

**AMANDA FERRARI UCELI**

**ANÁLISE DO RETORNO SALARIAL À EDUCAÇÃO NO BRASIL NO  
PERÍODO BIANUAL DE 1999 A 2011.**

Dissertação apresentada à  
Universidade Federal de  
Viçosa, como parte das  
exigências do Programa de  
Pós-graduação em Economia  
Aplicada, para a obtenção do  
título de *Magister Scientiae*.

VIÇOSA  
MINAS GERAIS - BRASIL  
2014

**Ficha catalográfica preparada pela Biblioteca Central da  
Universidade Federal de Viçosa - Câmpus Viçosa**

T

U16a  
2014 Uceli, Amanda Ferrari, 1987-  
Análise do retorno salarial à educação no Brasil no  
período bianual de 1999 a 2011 / Amanda Ferrari Uceli. -  
Viçosa, MG, 2014.  
xi, 86f. : il. (algumas color.) ; 29 cm.

Inclui anexos.  
Orientador : João Eustáquio de Lima.  
Dissertação (mestrado) - Universidade Federal de  
Viçosa.  
Referências bibliográficas: f.70-74.

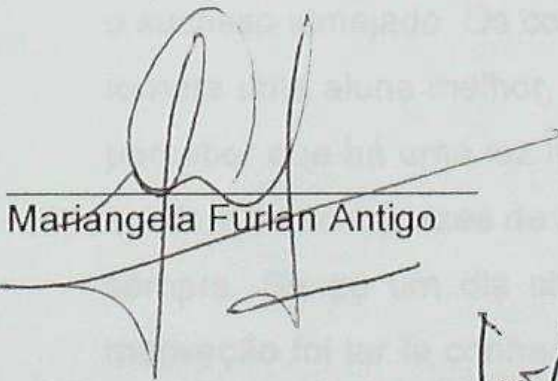
1. Educação. 2. Pseudo-painel. 3. Pesquisa Nacional de  
Amostra de Domicílio. I. Universidade Federal de Viçosa.  
Departamento de Economia Rural. Programa de  
Pós-graduação em Economia Aplicada. II. Título.

CDD 22. ed. 370.7

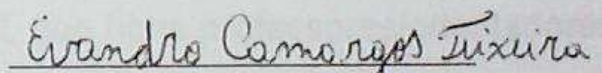
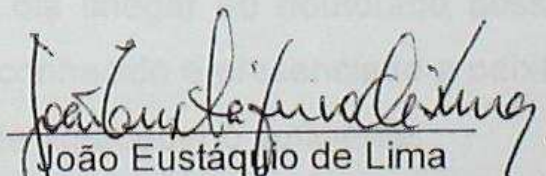
**ANÁLISE DO RETORNO SALARIAL À EDUCAÇÃO NO BRASIL NO  
PERÍODO BIANUAL DE 1999 A 2011.**

Dissertação apresentada à  
Universidade Federal de  
Viçosa, como parte das  
exigências do Programa de  
Pós-graduação em Economia  
Aplicada, para a obtenção do  
título de *Magister Scientiae*.

**Aprovada em: 17 de fevereiro de 2014**



Mariângela Furlan Antigo

  
Evandro Camargos Teixeira  
João Eustáquio de Lima  
(Orientador)

## **AGRADECIMENTOS**

Agradeço primeiramente aos meus pais, Marilza Ferrari Uceli e José Maria Uceli que são a fonte de força para todas as batalhas que eu encaro e encarei na vida. Essa conquista antes de ser minha é de vocês mãe e pai. Agradeço aos amigos que sempre me estimularam a não desanimar, mesmo quando as coisas pareciam bem.

À Deus, que me dá coragem todos os dias para seguir em frente.

À minha irmã e ao meu namorado, por sempre estarem ao meu lado.

Às meninas da Aloha, vocês foram verdadeiras irmãs e ouvidos preciosos na hora do desespero, obrigada por tudo. À Camila, minha amiga Cami de sempre, que entedia tudo e me apoiava em tudo, que acredita tanto em mim que me faz acreditar também.

Aos amigos da Economia Rural, que me trouxeram sempre novos e preciosos conhecimentos.

Especialmente agradeço à Professora Mariangela Furlan Antigo, pessoa iluminada que permitiu que esse projeto fosse possível e alcançasse o sucesso almejado. Os conhecimentos que aprendi com você professora me tornara uma aluna melhor, uma pesquisadora mais competente e me fizeram perceber que há uma luz no fim do túnel, que bons professores/orientadores existem e são capazes de mudar o curso da vida acadêmica de um aluno pra sempre. Se eu um dia chegar ao doutorado posso garantir que parte da motivação foi ter te conhecido e presenciado a paixão na docência em você. Nunca serei capaz de agradecer o suficiente.

À Universidade Federal de Viçosa e ao Departamento de Economia Rural pela oportunidade de aprendizado e desenvolvimento pessoal.

À Fundação de Amparo à Pesquisa do Estado de Minas Gerais pelo apoio financeiro.

## SUMÁRIO

<b>Lista de Tabelas .....</b>	<b>v</b>
<b>Lista de Gráficos .....</b>	<b>vi</b>
<b>Anexos .....</b>	<b>vii</b>
<b>Resumo .....</b>	<b>viii</b>
<b>Abstract .....</b>	<b>x</b>
<b>1 Introdução .....</b>	<b>1</b>
1.1 <i>Considerações Iniciais .....</i>	<i>1</i>
1.2 <i>O Problema e sua Importância .....</i>	<i>7</i>
1.3 <i>Hipótese .....</i>	<i>8</i>
1.4 <i>Objetivos: .....</i>	<i>8</i>
1.4.1 <i>Objetivo Geral: .....</i>	<i>8</i>
1.4.2 <i>Objetivos Específicos: .....</i>	<i>9</i>
<b>2 Evidências Empíricas .....</b>	<b>10</b>
2.1 <i>Retorno para a Educação no Mundo .....</i>	<i>10</i>
2.2 <i>Retorno para a Educação no Brasil .....</i>	<i>12</i>
<b>3 Referencial Teórico .....</b>	<b>18</b>
<b>4 Metodologia .....</b>	<b>23</b>
4.1 <i>Modelo Minceriano .....</i>	<i>23</i>
4.2 <i>Modelo de Pseudo-painel .....</i>	<i>30</i>
4.3 <i>Modelo Econométrico para Determinação do Retorno Salarial da Educação .....</i>	<i>38</i>
4.4 <i>Fontes de Dados .....</i>	<i>41</i>
<b>5 Resultados e Discussões .....</b>	<b>46</b>
5.1 <i>Análise descritiva dos dados .....</i>	<i>46</i>
5.1.1 <i>Dados Individuais .....</i>	<i>46</i>
5.1.2 <i>Dados Agrupados segundo Coorte de Nascimento, Raça, Gênero e Região Metropolitana .....</i>	<i>50</i>
5.2 <i>Retornos para a Educação no Brasil .....</i>	<i>53</i>
5.2.1 <i>Retorno à Escolaridade: Modelo de Efeitos Fixos com Escolaridade Contínua .....</i>	<i>55</i>
5.2.2 <i>Retorno à Escolaridade: Modelo de Efeitos Fixos com Escolaridade por Grupo .....</i>	<i>58</i>

5.2.3	Retorno à Escolaridade: Modelo de Efeitos Fixos com Escolaridade medida pela Posse de Diploma.....	62
<b>6</b>	<b>Conclusões .....</b>	<b>67</b>
<b>7</b>	<b>Referências Bibliográficas .....</b>	<b>70</b>
<b>8</b>	<b>Anexos .....</b>	<b>75</b>

## LISTA DE TABELAS

Tabela 1: Definição das Variáveis.....	39
Tabela 2: Média de Experiência, Escolaridade e Salário/hora para Homens Não Brancos* e Brancos* no Brasil, período bianual de 1999 a 2011.....	46
Tabela 3: Média de Experiência, Escolaridade e Salário/hora para Mulheres Não Brancas* e Brancas* no Brasil, período bianual de 1999 a 2011.....	47
Tabela 4: Proporção de Indivíduos por escolaridade no Brasil no período bianual de 1999 a 2011. ....	48
Tabela 5: Estimação por Efeitos Fixos da Equação de Rendimentos para o Brasil no período bianual de 1999 a 2011. Modelo 1.....	55
Tabela 6: Estimação por Efeitos Fixos da Equação de Rendimentos para o Brasil no período bianual de 1999 a 2011. Modelo 2.....	59
Tabela 7: Estimação por Efeitos Fixos da Equação de Rendimentos para o Brasil no período bianual de 1999 a 2011. Modelo 3.....	63

## LISTA DE GRÁFICOS

Gráfico 1: Percentual Amostrado por Grupo de Escolaridade no Brasil, nos anos do período Bianual de 1999 a 2011. ....	49
Gráfico 2: Escolaridade e Renda Médias para <i>Coortes</i> Femininas e Masculinas no Brasil no período bianual de 1999 a 2011.....	51
Gráfico 3: Escolaridade e Renda Médias para <i>Coortes</i> Brancas* e Não Brancas** no Brasil no período bianual de 1999 a 2011.....	51
Gráfico 4: Escolaridade e Renda Médias para <i>Coortes</i> residentes em Região Metropolitana e Não Metropolitana no Brasil no período bianual de 1999 a 2011.....	52



## ANEXOS

A1: <i>Coortes</i> por Período de Nascimento .....	75
A2: Número de Indivíduos por <i>Coorte</i> , em cada ano da amostra. ....	75
A 3: Teste de Hausman para determinação do Modelo mais adequado entre Efeitos Fixos e Aleatórios. Modelos 1, 2 e 3. ....	78
A 4: Teste de Chow para determinação do Modelo mais adequado entre Efeitos Fixos e <i>Pooled</i> . Modelos 1, 2 e 3. ....	79
A 5: Teste de Autocorrelação Serial para dados em painel de Wooldridge. Modelos 1, 2 e 3. ....	79
A 6: Teste de Wald Modificado para Heteroscedasticidade na unidade de <i>cross</i> , para dados em painel, sobre modelos de Efeitos Fixo. Modelos 1, 2 e 3. ....	80
A 7: Modelo 1 estimado pelo Método de Mínimos Quadrados Ordinários. ....	80
A 8: Modelo 1 estimado pelo Método de Efeitos Fixos para grupos diferenciados por: Raça, Gênero e Período de Nascimento. ....	81
A 9: Modelo 1 Completo. ....	82
A 10: Modelo 2 estimado pelo Método de Mínimos Quadrados Ordinários. ....	83
A 11: Modelo 2 estimado pelo Método de Efeitos Fixos para grupos diferenciados por: Raça, Gênero e Período de Nascimento. ....	83
A 12: Modelo 2 Completo. ....	84
A 13: Modelo 3 estimado pelo Método de Mínimos Quadrados Ordinários. ....	85
A 14: Modelo 3 estimado pelo Método de Efeitos Fixos para grupos diferenciados por: Raça, Gênero e Período de Nascimento. ....	85
A 15: Modelo 3 Completo. ....	86

## RESUMO

UCELI, Amanda Ferrari, M.Sc., Universidade Federal de Viçosa, fevereiro de 2014. **Análise do Retorno Salarial à Educação no Brasil no Período bianual de 1999 a 2011.** Orientador: João Eustáquio de Lima, Co-Orientadores: Viviani Silva Lírio e Leonardo Bornacki de Mattos.

A educação é considerada como o mecanismo central de desenvolvimento social e crescimento econômico, além de ser o meio mais eficaz de superação às desigualdades sociais e regionais no país. Por essa razão, na última década observou-se no Brasil um esforço crescente para que os investimentos em educação sejam suficientes para que as deficiências seculares desse setor possam ser superadas. Apesar disso, são escassos os estudos que determinem o ganho, mensurado pelos retornos salariais, desse investimento nos anos recentes para o Brasil. Mais que isso, poucos estudos consideraram que, sendo o Brasil um país em desenvolvimento, esse retorno é melhor mensurado a partir de uma análise longitudinal. O presente estudo propôs então a formação de uma base de dados com esse perfil, que capte o efeito da educação sobre a renda salarial ao longo do tempo. A partir dos dados da Pesquisa Nacional por Amostra de Domicílios (PNAD), é construída uma amostra de *pseudo-painel*, que agrupa os dados de *cross-sections* repetidas. O período considerado vai de 1999 à 2011, e se divide em intervalos bianuais. Essa base de dados foi desenvolvida seguindo os trabalhos de Deaton (1985), Moffitt (1993), Collado (1997), McKenzie (2004) e Verbeek e Vella (2005). As *coortes*, que são a unidade de análise nesse estudo, foram determinadas por gênero, raça, região de residência e período de nascimento. A hipótese feita foi de que o retorno salarial à educação quando medida em uma amostra longitudinal é menor do que aquele medido para dados de *cross-section*. Partindo de um modelo de Efeitos Fixos para dados em *pseudo-painel*, e contrapondo seus resultados àqueles obtidos pelo método de MQO, essa hipótese foi verificada. Os retornos para a escolaridade foram definidos em três formatos distintos: escolaridade contínua, por grupo de escolaridade e por diploma. Além disso, esses retornos foram diferenciados por gênero e raça. O retorno médio estimado para a educação,

em sua medida contínua, foi de 0,06%. Já os resultados por grupos educacionais, mostrou que o retorno por ano adicional de educação não é homogêneo entre os ciclos educacionais. No grupo com 5 à 8 anos de escolaridade o retorno à renda de um ano adicional de educação é de 0,18%, já no ciclo de 11 anos ou mais de escolaridade esse mesmo aumento de escolaridade eleva a renda em 0,83%. O retorno por diploma mostrou-se persistente em remunerar melhor os diplomas de maior escolaridade relativa. A obtenção do título de graduação representa, na amostra considerada, um ganho salarial de 1,42% em relação ao diploma do ensino fundamental de quatro anos.

## ABSTRACT

UCELI, Amanda Ferrari, M.Sc., Universidade Federal de Viçosa, february of 2014. **Analysis of Wage Return to Education in Brazil in the biennial period 1999-2011**. Adviser: João Eustáquio de Lima, Co-Advisers: Viviani Silva Lírio and Leonardo Bornacki de Mattos.

