propensity Score Matching (PSM)*, desenvolvida por Rosenbaum e Rubin (1983), tornou-se uma abordagem popular para estimar os efeitos causais de um tratamento. De acordo com esses autores, o *PSM* tenha sido formulado originalmente para aplicação em campos diversos da Economia, atualmente tem sido amplamente utilizado como uma importante ferramenta metodológica na avaliação de políticas de mercado de trabalho.

Desse modo, ao interpretar o gênero como um *tratamento*, Gravelle e Holle (2008) estimaram equações de salários pelo método *PSM* e, em seguida, utilizaram uma versão não linear da decomposição de Oaxaca (1973) e Blinder (1973) para decompor diferenças salariais entre homens e mulheres pertencentes à classe dos clínicos gerais na Grã-Bretanha.

Diante disso, serão apresentados alguns trabalhos empíricos que buscaram estimar efeitos discriminatórios no mercado de trabalho, cujo foco será essencialmente centrado no objetivo deste trabalho, qual seja: investigar a ocorrência de discriminação salarial em gênero no Brasil.

## **2.1 Estudos Empíricos sobre Diferenciais Salariais**

Uma das primeiras tentativas de se formular uma medida que captasse discriminação no mercado de trabalho foi formulada por Oaxaca (1973) por meio de um

coeficiente que decompõe as diferenças salariais entre homens e mulheres nos Estados Unidos. Os resultados indicam que o diferencial de salários entre homens e mulheres é bastante significativo, haja vista a considerável concentração de mulheres em empregos menos remunerados. Além disso, o autor sugere que uma substancial proporção desses diferenciais é atribuída aos efeitos da discriminação.

Na mesma linha de Oaxaca (1973), Blinder (1973) desenvolve uma metodologia para quantificar a discriminação no mercado de trabalho americano por meio de um sistema de equações simultâneas composto por 7 funções lineares, sendo a primeira delas representada pela equação *minceriana* constituída de variáveis *dummies* referentes à educação, ocupação, treinamento vocacional, tempo no presente trabalho, estado civil e *status* de veterano, treze variáveis de *background* familiar e mais um conjunto de variáveis exógenas. A partir da estimação do sistema, ele conclui que o diferencial de salário entre homens e mulheres é atribuído em cerca de 2/3 à discriminação absoluta no mercado de trabalho.

Polachek e Kim (1994) desenvolvem modelos específicos de inclinação para o indivíduo (em oposição aos modelos específicos de intercepto para o indivíduo) para considerar a heterogeneidade não observada na estimativa das diferenças salariais entre homens e mulheres. As estimativas desses modelos são comparadas com as de MQO tradicionais, bem como, com as de efeitos fixos e aleatórios para abordagens específicas de intercepto. Os autores verificaram que, inequivocamente, a heterogeneidade não observada responde por cerca de 50% da diferença salarial.

Para reavaliar o diferencial de salários entre homens e mulheres nos Estados Unidos, Choudhury (2001) utiliza a formulação de efeitos fixos e metodologia de Heckman para considerar o problema de omissão de variáveis, endogeneidade e seleção amostral. Um forte resultado que emerge é que o nível de experiência garantido no emprego, bem como, o nível de experiência alcançado no mercado de trabalho são os mais importantes determinantes dos salários. Os resultados das estimações das equações de salário mostram que existe uma diferença significativa nas estruturas de recompensa para experiência no mercado de trabalho entre homens e mulheres e é a fonte mais importante de discriminação. Além disso, para uma amostra de trabalhadores heterogênea, verificou-se que o salário das mulheres poderia aumentar em aproximadamente 12% na ausência de discriminação.

Levando em conta o viés de seletividade, Neuman e Oaxaca (2001) investigam as diferenças salariais em gênero entre duas etnias (ocidentais e orientais) de trabalhadores profissionais que residem em Israel por mais de dez anos. Inicialmente são realizadas estimações da equação *minceriana* de salários para proceder à decomposição de Oaxaca (1973) e, em seguida, corrigem o problema da seletividade utilizando o procedimento de Heckman em dois estágios. Os resultados indicam que as diferenças salariais entre homens e mulheres são, em média, maiores do que as diferenças salariais entre ocidentais e orientais, visto que, os homens judeus, tanto entre os ocidentais como entre os orientais, ganham 26% a mais (por hora) do que as mulheres judias, enquanto que os ocidentais, homens e mulheres, ganham 19% a mais do que os orientais. É também comprovado que essas diferenças salariais são superestimadas se a correção da seletividade não for considerada, o que indica, portanto, a presença do viés de seletividade amostral.

Otero, Lago e Sánchez (2004) apresentam as vantagens de levar em consideração a distribuição do diferencial salarial do indivíduo para medir a discriminação salarial feminina na Espanha a partir das metodologias de MQO e regressões nos quantis. Além disso, os autores discutem as abordagens anteriores<sup>23</sup> e as recentes propostas distributivas utilizando regressões quantílicas ou funções *contrafactuais*. No entanto, eles limitam a análise para o cálculo da diferença salarial média, impondo, desta forma, o mesmo peso em cada diferença salarial, independente da sua importância relativa ou valor dentro da distribuição salarial. Para isso, empregam a diferença salarial individual na medida de discriminação, como sugerido por Jenkins (1994). Concluem que, a discriminação absoluta aumenta com os salários observados, enquanto que em termos relativos, este resultado não é observado, pois somente as mulheres com salários muito baixos e muito altos, registraram uma significativa discriminação relativa<sup>24</sup>.

Objetivando propor uma solução para identificar discriminação no mercado de trabalho, Flabbi (2005) desenvolve um modelo de pesquisa combinando *matching*, negociação, e discriminação voluntária por parte dos empregadores. A idéia do autor é que, em equilíbrio, os empregadores tendem a discriminar as mulheres; ou seja, aqueles preconceituosos revelam suas preferências por trabalhadores do sexo masculino, que por sua

---

<sup>23</sup> Oaxaca (1973) e Blinder (1973).

<sup>24</sup> A noção de discriminação relativa baseia-se na razão entre a discriminação estimada e o salário sem discriminação, e não para o *gap* salário total, como é habitual na literatura.

vez gera efeitos colaterais sobre os não preconceituosos, piorando, assim, a posição de barganha das mulheres no mercado de trabalho. O modelo é caracterizado por quatro tipos de agentes: dois tipos de trabalhadores (masculino e feminino) e dois tipos de empregadores (preconceituosos e não preconceitos). Salários são determinados por meio de negociação entre empregadores e trabalhadores. Assim, o salário é uma função da produtividade, dos pontos de ameaça dos agentes envolvidos e do poder de barganha. As Funções de utilidade dos trabalhadores são lineares nos salários. Os resultados, da estimação por máxima verossimilhança, indicam que a produtividade das mulheres é 6,5% menor que a dos homens, e cerca da metade dos empregadores têm preconceito contra as mulheres. A partir da decomposição do diferencial de rendimentos, é constatado que o preconceito é o fator mais importante para explicar o diferencial de produtividade, vez que a presença de preconceito gera cerca de 2/3 do diferencial de rendimentos observado devido à discriminação, e o restante à diferença de produtividade.

Hsu (2006) investiga o papel do trabalho doméstico para explicar o diferencial de salário entre homens e mulheres. A partir de um sistema de equações simultâneas, os autores utilizam o tempo de alocação e modelos de utilidade da família, para testar empiricamente a endogeneidade do trabalho doméstico e analisar o impacto das horas empregadas em trabalhos domésticos no salário/hora da mulher. Os resultados indicam que as horas adicionais alocadas em trabalhos domésticos provocam uma elevada redução na produtividade da mulher no mercado de trabalho e, conseqüentemente, afetam negativa e significativamente os seus salários.

Em um estudo para o continente africano, Richard (2007) analisa os determinantes salariais de homens e mulheres e a discriminação de gênero no mercado de trabalho de Uganda nos anos de 2002 e 2003. As estimações das equações de salário são realizadas pelo método de Heckman para corrigir a seletividade na fase de entrada no mercado de trabalho. Além disso, utilizam-se as decomposições de Oaxaca (1973) e Neumark (1988) para decompor os diferenciais salariais entre os gêneros. Os resultados das estimações pelo método de Heckman indicam que a diferença de salários entre homens e mulheres é de 39%. Já os resultados das decomposições indicam que a maior parcela do diferencial salarial é atribuída à discriminação e que o maior componente não explicado da diferença salarial a favor dos homens decorre da desvantagem dos atributos femininos.

Para averiguar a existência de discriminação de gênero entre trabalhadores em uma mesma ocupação – clínicos gerais – Gravelle e Holle (2008) aplicam o método de decomposição de Oaxaca, obtida a partir da estimação de equações de salários por MQO, Mínimos Quadrados em Dois Estágios e *PSM*. Os autores propõem um novo teste para verificar a existência de discriminação de gênero nessa ocupação comparando as diferenças de renda em gênero, nas práticas em que todos os clínicos são do mesmo sexo, com as diferenças de renda, masculina e feminina, em práticas mistas. Verificam que discriminação, medida pela diferença não explicada do rendimento, varia entre 21% e 28%. Além disso, os testes não puderam rejeitar a hipótese nula de ausência de discriminação de gênero dentro do local de trabalho.

Considerando o mercado de trabalho formal brasileiro, Sachsida e Loureiro (1998) utilizam dados de séries temporais mensais sobre empregos e salários (masculinos e femininos) tabuladas pelo Ministério do Trabalho e Previdência Social de janeiro de 1986 a dezembro de 1996 para verificar a existência, ou não, de discriminação salarial por gênero. Nos moldes de Hamermesh e Grant (1979) e por Borjas (1983), os autores utilizam o método de elasticidade de substituição de Hicks para verificar se homens e mulheres são substitutos ou complementares no mercado de trabalho. Assim, é possível testar se o sinal do coeficiente da variável salário masculino na equação de demanda por mão de obra feminina é positivo (substitutibilidade) ou negativo (complementaridade), neste último estaria caracterizada discriminação. As estimativas apontaram para a existência de discriminação por gênero no mercado de trabalho formal brasileiro.

Utilizando dados da PNAD/2004 relativo ao conjunto das dez principais regiões metropolitanas do país<sup>25</sup>, Leone e Baltar (2006) se utilizam de estatística descritivas para analisar as mudanças ocorridas na década de 1990, no nível e na dispersão dos rendimentos dos ocupados, mostrando as diferenças segundo sexo, escolaridade e posição na ocupação.<sup>26</sup> Os autores verificam uma acentuada diferença de rendimentos por sexo entre os empregados formais, em todos os tipos de postos de trabalho, apesar de que, essas diferenças diminuíram na década de 1990. Além disso, concluem que as mulheres avançaram em termos de acesso

---

<sup>25</sup>As regiões metropolitanas consideradas são: Belém, Fortaleza, Recife, Salvador, Rio de Janeiro, Belo Horizonte, Curitiba, Porto Alegre, São Paulo e Distrito Federal.

<sup>26</sup> Além disso, realizaram um estudo específico das diferenças de rendimentos entre homens e mulheres com curso superior, conforme posição na ocupação (dirigentes em geral) e setor de atividade (profissionais das ciências e artes, técnicos de nível médio, trabalhadores de serviços administrativos, outras ocupações).

aos cargos mais bem remunerados, porém, nesses cargos, ainda predomina a presença masculina.

Santos e Ribeiro (2006) utilizam a PNAD de 1999 para aferir se existe diferencial de rendimentos entre homens e mulheres, controlando por atributos produtivos para o Brasil em todos os níveis de rendimentos, ao invés de focar o rendimento médio. Para que tal objetivo seja alcançado, os autores empregam a metodologia proposta por Machado e Mata (2005), visto que, esta permite a estimação de toda uma distribuição *contrafactual* com menos hipóteses do que os trabalhos existentes na literatura brasileira. Os resultados mostram que existem diferenciais não explicados pelos atributos produtivos, sendo bastante heterogêneos. As mulheres beneficiam-se destes diferenciais de tratamento nos rendimentos mais baixos e são prejudicadas nos rendimentos mais altos da distribuição. Dessa forma, eles argumentam pela existência do fenômeno de *glass ceiling*<sup>27</sup>, mas também de um possível *glass floors*<sup>28</sup> para o Brasil.

Carvalho, Neri e Silva (2006) decompõem o diferencial de salários por cor e sexo dos trabalhadores brasileiros usando os microdados da PNAD/2003. A metodologia consiste em estimar a equação de salários (Mincer, 1974) com a correção do viés de seleção das informações dos salários (Heckman, 1979). Em seguida, a decomposição do diferencial da média do logaritmo do salário/hora foi obtida pelo procedimento de Oaxaca (1973) apresentada em dois efeitos: características produtivas e discriminação. A análise empírica tem como foco o uso adequado de procedimentos de modelagem estatística em pesquisas, por amostragem complexa, conforme os trabalhos de Skinner e Smith (1989) e Pessoa e Silva (1998). Os resultados da análise controlada, a qual compara indivíduos com atributos semelhantes, exceto a cor e sexo, revelam que mulheres de cor branca e homens e mulheres de cor preta e parda ganham menos do que o grupo base (homens brancos), indicando, portanto, a presença de discriminação de cor e gênero no mercado de trabalho no Brasil.

---

<sup>27</sup> A literatura econômica sobre diferenciais de rendimento, muitas vezes encontra que o diferencial de rendimentos entre homens e mulheres é crescente. Especificamente, este diferencial apresenta uma aceleração para a parte final da função de distribuição. Este fenômeno é conhecido na literatura econômica como *glass ceiling*.

<sup>28</sup> delaRica, Dolado e Llorens (2005) verificaram que o diferencial de rendimentos ao longo da função de distribuição da Espanha se assemelha ao verificado para Grécia e Itália, onde o diferencial de rendimentos em favor dos homens é decrescente ao longo da distribuição de renda, aumentando apenas no final da distribuição. A este efeito os autores denominaram como *glass floors*.

Queiroz e Moreira (2009) visam analisar a inserção e a discriminação no mercado de trabalho brasileiro e os diferenciais de renda por estados. Utilizam informações da PNAD/2005 aplicadas a modelos de determinação da oferta de trabalho e de salários para analisar o diferencial de renda com base na decomposição de Oaxaca (1973). Os resultados mostram que os homens têm maior chance de se empregar e que as mulheres, embora sejam em média mais educadas, sua renda média é menor do que a dos homens. Além disso, verificou-se que esse hiato salarial a favor dos homens está presente em quase todos os estados brasileiros, à exceção do Amazonas, Roraima, Ceará, Sergipe e Distrito Federal. Essas evidências empíricas sugerem uma possível discriminação salarial em gênero, visto que, mesmo controlando pelas características individuais, os homens apresentam rendimentos médios maiores do que os das mulheres.

Com base nesta observação, este capítulo, visa contribuir com o debate presente na literatura acerca dos diferenciais de salários entre homens e mulheres, pois tem como objetivo avaliar o impacto do gênero nos rendimentos dos indivíduos, de forma a verificar se existe diferencial de salário entre gêneros e, por conseguinte, efeito discriminatório no mercado de trabalho brasileiro.

Neste sentido, se reserva a próxima seção para apresentar uma análise descritiva e o tratamento dos dados. Além disso, será descrito o modelo a ser estimado (equação *minceriana*) e os procedimentos metodológicos de estimação empregados para fins de comparação, quais sejam: MQO, Heckman em Dois Estágios, PSM (a partir dos procedimentos: um para um, vizinho mais próximo, radius (distância) e pela densidade de Kernel), bem como pela decomposição de Oaxaca-Blinder.