Education is regarded as the central mechanism of social development and economic growth, in addition to being the most effective way of overcoming social and regional inequalities in the country. For this reason, in the last decade has been observed in Brazil a growing effort done in investments in education sufficient to overcome secular deficiencies of this sector. Despite this, studies that determine the gain, measured by salary, investment returns in recent years for Brazil are scarce. More than that, few studies considered that, being the Brazil a developing country, this return is better measured from a longitudinal analysis. This study proposed the formation of a database with this profile, which captures the effect of education on wage income over time. From the data of the Pesquisa Nacional por Amostra de Domicílios (PNAD), is constructed a sample of pseudo-panel, which groups data from repeated *cross-sections*. The period considered going from 1999 to 2011, and is divided into twice-yearly intervals. This database was developed following the work of Deaton (1985), Moffitt (1993), Collado (1997), Mckenzie (2004) and Verbeeck and Vella (2005). The cohorts, that are the unit of analysis in this study, were determined by the gender, race, region of residence and period of birth. The hypothesis made was that the return on salary for education if measured in a longitudinal sample is smaller than that measured for *cross-section* data. Starting with a Fixed Effects model for data on *pseudo-painel*, and opposed its results to those obtained by the MQO method, this hypothesis was verified. The returns to schooling were set in three different formats: continuous education, group of years of schooling and diploma. Moreover, these returns have been differentiated by gender and race. The estimated average return for education, in its continuous measure, was 0.06%. Furthermore, the results by educational groups, showed that the return for additional year of education is not homogeneous between educational cycles. In the group with 5 to 8 years

of schooling the income return of an additional year of education is 0.18%, additionally, in cycle of 11 years or more of schooling that same increase in schooling raises the rent in 0.83%. The return by diploma was persistent in remunerate better higher school diplomas. The graduation title represents, in the sample considered, a wage gain of 1.42% in relation to the diploma of basic education of four years.

# **1 INTRODUÇÃO**

## **1.1 CONSIDERAÇÕES INICIAIS**

A relevância da qualidade da mão de obra para o crescimento econômico, da renda individual e da distribuição de renda de um país tem estado em foco nas análises econômicas desde que Solow (1956) observou que o crescimento do produto agregado norte-americano, em níveis superiores à taxa de crescimento populacional, da disponibilidade de terra e de capital, só poderia ser explicado por mudanças na tecnologia ou na mão de obra que elevou a produtividade dos insumos de produção.

Da necessidade de determinar o propulsor do crescimento econômico acima do crescimento do estoque de insumos produtivo é que tem início a Teoria do Capital Humano. Os trabalhos seminais de Schultz (1960), Becker (1962) e Mincer (1958, 1970) marcam o início das análises econômicas segundo os pressupostos desse corpo teórico, e neles é formulado e aprimorado o conceito de capital humano. Segundo Schultz (1960), a educação, o conhecimento e o treinamento são formas de capital na medida em que, quando adquiridos pelo indivíduo, fazem parte de quem ele é, não podendo ser mais retirados dele, ou comercializados separadamente de seu possuidor, como um bem material. Além disso, trata-se de um capital, pois possui a capacidade de gerar acréscimo de produtividade daquela mão de obra.

Mais do que apenas elevar a produtividade da mão de obra, os autores seminais da Teoria do Capital Humano apontam para a relevância desse tipo de investimento na determinação da renda individual para mostrar que as desigualdades de renda entre países eram mais bem explicadas por suas desigualdades de qualificação da mão de obra do que por suas diferenças de acumulação de capital físico.

O investimento em capital humano é definido, segundo Becker (1962), como a “atividade que influencia o futuro da renda real através da incorporação de recursos nas pessoas”. Por essa razão, não apenas a

educação formal ou a experiência no trabalho são consideradas como investimento produtivo no homem.

Ainda para esse autor, o bem estar do trabalhador é fonte de crescimento da produtividade, tendo em vista que, quando o indivíduo dispõe de condições adequadas de alimentação, saúde, lazer e conhecimento, se torna mais produtivo. Esse investimento pode assumir formas diversas, mas aquelas que recebem maior atenção em pesquisas econômicas são a escolaridade formal, a experiência no mercado de trabalho e o treinamento no trabalho.

A escolha dessas formas de acumulação de capital humano, para os estudos em economia, se justifica na viabilidade para mensuração quantitativa. Nesse sentido, muitas vezes, a escolaridade formal e a experiência<sup>1</sup> no trabalho são entendidas como sinônimos do capital humano, e assim será considerado no presente estudo.

De acordo com Schultz (1961), a educação tem o poder, em particular, de “aumentar a eficiência das instituições econômicas e políticas, e ainda gerar conhecimento científico avançado”. A educação formal é considerada como o mecanismo central de desenvolvimento social e crescimento econômico, além de ser o mecanismo mais eficaz de superação às desigualdades sociais e regionais Becker (1962), Mincer (1970), Heckman (2000). Ela possibilita que o indivíduo modifique a sociedade em que vive, determinando sua capacidade produtiva e sua capacidade dinamizadora do mercado (IPEA, 2008).

Uma vez que o investimento em capital humano é capaz de elevar a renda dos indivíduos, melhorar a distribuição dessa e promover o desenvolvimento social, como mostrado na extensa revisão de literatura sobre o tema em questão feita por Barbosa Filho e Pêsoa (2010), deve haver incentivo do Estado para que a oferta, a qualidade e a demanda por educação e treinamento sejam crescentes no país.

A Teoria do Capital Humano postula que o crescimento e o desenvolvimento econômico dependem, fundamentalmente, da participação

---

<sup>1</sup> Medida pelo número de anos vividos após findado o ensino básico de onze anos, considerando que o indivíduo não seja aposentado.

do Estado no processo de oferta de mão de obra qualificada. O motivo para a participação estatal, conforme essa teoria, é o alto custo que o indivíduo tem em adquirir conhecimento, seja na forma de educação, seja na forma de treinamento.

Para Mincer (1974), a acumulação de capital humano, com recursos privados, depende de dois fatores: custos e benefícios. Visto que acumular capital humano requer tempo, o trabalhador (efetivo ou potencial) deixa de participar de processos produtivos geradores de renda. Além de custo de oportunidade, associado ao tempo, o indivíduo tem também gastos financeiros com a aquisição de educação e aprimoramento de habilidades.

Uma consequência direta dos custos é que os indivíduos pertencentes às famílias mais pobres tendem a investir menos nesse tipo de capital, dada a restrição de renda. Mesmo para os indivíduos que dispõem dos recursos financeiros para adquirir a quantidade de capital humano demandada pelo mercado, arcar com esses gastos requer que os ganhos futuros, a eles associados, os compensem.

Portanto, a participação do Estado na oferta de educação e no incentivo a essa demanda é de fundamental importância, principalmente quando uma parcela considerável da população não dispõe de recursos próprios para elevar seu nível de capital humano, como no caso brasileiro. A educação compulsória juntamente com a oferta gratuita por parte do estado caracterizam a participação efetiva do governo no provimento de recursos para a formação de capital humano em um país.

Na visão do governo federal brasileiro, sem a devida valorização e o investimento adequado em educação, não será possível superar mazelas sociais e transpor os limites que estão postos para o desenvolvimento e o crescimento econômico no país (HADDAD, 2008). No entanto, é necessário considerar não só a quantidade de recursos ofertados, mas também a qualidade, a organização espacial e social do investimento realizado<sup>2</sup>.

Os objetivos das políticas educacionais e dos investimentos feitos serão alcançados à medida que a escolaridade média da população brasileira

---

<sup>2</sup> O presente trabalho não irá avaliar o efeito da qualidade da educação por não fazer parte do proposto.



se elevar e as disparidades salariais entre indivíduos com diferentes acumulações de capital humano forem reduzidas. Segundo Moretti (2004), essas reduções indicarão maior homogeneidade da oferta de mão de obra qualificada nos diversos níveis e uma parcela maior da população com todos os níveis escolares.

Muitos estudos econômicos tiveram como foco de análise o diferencial de retorno para os diversos níveis de educação, ao longo do último século, no Brasil. Dentre os mais recentes, destacam-se os trabalhos de Arbache, Dickerson e Green (2001); Sachsida, Loureiro e Mendonça (2004); Ferreira (2004); França, Gasparini e Loureiro (2005); Resende e Wyllie (2006); Moura (2008); Barbosa Filho e Pêssa (2008); Cangussu, Salvato e Nakabashi (2010); Santos e Camillo (2011); Pereira et al. (2013) e Teixeira e Menezes-Filho (2012).

Apesar de anterior a esses, o trabalho de Barros e Mendonça (1997) observa que há no Brasil um déficit secular de investimento em educação, em todos os níveis, que torna a mão de obra brasileira menos produtiva do que o potencial de que dispõe o país. Essa é também a conclusão dos estudos em destaque, nos quais, um resultado comum é o de um elevado retorno estimado para escolaridades superiores a nove anos. Evidencia-se, assim, que é persistente, no período analisado nos trabalhos supracitados, a deficiência no investimento em educação no Brasil.

A maioria desses estudos fundamentaram suas análises em períodos anteriores ao ano 2000. Entretanto, como será explicitado adiante, percebe-se um crescimento mais consistente do percentual do PIB investido em educação após o ano 2000. Nesse período, surgem ainda novas políticas e metas mais robustas de expansão da oferta de educação no país. Outra questão comum à maioria desses trabalhos é o tipo de dado considerado. De um modo geral, há a predominância de análises sobre dados de *cross section*, que desconsideram o efeito intertemporal da acumulação de capital humano entre os indivíduos. Observa-se então a existência de uma lacuna temporal e metodológica nos estudos sobre a Teoria do Capital Humano no Brasil.

Os estudos desse tipo, no Brasil e no restante do mundo, se baseiam majoritariamente no método proposto por Mincer (1970,1974). Porém, a

combinação de dados de seção cruzada com o método minceriano para a determinação dos retornos salariais à educação<sup>3</sup>, quando estimados por mínimos quadrados ordinários, lidam com vieses pela omissão de habilidade inatas aos indivíduos. Essa falha decorre do fato de considerar indivíduos heterogêneos como homogêneos, de modo que a diferenciação entre eles se dê exclusivamente pela quantidade de capital humano acumulado.

A maioria dos trabalhos sobre o tema no Brasil controla os vieses associados ao modelo de Mincer pelo método de variáveis instrumentais (VI). Desse modo, pressupõe-se que a inclusão do instrumento para a habilidade elimina o viés existente pela relação entre o nível de capital humano<sup>4</sup> acumulado e a habilidade intrínseca omitida no modelo. Entretanto, como destaca Card (2001), esse mecanismo é limitado e sua eficiência pode ser questionada. Entre os trabalhos que propõem formas diferentes para determinar os retornos ao capital humano, ganham atenção os trabalhos de Sachsida, Loureiro e Mendonça (2004) e Arbache, Dickerson e Green (2001), no Brasil; e, no restante do mundo, ganham destaque as análises de Heckman, Lochner e Todd(2003), Mora e Muro (2007) e Wanrunsiri e McNown (2010). Esses estudos propuseram uma análise longitudinal para a relação.

Como esses autores indicam, os vieses associados às heterogeneidades locais e individuais da amostra podem ser controlados de maneira eficiente por uma análise de longo prazo. Como apontado por Heckman, Lochner e Tood (2003), a análise a partir de informações intertemporal desse retorno é a mais adequada para países em desenvolvimento, onde o ciclo educacional de uma política econômica só terá seu efeito verificado em longo prazo, como é o caso do Brasil.

---

<sup>3</sup> Como é feito na maior parte dos estudos para o Brasil.

<sup>4</sup> Que na maioria das vezes é medido pela quantidade de anos de estudo do indivíduo.

Devido à ausência de uma base dados longitudinal que seja representativa do Brasil como um todo<sup>5</sup>, esse trabalho constroem, a partir dos dados anuais da Pesquisa Nacional por Amostra de Domicílios (PNAD), uma amostra temporal de *cross section* repetidas. Ambos seguem o estudo pioneiro de Deaton (1985), e constroem, a partir do período de nascimento dos indivíduos, grupos homogêneos para cada ano da amostra. Com as médias das variáveis de interesse, forma-se uma nova base de dados que permite considerações longitudinais, denominada painel. Sobre alguns pressupostos, dados dessa natureza permitem aplicar os métodos de estimação e tratar os resultados como em um painel verdadeiro, sem as limitações e os problemas associados aos dados que um painel real pode apresentar; tais qual a atrição, o condicionamento, entre outros que serão discutidos ao longo do presente trabalho.

Apesar dos ganhos, há desvantagens no uso de dados de *pseudo-painel*. As variáveis da amostra final são determinadas pelas médias das variáveis individuais e a unidade de análise deixa de ser o indivíduos, passando a ser a *coorte* homogênea. Entretanto, os resultados ainda são robustos e relevantes.

Dessa forma, este estudo dissertativo propõe atualizar os resultados acerca dos retornos salariais da educação no Brasil a partir de uma base de dados que forneça informações da amostra no decorrer dos anos, de modo que contribuirá com um avanço metodológico sobre os trabalhos já desenvolvidos no Brasil.

---

<sup>5</sup> Há apenas duas bases de dados verdadeiramente longitudinais no Brasil. São elas: a Pesquisa Mensal de Emprego (PME) e a Relação Anual de Informações Sociais (RAIS). A primeira se limita a seis regiões metropolitanas no país e compreende um acompanhamento de doze meses para um mesmo indivíduo. A segunda se aplica apenas sobre o setor formal da economia. Sendo assim, essas duas bases são limitadas para os objetivos propostos pelo presente estudo.

## **1.2 O PROBLEMA E SUA IMPORTÂNCIA**

O setor educacional se torna cada vez mais importante para o país que busca elevar seus níveis de desenvolvimento econômico. Na última década, o perfil educacional e social do Brasil mudou em função das demandas crescentes por uma mão de obra mais qualificada, que se adequasse aos novos paradigmas produtivos. Como destaca Saboia et al. (2010), o Brasil foi, durante anos, capaz de manter sua competitividade econômica por meio do uso de mão de obra com baixa qualificação e, portanto, remunerada a salários inferiores aos observados em países com produção concorrente. Porém, a rápida evolução tecnológica sobre os processos produtivos e sua difusão criaram maior competitividade no mercado mundial, forçando os países a se adequarem. Para que a competitividade se mantivesse, foi necessário investir em qualificação e capacitação dos trabalhadores.

Nesse sentido, novos caminhos têm sido traçados para o investimento estatal em educação. Dentre as políticas mais recentes nesse setor se destaca a criação, em 2007, do Plano de Desenvolvimento da Educação (PDE), que instituiu os mecanismos necessários para que a educação se torne uma ferramenta de propulsão econômica para o país. No mesmo ano, o Plano Nacional de Educação (PNE) formulou as bases aplicadas para propiciar o desenvolvimento econômico via educação. Uma medida quantitativa para o crescimento do investimento no setor educacional é a parcela do PIB investida nesse setor. No Brasil, no ano 2000, 4,7% do PIB eram destinados à educação pública. Já em 2011 esse percentual passou para 6,1%, sendo que a meta é de crescimento nos próximos anos. O PNE de 2007 definiu que até 2015 deve-se atingir 7% do PIB brasileiro destinado ao setor educacional (INEP, 2013).

Porém, como já foi mencionado, as medidas de retorno à escolaridade estão desatualizadas. A escassez de resultados que levem em consideração o crescimento recente do montante investido em educação no efeito da escolaridade sobre a renda limitam a ação da política pública. O resultado em questão fornece respostas para perguntas como: qual o nível escolar mais valorizado no mercado de trabalho? Qual o ganho, em termos salariais, para o indivíduo na conclusão de cada ciclo escolar? Como a população mais

jovem, que foi afetada diretamente pela elevação recente do investimento em educação, é remunerada?

Assim, o presente estudo acadêmico propõe determinar a magnitude dos retornos associados aos diferentes níveis educacionais no Brasil no período bianual de 1999 e 2011. Além de atualizar os resultados de investigações anteriores sobre esse tema, o trabalho contribui com o desenvolvimento de uma base dados pouco explorada para a análise longitudinal desse tema.

O estudo se apresenta dividido em oito partes principais, sendo a primeira a Introdução, que contempla o problema proposto, a justificativa, a hipótese e os objetivos do estudo. A segunda parte apresenta as Evidências Empíricas com informações trazidas pelos estudos acerca do tema desenvolvidos para o Brasil e para outros países. O Referencial Teórico é definido na terceira parte, norteando as ideias de todo o trabalho. Após isso, na quarta parte, é desenvolvido o Método adotado para alcançar os objetivos propostos. Já a quinta parte traz os Resultados e Discussões. A sexta seção apresenta as Conclusões do desenvolvimento deste estudo. A sétima identifica as Referências. Por fim, a oitava seção ocupa-se dos Anexos.

### **1.3 HIPÓTESE**

Espera-se que o retorno para a educação no Brasil, medido por uma análise longitudinal e sobre o período recente compreendido entre os anos de 1999 e 2011, seja menor do que aquele encontrado sobre uma perspectiva de curto prazo e para períodos anteriores a 1999.

### **1.4 OBJETIVOS:**

#### **1.4.1 Objetivo Geral:**

Mensurar e analisar o efeito da educação sobre os salários dos grupos de indivíduos que estão ocupados no mercado de trabalho e que são homogêneos por raça, gênero, idade e região de residência, no Brasil, no período bianual entre 1999 e 2011.

#### **1.4.2 Objetivos Específicos:**

- a) Determinar o efeito marginal da educação e da experiência sobre os ganhos salariais;
- b) Determinar o efeito diploma para os diferentes graus de escolaridade sobre os salários;
- c) Determinar sobre os quatro grupos escolares a magnitude ganho marginal de uma ano adicional de educação;
- d) Determinar os diferenciais de retorno à educação para os diferentes grupos da sociedade: homens, brancos, residentes em região metropolitana.

## 2 EVIDÊNCIAS EMPÍRICAS

### 2.1 RETORNO PARA A EDUCAÇÃO NO MUNDO

Muitos trabalhos já problematizaram os retornos associados à educação, e muitas formas de análise foram adotadas. A análise de Mincer (1974) é o marco e a principal referência para a estimação formal desses retornos. Seu trabalho buscou determinar como a renda se distribuía entre grupos diferenciados somente pela escolaridade. Para isso, propôs um modelo matemático de controle dos fatores determinantes do salário, permitindo isolar os retornos da educação sobre essa variável. A equação de salário proposta por Mincer (1970, 1974) vem sendo aplicada por diversos outros estudos sobre o tema.

Apesar de extensivamente usado, o método minceriano (1970, 1974) apresenta limitações e simplificações importantes. A análise a partir da equação de rendimentos proposta por esse autor requer que tanto o método quanto os dados sejam adequados, para que seja possível obter estimações não viesadas do retorno salarial do capital humano, viés esse que naturalmente emerge do modelo.

O trabalho de Heckman, Lochner e Todd (2003) aprimora a proposta da equação de Mincer (1970), e compara a estimação do retorno utilizando o formato minceriano e uma forma menos restrita desenvolvida por esses autores. Os pressupostos mincerianos de paralelismo da experiência<sup>6</sup>, linearidade da relação entre renda e escolaridade e perfeita informação conhecido pelos agentes são quebrados e Heckman, Lochner e Todd (2003), que diferenciam os indivíduos por grupo escolar, e estimam a equação de rendimentos segundo a formação desses grupos de maneira separada. Além de estimarem os retornos à escolaridade utilizando dados de *cross section* e de *coortes* em repetidas observações transversais, esses autores propõem formas funcionais variadas para a equação de rendimentos, e os resultados

---

<sup>6</sup> O paralelismo da experiência está relacionado com o fato de que, sem nenhuma consideração adicional, o retorno salarial determinado para dois indivíduos com mesma escolaridade, porém com experiência distinta, será o mesmo.

delas mostram que os pressupostos mincerianos não se verificam no período pós 1960 para dados norte-americanos.

O único pressuposto que se manteve válido foi o de que a relação entre rendimento e experiência apresenta o comportamento de “U” invertido. Segundo esse pressuposto, a relação entre a experiência (ou a idade) e a renda segue a trajetória de uma parábola invertida. Essa relação é explicada pelos rendimentos marginais decrescentes a essa forma de capital humano.

Os autores puderam relativizar entre os resultados que obtiveram com dados em *cross section* e séries temporais de dados agrupados, e concluíram que, quando há antecipação por parte dos agentes quanto ao retorno da escolaridade, os dados em *cross section* não são eficientes. Assim, a estabilidade da estimação em observações dos grupos é superior.

Heckman, Lochner e Todd (2003) observam ainda que as expectativas sobre os salários e os retornos da educação não são formados a partir de informações correntes. Os indivíduos antecipam o tamanho dos ganhos com o acumulo de educação, e escolhem o quanto acumularão de capital humano com base nessa informação. Por essa razão, não será viável a estimação dos retornos desse investimento com o uso de dados de secção cruzada.

Outra conclusão importante do trabalho de Heckman, Lochner e Todd (2003) é que dados longitudinais se adequam melhor para a determinação dos retornos salariais da educação em países em desenvolvimento. Nesse sentido, o trabalho de Mora e Muro (2007) utiliza uma base de dados em *pseudo-painel* para esse fim. Esse trabalho busca determinar o diferencial salarial, para a Colômbia entre 1996 e 2000, com a conclusão do ensino médio e do ensino superior. Os autores partem da hipótese de que há um prêmio salarial ao diploma. Os resultados mostram que obter o diploma de ensino médio eleva o salário em 14%, já o título de graduado eleva o salário em 26%.

Warunsiri e McNown (2010), por sua vez, utilizam o mesmo tipo de dado longitudinal para a análise semelhante feita sobre os indivíduos na Tailândia. Seus resultados mostram que entre 1946 e 1967 a taxa de retorno da educação variava entre 14% e 16%, e as mulheres apresentaram taxas superiores aos homens, assim como os indivíduos no meio urbano possuem maior retorno à educação do que aqueles que vivem no meio rural.



Apesar das diferenças nos controles usados para a determinação da taxa de retorno da escolaridade, a maior parte dos trabalhos que analisam essa relação é baseada na equação de rendimentos do modelo proposto por Mincer (1970). Não será diferente no presente estudo, porém novos controles serão propostos.

## **2.2 RETORNO PARA A EDUCAÇÃO NO BRASIL**

Os trabalhos empíricos sobre ganhos do investimento em educação no Brasil se fundam, principalmente, na análise de dados de corte transversal e na estimação de equações mincerianas. Quanto ao instrumento econométrico, predomina o uso do método de mínimos quadrados ordinários (MQO). Como visto, e mostrado por Heckman, Lochner e Todd (2003), a estimação do retorno salarial da educação segundo essa conformação apresenta vieses de seleção<sup>7</sup> e habilidade. Na literatura nacional esses desvios são contornados, principalmente, pelo uso do método de Heckman e de Variáveis Instrumentais (VI), respectivamente.

Arbache, Dickerson e Green (2001) propõem uma análise econométrica dos efeitos da educação sobre a renda no Brasil, após a abertura econômica do início da década de 90. O ponto de partida para a análise é a formação de um *pseudo-painel* com os dados da PNAD entre os anos de 1981 e 1999. Os dados individuais são agrupados segundo o ano de nascimento, em grupos divididos em intervalos de dez anos, e também segundo a codificação diferenciadora de setores para as indústrias. Consoante os autores, o uso de dados longitudinais é a maneira escolhida para contornar os vieses associados à equação (6).

Os resultados encontrados por Arbache, Dickerson e Green (2001), confrontam os setores de comércio e de não comércio, no sentido

---

<sup>7</sup> O procedimento proposto por Heckman (1974, 1979) diferencia a decisão de trabalho das mulheres e determina que o viés de seletividade ocorre quando os indivíduos possuem um salário limite (o salário reserva) que determina o valor mínimo pelo qual eles estão dispostos a ofertar o seu trabalho. Assim, não é apenas a demanda por trabalho que determinará o nível de renda, não é apenas a demanda por mão de obra e as características dessa, mas também as expectativas dos trabalhadores. Segundo Heckman (1979), esse viés muitas vezes associado pela omissão de uma variável relevante, na presença de seletividade, pode ser solucionada pela inclusão do princípio de seleção envolvido.

internacional, para os períodos anteriores e posteriores a 1992. Para evidenciar a qualidade da estimação com dados longitudinais, eles comparam os resultados obtidos com os métodos de MQO<sup>8</sup> e efeitos fixo para *pseudo-painel*. Essa comparação mostra que o viés de habilidade superestima os retornos para educação. Os resultados revelam que o setor de comércio sofreu um efeito negativo superior ao de não comércio, com a abertura econômica brasileira, concluída ao final de 1992. Apesar dessa redução, os retornos continuam positivos nos dois períodos.

Sachsida, Loureiro e Mendonça (2004) também fazem uso de uma análise longitudinal para determinar o retorno do investimento em educação, comparando essa forma de controle de vieses com as metodologias propostas por Heckman (1974), Garen<sup>9</sup> (1984) e MQO, sendo esse último método o controle da qualidade dos estimadores obtidos nas demais. Nesse estudo, os autores mostram que o viés de seletividade é característico no curto prazo. Já em uma análise longitudinal não se observa esse tipo de problema. O modelo de Garen (1984) mostra que existe uma quantidade crítica adquirida de anos de estudo, a partir da qual a disposição em investir e arcar com os custos de um ano adicional de estudo se eleva, pois o retorno associado à educação é crescente nessa faixa.

Segundo os resultados encontrados por Sachsida, Loureiro e Mendonça (2004), partindo do ponto crítico de 12 anos de escolaridade, percebe-se a reversão do incentivo (de negativo para positivo) em se investir em educação. Esse resultado é favorável às políticas públicas de incentivo a escolaridade para anos inferiores a esse limiar, pois não há incentivo privado para esse investimento. Esses resultados indicam que a estimação por efeitos fixos para dados de *pseudo-painel* e a inclusão de controles para os níveis educacionais são formas adequadas de controlar os vieses da estimação do retorno salarial da escolaridade por MQO.

Outrossim, a comparação dos modelos de *cross section* com o *pseudo-painel* mostra que não há ganhos no uso desse tipo de dado. Porém, esse

---

<sup>8</sup> Que, como mostrado na seção anterior, é viesado.

<sup>9</sup> Para superar o problema de endogeneidade da escolaridade, Garen (1984) propõe que ao determinar o nível ótimo de escolaridade que o indivíduo escolherá é preciso levar em consideração a característica dessa variável, ou seja, trata-se de uma variável contínua e ordenada.

estudo, assim como o Arbache, Dickerson e Green (2001), divide os grupos homogêneos em *coortes* definidas por período de nascimento com intervalos de 10 anos, sendo a amostra tomada ano a ano nas PNADs de 1992 e 1999. Entretanto, como mostra Ryder (1965), as *coortes* de nascimento devem ser definidas como sequências de intervalos igualmente distribuídas em relação aos períodos nos quais o fenômeno em análise é considerado.

Assim é possível que, ao considerar a formação das *coorte* com intervalos de tempo diferentes dos intervalos da amostra, o agrupamento dos indivíduos tenha afetado os resultados, dado que a escolaridade em dez anos é bastante distinta, principalmente entre as *coortes* nos extremos da distribuição etária.

Outro estudo que utiliza a formação de um *pseudo-painel*, formado por dados da Pesquisa Nacional por Amostra de Domicílios, para sua análise é o trabalho de França, Gasparini e Loureiro (2005). Quando analisam os retornos à escolaridade para a década de 1990, tendo em vista a qualidade do ensino no Brasil, encontraram resultados em que o déficit de qualidade está diretamente associado à estagnação dos retornos do capital humano. No mesmo período, observou-se uma expansão da oferta de vagas nas escolas brasileiras, que podem ser a justificativa dos retornos decrescentes encontrados anteriormente. O resultado do retorno para a escolaridade encontrado por eles, no modelo com mais controles, foi de 10,7% de aumento da renda (medida como logaritmo natural do salário/hora) por ano adicional de educação.

O trabalho de Resende e Wyllie (2006) inova ao usar dados referentes à Pesquisa sobre o Padrão de Vida (PPV) para o ano de 1996/97. Suas análises se preocupam principalmente com o viés de seleção, e seus resultados encontram retorno 12,6% para mulheres e 15,9% para homens a cada ano a mais de estudo. Esses autores ressaltam a presença de discriminação de gênero no país.

Outra maneira de determinar o retorno ao investimento em educação é a taxa interna de retorno, que é obtida como a diferença entre rendimentos e custos do processo de escolarização. Dois trabalhos foram desenvolvidos para o Brasil na última década sobre esse foco.

O primeiro deles é o estudo de Moura (2008), no qual o autor analisa o período de 1992 a 2004, por meio de regressões de MQO para dados de *cross section*, e encontra um elevado retorno para escolaridade, e um crescimento dos retornos para o ensino superior. Nessa mesma linha de análise, há o trabalho de Barbosa Filho e Pessoa (2008), que defendem a existência apenas de efeitos positivos para a educação sobre a renda. Eles estimam um retorno superior a 15% para pré-escola.

Em termos de análise por unidade de federação, Cangussu, Salvato e Nakabashi (2010) determinam o efeito da escolaridade média sobre o Produto Interno bruto (PIB) dos estados brasileiros. Em sua análise, os autores comparam os resultados segundo a modelagem de Mincer e Solow, encontrando um efeito marginal para a educação sobre o PIB estadual superior a 15%.

Ampliando a análise em termos regionais, Santos e Camillo (2011) avaliam o impacto do perfil do sistema educacional sobre o mercado de trabalho no Brasil. Esse trabalho é inovador em relação aos seus antecessores por analisar a correlação entre mercado de trabalho e a educação via análise fatorial que, segundo eles, permite medir “a interdependência entre todas as variáveis”. Os autores propõem diferentes formas funcionais para a equação de rendimentos, porém sempre utilizando os dois fatores construídos por eles a partir das variáveis explicativas selecionadas.

O primeiro desses fatores reúne as características de capital humano, como escolaridade, produtividade, analfabetismo funcional e qualidade da educação básica, e explica 72,7% das variações da renda. Já o segundo fator se relaciona mais à qualidade da educação média e básica, e também ao tempo de experiência do indivíduo, e, quando somado ao primeiro, eleva o poder explicativo para 87,12%. Como se nota, o resultado de Santos e Camillo (2011) sugere uma importância maior para a qualidade da educação e a quantidade da mesma sobre a renda salarial. Esse trabalho apresenta também as regressões a cada ano da amostra para a equação de rendimentos minceriana. Os retornos à escolaridade determinados dessa maneira, a cada ano, foram: 1,33 em 2006; 2,59 em 2007; 2,87 em 2008, e 2,27 em 2009.