### 3 METODOLOGIA

Esta seção se reserva a detalhar os fundamentos metodológicos utilizados para aferir o diferencial de salário entre homens e mulheres no mercado de trabalho brasileiro, e a partir desta constatação concluir se este é devido à presença de discriminação entre os gêneros. Dessa forma, primeiramente, são apresentados o modelo teórico e as estratégias de estimação da equação *minceriana* feitas pelos métodos MQO, Heckman em Dois Estágios/Razão Inversa de Mills e *PSM* e, por fim, o procedimento Oaxaca-Blinder de decomposição das diferenças salariais. Em seguida, é relatado o tratamento feito nos dados utilizados, bem como uma análise descritiva a fim de levantar algumas evidências empíricas.

#### 3.1 Modelo Empírico

Para se atestar a ocorrência de diferenças salariais entre homens e mulheres, e algum efeito discriminatório eventualmente decorrente no mercado de trabalho brasileiro, utiliza-se a equação *minceriana* expandida por um conjunto de características dos trabalhadores, como descrito pela seguinte equação:

$$\ln(w_i) = \alpha + \delta \text{Gênero} + X_i' \beta + \epsilon_i \quad (1)$$

Onde  $\ln(w_i)$  é o logaritmo da renda do trabalho do indivíduo  $i$ ;  $X_i$  é um conjunto de características inerentes e adquiridas do indivíduo, tais como: raça, idade, educação e região em que reside; *Gênero* é uma variável binária que assume valor igual a 1 se for homem e 0 caso seja mulher. Cabe realçar que o coeficiente  $\delta$  capta a existência de diferencial de salários entre homens e mulheres;  $\epsilon$  é o termo de erro aleatório.

Na equação (1), o vetor  $X_i$  é composto por variáveis contínuas e dummies. Dessa forma, a referida equação pode, portanto, ser representada da seguinte maneira:

$$\ln(w) = a + \sum_i b_i X_i + \sum_j c_j D_j + \epsilon_i \quad (2)$$

Onde  $X_i$  representa, agora, o vetor de variáveis contínuas e  $D_j$  é o vetor das binárias, dentre as quais encontra-se a variável de interesse – gênero. Retirando o logaritmo,

$$w = \exp[a + \sum_i b_i X_i] \cdot \exp[\sum_j c_j D_j] \quad (3)$$

Fazendo,  $e^{c_j} = 1 + g$ , então,  $g$  é o efeito relativo da variável binária sobre o salário ( $w$ ), e o coeficiente estimado na equação (1) é dado por:

$$c_j = \ln(1 + g) \quad (4)$$

Assim, o efeito relativo sobre  $w$  é:  $g = \exp(c_j - 1)$ , cuja medição em percentual é  $100g$ .

Para se alcançar o objetivo deste trabalho, verificar a existência de diferenças salariais em gênero no mercado de trabalho no Brasil, serão realizadas estimações da equação (1) pelos métodos de MQO, Heckman em dois Estágios e *PSM*, os quais serão descritos mais adiante, e em seguida, será feita a decomposição de Oaxaca, na tentativa de se aferir a diferença salarial é devida a presença de discriminação.

Devido à heterogeneidade presente entre as ocupações, optou-se por realizar as estimações supracitadas levando em consideração 4 estratos ocupacionais, quais sejam: todas as ocupações, dirigentes, profissionais da área de serviços e os profissionais das indústrias de produção e transformação, aqui especificados como técnicos. Note que, essa divisão é bastante representativa, no sentido que, contempla ocupações que exigem qualificações elevadas (dirigentes), médias (serviços) e baixas (técnicos). Portanto, ao se realizar essa estratificação, espera-se verificar se há efeitos diferenciados nos salários, ao se considerar ocupações que exigem diferentes qualificações dos seus empregados.

Definido o modelo a ser estimado, descrevem-se os dados em mais detalhes para, em seguida, explicitar as estratégias de estimação da equação que permitirão inferir sobre o diferencial de salários entre homens e mulheres no mercado de trabalho brasileiro, por ocupações e regiões.

### 3.2 Base de Dados

Para aferir se existe diferença nas rendas do trabalho entre homens e mulheres no Brasil em diversas ocupações, utilizam-se informações da PNAD-2009. No entanto, a amostra contempla homens e mulheres com idade entre 15 e 65 anos. Esse recorte se justifica por dois motivos: i) em virtude de a população em idade ativa (PIA) ser composta por todas as pessoas com 10 anos ou mais de idade; ii) a limitação nos 65 anos, deve-se ao fato que, a partir desta idade, os trabalhadores podem ser aposentados.

Convém mencionar que foram impostas algumas restrições à amostra original através dos seguintes filtros: 1) Exclusão de militares, pois a rigidez de suas atividades, e a determinação de seus rendimentos não são ditadas pelas “forças” de mercado; 2) Exclusão de trabalhadores rurais, haja vista as particularidades inerentes em suas atividades serem incomparáveis às dos trabalhadores urbanos; 3) Exclusão dos profissionais das ciências e das artes; 4) Exclusão dos indivíduos que estudam e trabalham, atendendo hipótese teórica do modelo de Mincer, no qual há um trade off entre educação e experiência, .

Assim, para aferir (e mensurar) se existe o diferencial de salários entre homens e mulheres nas ocupações (dirigentes, serviços e os profissionais das indústrias de transformação e produção) no Brasil optou-se por considerar algumas características demográficas, socioeconômicas e regionais como explicativas deste fenômeno. As variáveis utilizadas para captar estas características, bem como suas descrições e sinais esperados para os coeficientes estimados encontram-se no Quadro 2.1.

Quadro 2.1 - Descrição das Variáveis retiradas da PNAD.

<b>Variável Dependente</b>	<b>Descrição</b>	<b>Valores Assumidos</b>	<b>Sinal Esperado do Coeficiente</b>
<i>W</i>	Renda do Trabalho Principal	Contínuo	
<b>Variáveis Explicativas</b>		<b>Valores Assumidos</b>	
Atributos			
<i>G</i>	Gênero	1, se homem; 0, se mulher	+
<i>I</i>	Idade (em anos) <sup>29</sup>	15 a 65 anos	+
<i>R</i>	Raça/Cor da pele	1, se branco; 0, caso contrário	+
Educação			
<i>Sem Inst</i>	Sem Instrução	1 se não tem Instrução ou Ensino Fundamental completo 0 caso contrário?	Categoria de Referência
<i>EF</i>	Ensino Fundamental	1 se tem Ensino Fundamental completo 0 caso contrário?	+
<i>EM</i>	Ensino Médio	1 se tem Ensino Médio completo 0 caso contrário	+
<i>ES</i>	Ensino Superior	1, se possui ensino superior completo; 0, caso contrário	+
Áreas de Residência			
<i>U</i>	Urbana	1, se reside em área urbana; 0, caso contrário	+

Fonte: Elaboração Própria a Partir dos Dados da PNAD/2009

<sup>29</sup> Esta variável será utilizada como *proxy* para experiência no mercado de trabalho, conforme apresentado por Kusnetz e Thomas (1957 *apud* Connolly, 2004).

Após as estimações, espera-se, em conformidade com os resultados já verificados na literatura, que os coeficientes das variáveis aqui utilizadas apresentem os sinais relatados no Quadro 2.1. No entanto, convém mencionar que, o coeficiente da variável EF, na estimação referente ao grupo ocupacional que contempla os dirigentes, pode não apresentar o referido sinal, pois, espera-se, que as pessoas pertencentes à esta classe apresentem nível de escolaridade mais elevado.

### **3.2.1 *Análise Descritiva dos Dados***

Após apresentar a fonte e a descrição dos dados utilizados para aferir se existe diferença salarial entre homens e mulheres no mercado de trabalho brasileiro, tanto por nível ocupacional quanto regional, se faz pertinente realizar uma análise descritiva dos dados para se ter, inicialmente, algumas evidências sobre diferencial de salários entre gêneros no Brasil, regiões e ocupações.

A Tabela 2.1 contempla as estatísticas descritivas considerando a amostra completa, Brasil, bem como para as macrorregiões Sul-Sudeste e Nordeste<sup>30</sup> e, ainda, por ocupação (dirigentes, os do setor de serviços e das Indústrias de Produção e Transformação).

Observe que entre os profissionais considerando todas as ocupações, o maior salário observado é dos homens residentes nas regiões Sul e Sudeste, R\$ 1462,41, enquanto que as mulheres recebem, em média, R\$ 945,53 de renda do trabalho.

No tocante a idade média dos profissionais, verifica-se, entre todas as regiões, que esta é de 36/37 anos, com exceção dos Dirigentes que estão na faixa de 40 anos. Em relação ao local de residência, aproximadamente 94% dos profissionais são residentes da área Urbana.

Verifica-se que o percentual de mulheres com ensino superior supera o de homens em todas as regiões quando se consideram todas as ocupações. Entretanto, mesmo possuindo um maior nível de escolaridade, as mulheres auferem rendimentos inferiores ao dos homens, o que corrobora com Queiroz e Moreira (2009).

---

<sup>30</sup> Optou-se por realizar estimações para o Nordeste e Sul-Sudeste para tecer um comparativo entre as regiões. Vale ressaltar que no Sul-Sudeste encontram-se as pessoas que recebem, em média, os maiores salários, ao passo que, no Nordeste encontram-se os trabalhadores com remunerações inferiores. Portanto, esta distinção é para apenas verificar se há divergências na diferença salarial entre os gêneros em regiões consideradas ricas e pobres no Brasil, conforme realizado por Mariano e Arraes (2011).

Tabela 2.1 - Estatísticas Descritivas entre Homens e Mulheres, por Regiões e Ocupações, 2009.

Ocupações	Variáveis	Brasil				Nordeste				Sul-Sudeste			
		Homens		Mulheres		Homens		Mulheres		Homens		Mulheres	
		Média	DP	Média	DP	Média	DP	Média	DP	Média	DP	Média	DP
Todas as Ocupações	REN	1298.65	2423.25	866.75	1295.00	955.48	1596.05	659.40	996.60	1462.41	2844.41	945.53	1339.36
	R	0.46	0.50	0.49	0.50	0.28	0.45	0.31	0.46	0.62	0.49	0.64	0.48
	I	36.68	11.20	37.05	10.88	36.05	10.95	36.47	10.81	37.30	11.36	37.60	10.98
	U	0.93	0.25	0.93	0.25	0.92	0.27	0.92	0.28	0.95	0.22	0.95	0.22
	SEM INST	0.37	0.48	0.29	0.45	0.42	0.49	0.32	0.47	0.33	0.47	0.28	0.45
	EF	0.11	0.32	0.09	0.29	0.09	0.29	0.08	0.27	0.13	0.33	0.10	0.30
	EM	0.32	0.47	0.36	0.48	0.32	0.47	0.39	0.49	0.33	0.47	0.35	0.48
	ES	0.11	0.31	0.18	0.38	0.08	0.27	0.14	0.35	0.12	0.33	0.20	0.40
Dirigentes	REN	3390.49	6846.14	2252.92	2917.16	2627.77	2894.39	1770.35	1713.09	3684.69	8573.34	2340.51	3186.59
	R	0.66	0.47	0.68	0.47	0.49	0.50	0.45	0.50	0.79	0.40	0.84	0.37
	I	40.70	9.97	39.39	9.88	40.20	9.90	38.98	9.98	41.16	9.95	39.75	9.88
	U	0.95	0.22	0.97	0.16	0.94	0.24	0.97	0.17	0.96	0.19	0.98	0.14
	SEM INST	0.16	0.37	0.09	0.29	0.21	0.40	0.09	0.29	0.13	0.34	0.09	0.28
	EF	0.08	0.27	0.05	0.22	0.06	0.25	0.04	0.20	0.08	0.28	0.06	0.23
	EM	0.36	0.48	0.36	0.48	0.37	0.48	0.41	0.49	0.35	0.48	0.35	0.48
	ES	0.29	0.45	0.41	0.49	0.25	0.43	0.37	0.48	0.31	0.46	0.42	0.49
Serviços	REN	859.29	924.50	569.11	562.02	643.60	680.33	437.98	472.99	975.55	910.01	631.07	513.25
	R	0.42	0.49	0.43	0.50	0.27	0.44	0.27	0.45	0.60	0.49	0.58	0.49
	I	35.43	11.39	36.51	11.04	35.16	11.06	35.79	10.89	35.86	11.63	37.27	11.21
	U	0.94	0.24	0.93	0.25	0.93	0.25	0.92	0.27	0.95	0.22	0.95	0.22
	SEM INST	0.35	0.48	0.37	0.48	0.41	0.49	0.40	0.49	0.31	0.46	0.35	0.48
	EF	0.12	0.32	0.11	0.31	0.10	0.30	0.09	0.29	0.13	0.34	0.12	0.33
	EM	0.38	0.49	0.38	0.49	0.37	0.48	0.39	0.49	0.39	0.49	0.38	0.48
	ES	0.05	0.22	0.06	0.23	0.03	0.17	0.04	0.20	0.06	0.25	0.06	0.25
Técnicos	REN	1015.18	1087.10	801.60	982.20	754.64	911.33	603.93	835.31	1153.39	1069.06	870.00	910.11
	R	0.43	0.50	0.52	0.50	0.25	0.43	0.31	0.46	0.59	0.49	0.69	0.46
	I	36.63	11.19	37.38	10.99	35.92	10.87	37.13	10.93	37.29	11.38	37.36	11.04
	U	0.92	0.27	0.91	0.29	0.90	0.30	0.86	0.34	0.94	0.23	0.93	0.25
	SEM INST	0.44	0.50	0.24	0.43	0.50	0.50	0.25	0.43	0.40	0.49	0.25	0.43
	EF	0.13	0.33	0.10	0.29	0.10	0.30	0.09	0.28	0.14	0.35	0.11	0.31
	EM	0.30	0.46	0.48	0.50	0.29	0.45	0.52	0.50	0.32	0.47	0.45	0.50
	ES	0.04	0.19	0.10	0.30	0.02	0.15	0.07	0.25	0.05	0.21	0.12	0.32

Fonte: Elaboração Própria a partir dos dados da PNAD/2009.

Em relação aos profissionais que atuam como Dirigentes, observa-se que homens e mulheres são mais bem remunerados do que os profissionais do setor de serviço e técnicos (Profissionais das Indústrias de Produção e Transformação). Sendo que essa diferença, em favor dos Dirigentes, chega a ser de mais de três vezes. Além disso, observa-se que os rendimentos dos profissionais do setor de serviços são superiores aos dos técnicos.

Note também que na classe de Dirigentes estão às pessoas com idade mais avançada, na faixa de 40 anos, e com maior nível de escolaridade, onde, mais uma vez, se observa que as mulheres possuem maior nível de escolaridade, principalmente, quando se compara as pessoas com nível superior, visto que, 31% dos homens, ocupando cargos de dirigentes nas regiões Sul e Sudeste, possuem titulação de nível superior, enquanto que, entre as mulheres, este número é de 46%. Mesmo assim, em termos de salários, observa-se que os homens recebem salários superiores ao das mulheres.

Já os profissionais com menor nível de escolaridade (sem Ensino Fundamental) estão localizados entre os técnicos, seguidos pelos profissionais do setor de serviços. Além disso, nessas duas ocupações a idade média dos profissionais é de 36 anos. Portanto, os profissionais que ocupam cargos de técnicos e no setor de serviços, apresentam, em média, nível de escolaridade e idade inferior aos profissionais que ocupam cargos de dirigentes.