Já Pereira et al. (2013) compara os efeitos diferenciados da educação sobre as áreas urbana e rural no Brasil a partir de uma análise de *cross section* com dados de 2006. Esse trabalho controla o viés de seletividade por meio do procedimento proposto por Heckman (1979). Seus resultados apontam efeitos positivos nos dois meios. Para os moradores do meio urbano, o retorno salarial da educação alcança 13,73% nesse trabalho, enquanto no meio rural esse retorno é de 7,63%. Esses resultados parecem exaltar as diferenças tanto do mercado de trabalho quanto do setor educacional entre essas áreas censitárias.

Entre os trabalhos mais recentes, há ainda a análise feita por Teixeira e Menezes-Filho (2012) do efeito da lei 5692, de 1971, que aglutinou os quatro anos primários com os quatro anos do ginásio e retificou o ensino de primeiro grau com oito anos como obrigatório. O instrumento de leis compulsórias de educação são também usados na literatura internacional como instrumento de controle da endogeneidade dessa variável (ACEMOGLU, ANGRIST, 1999; BLÖNDAL, FIELD, GIROUARD, 2002).

Os resultados brasileiros para esses instrumentos são coerentes. Ou seja, os instrumentos controlaram de maneira satisfatórias os vieses de omissão associados à equação de rendimentos de Mincer. Prova disso é a redução progressiva da medida estimada para o retorno da escolaridade ao passo que são incluídos os controles no modelo. O modelo com controle para as mais diversas fontes de viés apresentou um retorno de 5% para um ano adicional de educação sobre os salários, contra 11% do modelo com menor quantidade de controles proposto.

Diante do exposto, percebe-se que os anos mais recentes não foram incluídos nas medidas disponíveis para o retorno desse investimento em educação no Brasil. Tendo em vista que os últimos anos foram marcados por um crescimento sustentado dos investimentos públicos, considera-se como necessário que se mensure novamente tal efeito. Portanto, espera-se poder elucidar resultados que orientem a formulação de políticas públicas para o setor educacional. Um dos efeitos possíveis da educação, entre outros, e que propicia uma indicação favorável ao investimento já feito e à elevação do nível

revertido futuramente, é a melhoria da distribuição de renda entre os grupos educacionais.

Outra contribuição proposta neste estudo é a análise dos efeitos desses investimentos baseando-se em dados longitudinais. A inovação proposta é, então, analisar o efeito de longo prazo do investimento em capital humano, que só é possível com dados dessa natureza. Pretende-se assim, contribuir para o entendimento de fenômenos sociais e econômicos de maneira continuada, não permitida com análises feitas a partir de informações *cross section*, dado que países em desenvolvimento, como o Brasil, ainda não consolidaram seus sistemas de ensino ou ainda não atingiram a estabilidade econômica esperada, para que sejam satisfatoriamente avaliados no curto prazo (ARBACHE, DICKERSON E GREEN, 2001).

É possível ainda que os dados longitudinais sejam mais eficientes quando as políticas em educação ainda não atingiram plenamente seus resultados, dado seu caráter cíclico com efeitos de longo prazo.

Nesse sentido, uma análise temporal a partir da construção de uma *pseudo-painel*, que considere um período recente de mudanças na forma como o investimento é feito, a partir dos dados da última década das PNADs, é a proposta do presente trabalho.

### 3 REFERÊNCIAL TEÓRICO

A Teoria do Capital Humano (TCH) tem sido a base das análises que buscam determinar como a aquisição de conhecimento, o aperfeiçoamento de habilidades inatas e o desenvolvimento das capacidades individuais afetam a renda, tanto no nível do indivíduo, quanto na economia do país. Essa teoria tem como trabalho seminais os estudos de Schultz (1961), Becker (1962) e Mincer (1970).

Essa corrente de pensamento econômico buscou inicialmente justificar porque o produto agregado, de países como os Estados Unidos, pode ser superior à disponibilidade de mão de obra, de terra e de capital físico para seus processos produtivos (SCHULTZ, 1961).

Para Mincer (1981), no sentido macroeconômico, o estoque de capital humano de um país explica grande parte do crescimento econômico registrado. No que tange ao indivíduo, a acumulação particular de capital humano é a fonte de grande parte da variação de renda observada e também da forma como essa renda se distribui entre os indivíduos.

Schultz (1960) destaca que o capital humano é um tipo diferente do capital físico, entre outros motivos, porque não pode ser transferido, vendido ou tratado como propriedade, no sentido de bem de mercado. O conhecimento do homem pode ser tratado como uma forma de capital quando o gasto em sua aquisição é transformado em fonte de renda e produtividade na economia. O mesmo autor afirma ainda que “os trabalhadores tornaram-se capitalistas (...) pois obtiveram conhecimento e habilidades com valor econômico” (SCHULTZ, 1961).

De acordo com Mincer (1981), as análises acerca do retorno associado ao investimento no homem devem se ocupar com as habilidades adquiridas, que são desenvolvidas de maneiras diversas por diferentes meios. Tudo que eleva o bem estar e, por consequência, a produtividade dos indivíduos deve ser visto, além de uma forma de consumo, como um tipo de investimento. Ou seja, investindo em si, as pessoas se colocam diante de uma variedade maior de oportunidades, com melhores perspectivas.

Porém, é na educação formal e na experiência<sup>10</sup>, que os estudos sobre os rendimentos salariais da acumulação de capital humano focam suas análises. Essa escolha está justificada na literatura por sua tangibilidade e por haver uma correspondência direta entre os diferenciais salariais e os diversos níveis de escolaridade alcançados (SCHULTZ, 1961).

O capital humano também sofre depreciação, porém a forma como se dá essa perda de valor é particular. Mincer (1981) atenta para a perda de valor do capital humano na forma de redução da capacidade física, que limita a capacidade produtiva do homem associada à idade, e da obsolescência do conhecimento adquirido no decorrer do tempo. Com o avanço tecnológico, o capital humano precisa ser renovado, no entanto, o custo dessa renovação para os trabalhadores mais velhos tende a ser maior do que para os mais jovens, que dispõem de um horizonte mais extenso para desfrutarem dos ganhos de se atualizarem (BECKER, 1962).

Entretanto, os indivíduos têm maior incentivo à investirem na agregação de conhecimento, que valorize sua força de trabalho, da infância até o início da vida ativa laboral. Dado que esse processo reduz a acumulação de renda – seja por menos tempo disponível para o trabalho, seja pela falta no mercado de trabalho –, e sabendo que a idade na qual o indivíduo sai do mercado é institucionalizada na maioria das vezes, seu horizonte de rendimentos é preestabelecido. Quanto menor o tempo de retorno para o investimento em conhecimento, menor a disposição dos indivíduos em investir na renovação desse capital. Esse comportamento faz com que a função que define a relação entre salário e experiência (ou idade) seja linear de segundo grau, com ponto de máximo<sup>11</sup>.

Por isso, observa-se nas séries históricas que o salário decresce em relação à idade, fixada a educação. Entretanto, quando a escolaridade de um indivíduo é maior que a de outro, o decrescimento da renda ao longo da vida para o primeiro tende a ser menos acelerado.

---

<sup>10</sup> Definida por Mincer (1974) como sendo os anos trabalhados, e considerando que durante o período escolar o indivíduo não trabalhe. Formalmente, considerando um indivíduo em idade economicamente ativa, com  $n$  anos de vida, que iniciou seus estudos aos seis anos de idade e estudou durante  $k$  anos, e atuando no mercado de trabalho, possuirá a experiência  $e$  dada por:  $e = n - k - 6$ .

<sup>11</sup> Formato de “U” invertido.



Análise de Becker (1962) indica que, durante o tempo de trabalho, os indivíduos aprimoram suas capacidades, podendo ser treinados de maneira ampla, absorvendo um conhecimento aplicável em função. Adquirindo habilidades específicas que os preparam para funções consoantes ao conhecimento apreendido. As habilidades genéricas são como inovações tecnológicas, que elevam a produtividade da economia como um todo, e causam uma elevação dos salários dos indivíduos que recebem esse treinamento. As aptidões específicas para uma determinada função, por sua vez, diferenciam o indivíduo dos demais, elevam a produção do empregador que o contrata, assim como seus salários.

Para mensurar o benefício líquido da acumulação de capital humano, há a necessidade de definir os custos desse investimento. Conforme Schultz (1960, 1961), o custo de educar-se não se restringe ao gasto com a aquisição de conhecimento, mas se apresenta, principalmente, como o custo da oportunidade de estar na escola em vez de estar no mercado de trabalho.

O custo de oportunidade se refere à necessidade de se ausentar da função laboral. Outro custo do investidor nesse processo é no sentido contábil, ou seja, o preço que o capital humano demanda. Esses conceitos são amplamente aceitos e usados em diversos trabalhos que abordam economicamente a Teoria do Capital Humano. Apesar de realista, a abordagem do custo segundo essa ótica é bastante intangível, o que predomina na literatura é, então, a determinação do retorno não pela diferença entre ganho e custo, mas pelo diferencial salarial relativo ao diferencial de capital humano acumulado.

Além dos trabalhos seminais, que delinearam o escopo da Teoria do Capital Humano, outros autores propuseram análises ou abordagens metodológicas diferentes. Spence (1973) propõe uma nova teoria, a Teoria da Sinalização, que contrapõe as conclusões da Teoria do Capital Humano. A ideia desse autor é a de que a escolaridade formal, nos diversos níveis, não torna de fato o indivíduo mais produtivo. A teoria dele sugere que o conhecimento formalizado no sistema educacional não fornece o conhecimento ou as habilidades demandadas pelo mercado de trabalho e capazes de elevar a produtividade da mão de obra.

Entretanto, Spence (1973) nega a ocorrência de melhor remuneração para aquele indivíduo com mais anos de estudo. Na Teoria da Sinalização, o que a educação formal sinaliza para o mercado é a capacidade produtiva que o indivíduo pode ter, mas não garante que de fato ele a terá. Assim, o indivíduo com diploma indica para o mercado que é capacitado para a função pretendida, sem que haja qualquer garantia de que ele de fato o seja.

A Teoria da Sinalização foi confrontada e refutada por muitos autores, no entanto, ela fornece uma contribuição relevante na análise do retorno à educação. A contribuição mais evidente dada por Spence (1973) é a concepção de que a compensação salarial para o indivíduo que obtém o diploma de um dado nível de escolaridade é tal que um ano a mais de educação possui remuneração diferenciada se esse ano adicional diz respeito ao fim de um ciclo educacional. Desde a proposta teórica de Spence (1973), tem sido crescente a inclusão de variáveis referentes à obtenção de diploma nas análises do retorno à educação.

Outro autor relevante e com contribuições amplamente reconhecidas para as pesquisas em Teoria do Capital Humano, Heckman (1974) propôs mudanças importantes na forma de estimar formalmente o retorno sobre os salários dos investimentos, especificamente, em educação. Esse autor observou que o processo de escolha do indivíduo – tanto quanto à quantidade de capital humano acumulado como quanto à colocação no mercado de trabalho – é econômico, no qual esse age racionalmente para maximizar o retorno futuro do investimento em educação. Por essa razão, a desconsideração do processo de escolha, das diferenças entre as decisões individuais, torna as estimativas do retorno para a educação viesadas.

Ao analisar o comportamento específico das mulheres, esse autor mostrou ainda que os indivíduos possuem um salário alvo, como um preço que limita a sua disponibilidade ao trabalho. Sendo assim, estar ou não empregado pode refletir não apenas o nível de habilidades desenvolvidas por uma pessoa, mas está diretamente ligado às expectativas salariais. Desse modo, Heckman (1974) propõe que a mensuração dos retornos salariais do investimento em educação ou treinamento deve considerar o processo de escolha anterior ao momento de busca por emprego e, a partir desse

momento, pois, caso contrário, as estimativas serão viesadas por não considerarem a escolha do indivíduo.

Mais um autor que contribuiu para expandir a relevância da Teoria do Capital Humano na economia foi Lucas (1988), que reformulou as teorias do crescimento exógeno preexistentes, como a proposta por Solow (1956), determinando um novo modelo que considere a capacidade humana como parte fundamental do processo de crescimento econômico de longo prazo. A partir de então, a economia parte de um modelo de crescimento endógeno, que considera o homem e suas habilidades como fundamental para esse crescimento, já que seu impacto é permanente (BARBOSA FILHO e PÊSSOA, 2010).

Apesar de todo o arcabouço teórico de Schultz, Becker, Spence, Heckman e Lucas, é a formalização de Mincer (1974) para a determinação dos retornos salariais do investimento em capital humano que guia os trabalhos sobre esse tema até a atualidade. Essa formalização será apresentada na seção seguinte que trata do método que dá sustentação ao presente estudo.

## 4 METODOLOGIA

Este estudo dissertativo busca determinar os retornos salariais da educação para o Brasil no período bianual compreendido entre 1999 e 2011, e fará uso da formalização proposta por Mincer (1974). A equação de rendimentos e o parâmetro de retornos salariais para a educação propostos por esse autor são a base para os resultados aqui trazidos. O avanço sobre os trabalhos descritos anteriormente está na utilização de dados longitudinais, a partir do agrupamento dos indivíduos por características homogêneas, para a determinação dos retornos salariais para a educação e para a experiência no Brasil, no período bianual iniciado em 1999.

Partindo de um modelo de Efeitos Fixos para dados em *pseudo-painel*, e contrapondo o resultado desses aos alcançados por Mínimos Quadrados Ordinário (MQO). Esta investigação será semelhante à proposta em Heckman, Lochner e Todd (2003), Mora e Muro (2007) e Warunsiri e McNown (2010). A base da análise do estudo realizado neste documento foi desenvolvida segundo os trabalhos de Deaton (1985), Moffitt (1993), Collado (1997), Mckenzie (2004) e Verbeck e Vella (2005).

### 4.1 MODELO MINCERIANO

Dentre os autores seminais da Teoria do Capital Humano, Mincer (1958, 1970, 1974) é o primeiro a propor uma medida formal para diferencial de renda que se verifica entre indivíduos que se distinguem por sua escolaridade e experiência<sup>12</sup>. A hipótese inicial desse modelo é a de que os indivíduos são semelhantes em suas capacidades e se deparam com as mesmas oportunidades ao longo da vida para se educarem e se inserirem no mercado de trabalho.

Partindo dessa hipótese, o autor propõe uma equação de rendimentos em que o ganho salarial é explicado pelo capital humano, acumulado em educação e experiência, tudo mais mantido constante. O fundamento para determinação dessa equação é o comportamento maximizador de utilidade do indivíduo racional. Diante de um *tradeoff* entre tempo no trabalho, gerador de

---

<sup>12</sup> Formas de capital humano usados pelo autor.

renda, e tempo dispendido com a aquisição de capital humano, que gera despesa no presente, o indivíduo maximizador determina intertemporalmente o nível ótimo de tempo gasto em cada uma dessas atividades, de modo que ao longo da vida a renda do trabalho seja a maior possível.

Sem perda de generalidade, o modelo simplificado de Mincer (1970) supõe que o custo de se educar<sup>13</sup> depende apenas do tempo gasto com educação e da consequente anulação de retorno salarial associado aos anos fora do mercado de trabalho (mensalidade nula, custo de materiais nulo ou negligenciável etc). Sendo a taxa de desconto dos rendimentos futuros definida por  $r$  e os anos de estudo determinados por  $s$ , o custo ( $C_s$ ) da educação será dado por:

$$C_s = a_s \int_0^l (e^{-rt}) dt = \frac{a_s}{r} (e^{-rs} - e^{-rl}) \quad (1)$$

em que  $a_s$  é o ganho anual de um indivíduo com  $s$  anos de estudo;  $l$  o tempo de vida profissional do indivíduo; e  $t$  número de anos do indivíduo no período em análise.

A partir da equação (1), é possível determinar a diferença nos rendimentos ( $E_s$ ) futuros de dois indivíduos, sendo que um possui  $d$  anos a mais de estudo do que o outro, como se segue:

$$k_{s,s-d} = \frac{E_s}{E_{s-d}} = \frac{e^{r(l+d-s)} - 1}{e^{r(l-s)} - 1} \quad (2)$$

A equação (2) fornece a proporção entre os rendimentos de indivíduos que se diferenciam por  $d$  anos de estudo,  $k_{s,s-d}$ . Aplicando o logaritmo natural à equação (2), obtém-se a relação entre os rendimentos como abaixo:

$$\ln(E_s) = \ln(E_{s-d}) + rd \quad (3)$$

Da equação (2) obtém-se que a inserção no mercado de trabalho e a renda futura serão determinadas pelo nível educacional que, por hipótese, é uma escolha individual. A decisão de se educar é vista assim como uma decisão econômica, pois determinará o fluxo de renda futura do indivíduo.

---

<sup>13</sup> No modelo proposto por Mincer (1970) a acumulação de capital humano se dá por duas vias: a educação formal e a experiência medida como tempo de trabalho. Por essa razão, considerar-se-á como sinônimos: acumulação de capital humano - educação e experiência. Além disso, treinamento torna-se sinônimo de experiência.

Apesar da relevância da equação de rendimentos minceriana, as equações (1), (2) e (3) restringem a análise a indivíduos com habilidades inatas e oportunidades semelhantes. Entretanto, é possível expandir essa análise de modo que seja considerada a heterogeneidade que se verifica entre os agentes. Outra fonte de diferenciação não considerada até o momento é a experiência no trabalho. Como o próprio Mincer (1958) destaca, “a experiência no trabalho é, frequentemente, a parte mais importante do processo de aprendizado” (MINCER, 1958 p.287).

Uma análise menos restrita é desenvolvida pelo mesmo autor e agrega as conclusões de diversos autores, com destaque para o trabalho de Becker (1962). As funções gerais dos rendimentos anuais bruto e líquido (MINCER, 1970), sendo no último descontados os custos do investimento em capital humano ( $C_{ij}$ ), são apresentadas abaixo, respectivamente:

$$E_{ij} = X_{ij} + \sum_{t=0}^{j-1} r_{ti} C_{ti} \quad (4)$$

$$Y_{ij} = X_{ij} + \sum_{t=0}^{j-1} r_{ti} C_{ti} - C_{ij} \quad (5)$$

em que  $i$  e  $j$  definem o indivíduo e o período de análise respectivamente;  $X_{ij}$  é vetor de características que determinam a renda do indivíduo posterior ao período de treinamento;  $Y_{ij}$  é o rendimento anual líquido;  $r_{ti}$  é a taxa de retorno individual anual para o investimento,  $C_{ti}$ , feito em educação. O somatório  $\sum_{t=0}^{j-1} r_{ti} C_{ti}$  fornece o componente da renda devido à acumulação de capital humano.

O modelo descrito por (3) torna-se assim um caso especial de (4), no qual a taxa de retorno da escolaridade é constante no tempo e entre os indivíduos, os investimentos em educação são restritos aos custos do tempo gasto em se educar e ao fato de serem as capacidades iniciais constantes ao longo do tempo para todos os indivíduos.

Com isso, percebe-se que o modelo descrito em (4) avança com relação ao modelo apresentado em (3), dado que consideram-se variações de retorno no tempo na medida em que os indivíduos tornam-se mais experientes e entre os indivíduos. Além disso, Mincer (1970) mostrou que a taxa de retorno possui variância positiva.

O modelo descrito em (5) é expandido ainda mais por Mincer<sup>14</sup>, permitindo que os retornos associados à educação e à experiência<sup>15</sup> sejam diferenciados no processo de determinação da renda. Essa separação mostra que a relação entre renda e experiência, ou a idade, é não linear, sendo melhor descrita por uma parábola côncava. Essa relação baseia-se na existência de retornos positivos decrescentes para a experiência. O que ocorre é que no início da vida laboral o tempo utilizado para a acumulação de capital com treinamentos tende a ser maior, visando elevar os rendimentos adquiridos ao longo dos anos que seguirão, maximizando o tempo de recebimento desses. Por outro lado, ao final da vida laboral, por idade avançada e possível redução das capacidades físicas do indivíduo, menos horas são dispensadas ao trabalho. Assim, observa-se um rendimento inferior no início e no fim do período da vida dedicado ao trabalho, o que reduz o rendimento nesses extremos.

A renda em função da escolaridade ( $S_i$ ) e da experiência ( $Exp_i$ ) do indivíduo  $i$ , é descrita como se segue:

$$\ln(Y_i) = \ln(y_0) + \beta_1 X_i + \beta_2 S_i + \beta_3 Exp_i + \beta_4 Exp_i^2 + \mu_i \quad (6)$$

Do lado direito da equação (6) tem-se o logaritmo da renda associada ao menor nível de escolaridade,  $\ln(y_0)$ ; características individuais que influenciam a renda como raça e gênero,  $X_i$ ; escolaridade do indivíduo em análise,  $S_i$ ; e a experiência no trabalho em termo linear e quadrático,  $Exp_i$ . A inclusão da experiência em nível e ao quadrado na equação se justifica na relação observada no ciclo de vida do rendimento individual, de modo que o sinal do termo ao quadrado é esperado negativo, refletindo o retorno marginal decrescente para essa variável.

---

<sup>14</sup> Partindo de (4), seja a proporção entre o rendimento bruto e o investimento no ano  $j$  dada por:  $k_j = \frac{E_j}{C_j} \Leftrightarrow C_j = k_j E_j$ . Considerando que no ano  $j$  o investimento em educação é nulo e que a taxa interna de retorno é constante por indivíduo, tem-se de (4) que:  $E_j = E_{j-1} + rC_{j-1} \Leftrightarrow E_j = E_{j-1}(1 + rk_{j-1})$ . Recursivamente, tem-se que:  $E_j = E_0 \prod_{i=0}^{j-1} (1 + r_i k_i)$ . Assumindo  $k_i \leq 1$  e  $r$  pequeno, tem-se que:  $\ln E_j = \ln E_0 + \sum_{i=0}^{j-1} (1 + r_i k_i)$ . Sendo  $k_j = 1$  durante os  $s$  anos escolares, tem-se:  $\ln E_j = \ln E_0 + rs + \sum_{i=s+1}^{j-1} (1 + r_i k_i)$ . Seja  $P_j = + \sum_{i=s+1}^{j-1} k_i$  as experiência pós-escolar e assumindo  $r_j$  constante para todo  $j$ , então tem-se (6) (MINCER, 1970).

<sup>15</sup> Medida como idade menos a escolaridade menos seis.

O coeficiente minceriano, derivado dessa relação, é dado pelo parâmetro  $\beta_2$ , que fornece a TIR da educação segundo esse modelo. O termo de erro  $\mu_i$  é pressuposto aleatório e não correlacionado com as variáveis explicativas do modelo segundo essa abordagem.

A obtenção de  $\beta_2$ , segundo o método dos mínimos quadrados ordinários (MQO)<sup>16</sup>, ainda que positivo e estatisticamente significativo, não garante a ocorrência de causalidade entre a escolaridade e a renda individual. As principais críticas ao parâmetro de Mincer recaem sobre as pressuposições de homogeneidade entre os indivíduos (quanto às habilidades, oportunidades e quanto ao meio). Havendo habilidades não estimadas e seletividade na determinação da escolaridade por parte dos indivíduos, o parâmetro em questão é viesado quando estimado segundo o MQO.

Os vieses que comprometem a estimação de (6) por MQO se devem à própria origem teórica do parâmetro  $\beta_2$ . O modelo proposto por Mincer (1970) deriva diretamente da Teoria do Capital Humano e de suas pressuposições, e, como a própria teoria prevê, os indivíduos mais habilidosos investirão mais para a aquisição de capital humano. Com isso, a renda será heterogênea não apenas por diferenciação nos níveis de escolaridade, mas por haver diferenças de habilidade intrínseca. Nesse caso, não é aleatória a população de indivíduos com mais educação, estes são selecionados segundo suas habilidades.

Apesar das diferenças em termos de oportunidades, os países desenvolvidos também apresentam viés de omissão da habilidade quando a estimação do parâmetro de Mincer por MQO não é tratada. Griliches (1977) mostrou que o viés associado à habilidade para o parâmetro  $\beta_2$  estimado sobre os dados norte-americanos subestima o retorno à escolaridade. Simplificando a determinação desse viés, propõe-se supor que a verdadeira equação de determinação da renda seja dada como se segue:

$$\ln(Y_i) = \ln(y_0) + \alpha_1 X_i + \alpha_2 S_i + \gamma H_i + \mu_i \quad (7)$$

---

<sup>16</sup> Como proposto por Mincer e seguido por diversos autores, como Leal e Werlang (1991).



em que, por simplificação, omitiu-se a experiência, e  $H$  mede a habilidade do indivíduo. Se ao invés de estimar (7) estima-se:

$$\ln(Y_i) = \ln(y_0) + a_1 X_i + b_2 S_i + \vartheta_i \quad (8)$$

então o parâmetro  $b_2$  será viesado na seguinte medida:

$$E(b_2) = \alpha_2 + \gamma \sigma_{HS} = \alpha_2 + \gamma \frac{\text{cov}(S,H)}{\sigma_S} \quad (9)$$

Consequentemente, sendo a habilidade e a escolaridade positivamente correlacionadas e sendo  $\gamma > 0$ , o parâmetro  $\beta_2$  estimado em (6) por MQO será superestimado, segundo o pressuposto de que os indivíduos são homogêneos. Além disso, Heckman, Lochner e Todd (2003) apontam para a deficiência do modelo minceriano ao considerar o paralelismo da experiência; ou seja, na equação (6), indivíduos com a mesma experiência possuem o mesmo retorno a cada ano adicional de trabalho, mesmo que tenham qualificações escolares diferentes. Desse modo, os autores propõem a inclusão de variáveis multiplicativas de escolaridade e de experiência que captem as diferenças no retorno à experiência ao longo do ciclo de vida por nível escolar.

A seletividade também é fonte de crítica ao modelo simplificado em (6), pois quem determina a escolaridade adquirida é o próprio indivíduo, que se baseia nas expectativas de renda futura para cada nível atingido de capacitação e nos custos necessários para obtê-la (CAHUC, 2004). Outra consideração a ser feita pelo indivíduo no ato da escolha é a observação histórica dos diferentes retornos para os diferentes graus de escolaridade. Ou seja, a escolha atual do indivíduo jovem, quanto à sua escolaridade, tende a ser baseada nas experiências passadas, e não somente nos retornos atuais dos grupos formados por indivíduos mais velhos. Com isso, o retorno à escolaridade deixa de ser constante ao longo dos anos, fazendo com que a análise com dados em *cross section* sejam menos eficientes do que aquela feita a partir de dados longitudinais.

O viés de seleção é ainda justificado segundo a Teoria da Sinalização de Spence (1973). Essa teoria postula que, além do mercado selecionar os indivíduos pelos sinais que o nível educacional fornece, a capacidade de se educar é determinada, entre outras coisas, pelas habilidades natas e

características próprias dos indivíduos. Sendo assim, mais do que um mecanismo usado pelo empregador para selecionar seus empregados, a educação pode ser um sinal da habilidade intrínseca. Se os indivíduos mais habilidosos usarem esse mecanismo para obter maior renda, então o investimento público em educação perde sua função social, pois não haverá ganho compartilhado com esse gasto do governo, apenas o indivíduo com mais escolaridade se beneficia do investimento público. A concepção teórica proposta por Spence (1973) foi amplamente criticada, como mostram Barbosa Filho e Pêssoa (2010). No entanto, para países em desenvolvimento essa concepção é relevante, à medida que a informação é, mais frequentemente, transacionada de forma imperfeita (FERREIRA, 2004).

Partindo da formulação básica como proposto na equação (6), a literatura<sup>17</sup> sugere que o controle da habilidade seja feito pelo uso do método de variáveis instrumentais, sendo esse capaz de controlar, além da habilidade, também a endogeneidade na escolha dos anos de estudo, dada a relação existente entre essas variáveis.

Como mostra o trabalho de Barbosa Filho e Pessôa (2010), diversas tentativas foram feitas, e muitos instrumentos testados, na esperança de controlar o viés de habilidade. No entanto, como ressalta Card (2001), a busca por um instrumento, que respeite o critério de não possuir correlação com o termo de erro da regressão a ser estimada e ser alta a correlação com a educação, tem sido infrutífera.

Esse autor indica que uma maneira mais apropriada de controlar esse viés está no uso de dados longitudinais. Por essa razão, o trabalho aqui produzido desenvolve suas análises a partir dessa proposição, especificamente usando dados de *pseudo-painel*. Segundo Moffit (1993), por ser a unidade em análise no *pseudo-painel* a *coorte* definida por características que tornam os indivíduos homogêneos dentro do grupo, o método de estimação com dados desse tipo é visto como um método de variáveis instrumentais, no qual o grupo acompanhado ao longo dos anos é instrumento de controle de si mesmo.

---

<sup>17</sup> Griliches (1977), Acemoglu e Angrist (1999), Card (2001), Moretti (2004, 2006).

Assim, a próxima seção apresenta a forma como os dados devem ser trabalhados para que sejam utilizados e gerem resultados coerente com o modelo teórico aqui proposto. São apresentadas também as condições para que a base de dados aqui formulada possa ter seus resultados avaliados como os oriundos de um verdadeiro painel.

#### **4.2 MODELO DE *PSEUDO-PAINEL***

Como foi explicitada, a equação de rendimentos minceriana deve ser usada com cautela e, para que não sofram com viés e sejam válidos, seus resultados demandam que as devidas correções sejam aplicadas ao modelo. Uma medida corretiva amplamente utilizada é o método de variáveis instrumentais. Porém, como ressaltam Barbosa Filho e Pêssoa (2010), esses instrumentos têm sido constantemente criticados e refutados.

Considerando a educação como um tratamento recebido por uma parcela da população e que a divide entre indivíduos tratados e não tratados, Angrist e Imbens (1995) sugerem que a análise por *VI* de efeitos de tratamento sobre os indivíduos determina somente o efeito sobre o grupo considerado pela especificação do instrumento. Assim, os resultados obtidos com variáveis instrumentais não devem, segundo esses autores, serem considerados como estimativas populacionais, mas, sim, estimativas para um subgrupo específico da população.