De uma maneira geral, se observa que os rendimentos, estão relacionados à idade, ao nível educacional e a região dos profissionais, visto que, em média, os indivíduos com mais idade, do gênero masculino, possuindo nível superior e residindo nas regiões Sul e Sudeste apresentam os maiores rendimentos.

Além disso, verifica-se um diferencial de salário por gênero, uma vez que as mulheres recebem, em média, rendimento inferior ao dos homens em todas as regiões e ocupações. Portanto, existem outros fatores determinantes da diferença salarial entre os gêneros, motivando, portanto, a análise em busca de se atestar a existência de discriminação entre os gêneros no mercado de trabalho brasileiro.

### **3.4 Avaliação do Impacto**

De uma maneira geral, as análises que envolvem avaliações de impacto se dedicam a computar o efeito da exposição de um conjunto de unidades (indivíduos, famílias,

idades, dentre outros) a um programa ou tratamento (programas educacionais, treinamentos no mercado de trabalho, leis ou regulações, tratamento médico, etc) com o fito de dimensionar o sucesso deste (Imbens e Wooldridge, 2009).

A literatura econométrica que trata a temática de avaliação de impacto tem focado, de forma mais ampla, na estimação do efeito médio do tratamento (EMT), cujo resultado médio é obtido para toda a população. Este efeito pode ser sintetizado na seguinte equação:

$$EMT = E[Y_1 - Y_0|X] = E[Y_1|X] - E[Y_0|X] \quad (5)$$

Onde  $Y_1$  e  $Y_0$  se referem aos resultados sobre as unidades quando expostas e não expostas ao tratamento, respectivamente, e  $X$  é um vetor de covariadas que determina a participação no tratamento e seu resultado.

O outro efeito médio em políticas de avaliação refere-se especificamente ao tratamento sobre os tratados (ETT), onde, neste caso, condiciona-se a média sobre o grupo dos tratados, o qual pode ser expresso pela equação:

$$ETT = E[Y_1 - Y_0|D = 1, X] = E[Y_1|D = 1, X] - E[Y_0|D = 1, X] \quad (6)$$

Onde  $D$  é uma variável binária, cujo sucesso é a ocorrência do tratamento (tratado).

Em muitos cenários de análise é mais apropriado se fazer uso do *ETT*. Se o interesse fosse avaliar a participação em um programa de treinamento de trabalho, não seria adequado considerar o efeito do programa sobre indivíduos com estabilidade, que ocupam cargos com altos salários, visto que, estes não se sentiriam estimulados a entrar em tal programa de treinamento, conforme exemplificam Imbens e Woodridge (2009). Dessa forma, este artigo estimará o *ETT*, visto que a análise conduzida para avaliar o impacto do gênero no salário envolve apenas para a parcela da população que está inserida no mercado de trabalho.

Para a estimação de  $E[Y_1 - Y_0|D = 1, X]$  seria necessário ter-se a informação acerca de  $E[Y_0|D = 1, X]$ , o que é impossível, visto não se poder computar o resultado médio dos não tratados, caso tivessem sido expostos ao tratamento. Para Holland (1986), esse é o problema fundamental da inferência causal em avaliações de impacto, haja vista que, para se avaliar o efeito do tratamento seria necessário comparar as diferentes unidades recebendo os

diferentes níveis de tratamento. Então, a estratégia para estimar *ETT* consiste em utilizar o *contrafactual*, isto é,  $E[Y_0|D = 0, X]$ .

No entanto, Imbens e Woodridge (2009) argumentam que os indivíduos que escolhem participar de um determinado tratamento são, por definição, diferentes daqueles que optam por não participar. Essas diferenças, se elas influenciam a resposta, podem invalidar comparações causais dos resultados por *status* de tratamento, mesmo depois do ajustamento pelas covariadas observadas. Portanto, a utilização de  $E[Y_0|D = 0, X]$  em substituição a  $E[Y_0|D = 1, X]$  pode gerar viés de seleção<sup>31</sup>.

De acordo com Rosenbaum e Rubin (1983), em experimentos aleatórios, os resultados nos dois grupos, tratamento e controle, podem ser diretamente comparados porque suas unidades são provavelmente similares, caso contrário, tal comparação direta poderia ser errônea, haja vista que as unidades expostas ao tratamento geralmente diferem das unidades não expostas. Portanto, em experimentos aleatórios, o viés de seleção, que surge pela utilização de um grupo de comparação inadequado, é eliminado, como decorrência do atendimento das seguintes hipóteses:

- a) ignorabilidade ao tratamento. Os resultados potenciais são independentes ao tratamento condicionado nas variáveis observadas  $[(Y_1, Y_0) \perp D|X]$ ;
- b) suporte comum, ou seja,  $0 < Pr(D = 1|X) = p(X) < 1 \forall X \in \Gamma$ , onde  $\Gamma$  é o suporte<sup>32</sup>.

A satisfação dessas hipóteses garante a comparabilidade entre tratados e controles, visto que, não há valor de  $X$  para o qual se possa inferir com certeza a que grupo, tratado ou controle, o indivíduo pertença, usando apenas informação obtida pelo  $X$ .

Não obstante o suporte dessas hipóteses, Imbens e Woodridge (2009) atestam que existem poucos casos de experimentos aleatórios em economia. Alternativamente,

---

<sup>31</sup> A intuição indica que o grupo de controle deve ser o mais similar possível ao grupo de tratamento em termos dos elementos observáveis antes que o tratamento se desenvolva (assumindo que não há diferenças nos não observáveis). Assim, os métodos de pareamento são técnicas para construir grupos de controle com base em observáveis.

<sup>32</sup> Essa hipótese exige que se tenham unidades em ambos os grupos, tratamento e controle, para cada característica  $X$  que se deseja comparar e, segundo Heckman, Lalonde e Smith (1999), esta condição garante que existam indivíduos no grupo de tratamento que possam ser pareados com os do grupo de controle, tendo o mesmo vetor de características  $X$ . No entanto, a violação dessa hipótese tende a gerar viés devido a não-existência de participantes no grupo de controle que possam ser comparados com os que estão no grupo de tratamento e vice-versa.

Rosenbaum e Rubin (1983) desenvolveram a metodologia do *PSM* para agrupar tratados e controles, tal que, a comparação direta possa ser feita de maneira significativa.

### 3.4.1 *Propensity Score Matching*

A ideia do procedimento por *PSM* é utilizada para agrupar os indivíduos pertencentes aos grupos de tratamento e de controle com base em características observáveis e, ainda, que os do grupo de controle tenham a probabilidade *a priori* de este pertencer ao grupo de tratamento<sup>33</sup>.

Para estimar o *propensity score*, primeiramente estima-se um modelo de escolha binária que tem como variável dependente a participação do tratamento condicionada às características que afetam a participação no tratamento.<sup>34</sup> Ou seja, a variável dependente é binária que assume os valores “1” ou “0”, caso o indivíduo pertença ao grupo de tratamento ou de controle, respectivamente<sup>35</sup>.

Rosenbaum e Rubin (1983) mostram que ao se considerar duas pessoas, uma pertencente ao grupo de tratamento e a outra ao de controle, com o mesmo valor do *propensity score*, então, pode-se dizer que os dois grupos possuem a mesma distribuição em observadas. Neste caso, pode-se assumir que a alocação dos indivíduos nos grupos de tratamento e controle se deu de maneira aleatória. Portanto, a diferença em suas respostas é o efeito do tratamento mais um elemento aleatório e, a média sobre o conjunto destas diferenças estima o efeito do tratamento.

De uma maneira geral, e segundo Becker e Ichino (2002), pode-se dizer que o *PSM* tende a dirimir o problema de viés gerado pela ausência de suporte comum e também o viés proveniente dos observáveis<sup>36</sup> que podem estar presentes na estimação por regressão<sup>37</sup>.

---

<sup>33</sup> Assim, o  $P(x)$  remete a probabilidade do indivíduo pertencente ao grupo de controle ser do grupo de tratado a partir das características observáveis. Ou seja, o pareamento é realizado com base nos valores das covariadas.

<sup>34</sup> O impacto estimado obtido por *PSM* deve sempre depender das variáveis usadas para o pareamento. Assim, Se a escolha das covariadas não inclui importantes determinantes do recebimento do tratamento, a presença destas características não observáveis fará com que *PSM* não seja capaz de reproduzir os resultados de um experimento aleatório. No entanto, se as variáveis determinantes do recebimento do tratamento são bem selecionadas o *PSM* consegue eliminar o viés proveniente dos observáveis.

<sup>35</sup> Nesta pesquisa, especificamente, assume valor igual 1 se o indivíduo for do sexo masculino e 0 se for feminino.

<sup>36</sup> No entanto, caso o viés seja proveniente de diferenças nas características não-observáveis, pode-se dirimir esse viés dependendo da qualidade das variáveis de controle com as quais o *propensity score* é calculado e o *matching* é realizado.

Assim, para estimar o pareamento, optou-se por considerar quatro métodos, a saber: *Um para Um*, *Vizinho mais próximo*, *Distância* e *por Densidade de Kernel*. Vale ressaltar que a opção por considerar esses métodos consiste na busca por robustez dos resultados.

#### 3.4.1.1 *Um para Um*

No pareamento um para um, *one-to-one*, um indivíduo do grupo de controle é pareado com um indivíduo do grupo de tratados, ou seja, a avaliação dos efeitos do tratamento é realizada mediante a análise da diferença das médias das variáveis de interesse de cada par tratado-não tratado. Portanto, quando se utiliza o pareamento um para um procura-se que, os indivíduos pertencentes ao grupo de tratamento sejam comparáveis aos indivíduos do grupo de controle.

De acordo com Freeman, Rossi e Wright (1980) o pareamento um por um parece ser mais preciso e proporciona resultados mais confiáveis que os métodos de pareamento por grupo. Embora, as aplicações desta técnica de pareamento sejam estatisticamente mais desejáveis que o de pareamento por grupo, em geral os estudos sobre avaliações de impacto utilizam o método agregado.

#### 3.4.1.2 *Pareamento por Vizinho mais Próximo*

Este método consiste em procurar para cada indivíduo tratado um não tratado com o escore de propensão mais próximo. Assim sendo, o pareamento é realizado de forma a minimizar a diferença absoluta entre o escore de propensão da unidade tratada e não tratada.

De uma maneira mais formal, considere que  $p_T$  e  $p_{NT}$  denotem o escore de propensão das unidades tratadas e não tratadas, respectivamente. O conjunto de unidades não tratadas pareadas com as unidades tratadas é dado por:

$$C(T) = \min \|p_T - p_{NT}\| \quad (7)$$

A vantagem deste método é que, para cada unidade tratada sempre será encontrado um par, não tratado, evitando, desta forma, a exclusão de observações tratadas.

---

<sup>37</sup> Isto decorre da estimação do PSM, uma vez que para cada indivíduo  $i$  do grupo de tratamento são calculadas as diferenças de *propensity score* com todos os indivíduos do grupo de controle e os indivíduos com as  $k$  menores diferenças fazem o *matching* com o indivíduo  $i$ .

Todavia, pode-se gerar pareamento de indivíduos com escores de propensão muito diferentes, pois o vizinho mais próximo pode não ser tão próximo assim. Para corrigir esse problema pode-se utilizar os métodos de pareamento por Distância (Raio) e de *Kernel*, descritos a seguir.

### 3.4.1.3 Distância (Raio)

No pareamento por Distância, todos os escores de propensão dos indivíduos pertencentes ao grupo de controle que estão próximos, dentro de um raio de tamanho  $r$ , ao escore de propensão de um indivíduo  $i$ , do grupo de tratamento, são compatíveis com este indivíduo. Portanto, o conjunto de unidades não tratadas pareadas com a unidade tratada é dado por:

$$C(i) = \{p_j \mid \|p_j - p_i\| < r\} \quad (8)$$

### 3.4.1.4 Kernel Matching

O efeito do tratamento sobre os tratados é aferido a partir da comparação do rendimento do trabalho (salários) dos homens (grupo de tratados, denotado por  $Y_1$ , indexado por  $I_1$ ) *vis-à-vis* o salário das mulheres (denotados por  $Y_0$ , indexados por  $I_0$ ), por meio da seguinte equação:

$$E[(Y_1 - Y_0) \mid D = 1, P(X)] = \frac{1}{N_1} \sum_{i \in I_1} \left[ Y_{1i} - \sum_{j \in I_0} W_{N_0 N_1}(i, j) Y_{0j} \right] \quad (9)$$

Onde  $W_{N_0 N_1}(i, j)$  é uma matriz de pesos positivos, definida de modo que cada  $i \in I_1$ ,  $\sum_{j \in I_0} W_{N_0 N_1}(i, j) = 1$ , onde  $N_0$  e  $N_1$  são os números de observações em  $I_0$  e  $I_1$ , respectivamente.

O estimador de *kernel* define os pesos em termos das distâncias medidas por  $|P(X_i) - P(X_j)|$ , neste caso, quanto menor for a distância, maior será o peso recebido; ou seja, é uma medida ponderada a partir de uma função *kernel*. Conforme Härdle e Linton (1994) essa função deve ser contínua, simétrica em torno da origem e a integral deve somar um, como descrito pela equação abaixo:

$$K(u) = K(-u) = \int_{-1}^1 K(u) du = 1 \quad (10)$$

Uma forma funcional frequentemente empregada é a *biweight* (ou quártica), expressa da seguinte maneira:

$$K(u) = \begin{cases} \frac{15}{16}(u^2 - 1)^2 & \text{para } |u| < 1 \\ 0 & \text{caso contrário} \end{cases} \quad (11)$$

$$\text{Onde } u = \frac{P(X_i) - P(X_j)}{h}.$$

Para realizar a estimação por meio da função *kernel* deve-se escolher uma banda (*bandwidth*) apropriada ( $h$ ). Quanto menor  $h$ , menor será o peso dado para maiores distâncias, e maior será o peso dado para as observações mais próximas umas das outras. A consistência de estimadores não paramétricos requer que a banda se aproxime de zero a medida que o tamanho da amostra cresça, mas não necessariamente à mesma velocidade. A aproximação da distribuição do *score*, por meio da função *kernel*, é dada por:

$$\hat{f}_h(P(X)) = \frac{1}{n} \sum_{i=1}^n K_h(P(X) - P(X_i)) = \frac{1}{n} \sum_{i=1}^n \frac{1}{h} \left( \frac{P(X) - P(X_i)}{h} \right) \quad (12)$$

### 3.5 Decomposição dos Diferenciais Salariais

Para verificara existência de uma possível discriminação de gênero no mercado de trabalho brasileiro será utilizado o arcabouço teórico desenvolvido por Oaxaca (1973) e Blinder (1973), pois, conforme argumentado anteriormente, este método permite realizar uma decomposição das diferenças salariais em dois efeitos: da discriminação (não explicada) e das diferenças entre as características individuais (explicada).