Sachsida, Loureiro e Mendonça (2004) e Moretti (2004), sugerem que, sendo a habilidade uma característica constante a cada indivíduo, o uso de dados longitudinais e resultados estimados pelo método de efeitos fixos individuais seriam suficientes para controlar o viés de omissão, sem a necessidade de adoção de um instrumento, já que esse poderá ser questionado quanto à sua eficiência.

O mesmo é verificado por Card (2001) que, ao testar a presença de viés para o parâmetro de Mincer, tanto utilizando MQO quanto *VI*, para os Estados Unidos, aferiu que o retorno estimado no primeiro desses modelos é superior ao verdadeiro retorno para educação. Já o parâmetro obtido sobre o método de variáveis instrumentais é ainda maior. Assim, a comparação entre a estimação por MQO, sabidamente viesada, e por *VI*, que está sendo testada

quanto à sua eficiência, leva o autor a concluir que o instrumento pode causar viés superior ao uso de mínimos quadrados nesse tipo de análise.

Além disso, segundo Oliveira (2002), o uso de dados longitudinais em análises econômicas é importante, pois esse tipo de dado traz para discussão proposta fatores e fenômenos econômicos que são homogêneos em nível do indivíduo, e que, no entanto, não são captados em análises feitas a partir de dados de seção cruzada, onde as unidades de interesse são heterogêneas. Deaton (1985) ressalta ainda que outra vantagem de seguir, no tempo, os mesmos indivíduos é que eles podem ser usados como seus próprios controles, podendo seus comportamentos serem avaliados antes e depois de um determinado acontecimento, e permitindo ao pesquisador avaliar, portanto, o efeito desse acontecimento sobre o comportamento da pessoa.

Todavia, no Brasil apenas a Pesquisa Mensal de Emprego (PME), promovida pelo IBGE, segue o caráter longitudinal, mas não fornece informações suficientes para uma análise de longo prazo, dado que o mesmo indivíduo é acompanhado durante o horizonte temporal de um ano. Além disso, a PME não fornece informação a todos os estados do país, apenas representa a dinâmica metropolitana de seis deles: Recife, Salvador, Belo Horizonte, São Paulo, Rio de Janeiro e Porto Alegre.

Outra fonte de informações que faz um acompanhamento longitudinal no Brasil é a Relação Anual de Informações Sociais (RAIS), que tem por objetivo gerar dados anuais sobre a atividade de empregados e empregadores formais, privado e públicos, no país. No entanto, assim como a PME, trata-se de uma amostra limitada. Essa base de dados não abrange os indivíduos empregados no setor informal, excluindo a parcela da população economicamente ativa e ocupada que está empregada no setor informal no país. A importância relativa dessa parcela da economia brasileira não pode ser negligenciada, dado que em 2011 o percentual do PIB produzido informalmente no país foi de 17% (ETCO e IBRE/FGV, 2013).

Apesar de haver argumentos favoráveis ao uso de painéis verdadeiros, como a PME e a RAIS, em detrimento de séries de dados de seção cruzada ou *pseudo-painéis*, o uso de um painel real pode acarretar alguns problemas à pesquisa. As questões mais debatidas na literatura são a atribuição não

aleatória, o condicionamento, o fato de painéis verdadeiros não acompanharem a composição populacional e não considerarem mudanças sociais que alteram a interpretação de termos e palavras das pesquisas (OLIVEIRA, 2002).

Especificamente, a *atrição* significa que, de uma pesquisa à outra, alguns indivíduos podem ser perdidos da amostra, seja por morte ou por não se disporem a participar da pesquisa novamente. Essa perda de observações gera no painel um desbalanceamento, o que causará viés na estimação (OLIVEIRA, 2002).

Já o condicionamento é a mudança no comportamento do indivíduo no ato de responder à pesquisa pelo simples fato de já tê-lo feito no período anterior, o que pode gerar dependência de um período ao outro. Como dados em painel consideram os mesmos indivíduos em cada período, esses dados não são capazes de considerar a entrada e saída de diferentes *coortes* de nascimento na população. Com isso, a estrutura etária pode se alterar ao longo dos anos sem que a amostra analisada seja refeita, o que pode reduzir sua representatividade. (OLIVEIRA, 2002).

O último desses problemas diz respeito à mudança de conceitos relacionados aos termos da pesquisa dentro da população analisada. Oliveira (2002) exemplifica esse problema com a classificação de raças e as similaridades e diferenças de interpretações de palavras como “negro” e “preto” que podem gerar erros de medidas em painéis.

Segundo Moffitt (1993), mesmo quando se tem disponível uma base de dados em painel, os dados de *cross sections* analisados ao longo do tempo podem ser mais completos. Além disso, ainda que os problemas acima mencionados não ocorram, por sua complexidade, dados em painel configuram, normalmente, uma base de dados na qual poucos indivíduos são observados.

Diante do exposto, pretende-se neste trabalho agregar os dados de sete biênios de amostras de *cross sections* das PNADs<sup>18</sup> de 1999 a 2011,

---

<sup>18</sup> O delineamento da amostra do PNAD como uma amostra complexa altera as análises, porém só pode ser considerado quando se utilizam os dados domiciliares, uma vez que as variáveis da amostra que caracterizam-na como tal estão presentes apenas nos dados domiciliares. Como no presente estudo apenas os dados individuais são considerados, a amostra não é tratada como complexa.

considerando intervalos de dois anos entre um período e outro da amostra, para formar uma série de dados de secção cruzada repetida, ou um *pseudo-painel*, que apresente características tão adequadas para a análise proposta quanto um painel real.

Os dados assim ordenados possuem características que, sobre a luz de alguns pressupostos, os tornam tão adequados quanto um painel verdadeiro, por isso serão usados para contornar o viés de habilidade para a estimação dos retornos salariais da educação (DEATON, 1985). Esse procedimento coincide com o proposto por Arbach, Dickerson e Green (2001); Sachsida, Loureiro e Mendonça (2004) e outros.

Esse método para delimitação da amostra foi primeiramente proposto e aplicado por Deaton (1985). Assim como seus sucessores, esse autor avulta, como uma característica favorável ao uso dos dados em *pseudo-painel*, que esses dados têm a capacidade de avaliar tanto fatores agregados dos dados, como fatores microeconômicos. Isso porque esse tipo de dado contém tanto características de dados de secção cruzada, curto prazo, que permite a comparação entre unidades diferentes em cada período de tempo; como de dados longitudinais, longo prazo, que permitem a comparação de uma mesma unidade homogênea ao longo dos anos. Além disso, o ciclo de vida pode ser observado no uso desse tipo de dado, de modo que é possível analisar o efeito do envelhecimento sobre a variável de interesse do estudo (MOFFITT, 1993; COLLADO, 1997; MCKENZIE, 2000; OLIVEIRA, 2002).

A determinação das *coortes* é feita segundo uma ou mais características ou fenômeno que são comuns aos indivíduos que definirão o grupo homogêneo. Na presente pesquisa, são considerados, para isso, o período de nascimento, o gênero, a raça e a região de residência, entre metropolitana e não metropolitana, como características homogêneas de determinação das *coortes*, e que diferenciam uma da outra. É possível que os grupos se definam apenas pelo período de nascimento, entretanto.

Partindo dessa conformação, os grupos formados passam a ser a unidade em análise a ser acompanhada ao longo dos anos. Já as variáveis do modelo são determinadas para cada *coorte*, como sendo a média das variáveis observadas em cada ano para os indivíduos que pertencem àquela

*coorte*. Contudo, as análises dos resultados de dados de *pseudo-painel* são feitas como usual para dados de um painel verdadeiro.

Apesar de suas vantagens, há críticas quanto a possibilidade de erros de medida pela utilização de médias das variáveis em cada *coorte* e em cada ano, sendo que o tamanho das *coorte* não é o mesmo de um ano a outro devido a aleatoriedade da definição da amostra da pesquisa de *cross sections*. Porém, como sugere Deaton (1985) e seus sucessores<sup>19</sup>, se o tamanho da amostra por grupo em cada ano for suficientemente grande, pelo teorema do limite central, erros desse tipo não são relevantes e não viesam os estimadores do modelo.

Mckenzie (2000) parte das mesmas pressuposições e considera que elas validam os parâmetros estimados para esse tipo de dado e garantem as características assintóticas que permitem o uso da lei dos grandes números para o modelo. O autor também chama a atenção para o fato de que, em painéis verdadeiros, a lei dos grandes números e o teorema do limite central são garantidos pela extensão do número de indivíduos sobre o número de períodos que se dispõe para análise. Nos dados em *pseudo-painel*, no entanto, para que essas propriedades se verifiquem, é necessário que se tenham muitas observações por *coorte* em cada período.

Assim, assume-se, visando garantir a consistência dos parâmetros estimados, que o tamanho da *coorte* tende a infinito e o número de *coortes* e períodos é fixo. Essa relação decorre da própria formação dos grupos. Dado que a amostra para determinação dos grupos é definida, se o número de grupos formados é grande, então o tamanho deles será limitado, e vice-versa. Essa consideração difere da feita por Collado (1997), na qual parâmetros consistentes dependerão de um número infinito de grupos no limite, com períodos e indivíduos por *coorte* relativamente menores.

Um fator ressaltado na literatura como sendo legitimador das condições delimitadoras da amostra utilizada nesse trabalho é a de que o tamanho dos grupos é constante ao longo das diferentes amostras de *cross section*. Para contornar essa dificuldade, a literatura sugere que, nesse caso, a ponderação das variáveis de *pseudo-painel* pela raiz quadrada do tamanho da *coorte*, a

---

<sup>19</sup> Para mais detalhes, ver Moffit (1993), Collado (1997), Mckenzie (2000), Oliveira (2002).

cada ano (ARBACHE, DICKERSON e GREEN, 2001; WARUNSIRI e McNOWN, 2010). Ou seja, cada variável presente no modelo final, e na amostra final do *pseudo-painel*, é determinada como a média ponderada pelo tamanho da *coorte*.

O uso de dados de secção cruzada independentes, acompanhados em séries temporais, gera estimadores consistentes desde que alguns pressupostos sejam atendidos. Moffitt (1993) afirma que modelos lineares com efeitos fixos são adequados para estimação com dados em *pseudo-painel* e geram estimadores consistentes e identificados:

- i. Os estimadores serão mais eficientes se as *coortes* forem definidas (diferenciadas) por variáveis constantes no tempo (MOFFITT, 1993);
- ii. Quando o número de períodos analisados é finito, existirá um *trade-off* entre o tamanho das *coortes* e a quantidade de *coortes* a disposição da análise. Nesse sentido, quanto maior o número de *coortes* formadas (tendendo assintoticamente a infinito), a estimação de efeitos fixos para o *pseudo-painel* é factível (COLLADO, 1997);
- iii. Para que o modelo seja identificado, é necessário que as médias das variáveis diversifiquem no tempo e sejam particulares entre diferentes *coortes* (COLLADO, 1997);
- iv. Quanto maior o número de indivíduos em cada *coorte*, em cada período, menor o risco de erros de medida das médias das variáveis, que decorrem do fato de que, a cada avaliação de secção cruzada, os indivíduos em consideração em cada *coorte* não são os mesmos (COLLADO, 1997).

Além das especificidades de cada análise proposta para uma amostra de *pseudo-painel*, por seus objetivos e dados, as médias das variáveis individuais tornam a estimação não eficiente caso não seja possível garantir a estabilidade do tamanho dos grupos a cada ano amostrado. Ou seja, havendo a cada ano um número diferente de indivíduos por grupo faz-se necessária uma ponderação dos grupos segundo os tamanhos dos mesmos para que se garanta a validade dos resultados obtidos da amostra de *pseudo-painel* (DEATON, 1985).



A equação a ser estimada a partir de dados em *pseudo-painel* sofre algumas modificações em seus parâmetros e variáveis, em relação à equação (6). Os indicadores de indivíduos dos termos da equação devem indicar, agora, que estão sendo considerados como unidades de avaliação as *coortes* ( $c$ ) em cada período analisado ( $t$ ). Além disso, as variáveis do modelo são, como dito, as médias ponderadas de cada variável por grupo.

Ou seja, a variável  $x$ , referente ao grupo  $n$ , formado por  $N$  indivíduos homogêneos, é dada em cada ano como a média dos valores de  $x$  para cada indivíduo em relação ao tamanho do grupo, e o fator de ponderação é a raiz quadrada do tamanho do grupo. Essa ponderação garante a existência de erros homoscedásticos no modelo final de regressão. Elas são representadas matematicamente com um traço superior, que indica essa condição. Assim, o modelo de dados em *pseudo-painel*, com efeitos fixos por *coorte* e por período, a ser estimado será:

$$\overline{\log(w_{ct})}^{i(t)} = \beta \overline{X_{ct}}^{i(t)} + \pi_{ct} \overline{P_{ct}}^{i(t)} + \alpha \overline{Z_{ct}}^{i(t)} + d_c + \overline{u_{ct}}^{i(t)} \quad (10)$$

em que o índice  $i(t)$  indica que as médias são calculadas sobre os indivíduos de cada *coorte* ( $c$ ) e  $t$  a unidade de tempo. Além disso,  $X_{ct}$  é o vetor de características médias individuais em cada grupo ou *coorte*.  $P_{ct}$  é a escolaridade,  $Z_{ct}$  é a experiência, ambas tomadas pelas médias dos indivíduos nas *coortes* homogêneas. Mantém-se o termo de erro composto por uma parcela independente e identicamente distribuída ( $\overline{\varepsilon_{ct}}^{i(t)}$ ) e um termo de choques aleatórios ( $v_t$ ):

$$\overline{u_{ct}}^{i(t)} = v_t + \overline{\varepsilon_{ct}}^{i(t)} \quad (11)$$

Para Baltagi (2008), a consideração de efeitos aleatórios é importante no caso de haver correlação não prevista na amostra, que causaria inconsistência nos parâmetros estimados. Já os efeitos fixos, controlam problemas de identificação, desde que o termo de erro da *coorte* não seja constante no tempo o que causaria viés nos estimadores devido a correlação desses dois parâmetros (efeitos fixos e termo de erro). Desse modo, a escolha do método dependerá dos resultados estatísticos do teste de Hausman que

compara as duas formas de estimação e, por uma estatística de Qui-quadrado, indica o método mais adequado à amostra.

Como forma de verificar a eliminação do erro de medida pelo uso do *pseudo-painel* com um número suficientemente grande de indivíduos, Antman e McKenzie (2005) partem de um modelo simples de estimação pelos valores médios das variáveis de interesse como a seguir:

$$\overline{\log(w_{ct})}^{i(t)} = \beta \overline{X_{ct}}^{i(t)} + \pi_{ct} \overline{P_{ct}}^{i(t)} + \overline{u_{ct}}^{i(t)} \quad (12)$$

onde  $\overline{X_{ct}}^{i(t)}$  é medido com erro, de modo que:

$$\overline{X_{ct}}^{i(t)} = \overline{x_{ct}}^{i(t)} + \lambda_{ct}^{i(t)} \quad (13)$$

$$\overline{x_{ct}}^{i(t)} - \overline{X_{ct}}^{i(t)} = \lambda_{ct}^{i(t)} \quad (13.1)$$

em que  $\lambda_{ct}^{i(t)}$  é a medida do erro e  $\overline{x_{ct}}^{i(t)}$  a verdadeira média da *coorte* populacional. Sendo  $\overline{X_{ct}}^{i(t)} = \frac{\sum_{i=1}^n X_{it}}{n}$  e  $\overline{x_{ct}}^{i(t)} = \frac{\sum_{i=1}^n x_{it}}{n}$ , pode-se reescrever (13.1) como:

$$\lambda_{ct}^{i(t)} = \frac{\sum_{i=1}^n (x_{it} - X_{it})}{n} \quad (13.2)$$

Assim, a razão apresentada em (13.2) será negligenciável se o tamanho  $n$  da *coorte* for suficientemente grande. Desse modo, a estimação de (10) por MQO, controlando para efeitos fixos nas *coortes*, é consistente (ANTMAN e MCKENZIE, 2005).

Apesar dos ganhos associados ao uso de dados longitudinais que o *pseudo-painel* concede ao pesquisador, é necessário levar em consideração também as limitações desse tipo de dado. A primeira delas está na necessidade de obter um número de indivíduos por grupo suficientemente grande. Collado (1997) adverte que, se os grupos tiverem no mínimo 100 indivíduos, garante-se a qualidade dos estimadores a partir desses grupos. O presente estudo usará a medida de tamanho consistente para os grupos sugerida por Collado (1997). Essa condição impede que muitas análises sejam desenvolvidas com esse tipo de dado, o que seria possível com um painel verdadeiro.

Um exemplo é a divisão, para dados da PNAD, de grupos homogêneos por *coorte* de nascimento e por unidade de federação, que permitiria uma comparação locacional dos resultados. Essa análise pode não ser viável quando os dados são filtrados para atender às necessidades da pesquisa, pois o número de indivíduos por grupo não garante a ausência de erros de medida nas médias dos grupos. Essa limitação foi verificada ao longo do desenvolvimento deste trabalho. Ademais, uma fonte de crítica sobre dados de *pseudo-painel* é a impossibilidade de análise em termos individuais, sendo possível apenas a comparação entre grupos homogêneos.

Diante das considerações acerca das especificidades dos dados de uma amostra de *pseudo-painel* e as ponderações sobre o uso da mesma para análise econométrica, passa-se ao modelo que será analisado na seção de resultados.

#### 4.3 MODELO ECONOMETRICO PARA DETERMINAÇÃO DO RETORNO SALARIAL DA EDUCAÇÃO

Partindo da equação de salários minceriana ampliada:

$$\overline{\log(w_{ct})}^{i(t)} = \beta \overline{X_{ct}}^{i(t)} + \pi_{ct} \overline{P_{ct}}^{i(t)} + \alpha \overline{Z_{ct}}^{i(t)} + d_c + \overline{u_{ct}}^{i(t)} \quad (14)$$

em que:  $c$  e  $t$  indicam, respectivamente, *coorte* e tempo na amostra. Além disso,  $X_{ct}$  é o vetor de características médias individuais em cada grupo ou *coorte*.  $P_{ct}$  é a escolaridade,  $Z_{ct}$  é a experiência, ambas tomadas pelas médias dos indivíduos nas *coortes* homogêneas.

Sobre a equação (14), o retorno para o investimento em educação é mensurado por  $\pi$ . Para capturar efeitos fixos da amostra para cada *coorte*, há o termo  $d_c$ , que será válido para controlar possíveis efeitos fixos sobre os indivíduos quando o número de indivíduos por grupo for suficientemente grande. O tamanho por *coorte* considerado para isso neste estudo, seguindo o proposto por Collado (1997), foi de, no mínimo, 100 indivíduos por grupo. Nesse sentido, o modelo permite que se controlem características fixas que podem influenciar os retornos à educação, como as diferenças permanentes no capital humano geradas por habilidades que não são observáveis ou

mensuráveis. O termo de erro  $u_{ct}$  é assumido como independente e identicamente distribuído,

Aqui, o modelo de estimação para dados em painel utilizado será escolhido com base nos testes de Hausman, de Chow e o teste LM de Breush-Pagan. Esses testes são amplamente utilizados para determinar, entre as possíveis formas de tratamento de dados em painel<sup>20</sup>, aquela que melhor se aplica à amostra (GREENE, 2008). As variáveis presentes no vetor  $X_{ct}$  foram selecionadas de acordo com os preceitos teóricos da seção 2.1 e segundo indicado em trabalhos precedentes da seção 2.2. A descrição e os sinais esperados dessas variáveis se encontram na Tabela 1 abaixo:

**Tabela 1: Definição das Variáveis**

Variável	Descrição	Sinal Esperado
<b>Salário/hora</b>	Média do logaritmo natural do salário/hora deflacionado pelo IPC-FGV com base em setembro de 2011.	Variável Dependente
<b>Experiência</b>	Média da experiência individual $i$ definida como: idade atual subtraída dos anos de estudo menos seis.	+
<b>Experiência<sup>2</sup></b>	Quadrado da Experiência.	-
<b>Escolaridade</b>	Média dos anos de estudo.	+
<b>Esco. 5-8 anos</b>	Média da <i>dummy</i> que assume valor 1 na amostra individual quando o indivíduo possui de 4 à 7 anos de escolaridade, e 0 caso contrário.	+
<b>Esco. 9-11 anos</b>	Média da <i>dummy</i> que assume valor 1 na amostra individual quando o indivíduo possui de 8 à 10 anos de escolaridade, e 0 caso contrário.	+
<b>Esco. 11 ou mais</b>	Média da <i>dummy</i> que assume valor 1 na amostra individual quando o indivíduo possui o 11 anos ou mais completos de escolaridade, e 0 caso contrário.	+
<b>Esco. 8 anos</b>	Média da <i>dummy</i> que assume valor 1 se o indivíduo possui ensino primeiro grau completo (8 anos de estudo).	+

<sup>20</sup> Modelo *Pooled*, Modelo de Efeitos Fixos e Modelo de Efeitos Aleatórios.

<b>Esco. 11 anos</b>	Média da <i>dummy</i> que assume valor 1 se o indivíduo possui ensino médio completo (11 anos de estudo).	+
<b>Esco. 15 ou mais</b>	Média da <i>dummy</i> que assume valor 1 se o indivíduo possui terceiro grau completo ou mais (15 anos ou mais de estudo).	+
<b>Homem</b>	Média da <i>dummy</i> que assume valor 1 se o indivíduo é do sexo masculino.	+
<b>Branco</b>	Média da <i>dummy</i> que assume valor 1 se o indivíduo é branco	+
<b>Metrópole</b>	Média da <i>dummy</i> que assume valor 1 se o indivíduo reside em região metropolitana.	+
<b>t1 a t6</b>	<i>Dummy</i> de ano (1999 à 2009, respectivamente).	*
<b>Escolaridade* Homem</b>	Média da variável multiplicativa de gênero e escolaridade.	+
<b>Escolaridade*Branco</b>	Média da variável multiplicativa de raça e escolaridade.	+
<b>Escolaridade*Homem*Branco</b>	Média da variável multiplicativa de gênero, raça e escolaridade.	+
<b>Branco*Esco. 5-8</b>	Média da variável multiplicativa de raça e a variável classificadora de escolaridade que inclui quem possui entre 5 e 8 anos de escolaridade.	+
<b>Branco*Esco.9-11</b>	Média da variável multiplicativa de raça e a variável classificadora de escolaridade que inclui quem possui entre 9 e 11 anos de escolaridade.	+
<b>Branco*Esco. 11 ou mais</b>	Média da variável multiplicativa de raça e a variável classificadora de escolaridade que inclui quem possui entre 11 anos ou mais de escolaridade.	+
<b>Homem*Esco. 5-8</b>	Média da variável multiplicativa de gênero e a variável classificadora de escolaridade que inclui quem possui entre 5 e 8 anos de escolaridade.	+
<b>Homem*Esco. 9-11</b>	Média da variável multiplicativa de gênero e a variável classificadora de escolaridade que inclui quem possui entre 9 e 11 anos de escolaridade.	+
<b>Homem*Esco. 11 ou mais</b>	Média da variável multiplicativa de gênero e a variável classificadora de escolaridade que inclui quem possui entre 11 anos ou mais de escolaridade.	+

<b>Branco*Esco. 8</b>	Média da variável multiplicativa de raça e a variável classificadora de escolaridade que inclui quem possui 8 anos de escolaridade.	+
<b>Branco*Esco.11</b>	Média da variável multiplicativa de raça e a variável classificadora de escolaridade que inclui quem possui 11 anos de escolaridade.	+
<b>Branco*Esco. 15 ou mais</b>	Média da variável multiplicativa de raça e a variável classificadora de escolaridade que inclui quem possui 15 anos ou mais de escolaridade.	+
<b>Homem*Esco. 8</b>	Média da variável multiplicativa de gênero e a variável classificadora de escolaridade que inclui quem possui 8 anos de escolaridade.	+
<b>Homem*Esco. 11</b>	Média da variável multiplicativa de gênero e a variável classificadora de escolaridade que inclui quem possui entre 11 anos de escolaridade.	+
<b>Homem*Esco. 15 ou mais</b>	Média da variável multiplicativa de gênero e a variável classificadora de escolaridade que inclui quem possui 15 anos ou mais de escolaridade.	+
<b>período_*</b>	Média dos controles para os períodos de nascimento formadores das <i>coorte</i> (de 1936 à 1983, em intervalos bianuais).	*

As variáveis *dummies* e *dummies* multiplicativas excluídas do modelo são as consideradas como grupo de comparação. São elas: escolaridade de 0 a 4 anos, escolaridade 4 anos de estudos concluídos, *dummy* para 2011, mulher, não-branco, e residente em região não metropolitana. \*Sinal incerto.

A definição das *coortes* para o presente trabalho é descrita a seguir.

#### 4.4 FONTES DE DADOS

Os dados a serem utilizados nesta investigação foram retirados da Pesquisa Nacional por Amostra de Domicílios (PNAD), disponibilizados pelo Instituto Brasileiro de Geografia e Estatística (IBGE), para os anos em que houve a pesquisa no período bianual compreendido entre 1999 e 2011, inclusive. No total, serão utilizados sete anos<sup>21</sup> para a confecção de resultados e respostas ao problema proposto. Todos os 26 (vinte e seis) estados, mais o Distrito Federal, serão considerados na amostra.

<sup>21</sup> 1999, 2001, 2003, 2005, 2007, 2009 e 2011.

A escolha do período em questão é justificada pelos diversos acontecimentos econômicos e políticos vividos no Brasil nesse período, que influenciaram de maneira significativa as políticas de fornecimento e expansão da rede pública de ensino. Foi considerado o período pós-1994 pela característica econômica da época, com a mudança de moeda e estabilização econômica. Acredita-se, assim, que os dados sobre o nível de educação da população tenham mudado a partir de então, tendo em vista as mudanças conjunturais no país.

Apesar de satisfazer os pressupostos do trabalho, os anos de 1995 e 1997 não foram incluídos na amostra, pois os grupos homogêneos delimitados para esses anos não apresentaram o comportamento adequado para a inclusão na amostra. Nesses anos, a maior parte dos grupos não conteve o número de indivíduos necessários para a inclusão dos mesmos no modelo. Dessa maneira, a junção dos anos em questão na amostra causaria ao modelo uma perda considerável de grupos observados, uma vez que se exclui do modelo o grupo que não possui tamanho significativo.

Já a escolha dos biênios é justificada porque, nos anos 2000 e 2010, devido a pesquisa do Censo ter sido realizada nesses anos, não foram produzidas as pesquisas referentes à PNAD. Além disso, quando agrupados por biênios de nascimento, os grupos se tornam maiores e, conseqüentemente, mais robustos.

Serão considerados apenas os indivíduos economicamente ativos, 25 a 65 anos, que declararam estarem ocupadas no ano da pesquisa, e que não estavam estudando. A faixa etária da amostra foi assim definida com o intuito de incluir indivíduos com níveis mínimos de escolaridade e com alguma experiência no mercado de trabalho.

Além disso, como sugere a literatura nacional sobre retornos à educação (SACHSIDA, LOUREIRO e MENDONÇA, 2004; FALCÃO e SILVEIRA NETO, 2007), são excluídos da amostra os trabalhadores do setor público, dada as características que particularizam os indivíduos desse setor, como a estabilidade dos salários. Outro filtro usado foi a condição de

residentes no meio urbano, pois o meio rural dos estados da Região Norte do país foi incluído na pesquisa da PNAD apenas a partir do ano de 2004<sup>22</sup>.

Mais um critério de seleção da amostra foi o rendimento. Não foram considerados na análise os indivíduos que tivessem rendimento nulo ou não declarado. Haja vista que os dados de um *pseudo-painel* consideram como variável de análise a média das variáveis individuais, a ocorrência de valores nulo ou *missing values*<sup>23</sup> anula a média da *coorte* como um todo, comprometendo a análise. Por essa razão, esses indivíduos não foram incluídos na amostra final.

Além dos dados da PNAD, buscou-se ainda informações sobre o índice de preços ao consumidor, determinado pela Fundação Getúlio Vargas (FGV) e divulgada pelo IPEA. Essa série foi usada como deflator dos salários/hora sendo o período de referência o mês de setembro de 2011. A escolha do mês de referência se justifica por ser o período de coleta dos dados da PNAD anualmente.

Tendo sido selecionadas as amostras individuais para os intervalos bianuais entre 1999 a 2011, para determinar grupos homogêneos, unidade de análise da metodologia proposta, os indivíduos foram divididos segundo características que se mantêm constantes no decorrer do tempo, quais sejam: gênero, raça, região de moradia (metropolitana e não metropolitana)<sup>24</sup> e biênio de nascimento.

Apesar de seu caráter transitório, a divisão dos grupos por região de moradia se deve a características próprias de mercado de trabalho e ao acesso diferenciado à educação nas regiões metropolitanas do Brasil. Procedimento semelhante foi adotado por Warunsiri e McNown (2010) para a avaliação dos retornos à escolaridade na Tailândia.

Serão apresentados adiante os resultados estimados para o retorno à educação conforme uma amostra diferente, na qual os grupos foram formados

---

<sup>22</sup> A partir de 2004, a PNAD passou a ser feita também nas áreas rurais de Rondônia, Acre, Amazonas, Roraima, Pará e Amapá e desde então cobre completamente o território nacional.

<sup>23</sup> Quando não se observa uma das variáveis, definida no dicionário da PNAD como valor não definido, tem-se um *missing value*. Os softwares de análise estatística e econométrica excluem esses valores ou não os consideram na análise, portanto, é preciso filtrá-los da amostra.