A metodologia de Oaxaca (1973) e Blinder (1973) é dividida em duas partes. Primeiramente, para analisar a diferença de remuneração entre homens e mulheres, estima-se uma equação de rendimento para cada um destes dois grupos de comparação.

$$\begin{aligned} \ln W_h &= X'_h \beta_h + \varepsilon_h \\ \ln W_m &= X'_m \beta_m + \varepsilon_m \end{aligned} \quad (13)$$

Onde  $W, X$  e  $\varepsilon$  representam a renda, o vetor de variáveis explicativas e o termo de erro respectivamente. Os subscritos  $h$  e  $m$  indicam os dois grupos de comparação, ou seja, homens e mulheres. Convém mencionar que, para corrigir o problema do viés de seletividade amostral, as equações acima serão estimadas pelo método de Heckman em Dois Estágios. Em seguida, decompõe-se o diferencial de salários médios em partes atribuíveis às diferenças nas médias de variáveis explicativas e de diferenças nos coeficientes estimados.

$$\ln \bar{W}_m - \ln \bar{W}_h = (\bar{X}_m - \bar{X}_h)\hat{\beta}_h + \bar{X}_h(\hat{\beta}_m - \hat{\beta}_h) \quad (14)$$

A primeira parte no lado direito da equação (4), devido às diferenças nas variáveis explicativas, é a diferença de salário explicada e a segunda parte, devido às diferenças nos coeficientes, é a diferença não explicada. A proporção da diferença não explicada em relação à diferença global é utilizada como uma medida de discriminação.

Diante do exposto, optou-se por estimar o diferencial de rendimento entre gênero por ocupação e região do Brasil a partir da equação *minceriana* de salários pelo método de Heckman em dois estágios por meio da decomposição de Oaxaca (1973) e Blinder (1973).

Conforme argumentado no primeiro capítulo, a escolha do método de Heckman em dois estágios se faz necessária para corrigir o viés de seleção da informação dos salários. Portanto ao empregar essa técnica objetiva-se melhorar a qualidade das estimativas das equações de salários e avaliar adequadamente as medidas de discriminação, visto que, em Carvalho, Neri e Silva (2006) a estimativa do coeficiente de discriminação salarial entre homens e mulheres de cor branca é 0,37, sem a correção do viés, e 0,30, com a correção do viés de seleção, das informações dos salários.

Além disso, objetivando homogeneizar o máximo possível a amostra de valores de modo a obter resultados mais confiáveis, as estimações e decomposições nos dois grupos de comparações (homens e mulheres) serão realizadas em três níveis regionais (Brasil, Nordeste e Sul-Sudeste) e quatro níveis ocupacionais (todas as ocupações, dirigentes, serviços e profissionais das indústrias de transformação e produção).

#### 4 ANÁLISE E DISCUSSÃO DOS RESULTADOS

Esta seção se reserva a apresentação e análise dos resultados a partir das estimações da equação *minceriana* de salários, equação (1), com o propósito de se verificar o impacto do gênero nas remunerações dos indivíduos no Brasil. Inicialmente, são reportadas as estimações obtidas utilizando o estimador MQO e os resultados aferidos a partir da metodologia de Heckman em dois estágios, bem como os obtidos por *PSM*. Em seguida são apresentados os diferenciais de salários observados ao empregar a decomposição de Oaxaca-Blinder combinado com o de Heckman em dois estágios.

Convêm mencionar que foram feitas algumas divisões na amostra, visando homogeneizá-la de forma a dar uma maior robustez aos resultados. Desse modo, as estimativas supracitadas foram realizadas levando-se em consideração todos os níveis ocupacionais e, em três subamostras ocupacionais: Dirigentes, Serviços e os profissionais das indústrias de transformação e produção, denominados como Técnicos. Além disso, cada uma dessas estimações foi realizada em três níveis regionais: Brasil, Nordeste e Sul e Sudeste.

Além disso, ainda optou-se por apresentar os resultados para cada uma das estratégias adotadas (MQO, Método de Heckman em dois estágios, *PSM* e Oaxaca-Blinder) e, ao final desta seção, é reportada uma síntese dos resultados de modo a comparar o diferencial de salários entre homens e mulheres na subamostras supracitadas pelos respectivos métodos de estimação.

#### 4.1 Estratégia 1: Estimações por MQO

Diante do exposto, foram, inicialmente, realizadas 12 estimações da equação *minceriana* de salários pelo método de MQO, considerando o erro padrão robusto, as quais estão dispostas na tabela 2.2.

Em relação às estimações obtidas, verifica-se que os coeficientes estimados apresentam sinais condizentes com a literatura em todas as estimações realizadas. No entanto, vale destacar que entre as subamostras regionais e ocupacionais, a magnitude dos coeficientes de uma mesma variável explicativa apresenta mudanças consideráveis, indicando, portanto, que existe certa heterogeneidade entre os indivíduos ao se considerar tanto as regiões brasileiras quanto os níveis ocupacionais.

No tocante a significância estatística, verifica-se que, com exceção dos coeficientes da variável Ensino Fundamental na subamostra de Dirigentes, todos os coeficientes das demais variáveis se mostraram estatisticamente significantes em todas as estimações realizadas. Pode-se dizer que, de certa forma, esse resultado era esperado, visto que, entre os dirigentes é de se esperar que as pessoas apresentem nível de escolaridade mais elevado (nível superior).

Em todas as estimações, verifica-se que existe diferença salarial entre homens e mulheres, e que, esta é sempre favorável aos indivíduos do sexo masculino. Além do mais, percebe-se que na região Nordeste a diferença é mais acentuada do que no Brasil e na região Sul/Sudeste, com exceção para os cargos de Dirigentes.

Em relação à diferença salarial, verifica-se, a menor diferença ocorre entre os dirigentes nordestinos, mesmo assim, esta é de 53,7% em favor dos homens. Já a maior diferença salarial foi observada entre os técnicos nordestinos, onde se verificou que os homens ganham, em média, 64,5% a mais que as mulheres.

No tocante às características dos indivíduos, observe que os brancos ganham mais do que os não brancos, no entanto, essa diferença é menor na Região Nordeste. Em relação à idade, pode-se dizer que, esta contribui positivamente para a renda do trabalho, no entanto, a relação se dá a partir de uma forma de U-invertido, isto é, tanto os mais jovens quanto os mais velhos recebem, em média, salários menores.

O fato de residir em área Urbana, também afeta positivamente a renda do trabalho. No caso dos dirigentes nordestinos, por exemplo, a diferença chega a 58,2%. Já em relação aos residentes em áreas urbanas localizadas nas regiões Sul/Sudeste, a diferença é bem menor do que no Brasil e no Nordeste, entretanto, supera os 43% em todas as estimações.

Em relação ao nível educacional, observa-se que quanto maior o nível de escolaridade maior é a diferença salarial. No entanto, entre os Dirigentes, verifica-se que não existe diferença salarial entre os indivíduos sem ensino fundamental e os que possuem ensino fundamental, tanto no Brasil, como nas regiões Nordeste e Sul/Sudeste. Vale destacar que este resultado não chega a ser surpreendente, haja vista que, para se tornar Dirigente, espera-se que, dentre as qualificações dos indivíduos, a educação seja condição necessária. Tanto é que, com nível superior a diferença salarial é superior a 69%.

Diante desses resultados, observa-se que de fato há indícios de um diferencial de salários acentuado entre gêneros no mercado de trabalho brasileiro, visto que, a diferença em favor dos homens varia entre 53,7% para os técnicos nordestinos e 64,5% para os dirigentes situados na Região Nordeste. Apesar dessas diferenças serem significativas, não se pode afirmar que, seja essa a diferença salarial entre os gêneros, visto que, esses valores foram alcançados a partir da estimação de uma equação *a lá* Mincer pelo método de MQO e, de acordo com a literatura, fornece estimativas viesadas. Portanto, utiliza-se do método de Heckman em dois estágios para corrigir o viés.

Tabela 2.2 - Estimções da Equação *Minceriana* por MQO por Regiões e Ocupações

Variáveis Explicativas	Todas as Ocupações			Dirigentes			Serviços			Técnicos		
	Brasil	Nordeste	Sul/Sudeste									
G	<b>0.535</b> (0.000)	<b>0.559</b> (0.000)	<b>0.526</b> (0.000)	<b>0.397</b> (0.000)	<b>0.379</b> (0.000)	<b>0.417</b> (0.000)	<b>0.427</b> (0.000)	<b>0.465</b> (0.000)	<b>0.426</b> (0.000)	<b>0.496</b> (0.000)	<b>0.562</b> (0.000)	<b>0.456</b> (0.000)
R	0.246 (0.000)	0.159 (0.000)	0.193 (0.000)	0.234 (0.000)	0.219 (0.000)	0.241 (0.000)	0.204 (0.000)	0.0988 (0.000)	0.152 (0.000)	0.238 (0.000)	0.139 (0.000)	0.161 (0.000)
I	0.0574 (0.000)	0.0548 (0.000)	0.0599 (0.000)	0.0759 (0.000)	0.0643 (0.000)	0.0907 (0.000)	0.0452 (0.000)	0.0477 (0.000)	0.0462 (0.000)	0.0607 (0.000)	0.0594 (0.000)	0.0626 (0.000)
I <sup>2</sup>	-0.0006 (0.000)	-0.0005 (0.000)	-0.0006 (0.000)	-0.0007 (0.000)	-0.0006 (-0.001)	-0.0009 (0.000)	-0.0005 (0.000)	-0.0005 (0.000)	-0.0005 (0.000)	-0.0006 (0.000)	-0.0006 (0.000)	-0.0007 (0.000)
U	0.234 (0.000)	0.28 (0.000)	0.188 (0.000)	0.322 (0.000)	0.459 (0.000)	0.269 (0.000)	0.214 (0.000)	0.253 (0.000)	0.175 (0.000)	0.223 (0.000)	0.271 (0.000)	0.157 (0.000)
EF	0.177 (0.000)	0.215 (0.000)	0.126 (0.000)	0.00972 (-0.793)	0.0837 (-0.329)	-0.0303 (-0.521)	0.186 (0.000)	0.20 (0.000)	0.141 (0.000)	0.176 (0.000)	0.243 (0.000)	0.116 (0.000)
EM	0.427 (0.000)	0.526 (0.000)	0.372 (0.000)	0.171 (0.000)	0.18 (0.000)	0.168 (0.000)	0.40 (0.000)	0.492 (0.000)	0.342 (0.000)	0.395 (0.000)	0.517 (0.000)	0.342 (0.000)
ES	1.288 (0.000)	1.518 (0.000)	1.126 (0.000)	0.679 (0.000)	0.739 (0.000)	0.639 (0.000)	1.008 (0.000)	1.217 (0.000)	0.811 (0.000)	1.086 (0.000)	1.361 (0.000)	0.929 (0.000)
Constante	4.337 (0.000)	4.051 (0.000)	4.542 (0.000)	4.777 (0.000)	4.751 (0.000)	4.592 (0.000)	4.630 (0.000)	4.266 (0.000)	4.863 (0.000)	4.367 (0.000)	3.990 (0.000)	4.648 (0.000)
N	127.299	35.049	63.237	7.595	1.735	4.107	59.482	17.527	27.924	48.405	12.937	24.976
R <sup>2</sup>	0.372	0.359	0.369	0.236	0.248	0.229	0.243	0.234	0.241	0.248	0.241	0.243
R <sup>2</sup> ajustado	0.372	0.359	0.369	0.236	0.245	0.228	0.242	0.233	0.241	0.248	0.24	0.243
RMSE	0.701	0.749	0.643	0.741	0.736	0.723	0.655	0.711	0.584	0.657	0.719	0.586
F	7554.4	1983.1	3625.7	295.1	74.84	149.4	2027	608.1	943.6	1508.6	367.7	795
Valor-p	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Fonte: Elaborado pelo autor a partir dos da PNAD/2009.

Nota: p-valor entre parênteses.

#### 4.2 Estratégia 2: Método de Heckman em dois estágios, razão inversa de Mills

Os resultados obtidos a partir das estimações da equação *minceriana* pelo método de MQO devem ser observados com cautela, uma vez que podem estar viesados. Assim, optou-se por estimar a referida equação pelo procedimento proposto por Heckman em 1979. O qual considera a estimação em dois estágios, onde no primeiro se obtém a probabilidade de ter renda por meio de um modelo Probit, enquanto que o segundo se reserva a estimar a equação de salários considerando o coeficiente da variável razão inversa de Mills ( $\lambda$ ). Os resultados dessas estimações encontram-se na Tabela 2.3.

Primeiramente, observe que o coeficiente de Mills é estatisticamente significativo na grande maioria dos casos<sup>38</sup>, indicando que, devido à seletividade amostral, a não inclusão desta variável, em uma equação de salários, fornecerá estimativas viesadas.

Assim como nas estimações obtidas pelo método de MQO, verifica-se que, os coeficientes estimados apresentaram o mesmo sinal em todas as estimações realizadas, que estes se mostraram condizentes com a literatura e que as suas magnitudes, para uma mesma variável, chegam a apresentar mudanças consideráveis entre as subamostras consideradas, indicando, portanto, que existe heterogeneidade entre os indivíduos ao se considerar tanto as regiões brasileiras quanto os níveis ocupacionais.

No tocante a significância estatística, verifica-se que com exceção do Ensino Fundamental na subamostra de Dirigentes, todas as demais variáveis se mostraram estatisticamente significantes em todas as estimações realizadas. Pode-se dizer que, de certa forma, esse resultado era esperado, visto que, entre os dirigentes é de se esperar que as pessoas apresentem nível de escolaridade mais elevado.

Como os coeficientes das *dummies* de ensino fundamental (EF), ensino médio (EM) e ensino superior (ES) apresentam sinais positivos, em todas as estimações realizadas, indica que indivíduos com estes níveis de escolaridade recebem, em média, salários maiores que os indivíduos sem ensino fundamental. Além disso, verifica-se que, a magnitude dos coeficientes é diretamente proporcional ao nível de escolaridade, ou seja, o impacto da obtenção do título de nível superior é maior que o impacto da obtenção do título de nível

---

<sup>38</sup> Salvo no caso das estimações considerando Todas as Ocupações, Serviços e Técnicos no Brasil e Dirigentes na região Nordeste.

médio que, por sua vez, é superior ao impacto da obtenção do título de ensino fundamental, sobre o salário dos indivíduos.

Os coeficientes das *dummies* de raça (R) e urbana (U) e da variável referente à idade (I) apresentam, sinais positivos em todas as estimações, indicando que, *ceteris paribus*, indivíduos brancos, residentes em áreas urbanas e de maiores idades, apresentam, em média, os maiores salários. No entanto, o coeficiente da variável de idade ao quadrado ( $I^2$ ) apresentou sinal negativo em todos os modelos estimados, indicando que, apesar de a idade apresentar impacto positivo sobre o salário, este vai decrescendo ao longo do tempo.

Em relação ao principal objetivo deste artigo, avaliar uma possível diferença salarial entre os indivíduos quanto ao gênero, pode-se inferir a partir dos resultados expostos na Tabela 2.3 que, realmente existe uma discrepância salarial entre homens e mulheres no Brasil, visto que, em todas as estimações, o coeficiente da *dummy* de gênero apresentou sinal positivo, indicando, portanto, que em média, os homens recebem maiores salários.

Note que, na mesma região e no mesmo nível ocupacional, em alguns casos, a divergência salarial chega a ser bastante significativa como, por exemplo, no caso dos técnicos no Nordeste, onde em média, os homens recebem um salário cerca de 61% maior que o das mulheres. A menor diferença foi verificada entre os dirigentes na região Nordeste, no entanto, esta supera os 56%, ou seja, entre os dirigentes nordestinos, os homens ganham, em média, 56,1% a mais que as mulheres.

Portanto, de acordo com as estimações da equação *minceriana* de salários pelo método de Heckman em dois estágios, pode-se inferir que existe uma considerável discrepância salarial entre homens e mulheres no Brasil, onde esta diferença varia de 56,1% a 65,1%, dependendo da região e do nível ocupacional considerado. Percebe-se que a menor diferença salarial ainda é verificada entre os dirigentes da região Nordeste, no entanto, esta é bem superior àquela verificada pelo método de MQO (53,7%).

Portanto, de uma maneira geral, verifica-se que, no mínimo, os homens recebem, em média, 53% a mais de salários que as mulheres. No entanto, esses resultados não podem ser utilizados para caracterizar a presença de discriminação no mercado de trabalho, visto que, a comparação não está sendo feita entre pares semelhantes. Dessa forma, a próxima seção

apresentará os resultados por *PSM*, onde, neste caso, a comparação é feita entre homens e mulheres semelhantes, onde a única diferença observável entre eles será o gênero.

Tabela 2.3 - Estimções da Equação *Minceriana* de Salários pelo Método de Heckman em Dois Estágios, por Regiões e Ocupações

Variáveis	Todas as Ocupações			Dirigentes			Serviços			Técnicos		
	Brasil	Nordeste	Sul e Sudeste									
G	<b>0.556</b> (0.000)	<b>0.547</b> (0.000)	<b>0.572</b> (0.000)	<b>0.459</b> (0.000)	<b>0.423</b> (0.000)	<b>0.473</b> (0.000)	<b>0.428</b> (0.000)	<b>0.442</b> (0.000)	<b>0.459</b> (0.000)	<b>0.490</b> (0.000)	<b>0.512</b> (0.000)	<b>0.476</b> (0.000)
R	0.245 (0.000)	0.159 (0.000)	0.190 (0.000)	0.230 (0.000)	0.218 (0.000)	0.237 (0.000)	0.204 (0.000)	0.100 (0.000)	0.150 (0.000)	0.238 (0.000)	0.141 (0.000)	0.161 (0.000)
I	0.0666 (0.000)	0.0495 (0.000)	0.0804 (0.000)	0.102 (0.000)	0.0820 (0.000)	0.115 (0.000)	0.0456 (0.000)	0.0362 (0.000)	0.0620 (0.000)	0.0585 (0.000)	0.0401 (0.000)	0.0701 (0.000)
I <sup>2</sup>	-0.0007 (0.000)	-0.0005 (0.000)	-0.0009 (0.000)	-0.001 (0.000)	-0.0008 (0.002)	-0.0012 (0.000)	-0.0005 (0.000)	-0.0004 (0.000)	-0.0007 (0.000)	-0.0006 (0.000)	-0.0004 (0.000)	-0.0008 (0.000)
U	0.247 (0.000)	0.273 (0.000)	0.218 (0.000)	0.359 (0.000)	0.484 (0.000)	0.306 (0.000)	0.215 (0.000)	0.236 (0.000)	0.197 (0.000)	0.220 (0.000)	0.245 (0.000)	0.169 (0.000)
EF	0.181 (0.000)	0.213 (0.000)	0.134 (0.000)	0.0216 (0.560)	0.0915 (0.288)	-0.0188 (0.690)	0.186 (0.000)	0.197 (0.000)	0.147 (0.000)	0.175 (0.000)	0.238 (0.000)	0.119 (0.000)
EM	0.441 (0.000)	0.518 (0.000)	0.404 (0.000)	0.214 (0.000)	0.207 (0.000)	0.208 (0.000)	0.400 (0.000)	0.474 (0.000)	0.367 (0.000)	0.392 (0.000)	0.489 (0.000)	0.353 (0.000)
ES	1.307 (0.000)	1.507 (0.000)	1.169 (0.000)	0.743 (0.000)	0.782 (0.000)	0.698 (0.000)	1.009 (0.000)	1.194 (0.000)	0.841 (0.000)	1.082 (0.000)	1.322 (0.000)	0.944 (0.000)
Mills	2.389 (0.000)	-1.327 (0.125)	5.433 (0.000)	11.15 (0.003)	7.786 (0.289)	10.05 (0.034)	0.104 (0.856)	-2.423 (0.019)	3.432 (0.000)	-0.676 (0.467)	-5.624 (0.004)	2.289 (0.027)
Constante	3.423 (0.000)	4.563 (0.000)	2.474 (0.000)	0.907 (0.481)	2.055 (0.423)	1.086 (0.509)	4.588 (0.000)	5.239 (0.000)	3.497 (0.000)	4.617 (0.000)	6.086 (0.000)	3.800 (0.000)
N	127.299	35.049	63.237	7.595	1.735	4.107	59.482	17.527	27.924	48.405	12.937	24.976
R <sup>2</sup>	0.372	0.359	0.370	0.237	0.249	0.230	0.243	0.234	0.242	0.248	0.242	0.243
F	6733.9 0.000	1764.4 0.000	3239.5 0.000	266.3 0.000	66.81 0.0000	136.4 0.0000	1805.7 0.000	541.3 0.000	846.9 0.000	1357.3 0.000	331.8 0.000	710.7 0.000

Fonte: Elaboração Própria a partir dos dados da PNAD/2009.

Nota: Valor-p entre parênteses.

### 4.3 Estratégia 3: Propensity Score Matching

Esta seção contempla os resultados das estimações realizadas pelo método de *propensity score matching* a partir de quatro formas de pareamento (*Um para Um*, *Vizinho mais próximo*, *Distância* e *por Densidade de Kernel*) na tentativa de se mitigar algum viés que possa ocorrer pela comparação entre homens e mulheres de forma inadequada, uma vez que estes são diferentes por natureza. Assim, se faz apropriado que as características observáveis como, idade, raça e educação, por exemplo, sejam parecidas a fim de homogeneizar as características e aferir a real diferença salarial entre homens e mulheres.

Dessa forma, pode-se dizer que a partir do pareamento, a comparação entre homens e mulheres (aqui, denominados como tratados e controles, respectivamente) é mais homogênea, dado que as amostras são pareadas com base em características observadas. E, por isso, pode-se inferir que o diferencial de renda (caso exista) se deve, exclusivamente, ao gênero.

A Tabela 2.4 contempla os rendimentos médios de homens e mulheres com e sem pareamento. Observe que em todas as divisões amostrais, regionais e ocupacionais, há uma mudança na renda média dos não tratados quando se compara os pareados (ETT) com os não pareados (*unmatching*). Dessa forma, pode-se concluir que, no grupo de controle, existiam algumas mulheres com características diferentes dos homens e, portanto, o diferencial de renda média entre tratados e controles, antes do pareamento, não pode ser atribuído exclusivamente ao gênero.

Além disso, após o pareamento, verificou-se que na grande maioria dos casos, exceto para os profissionais da área de serviços no Brasil (estimação um para um) e no Sul/Sudeste (estimações por Densidade de Kernel, vizinho mais próximo e distância), a renda média dos não tratados diminuiu após o pareamento, aumentando ainda mais a disparidade de renda entre homens e mulheres. Portanto, diferenças entre rendimentos de homens e mulheres sem pareamento podem ser subestimadas, visto que, a renda média das mulheres, sem o pareamento, apresenta viés positivo.

Assim como nas estimativas reportadas nas Tabelas 2.2 e 2.3, verifica-se, pela Tabela 2.4, que os diferenciais de rendimento variam entre as subamostras, isto é, entre as

regiões e as ocupações, corroborando, portanto, que existe heterogeneidade entre os indivíduos ao se considerar tanto as regiões brasileiras quanto as ocupações.

Considerando todas as ocupações, verifica-se que as menores remunerações são dos residentes na região Nordeste, tanto para homens quanto para as mulheres. No entanto, nessa região verificaram-se as maiores diferenças salariais entre os homens e mulheres.

Em relação ao salário dos dirigentes, primeiramente, observa-se que eles recebem maiores salários do que os trabalhadores de serviços e técnicos. Além disso, a diferença salarial é menor do que ao se considerar todas as ocupações. Mesmo assim, a diferença entre homens e mulheres é de no mínimo 42,78%. Note ainda que, as maiores diferenças são entre os Dirigentes que residem na região Sul/Sudeste e que a mulheres nordestinas nos cargos de dirigentes são as que recebem os menores salários.

Já entre os profissionais do setor de serviços e das Indústrias de Produção e Transformação (Técnicos) os valores são inferiores aos já observados, porém, também se verifica que existe diferença entre as remunerações de homens e mulheres no mercado de trabalho brasileiro, tanto a nível nacional como regional. Nestes dois setores verifica-se que a maior diferença se dá entre os nordestinos.

Portanto, diante do exposto, e dos resultados apresentados na Tabela 2.4, pode-se dizer que existe diferencial de salários entre os gêneros no mercado de trabalho brasileiro, e que este varia entre 42,78% e 118,7% em favor dos homens. Este diferencial se torna menos acentuado nos cargos de maiores remunerações, visto que, na classe dos dirigentes observou-se a menor discrepância entre os salários de homens e mulheres. Além disso, a região Nordeste remunera menos, em média, do que no Brasil e no Sul/Sudeste entre as ocupações consideradas.

Tabela 2.4 - Diferenciais de Renda entre Tratados e Controles com e sem Pareamento.

Ocupação/ Região	Método	Amostra	Tratados	Controles	Diferença	Erro Padrão	Estatística t
Todas as Ocupações / Brasil		<i>Unmatched</i>	6.7686	6.3591	<b>0.4095</b>	0.0049	84.31
	Um p/ um	ATT	6.7686	6.3165	<b>0.4521</b>	0.0512	8.83
	Vizinho	ATT	6.7686	6.2311	<b>0.5375</b>	0.0051	105.33
	Distância						
	Kernel						
Todas as Ocupações / Nordeste		<i>Unmatched</i>	6.4444	6.0384	<b>0.4059</b>	0.0098	41.37
	Um p/ um	ATT	6.4444	5.8431	<b>0.6014</b>	0.0673	8.93
	Vizinho	ATT	6.4444	5.8752	<b>0.5693</b>	0.0107	53.44

	Distância	ATT	6.4443	5.8771	<b>0.5672</b>	0.0105	54.2
	Kernel	ATT	6.4444	5.8901	<b>0.5543</b>	0.0103	53.69
		<i>Unmatched</i>					
Todas as Ocupações / Sul/Sudeste	Um p/ um	ATT	6.9364	6.3540	<b>0.5824</b>	0.0518	11.24
	Vizinho	ATT	6.9364	6.4172	<b>0.5192</b>	0.0066	79.3
	Distância	ATT	6.9364	6.4263	<b>0.51</b>	0.0064	79.44
	Kernel	ATT	6.9364	6.4109	<b>0.5255</b>	0.0064	82.22
		<i>Unmatched</i>		7.7129	7.3915	<b>0.3214</b>	0.0199
Dirigentes/ Brasil	Um p/ um	ATT	7.7133	7.1961	<b>0.5172</b>	0.0545	9.5
	Vizinho	ATT	7.7133	7.3212	<b>0.3921</b>	0.0235	16.71
	Distância	ATT	7.7229	7.3356	<b>0.3874</b>	0.0213	18.18
	Kernel	ATT	7.7133	7.3359	<b>0.3775</b>	0.0204	18.48
		<i>Unmatched</i>		7.4955	7.1953	<b>0.3002</b>	0.0417
Dirigentes/ Nordeste	Um p/ um	ATT	7.4992	7.1107	<b>0.3885</b>	0.0671	5.79
	Vizinho	ATT	7.4992	7.1423	<b>0.3568</b>	0.0532	6.7
	Distância	ATT	7.5549	7.1693	<b>0.3857</b>	0.0503	7.67
	Kernel	ATT	7.4992	7.1431	<b>0.3561</b>	0.0437	8.15
		<i>Unmatched</i>		7.7974	7.4438	<b>0.3536</b>	0.0262
Dirigentes/ Sul/Sudeste	Um p/ um	ATT	7.7974	7.4226	<b>0.3748</b>	0.0699	5.36
	Vizinho	ATT	7.7974	7.3935	<b>0.4039</b>	0.0305	13.26
	Distância	ATT	7.8094	7.3824	<b>0.427</b>	0.0279	15.26
	Kernel	ATT	7.7974	7.4012	<b>0.3961</b>	0.0266	14.91
		<i>Unmatched</i>		6.5224	6.1166	<b>0.4057</b>	0.0061
Serviços / Brasil	Um p/ um	ATT	6.5224	6.1585	<b>0.3639</b>	0.0508	7.16
	Vizinho	ATT	6.5224	6.1025	<b>0.4198</b>	0.0061	68.63
	Distância	ATT	6.5224	6.1113	<b>0.4111</b>	0.006	68.29
	Kernel	ATT	6.5224	6.1087	<b>0.4137</b>	0.006	68.85
		<i>Unmatched</i>		6.2475	5.8124	<b>0.4352</b>	0.012
Serviços / Nordeste	Um p/ um	ATT	6.2474	5.7486	<b>0.4988</b>	0.0686	7.27
	Vizinho	ATT	6.2474	5.7843	<b>0.4631</b>	0.0121	38.13
	Distância	ATT	6.2473	5.7732	<b>0.4741</b>	0.0119	39.98
	Kernel	ATT	6.2474	5.7756	<b>0.4718</b>	0.0117	40.3
		<i>Unmatched</i>		6.6904	6.2715	<b>0.4189</b>	0.0079
Serviços / Sul/Sudeste	Um p/ um	ATT	6.6905	6.2172	<b>0.4733</b>	0.0478	9.91
	Vizinho	ATT	6.6905	6.2779	<b>0.4125</b>	0.0081	50.93
	Distância	ATT	6.6905	6.2799	<b>0.4106</b>	0.0079	52.2
	Kernel	ATT	6.6905	6.2904	<b>0.4001</b>	0.0078	51.02
		<i>Unmatched</i>		6.6653	6.3357	<b>0.3296</b>	0.0083
Técnicos / Brasil	Um p/ um	ATT	6.6651	6.1728	<b>0.4924</b>	0.0661	7.44
	Vizinho	ATT	6.6652	6.1282	<b>0.537</b>	0.0116	46.27
	Distância	ATT	6.6649	6.1351	<b>0.5299</b>	0.0113	46.8
	Kernel	ATT	6.6653	6.1399	<b>0.5255</b>	0.0108	48.87
		<i>Unmatched</i>		6.3366	5.9775	<b>0.3591</b>	0.0169
Técnicos / Nordeste	Um p/ um	ATT	6.3364	5.5538	<b>0.7826</b>	0.0858	9.13
	Vizinho	ATT	6.3364	5.7178	<b>0.6187</b>	0.0273	22.66
	Distância	ATT	6.3360	5.7085	<b>0.6275</b>	0.0262	23.97
	Kernel	ATT	6.3364	5.7029	<b>0.6335</b>	0.0242	26.23
		<i>Unmatched</i>		6.8333	6.5004	<b>0.3329</b>	0.0101
Técnicos / Sul/Sudeste	Um p/ um	ATT	6.8336	6.3173	<b>0.5163</b>	0.0522	9.89
	Vizinho	ATT	6.8336	6.3615	<b>0.4721</b>	0.0133	35.54
	Distância	ATT	6.8332	6.3597	<b>0.4736</b>	0.0126	37.73
	Kernel	ATT	6.8336	6.3711	<b>0.4625</b>	0.0119	38.55

Fonte: Elaboração própria a partir dos dados da PNAD/2009.

#### 4.4 Decomposição de Oaxaca-Blinder

A partir das estimações da equação *minceriana* pelos métodos de MQO, Heckman em dois estágios e por *PSM*, verifica-se que existe diferencial de salários entre homens e mulheres e, que isto, pode ser em decorrência de diversas razões, inclusive por discriminação por gênero. Neste sentido, a decomposição de Oaxaca permite que sejam comparados os rendimentos de homens e mulheres e, ainda, que o diferencial seja exclusivamente em decorrência do gênero e, isto, seria um indicativo de discriminação.

Neste sentido, optou-se por aplicar a decomposição de Oaxaca-Blinder combinado com o método de Heckman em dois estágios, uma vez que este método tende a dirimir o viés causado pela seletividade amostral. Além do mais, a não inclusão desta variável em uma equação de salários pode subestimar as estimativas.

A Tabela 2.5 reporta os resultados da decomposição por ocupações (Todas as Ocupações, Dirigentes, Serviços e Técnicos) e regiões censitárias (Brasil, Nordeste e Sul/Sudeste). Note que, independentemente da ocupação e da região, existe diferença salarial entre homens e mulheres de no mínimo 36,32%.

Essa diferença é observada para os Dirigentes no Brasil. Portanto, nas demais ocupações e regiões, a diferença salarial entre homens e mulheres, devida exclusivamente ao gênero, é superior a 36%.

No setor de serviços a diferença salarial entre homens e mulheres é superior à verificada entre os dirigentes. Em termos regionais, mais uma vez, se observou que a maior diferença salarial persiste na região Nordeste.

Verifica-se ainda que a média do salário dos homens (em logaritmo) é superior à média das mulheres em todas as ocupações e regiões.<sup>39</sup> Portanto, existe de fato um diferencial de salários a favor dos homens no Brasil, que varia entre 31% e 48%.<sup>40</sup>

A partir de agora, a análise dar-se-á separadamente pelos componentes da decomposição denominados como endógeno e coeficientes (discriminação). O primeiro

---

<sup>39</sup> Vale ressaltar que, esses resultados corroboram com as estimações por MQO, Heckman em dois estágios e por *PSM*.

<sup>40</sup> O sinal dessas diferenças é negativo, pois se utilizou as mulheres como grupo base.

componente é calculado a partir de  $(\bar{X}_m - \bar{X}_h)\hat{\beta}_h$ , e refere-se às dotações dos indivíduos. Ou seja, às características produtivas, o que indica que, a diferença salarial pode ser atribuída às diferenças nas características. Portanto, se homens e mulheres fossem remunerados da mesma forma, como os homens ( $\hat{\beta}_h$ ), então as mulheres receberiam salários superior, em média, que o dos homens. Esse diferencial varia de 0,0146 a 0,156. Nesse caso, pode-se inferir que as mulheres possuem características superiores às dos homens [se  $\hat{\beta}_h$  é positivo e  $[(\bar{X}_m - \bar{X}_h)] > 0$  também].

O segundo termo, calculado a partir de  $\bar{X}_h(\hat{\beta}_m - \hat{\beta}_h)$ , reporta o diferencial no caso dos coeficientes da regressão para as mulheres são aplicados sobre as características dos homens, ou seja, se as mulheres tivessem as mesmas características dos homens. Neste caso, de quanto seria o diferencial?

Em todas as estimações, verifica-se que o sinal desse termo é negativo, o que confirma que se, homens e mulheres tivessem as mesmas características, as mulheres estariam recebendo salários ainda menores, pois  $(\hat{\beta}_m - \hat{\beta}_h) < 0$  dado  $\bar{X}_h > 0$ . Diante disso, esse diferencial de coeficiente remete ao efeito da discriminação, haja vista que o *contrafactual* (isto é, mulheres com as mesmas características dos homens) é remunerado de maneira diferente por conta de uma característica observável, o gênero. No presente caso, as mulheres, apresentando as mesmas características que os homens, recebem, em média, menores salários.

Tabela 2.5 - Decomposição de Oaxaca-Blinder pelo método de Heckman em dois estágios

Ocupações	Região	Variáveis	Coefficiente	Erro Padrão	Estatística – Z	P-Valor	
Todas as ocupações	Brasil	Mulher	6.3591	0.0038	1664.35	0.0000	
		Homem	6.7667	0.0048	1415.23	0.0000	
		Diferença	<b>-0.4076</b>	0.0061	-66.60	0.0000	
		Endógeno	0.1231	0.0030	40.89	0.0000	
		Coefficientes	-0.5348	0.0055	-97.58	0.0000	
		Interação	0.0042	0.0014	3.04	0.0020	
	Nordeste	Mulher	6.0384	0.0080	756.19	0.0000	
		Homem	6.4716	0.0093	696.88	0.0000	
		Diferença	<b>-0.4332</b>	0.0122	-35.37	0.0000	
		Endógeno	0.1426	0.0058	24.56	0.0000	
		Coefficientes	-0.5938	0.0110	-53.75	0.0000	
		Interação	0.0181	0.0031	5.76	0.0000	
	Sul Sudeste	Mulher	6.5091	0.0048	1351.63	0.0000	
		Homem	6.9309	0.0067	1028.75	0.0000	
		Diferença	<b>-0.4219</b>	0.0083	-50.94	0.0000	
		Endógeno	0.0989	0.0038	25.74	0.0000	
		Coefficientes	-0.5199	0.0075	-68.99	0.0000	
		Interação	-0.0008	0.0017	-0.46	0.6430	
Dirigentes	Brasil	Mulher	7.3915	0.0156	474.61	0.0000	
		Homem	7.7013	0.0261	295.14	0.0000	
		Diferença	<b>-0.3098</b>	0.0304	-10.20	0.0000	
		Endógeno	0.0824	0.0125	6.58	0.0000	
		Coefficientes	-0.3742	0.0296	-12.65	0.0000	
		Interação	-0.0180	0.0095	-1.90	0.0570	
	Sul Sudeste	Mulher	7.4438	0.0204	364.50	0.0000	
		Homem	7.8043	0.0260	300.50	0.0000	
		Diferença	<b>-0.3604</b>	0.0330	-10.91	0.0000	
		Endógeno	0.0629	0.0147	4.29	0.0000	
		Coefficientes	-0.4183	0.0318	-13.14	0.0000	
		Interação	-0.0050	0.0106	-0.47	0.6400	
	Serviços	Brasil	Mulher	6.1166	0.0039	1554.82	0.0000
			Homem	6.5407	0.0075	866.95	0.0000
			Diferença	<b>-0.4241</b>	0.0085	-49.84	0.0000
			Endógeno	0.0221	0.0027	8.17	0.0000
			Coefficientes	-0.4425	0.0081	-54.44	0.0000
			Interação	-0.0036	0.0012	-3.02	0.0030
Nordeste		Mulher	5.8124	0.0083	698.67	0.0000	
		Homem	6.2926	0.0159	395.45	0.0000	
		Diferença	<b>-0.4802</b>	0.0180	-26.74	0.0000	
		Endógeno	0.0240	0.0043	5.57	0.0000	
		Coefficientes	-0.5121	0.0174	-29.41	0.0000	
		Interação	0.0080	0.0025	3.22	0.0010	
Sul Sudeste		Mulher	6.2715	0.0049	1277.07	0.0000	
		Homem	6.6991	0.0119	561.07	0.0000	
		Diferença	<b>-0.4276</b>	0.0129	-33.12	0.0000	
		Endógeno	0.0146	0.0037	3.91	0.0000	
		Coefficientes	-0.4298	0.0125	-34.31	0.0000	
		Interação	-0.0123	0.0021	-5.86	0.0000	
Técnicos	Brasil	Mulher	6.3357	0.0088	720.13	0.0000	
		Homem	6.6764	0.0061	1090.60	0.0000	
		Diferença	<b>-0.3407</b>	0.0107	-31.79	0.0000	
		Endógeno	0.1560	0.0045	34.57	0.0000	
		Coefficientes	-0.5242	0.0106	-49.28	0.0000	
		Interação	0.0274	0.0052	5.33	0.0000	

Fonte: Elaboração Própria a partir dos dados da PNAD.

#### 4.5 Sínteses dos Resultados

Esta seção se dedica a apresentar uma síntese dos resultados alcançados neste ensaio, isto é, os diferenciais de salários entre homens e mulheres no Brasil, Nordeste e Sul/Sudeste, considerando Todas as Ocupações, Dirigentes, Serviços e Técnicos.

Deste modo, optou-se por apresentar na Tabela 2.6 as estimativas para os diferenciais de salários aferidos pelos métodos: MQO, Heckman em dois estágios, *PSM* (Um para Um, Vizinho mais próximo, Distância e Kernel), bem como pela decomposição de Oaxaca-Blinder.

Primeiramente, verifica-se que, ao se estimar uma equação *minceriana* de salários pelo método de MQO, a diferença salarial entre os gêneros varia entre 53,7% e 64,5%, já pelo método de Heckman, esta diferença varia entre 56,1% e 65,1%. Portanto, deve-se utilizar o método de Heckman em dois estágios, para corrigir o viés de seleção amostral, pois a não utilização deste método poderia estar subestimando a diferença salarial entre os gêneros.<sup>41</sup>

No entanto, mesmo corrigindo o problema do viés de seleção amostral, essas estimativas foram obtidas a partir de uma comparação entre homens e mulheres com características diferentes. Dessa forma, esses resultados não podem ser utilizados para caracterizar a presença de discriminação no mercado de trabalho, visto que, a comparação não está sendo feita entre pares semelhantes. Diante disso, optou-se por utilizar quatro metodologias de *PSM*, para obter um diferencial de salário entre homens e mulheres com característica observáveis semelhantes à exceção do gênero.

A partir das técnicas de pareamento (Um para Um, Vizinho mais próximo, Distância e Kernel) observa-se que o diferencial de salários entre homens e mulheres varia entre 42,7% e 118,7%, sempre em favor dos homens. Note que os métodos de Vizinho mais próximo, Distância e Kernel apresentam resultados com uma diferença menor em relação a técnica de *Um para Um*. Todavia, após parear as amostras, verifica-se que, em média, os homens continuam recebendo salários maiores que os das mulheres, entretanto, a diferença é menor.

---

<sup>41</sup> Além disso, verifica-se que apenas em 4 das 12 estimações (Todas as Ocupações no Nordeste, Serviços no Nordeste e Técnicos no Brasil e no Nordeste), o diferencial de salário obtido pelo método de MQO foi superior ao obtido pelo método de Heckman.

A partir das estimativas alcançadas pela decomposição de Oaxaca-Blinder, também se observa diferença salarial entre gêneros no Brasil, entretanto, a diferença varia de 36,3% e 61,6%.

Em relação aos extratos ocupacionais, observa-se que, em média, as menores diferenças salariais entre homens e mulheres são verificadas entre os que ocupam cargo de dirigente. Isto, provavelmente, deriva da necessidade de se ter maior qualificação para ocupar o referido cargo, o que tende a reduzir o diferencial de salário. Além disso, entre os dirigentes e os profissionais da área de serviços, observam-se que as diferenças são próximas.

No outro extremo, estão às diferenças salariais entre os Técnicos, principalmente entre os residentes na região Nordeste do Brasil. Vale ressaltar ainda que, as maiores diferenças salariais entre homens e mulheres são observadas nesta região.

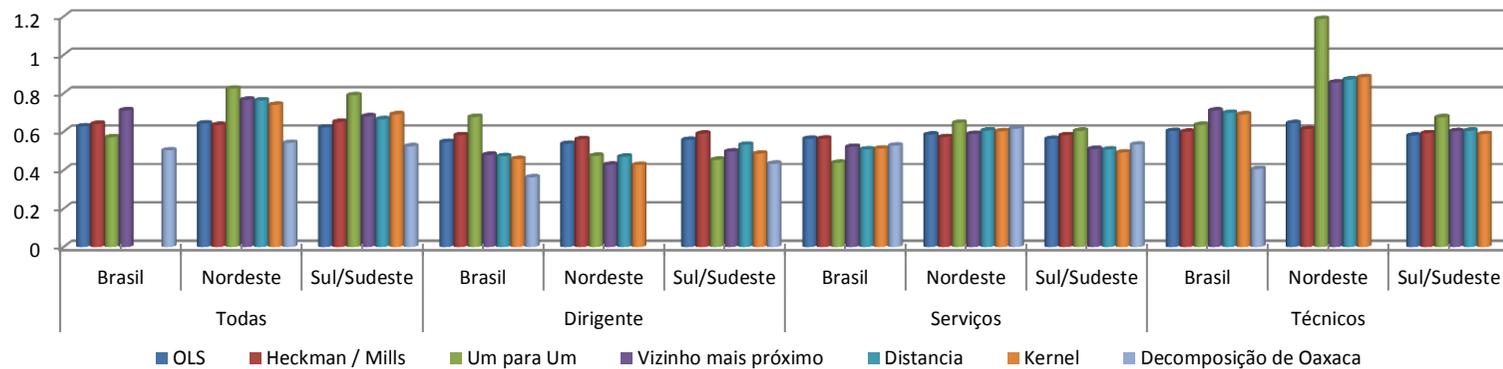
Com esteio nestas observações, percebe-se que, independentemente do método de estimação utilizado, de fato existe diferença salarial entre gêneros no Brasil, sendo que os homens sempre dispõem de uma situação mais favorável. E, ainda, que ao controlar por ocupação e região, percebe-se que, a diferença salarial entre homens e mulheres é de no mínimo 36,31%, em média.

Tabela 2.6 - Síntese dos Resultados das Diferenças Salariais Por Gênero.

Ocupações	Região	Método de Estimação						Decomposição de Oaxaca
		OLS	Heckman / Mills	PSM				
				Um para Um	Vizinho mais próximo	Distancia	Kernel	
Todas	Brasil	0.6281	0.6415	0.5716	0.7117			0.5032
	Nordeste	0.6434	0.6357	0.8245	0.7669	0.7633	0.7407	0.5422
	Sul/Sudeste	0.6225	0.6518	0.7903	0.6807	0.6655	0.6913	0.5247
Dirigente	Brasil	0.5472	0.5822	0.6773	0.4801	0.4730	0.4585	0.3632
	Nordeste	0.5374	0.5616	0.4748	0.4289	0.4705	0.4278	
	Sul/Sudeste	0.5582	0.5904	0.4547	0.4977	0.5327	0.4862	0.4340
Serviços	Brasil	0.5638	0.5644	0.4389	0.5218	0.5085	0.5124	0.5282
	Nordeste	0.5857	0.5724	0.6467	0.5890	0.6066	0.6029	0.6164
	Sul/Sudeste	0.5633	0.5822	0.6053	0.5107	0.5077	0.4920	0.5336
Técnicos	Brasil	0.6041	0.6005	0.6361	0.7109	0.6986	0.6911	0.4059
	Nordeste	0.6453	0.6139	1.1872	0.8563	0.8729	0.8842	
	Sul/Sudeste	0.5804	0.5921	0.6758	0.6034	0.6056	0.5880	

Fonte: Elaboração Própria a partir dos dados da PNAD/2009.

Gráfico 1.1 - Síntese dos resultados para diferença salarial entre homens e mulheres, por regiões e ocupações.



Fonte: Elaborado pelo autor.

## 5 CONSIDERAÇÕES FINAIS

Com o objetivo de aferir se existe diferencial de salários entre homens e mulheres no mercado de trabalho brasileiro, optou-se por utilizar as informações da PNAD de 2009, para estimar equações de rendimentos como proposto por Mincer (1974) por meio dos métodos de MQO, Heckman em dois estágios, *Propensity Score Matching* por decomposição de Oaxaca-Blinder.

Além disso, optou-se ainda por realizar as estimativas em três níveis regionais (Brasil, Nordeste e Sul e Sudeste) e em quatro níveis ocupacionais (todas as ocupações, dirigentes, serviços e profissionais das indústrias de produção e transformação) visando gerar maior robustez das estimativas através da homogeneização da amostra.

A partir das estimações percebe-se que os homens ganham mais do que as mulheres em todas as subamostras consideradas. No entanto, como os coeficientes apresentaram variações em suas magnitudes entre as subamostras utilizadas, pode-se dizer que a escolha por estratificar a amostra gera um resultado mais preciso sobre o diferencial de salários entre homens e mulheres no Brasil. Portanto, pode-se concluir que existe diferencial de salários em favor dos homens, sendo que, no mínimo a diferença é de 36% dependendo da região e da ocupação.

Em relação aos resultados obtidos a partir do pareamento<sup>42</sup>, verifica-se que na grande maioria dos casos, exceto para os dirigentes no Sul e Sudeste, a renda média das mulheres diminuiu após o pareamento, aumentando ainda mais a disparidade de renda em relação aos homens.

Além do mais, ao parear a amostra com base em características observáveis como idade, idade ao quadrado, raça, região censitária (urbana) e nível de escolaridade, pode-se associar a diferença salarial ao efeito discriminatório, isto é, ao gênero. Esta análise decorre do fato de que ao se controlar pelas características supracitadas, a única diferença entre os trabalhadores considerados na amostra se dá em função do gênero e, por isso, pode-se concluir que o mercado remunera mais aos homens do que as mulheres com as mesmas características.

---

<sup>42</sup> Utilizado para mitigar algum viés que possa ocorrer a partir da comparação entre homens e mulheres.

De uma maneira geral, pode-se concluir que o mercado de trabalho brasileiro remunera homens e mulheres de forma diferenciada. No entanto, a discrepância é menor entre os trabalhadores que ocupam cargos de dirigentes. Por outro lado, os trabalhadores ocupados na produção de bens e serviços e de reparação e manutenção e na região Nordeste apresentam as maiores diferenças salariais em gênero.

## REFERÊNCIAS

AEBI, R.; NEUSSER, K.; STEINER, P. **Improving models of income dynamics using cross-section information**. University of Berne: Working paper, 2001.

ANGRIST, J. D.; KRUEGER, A. B. The effect of age at school entry on educational attainment: an application of instrumental variables with moments from two samples. **Journal of the American Statistical Association**, 87, 328-336, June 1992.

ANTIGO, M. F. **Mobilidade de Rendimentos no Brasil: Uma Análise a Partir de Dados Cross-Section e Longitudinais**. Disponível em: [http://www.cedeplar.ufmg.br/economia/teses/2010/Mariangela\\_Furlan\\_Antigo.pdf](http://www.cedeplar.ufmg.br/economia/teses/2010/Mariangela_Furlan_Antigo.pdf). Acesso em: 30 ago. 2011.

ARELLANO, M.; MEGHIR, C. Female labor supply and on-the-job search: an empirical model estimated using complementary data set. **Review of Economic Studies**, 59, 537-559, July 1992.

AYDEMIR, A.; CHEN, Wen-Hao; CORAK, M. Intergenerational Earnings Mobility Among the Children of Canadian Immigrants. Ottawa: Statistics Canada; **Analytical Studies Branch**. 2005. (Analytical Studies Branch Research Paper Series, 2005267).

BARROS, R. P. ET AL. **A Queda Recente da Desigualdade de Renda no Brasil**. Disponível em: <http://www.plataformademocratica.org/Publicacoes/1277.pdf>. Acesso em: 22 ago. 2013.

BARROS, R. P.; MENDONÇA, R.; DOS SANTOS, D. D.; QUINTAES, G. Determinantes do desempenho educacional no Brasil. **Pesquisa e Planejamento Econômico**, Rio de Janeiro, v. 31, n. 1, p. 1-42, 2001.

BECKER, G. S. **A Treatise on the Family**. Cambridge: Harvard University Press, 1981.

BECKER, G. S.; TOMES, N. An Equilibrium Theory of the Distribution of Income and Intergeneration Mobility. **Journal of Political Economy**, 87, p. 1153-1189, 1979.

BECKER, G. S.; TOMES, N. Human capital and the rise and fall of families, **Journal of Labor Economics**, Chicago, v. 4, p. 1-39, Jul. 1986.

BECKER, G. **The economics of discrimination**. Chicago: University of Chicago Press, 1957.

BEHRMAN, J. R.; GAVIERIA, A. U.; SZEKELY, M. S. **Intergenerational Mobility in Latin America**. Washington: Inter-American Development Bank; Research Department, 2001 (Working Papers Series, 452).

BENABOU, R.; OK, E. A. Mobility as progressivity: ranking income processes according to equality of opportunity. **National Bureau of Economic Research**, 2001 (Working Paper, 8.431).

BEN-PORATH, Y. The Production of Human Capital and the Life Cycle of Earnings. **Journal of Political Economy**, v. 75, n.4, p. 352-365, 1967.

BIGARD, A.; GUILLOTIN, Y.; LUCIFORA, C. Earnings mobility: an international comparison of Italy and France. **Review of Income and Wealth**, v. 44, n. 4, p. 535-554, 1998.

BJÖRKLUND, A.; JÄNTTI, M. Intergenerational Income Mobility in Sweden Compared to the United States. **American Economic Review**, 87(5), p. 1009-18, 1997.

BJÖRKLUND, A.; JÄNTTI, M. Intergenerational Mobility of Socio-Economic Status in Comparative Perspective. **Nordic Journal of Political Economy**, v. 26, p. 3–32. 2000.

BLAU, P. M.; DUNCAN, O. D. The American Occupational Structure. New York: John Wiley and Sons, 1967.

BLINDER, Alan S. Wage Discrimination: Reduced Form and Structural Estimates. **Journal of Human Resources**, v.8, p.436-455. 1973.

BORJAS, George J. The substitutability of black, hispanic and white labor. **Economic Journal of Human Resources**, v.8, p.436-455. 1973.

BROWN, R.; MOON, M.; ZOLOTH, B. S. Incorporating Occupational Attainment in Studies of Male-Female Earnings Differentials. **Journal of Human Resources**, 15(1), 3-28. 1980.

CALIENDO, M.; KOPEINIG, S. Some Practical Guidance for the Implementation of Propensity Score Matching. **Journal of Economic Surveys**, v. 22, n 1, p. 3172, 2008.

CAMERON, A. Colin; TRIVEDI, Pravia K. **Microeconometrics: Methods and Applications**. Cambridge University Press, Cambridge, 2005, 1058p.

CARVALHO, A. P.; NERI, M. C. ; SILVA, D. B. do N. . Diferenciais de Salários por Raça e Gênero: Aplicação dos procedimentos de Oaxaca e Heckman em Pesquisas Amostrais Complexas. In: XV Encontro Nacional de Estudos Populacionais, 2006, Caxambu - MG. **Anais do XV Encontro Nacional de Estudos Populacionais**, 2006.

CELHAY, P.; SANHUEZA, C.; ZUBIZARRETA, J. R. **Intergeneration Mobility of Income: The Case of Chile**. Disponível em: <http://fen.uahurtado.cl/wp/wp-content/uploads/2010/07/inv237.pdf>. Acesso em: 5 ago. 2011.

CHOUDHURY, S. **Reassessing the Male-Female Wage Differential: A Fixed Effect Approach**. 2001. Disponível em: <http://www.uh.edu/~adkugler/Choudhury.pdf>. Acesso em: 5 fev. 2013.

CONNOLLY, M. P. Human Capital and Growth in the Post-Bellum South: A Separate but Unequal Story. **Journal of Economic History**. 64(2), 1-39, 2004.

CORAK, M. **Do Poor Children Become Poor Adults?** Lessons for Public Policy from a Cross-Country Comparison of Generational Earnings Mobility. Disponível em: <http://www.iza.org/en/papers/Corak280904.pdf>. Acesso em: 10 jun. 2012.

COSTA, L. Aumento da Participação Feminina: Uma Tentativa de Explicação. **Anais do VII Encontro Nacional de Estudos Populacionais**, Abep, v. 2, p. 231-243, 1990.

DAVIDSON, Russel & MACKINNON, James G. **Econometric Theory and Methods**. New York: Oxford University Press, 2004.

DEARDEN, L.; MACHIN, M.; REED, H. Intergenerational Mobility in Britain. **The Economic Journal**, 107(440), pp. 47-66, 1997.

DEHEJIA, R. Practical Propensity Score Matching: A Reply to Smith and Todd. **Journal of Econometrics**, 2003.

DEHEJIA, R. Program Evaluation as a Decision Problem. **Journal of Econometrics**, v. 125, p. 141-173, 2005.

DEHEJIA, R.; WAHBA, S. Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs. **Journal of the American Statistical Association**, n. 94, p. 1053-1062, 1999.

DUARTE, A. J. M. **Disparidades regionais ou educacionais?** Um exercício contra factual. Rio de Janeiro: EPGE/FGV, 2003.

DUCLOS, Jean-Yves; ESTEBAN, J.; RAY, D. Polarization: concepts, measurement, estimation. **Econometrica**, v. 72, 2004.

DUNN, C. The intergenerational transmission of earnings: evidence from Brazil. The B.E. **Journal of Economic Analysis & Policy**, Vol. 7: Ed. 2. 2007.

ERIKSON, R.; GOLDTHORPE, J.H. Intergenerational Inequality: A Sociological Perspective. **Journal of Economic Perspectives**, Pittsburgh, v. 16, n. 3, p. 31-44, Aug. 2002.

ERMISCH J.; NICOLETTI, C. **Intergenerational Earnings Mobility: Changes Across Cohorts in Britain**, Colchester: University of Essex, 2005 (ISER Working Paper, 2005-19).

ERMISCH, J.; FRANCESCONI, J. **Intergenerational Mobility in Britain: New Evidence from the BHPS**. Acesso em: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.199.9951&rep=rep1&type=pdf>. Disponível em: 10 jun. 2012.

ERMISCH, J.; FRANCESCONI, J. Intergenerational mobility in Britain: new evidence from the British household panel study. In: CORAK, M. (Ed.) **Generational income mobility in North America and Europe**. New York: Cambridge University Press, 2004.

- FERREIRA, E. A.; VELOSO, F. A. Intergeneration Mobility of Wages in Brazil. **Brazilian Review of Econometrics**, v. 26, p. 181-211, 2006.
- FIELDS, G.; OK, E. The meaning and measurement of income mobility. **Journal of Economic Theory**, San Diego, v. 71, n.2, p. 349-377, Nov. 1996.
- FIGUEIREDO, E. A. Mobilidade Intrageracional de Renda no Brasil. **Nova Economia**, Belo Horizonte, v. 20 (3), p. 427-455, 2010.
- FLABBI, L. **Gender Discrimination Estimation in a Search Model with Matching and Bargaining**. 2005. Disponível em: <http://ftp.iza.org/dp1764.pdf>. Acesso em: 5 fev. 2013.
- FLEISCHHAUER, K. **A Review of Human Capital Theory**: Microeconomics. University of St. Gallen, Discussion Paper n. 01, 2007.
- FORTIN, N. M.; LEFEBVRE, S. **Intergeneration Income Mobility in Canada**. Disponível em: <http://faculty.arts.ubc.ca/nfortin/chapt4a.pdf>. Acesso em: 9 ago. 2011.
- FREEMAN, H.; ROSSI, P. Y.; WRIGHT, S. **Evaluating social projects in developing countries**. Paris: Development Centre/Organisation for Economic Co-operation and Development, 1980.
- GRAVELLE, H. HOLE, A. R. **Measuring and Testing for Gender Discrimination in Professions: The Case of English Family Doctors**. Disponível em: <http://www.york.ac.uk/media/economics/documents/discussionpapers/2008/0827.pdf>. Acesso em: 10 fev. 2013.
- GRAWE, N. Intergenerational Mobility for Whom? The Experience of High- and Low-Earnings Sons in International Perspective. In: Miles Corak, ed. **Generational Income Mobility in North America and Europe**. Cambridge: Cambridge University Press, 2004.
- GRAWE, N. **Intergenerational Mobility in the US and Abroad: Quantile and Mean Regression Measures**. PhD Dissertation, (University of Chicago), 2001.
- GUO, S.; FRASER, W. M. **Propensity Score Analysis: Statistical Methods and Applications**. Thousand Oaks, CA: Sage Publications, 2009.
- HAMERMESH, D. S.; GRANT, J. Econometric studies of labor-labor substitution and their implications for policy. **Journal of Human Resources**, v. XIV, n.4, 1979.
- HÄRDLE, W.; LINTON, O. Applied Nonparametric Methods. In: **Handbook of Econometrics**, v. 4, ed. R. F. Engle and D. McFadden. Amsterdam: North Holland, p. 2295–2339, 1994.
- HECKMAN, J.J. et al. **Earning functions, rates of return and treatment effects: the Mincer Equation and Beyond**. Chicago: USA, Discussion Paper n.1700, 200 p. August, 2005.

HECKMAN, J.J. Sample Selection Bias as a Specification Error. **Econometrica**, v. 47, n. 1, p. 153-161, Jan. 1979.

HECKMAN, J.J.; URZUA, S. **Comparing IV with Structural Models: What Simple IV Can and Cannot Identify**. NBER Working Paper, n. 14.706, 33 p., 2009.

HECKMAN; J. J. LOCHNER; L. J., TODD; P. E. Earnings equations and rates of return: The Mincer equation and beyond. In E. A. Hanushek and F. Welch (Eds.), **Handbook of the Economics of Education**. Amsterdam: Elsevier, cap. 7, p. 307- 458. 2006.

HIRANO, K.; IMBENS, G.; RIDDER, G. Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score. **Econometrica**, v. 71, n. 4, p. 1161-1189, 2003.

HIRSCH, Barry T.; MACPHERSON, David A. **Wages, Racial Composition, and Quality Sorting in Labor Markets**. Discussion Paper no. 1038-94, Madison, WI: Institute for Research on Poverty. 1994.

HIRSCH, Barry T.; SCHUMACHER, Edward J. Labor Earnings, Discrimination, and the Racial Composition of Jobs. **Journal of Human Resources**, 27, 426-470. 1992.

HIRVONEN, L. **Accounting for Intergeneration Earnings Persistence: Can We Distinguish Between Education, Skills, and Health?**. Disponível em: <http://www.sofi.su.se/content/1/c6/03/09/74/WP10no2.pdf>. Acesso em: 5 ago. 2011.

HOFFMANN, R.; LEONE, E. T. Participação da mulher no mercado de trabalho e desigualdade da renda domiciliar per capita no Brasil: 1981-2002. **Nova Economia**, Belo Horizonte, v. 14, n. 2, p. 35-58, ago. 2004.

HOLLAND, P. W. Statistics and Causal Inference. **Journal of the American Statistical Association**, v. 81, p. 945-70, 1986.

HSU, Mei. **Housework and Male-Female Wage Differential**. University of Washington. APEA Conference, Seattle, WA: 29-30 July 2006. N. p. Unpublished conference paper. Disponível em: <http://dspace.cigilibrary.org/jspui/bitstream/123456789/29983/1/Research%20Series%20No.%2050%20%20Wage%20Determination%20and%20Gender%20Discrimination%20in%20Uganda.pdf?1>. Acesso em: 15 fev. 2013.

IBGE. **Mulher no Mercado de Trabalho: Perguntas e Respostas**. Disponível em: [http://www.ibge.gov.br/home/estatistica/indicadores/trabalhoerendimento/pme\\_nova/Mulher\\_Mercado\\_Trabalho\\_Perg\\_Resp\\_2012.pdf](http://www.ibge.gov.br/home/estatistica/indicadores/trabalhoerendimento/pme_nova/Mulher_Mercado_Trabalho_Perg_Resp_2012.pdf). Acesso em: 13 fev. 2013.

ICHIMURA, Hidehiko. Semiparametric Least Squares (SLS) and Weighted SLS Estimation of Single-Index Models. **Journal of Econometrics**, 58.1-2, pp. 71-120. 1993.

INOUE, A.; SOLON, G. Two-Sample Instrumental Variables Estimators. **The Review of Economic and Statistic**. V. 92, n. 3, p. 557 – 561, aug. 2010.

JENCKS, C. What is the true rate of social mobility? In: BREIGER, R.L. (Eds.) **Social mobility and social structure**. Cambridge: Cambridge University Press, 1990.

JENCKS, C. ET AL. **Who Gets Ahead?** The Determinants of Economic Success in America. New York: Basic Books, Inc., 1979.

KLEIN, Roger; VELLA Francis. Estimating a Class of Triangular Simultaneous Equations Models Without Exclusion Restrictions. **Journal of Econometrics**, 154.2, pp. 154-164. 2010.

KOENKER, R., BASSET, G. Regression quantiles. **Econometrica**, v. 46, p. 33-50, 1978.

LAM, D.; SCHOENI, R. Effects of Family Background on Earnings and Returns to Schooling: Evidence from Brazil. **Journal of Political Economy**, v.101, n.4, p.710-739, aug., 1993.

LAM, D; LEVINSON, D. Declining inequality of schooling in Brazil and its effects on inequality of wages. **Journal of Development Economics**, n.37, p.199-225, 1992.

LANGONI, C. G. **Distribuição da renda e desenvolvimento econômico do Brasil**. Rio de Janeiro: Expressão e Cultura, 1973.

LANGONI, C. G. **Distribuição de Renda e Desenvolvimento Econômico no Brasil**. 3. Ed. Rio de Janeiro: FGV, 2005.

LEAL, C. I. S.; WERLANG, S. R. C. Educação e Distribuição de Renda. In: CAMARGO, J. M.; GIAMBIAGI, F. (ORG). **Distribuição de Renda no Brasil**. Rio de Janeiro: Paz e Terra, 2000.

LEAL, C. I. S.; WERLANG, S. R. C. Retornos em Educação no Brasil: 1976/89. **Pesquisa e Planejamento Econômico**, Rio de Janeiro, v.21, n.3, p.559-574, dez., 1991.

LEE, C; SOLOW, G. **Trends in Intergeneration Income Mobility**. Disponível em: <http://www.nber.org/papers/w12007.pdf>. Acesso em: 8 ago. 2011.

LEUVEN, E.; SIANESI, B. **Psmatch2**: Stata Module to Perform Full Mahalanobis and Propensity Score Matching, Common Support Graphing, and Covariate Imbalance Testing. 2003. Disponível em: <http://ideas.repec.org/c/boc/bocode/s432001.html>. Acesso em: 28 jan. 2013.

MACHADO J.A.F.; MATA J. Counterfactual Decomposition of Changes in Wage Distributions Using Quantile Regression. **Journal of Applied Econometrics**. 2005.

MACPHERSON, David A.; HIRSCH, Barry T. Wages and Gender Composition: Why Do Women's Jobs Pay Less?. **Journal of Labor Economics**, 13, 426-471. 1995.

MARIANO, F. Z. ; ARRAES, R. A . Endogeneidade da Educação na Previsão da Taxa de Retorno: Avaliação Metodológica e Aplicação para Regiões Brasileiras e Estados Selecionados.. In: VII Encontro da Economia do Ceará em Debate, 2011, Fortaleza.

**Anais do VII Encontro da Economia do Ceará em Debate.** Fortaleza: IPECE, 2011. p. 1-25.

MENEZES-FILHO, N. A. **A evolução da educação no Brasil e seu impacto no mercado de trabalho.** São Paulo: USP/ Departamento de Economia, março, 2001.

MENEZES-FILHO, N. A. Educação e Desigualdade. In: MENEZES-FILHO, N.; LISBOA, M. (ORG). **Microeconomia e Sociedade no Brasil.** Rio de Janeiro: EPGE-FGV, 2001b.

MENEZES-FILHO, N. A. Equações de Rendimentos: Questões Metodológicas. In: CORSEUIL, C. H. (EDITOR). **Estrutura Salarial:** aspectos conceituais e novos resultados para o Brasil. Disponível em:

[http://www.ipea.gov.br/agencia/images/stories/PDFs/livros/capitulo2\\_equacoes.pdf](http://www.ipea.gov.br/agencia/images/stories/PDFs/livros/capitulo2_equacoes.pdf).

Acesso em: 5 jun. 2002.

MENEZES-FILHO, N. A.; FERNANDES, R.; PICCHETTI, P. A Evolução da Distribuição de Salários no Brasil: Fatos Estilizados para as Décadas de 80 e 90. In: HENRIQUES, R. (ORG). **Desigualdade e Pobreza no Brasil.** Rio de Janeiro: IPEA, 2000.

HENRIQUES, R. (ORG). **Desigualdade e Pobreza no Brasil.** Rio de Janeiro: IPEA, 2000.

MENEZES-FILHO, N. A.; FERNANDES, R.; PICCHETTI, P. Rising Human Capital but Constant Inequality: the Education Composition Effect in Brazil. **Revista Brasileira de Economia**, Rio de Janeiro, v.60, n.4, p.407-424, out-dez., 2006.

MENEZES-FILHO, N.; ZYLBERSTAJN, H.; CHAHAD, J. P. Z. (ORG). Os Efeitos dos Sindicatos sobre o Salário Médio das Firms Brasileiras. In: CHAHAD, J. P. Z.; MENEZES-FILHO, N. **Mercado de Trabalho no Brasil.** São Paulo: LTr, 2002.

MENEZES-FILHO, N. **Mercado de Trabalho no Brasil.** São Paulo: LTr, 2002.

MILLER, Paul, W. The Effect of the Occupational Segregation of Women in Britain. **The Economic Journal**, 97, 885-896. 1987.

MINCER, J. B. **Schooling, experience and earnings.** New York: NBER, 1974. 152p.

MINCER, J. Investment in Human Capital and Personal Income Distribution. **Journal of Political Economy**, v.6, n.4, p 281-302, 1958.

MOCETTI, S. **Intergeneration Income Mobility in Italy.** Disponível em: <http://www-3.unipv.it/websiep/wp/200703.pdf>. Acesso em: 9 ago. 2011

NETTO JUNIOR, J. L. da S.; RAMALHO, Hilton Martins de Brito; ARAGON, E. K. B. Transmissão Intergeracional de Educação e Mobilidade de Renda no Brasil. In: XXXVIII Encontro Nacional de Economia - ANPEC, 2010, Salvador, BA. **Anais do XXXVIII Encontro Nacional de Economia - ANPEC, 2010.**

NEUMAN, S.; OAXACA, Ronald L. **Estimating Labor Market Discrimination with Selectivity-Corrected Wage Equations: Methodological Considerations and an Illustration from Israel.** 2001. Disponível em:

<http://www.cepr.org/meets/wkcn/3/3510/papers/neuman.pdf>. Acesso em: 5 fev. 2013.

NEUMAN, S.; SILBER, Jacques. Wage Discrimination Across Ethnic Groups: Evidence from Israel. **Economic Inquiry**, 34(4), 648-661. 1996.

NEUMARK, David. Employers' Discriminatory Behavior and the Estimation of Wage Discrimination. **Journal of Human Resources**, 23, 279-295. 1988.

OAXACA, Ronald L. Male-Female Wage Differentials in Urban Labor Markets. **International Economic Review**, 14, 693-709. 1973.

OAXACA, Ronald L.; RANSOM, Michael R. On Discrimination and the Decomposition of Wage Differentials. **Journal of Econometrics** 61, 5-21. 1994.

OAXACA, Ronald L.; RANSOM, Michael R. Searching for the Effect of Unionism on the Wages of Union and Nonunion Workers. **Journal of Labor Research** 9, 139-148. 1988.

OECD, Relatório Territorial da OCDE: Brasil 2013, OECD Publishing.

OTERO, Coral del Río; LAGO, Carlos Gradín; SÁNCHEZ, Olga Cantó. **The Measurement of Gender Wage Discrimination: The Distributional Approach Revisited.** 2004. Disponível em: <http://www.fedea.es/pub/eee/eee192.pdf>. Acesso em: 8 fev. 2013.

PASTORE, J. & VALLE SILVA, N. **Mobilidade social no Brasil.** São Paulo: Makron, 2000.

PERO, V.; SZERMAN, D. Mobilidade Intergeracional de Renda no Brasil. **Pesquisa e Planejamento Econômico**, Rio de Janeiro, v. 38, p. 20-40, 2008.

PESSOA, D.G.C.; SILVA, P.L.N. **Análise de dados amostrais complexos.** São Paulo: ABE, (13º Simpósio Nacional de Probabilidade e Estatística – Caxambu) 187p. 1998.

PETERS, H. E. Patterns of Intergeneration Mobility in Income and Earnings. **The Review of Economic and Statistics**, v. 74, p. 456-466, 1992.

POLACHEK, S. W. e KIM, M. Panel Estimates of the Gender Earnings Gap: Individual-Specific Intercept and Individual Specific Slope Models. **Journal of Econometrics**, v.61, n.1, p.23-43, 1994.

PRAIS, S. Measuring social mobility. **Journal of Royal Statistical Society**, serie A, p. 56-66, 1955.

QUEIROZ, V. S.; MOREIRA, I. T. **Uma Análise dos Diferenciais de Renda e Discriminação no Mercado de Trabalho Brasileiro.** Disponível em: <http://aplicativos.fipe.org.br/enaber/pdf/130.pdf>. Acesso em: 10 fev. 2013.

REILLY, Barry. Occupational Segregation and Selectivity Bias in Occupational Wage Equations: An Empirical Analysis using Irish Data. **Applied Economics**, 23, 1-7. 1991.

REIS, J.; BARROS, R. Desigualdade Salarial e Distribuição de Educação: A Evolução das Diferenças Regionais no Brasil. **Pesquisa e Planejamento Econômico**, Rio de Janeiro, v.20, n.3, p.415-478, dez., 1990.

REIS, J.; BARROS, R. Wage Inequality and Distribution of Education. **Journal of Development Economics**, v.36, n.1, p.117-143, jul., 1991.

RICHARD, S. **Wage Determination and Gender Discrimination in Uganda**. 2007. Disponível em: [http://www.eprc.or.ug/pdf\\_files/researchseries/series50.pdf](http://www.eprc.or.ug/pdf_files/researchseries/series50.pdf). Acesso em: 23 maio 2013.

ROSENBAUM, P.; RUBIN, D. The Central Role of the Propensity Score in Observational Studies for Causal Effects. **Biometrika**, v. 70, p. 41-55, 1983a.

ROSHOLM, Michael; SMITH, Nina. The Danish Gender Wage Gap in the 1980s: A Panel Data Study. **Oxford Economic Papers**, 48, 254-279. 1996.

RUBIN, D. B. Estimating Causal Effects of Treatments in Randomized and Nonrandomized Experiments. **Journal of Educational Psychology**, v. 66, p. 688-701, 1974.

SACHSIDA, A. LOREIRO, P. R. **Homens x Mulheres: Substitutos ou complementares no Mercado de Trabalho?** Brasília: IPEA, 1998 (Texto para Discussão N° 595).

SALVATO, M. A.; SILVA, D. G. **O Impacto da educação nos rendimentos do trabalhador: uma análise para a região metropolitana de Belo Horizonte**. 2008. Disponível em: < [http://www.marciosalvato.com/pdf/marcio\\_denis.pdf](http://www.marciosalvato.com/pdf/marcio_denis.pdf)> Acesso em: 10 Jul. 2011.

SANITER, N. **Estimating Heterogeneous Returns to Education in Germany via Conditional Heteroskedasticity**. Disponível em: [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2102484](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2102484). Acesso em: 5 ago. 2012.

SANTOS, R. V.; RIBEIRO, E. P. **Diferenciais de rendimentos entre homens e mulheres no Brasil revisitado: explorando o “Teto de Vidro”**. Disponível em: [www.eg.fjp.mg.gov.br/seminarioiv/download/vale.pdf](http://www.eg.fjp.mg.gov.br/seminarioiv/download/vale.pdf). Acesso em: 5 fev 2013.

SCHULTZ, T. W. Investment in Human Capital. **The American Economic Review**, v. 51, n. 1, p. 1-17, mar., 1961.

SCORZAFAVE, L. G.; MENEZES-FILHO, N. Caracterização da participação feminina no mercado de trabalho: uma análise de decomposição. **Economia Aplicada**, v. 10, n. 1, p. 41-55, 2006.

SCORZAFAVE, L. G.; MENEZES-FILHO, N. Participação feminina no mercado de trabalho brasileiro: evolução e determinantes. **Pesquisa e Planejamento Econômico**, Rio de Janeiro, v. 31, n. 3, p. 441-478, dez. 2001.

SHORROCKS, A. The measurement of mobility. **Econometrica**, v. 46, p. 1013-1024, 1978.

SKINNER, C.J.; HOLT, D e SMITH, T.M.F. **Analysis of Complex Surveys**. Chichester: John Wiley & Sons. 1989.

SMITH, J. A.; TODD, P. E. Does Matching Address LaLonde's Critique of Non experimental Estimators. **Journal of Econometrics**, v.125, n. 1-2, p. 305-353, 2005.

SMITH, J. A.; TODD, P. E. Reconciling Conflicting Evidence on the Performance of Propensity-Score Matching Methods. **American Economic Review**, v. 91, p. 112-118, 2001.

SOLON, G. Intergenerational Mobility in the Labor Market. In Orley Ashenfelter and David Card (eds.), **Handbook of Labor Economics**, v. 3A, p. 1761-1800, Amsterdam: North-Holland, 1999.

SOUZA, L. R. de; RIOS-NETO, E. L. G.; QUEIROZ, B. L. A relação entre parturição e STOCK, H. J.; WATSON, W. M. **Econometria**. Pearson, Addison Wesley, 2004.

SULIANO, D. C.; SIQUEIRA, M. L. **Um estudo do retorno da educação na região nordeste: análise dos estados da Bahia, Ceará e Pernambuco a partir da recente queda da desigualdade**. (texto para discussão, n. 72) Fortaleza, 2010.

trabalho feminino no Brasil. **Revista Brasileira de Estudos de População**, v. 28, n. 1, p. 57-79, jun. 2011.

VIEIRA, C. R.; BAGOLIN, I. P. Mobilidade Intrageracional e Intergeracional da Renda na Região Sul do Brasil nos Períodos Pré-Estabilização e Pós-Estabilização Econômica. In: XI Encontro de Economia da Região Sul - ANPEC SUL, 2008, Curitiba. **Anais do XI Encontro de Economia da Região Sul**. Curitiba: Boletim de Conjuntura - UFPR, 2008.

WOOLDRIDGE, J. M. **Econometric Analysis of Cross Section and Panel Data**. The MIT Press, Cambridge, MA, 2002.