<sup>24</sup> Considerando que a migração é nula entre região metropolitana e não metropolitana.



apenas de acordo com o sexo, raça e período de nascimento. Essa estimação será usada como forma de comparação e validação dos resultados obtidos com a amostra dos grupos formados por período, sexo, raça e região metropolitana.

O gênero foi codificado como uma variável *dummy* que recebeu valor um para os homens. Já a variável de raça, divide os indivíduos entre brancos e não brancos, assumindo valor um quando o indivíduo se declara branco ou amarelo, e zero caso contrário.

As *coortes* de nascimento foram divididas de modo que todos os grupos formados fossem acompanhados ao longo da série em pelo menos dois anos seguidos. Assim, a *coorte* mais velha da amostra é formada por indivíduos que nasceram entre os anos de 1936 e 1938, e a *coorte* mais jovem inclui os indivíduos nascidos entre 1984 e 1986. Um exemplo de *coorte* formada é o grupo de homens, não brancos, residentes em região metropolitana, nascidos entre 1951 e 1953. Esse grupo, por sua vez, é acompanhado ao longo de toda o período de análise, uma vez que, em 1999, a idade no grupo variou entre 46 e 48 anos, e, em 2011, entre 58 e 60 anos. Assim, as *coortes* serão ditas “mais velhas” quando no ano considerado forem formadas por indivíduos nascidos nos primeiros períodos considerados, e as *coortes* “mais jovens” aquelas que nasceram nos períodos mais recentes em relação à amostra. A tabela que apresenta a relação idade, período e *coorte* está apresentada na Tabela A1 em Anexo.

Assim, foram formados 128 grupos, sendo que, desses, dois grupos foram retirados por não apresentarem observações suficientes no critério de confiabilidade adotado neste estudo (100 indivíduos por grupo). Alguns grupos apresentaram número insuficiente de indivíduos e não foram excluídos. A justificativa para tanto está no fato de que, no ano em que os grupos em questão não respondiam a esse critério, eles também não faziam parte da amostra. A formação dos grupos, quanto aos valores assumidos por cada variável de controle, e o número de indivíduos por grupo estão descritos na Tabela A2 do Anexo.

Outras variáveis relevantes do modelo são as variáveis independentes. Dentre essas, ressalta-se a escolaridade, que foi medida segundo duas vias.

Diretamente, a variável escolaridade participa do modelo em sua forma contínua, como a média os anos de estudo dos indivíduos. Agregada, divide-se a população segundo quatro grupos de escolaridade. O primeiro grupo, classificado com *esco1*, inclui os indivíduos com escolaridade entre 0 e 3 anos de estudo. O segundo grupo, *esco2*, agrega os indivíduos com 4 a 7 anos de estudo. O terceiro, *esco3*, considera a escolaridade entre 8 a 10 anos de estudo. O quarto e último, *esco4*, dispõe dos indivíduos com 11 ou mais anos de escolaridade. Essas medidas alternativas de escolaridade não foram simultaneamente utilizadas na estimação dos modelos para evitar a multicolinearidade.

Uma variável relevante na análise dos retornos à educação é a experiência. Buscando captar o efeito do ciclo de vida sobre os retornos a esse tipo de capital humano, incluiu-se no modelo não só a forma linear da experiência, mas também sua forma quadrática. A medida de experiência utilizada para esse estudo segue a proposta de Mincer (1970) e considera que os anos trabalhados têm início após o fim do período de escolarização e que todos iniciam a educação formal aos seis anos de idade, como mostra a equação (15):

É importante frisar que as análises apresentadas na seção de resultados são feitas para as *coortes*, e as variáveis em análise são calculadas como a média ponderada das variáveis individuais em cada ano. Apesar dessa singularidade, as análises feitas sobre os dados de *pseudo-painel*, quando seguidos os devidos pressupostos, são as usuais de um modelo com dados em painel.

O software utilizado para estimar os resultados desse estudo foi o Stata 11. As próximas seções se ocupam de analisar estatisticamente os dados em questão e de apresentar os resultados obtidos para os objetivos aqui traçados. As conclusões e discussões acerca desses resultados também serão expostas nas seções seguintes.

## 5 RESULTADOS E DISCUSSÕES

### 5.1 ANÁLISE DESCRITIVA DOS DADOS

#### 5.1.1 Dados Individuais

Após a amostra ter sido filtrada pelas características de interesse sobre os indivíduos, foi possível determinar as médias totais das variáveis de interesse em cada ano. As Tabelas 2 e 3 mostram os valores médios das variáveis experiência, escolaridade e salário/hora brancos e não brancos para os homens, e para as mulheres brancas e não brancas, respectivamente.

**Tabela 2: Média de Experiência, Escolaridade e Salário/hora para Homens Não Brancos\* e Brancos\* no Brasil, período bianual de 1999 a 2011.**

	experiência		escolaridade		salário/hora	
	Homem Não Branco	Homem Branco	Homem Não Branco	Homem Branco	Homem Não Branco	Homem Branco
<b>1999</b>	32.23154	32.42486	5.513221	7.769032	2.673471	3.251159
<b>2001</b>	33.45016	33.89996	5.532098	7.761678	2.962638	3.543873
<b>2003</b>	32.59266	33.04548	5.98101	8.166342	3.296013	3.85279
<b>2005</b>	31.79765	32.33521	6.478691	8.508001	3.600019	4.111149
<b>2007</b>	33.5008	34.2078	6.660975	8.640042	3.897376	4.387278
<b>2009</b>	32.78957	33.41959	7.215288	9.04239	4.184024	4.641162
<b>2011</b>	34.42426	35.06121	7.20577	8.97071	4.574229	4.970948

Fonte: Elaborado pela autora a partir de dados das PNAD do intervalo bianual de 1999 a 2011. \*Sendo brancos aqueles indivíduos que declaram-se das raças branca e amarela, e não brancos são os indivíduos que declararam-se pretos, pardos, mulatos ou não declararam a raça.

**Tabela 3: Média de Experiência, Escolaridade e Salário/hora para Mulheres Não Brancas\* e Brancas\* no Brasil, período bianual de 1999 a 2011.**

	experiência		escolaridade		salário/hora	
	Mulher Não Branca	Mulher Branca	Mulher Não Branca	Mulher Branca	Mulher Não Branca	Mulher Branca
<b>1999</b>	11.58006	31.32466	2.30276	8.450259	0.874245	2.944954
<b>2001</b>	10.6969	32.84341	2.011425	8.403791	0.895848	3.287573
<b>2003</b>	13.59277	31.9837	2.909486	8.840918	1.304771	3.552598
<b>2005</b>	18.27332	31.20118	4.308528	9.276116	1.981109	3.846811
<b>2007</b>	19.25261	32.80022	4.435456	9.329285	2.161949	4.086959
<b>2009</b>	28.9008	32.2231	7.174174	9.793956	3.525155	4.353902
<b>2011</b>	30.89592	33.60569	7.555798	9.784787	3.980456	4.703969

Fonte: Elaborado pela autora a partir de dados das PNAD do intervalo bianual de 1999 a 2011. \*Sendo brancos aqueles indivíduos que declaram-se das raças branca e amarela, e não brancos são os indivíduos que declararam-se pretos, pardos, mulatos ou não declararam a raça.

A primeira característica trazida na Tabela 2 é a média da experiência no mercado de trabalho, mostrando que os homens brancos, apesar de, em média, estudarem mais do que os não brancos, apresentam mais anos de experiência. Esse comportamento é verificado, até 2009, muito mais intenso entre as mulheres. No início do período levantado, a experiência das mulheres não brancas que foram incluídas na amostra é menor do que a metade daquela que se observa entre as mulheres brancas que permaneceram na amostra. Porém, a partir de 2009. Esse resultado pode ser explicado na mudança de acesso ao mercado de trabalho para as mulheres, principalmente às não brancas.

Quanto à educação, os dados da PNAD aferem que, em média, as mulheres brancas estudam por mais tempo do que os homens da mesma raça em todos os anos da amostra. Já entre os não brancos, percebe-se uma mudança no perfil educacional ao longo desse período. A escolaridade feminina entre os não brancos cresce, proporcionalmente, mais do que a dos homens entre 1999 e 2011. Até 2011, a escolaridade média dos homens não brancos supera a das mulheres do mesmo grupo racial. No início dos períodos essa diferença é superior a 100%. Entretanto, em 2011, essa realidade se altera, e escolaridade média das mulheres passa a ser maior do que a dos

homens. Esse comportamento reflete uma tendência de maior acesso para as mulheres, de todas as raças, à educação formal no Brasil.

A diferenciação salarial entre raças está presente na amostra filtrada para atender aos pressupostos deste estudo, tanto entre os homens quanto entre as mulheres. Porém, a diferença foi decrescente ao longo dos anos considerados, para a amostra filtrada. Além da diferença entre raças, a renda média das mulheres persiste inferior em relação a dos homens. Esse resultado persiste ainda que a escolaridade entre os brancos seja superior para as mulheres, na média. Mais do que isso, a média salário/hora das mulheres brancas, que possuem a maior escolaridade média da amostra, é muito próximo do verificado para os homens não brancos.

Além das médias, existe uma visão mais ampla da amostra quando se observa a evolução da proporção de indivíduos em cada nível de escolaridade. A Tabela 4 exibe o percentual de indivíduos da amostra filtrada para cada ano de educação adquirida por cada ano amostrado:

**Tabela 4: Proporção de Indivíduos por escolaridade no Brasil no período bianual de 1999 a 2011.**

	1999	2001	2003	2005	2007	2009	2011
<b>0</b>	29.93%	34.22%	26.88%	18.53%	18.48%	7.93%	10.10%
<b>1</b>	1.71%	1.73%	1.63%	1.66%	1.73%	1.62%	1.27%
<b>2</b>	3.28%	3.04%	2.87%	2.73%	2.54%	2.65%	2.12%
<b>3</b>	5.04%	4.42%	4.39%	4.18%	3.89%	3.72%	2.94%
<b>4</b>	10.90%	9.90%	9.79%	9.74%	9.40%	9.23%	8.17%
<b>5</b>	6.60%	5.96%	6.46%	6.50%	6.44%	6.46%	5.43%
<b>6</b>	3.34%	3.14%	3.36%	3.63%	3.43%	3.73%	3.11%
<b>7</b>	3.61%	3.34%	3.71%	4.14%	3.57%	4.04%	3.92%
<b>8</b>	8.43%	7.94%	8.68%	9.32%	10.21%	10.06%	11.53%
<b>9</b>	1.58%	1.45%	1.87%	2.23%	2.15%	2.72%	2.28%
<b>10</b>	2.07%	1.93%	2.27%	2.52%	2.38%	2.87%	2.65%
<b>11</b>	14.16%	14.14%	17.66%	22.41%	22.88%	28.59%	29.61%
<b>12</b>	1.02%	1.03%	1.26%	1.47%	1.51%	1.95%	1.99%
<b>13</b>	0.82%	0.79%	1.09%	1.41%	1.27%	1.71%	1.59%
<b>14</b>	0.90%	0.82%	0.98%	1.29%	1.33%	1.71%	1.82%
<b>15</b>	6.61%	6.17%	7.11%	8.24%	8.77%	11.02%	11.46%

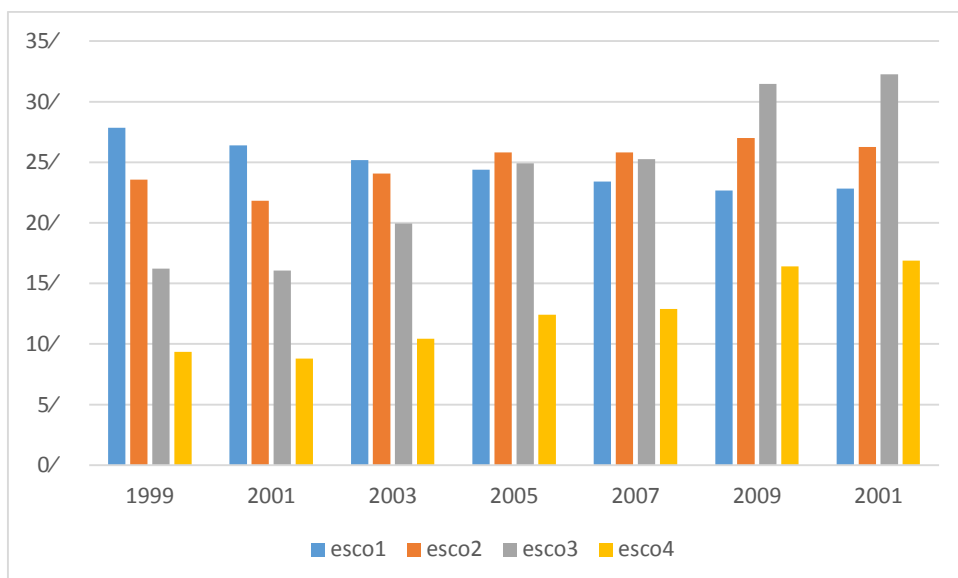
Fonte: Elaborado pela autora a partir de dados das PNAD do intervalo bianual de 1999 a 2011.

A Tabela 4 fornece indicações quanto a evolução educacional da população brasileira entre 1999 e 2011 na amostra de indivíduos empregados e residentes no meio urbano. Percebe-se uma redução dos mais de 29% da amostra de analfabetos, em 1999, para pouco mais de 10% em 2011.

Os grupos que apresentaram maior crescimento em postos percentuais foram os formados por indivíduos com 11 e 15 anos de escolaridade. Dentro da amostra é percebido, assim, um crescimento dos grupos mais escolarizados. Todavia, apenas a partir de 2009 a amostra apresentou mais de 10% de indivíduos com ensino superior. Em 2011, o que se observa é um grupo de semelhante participação percentual na amostra com nenhuma escolaridade ou com ensino superior.

Uma maneira de observar essas mudanças é comparar o percentual da amostra que, a cada ano, se encontra em cada um dos grupos de escolaridade determinados como: esco1 que inclui indivíduos com escolaridade completa de 0 a 4 anos de estudo; esco2 agrega indivíduos com 5 a 8 anos de estudo; esco3, de 9 a 11 anos de estudo; e esco4, de 11 a 15 anos de estudo. Esse resultado é apresentado a seguir no Gráfico 1.

**Gráfico 1: Percentual Amostrado por Grupo de Escolaridade no Brasil, nos anos do período Bianual de 1999 a 2011.**



Observando o Gráfico 1, é fácil perceber que, para a amostra que satisfaz o presente trabalho dissertativo, ouve um efeito oposto entre os

grupos que incluem apenas indivíduos com, no máximo, ensino fundamental e aquele composto por indivíduos com, pelo menos, ensino médio completo. O primeiro desses grupos era a maioria da amostra de 1999. Já em 2009 e 2011, o maior grupo da amostra é o caracterizado por ter entre 9 e 11 anos de escolaridade completos.

Apesar de uma melhora na distribuição da escolaridade dentro da amostra, o grupo com menor participação, em todos os anos, é aquele com indivíduos com no mínimo 12 anos de escolaridade completados. Por outro lado, o grupo de menor escolaridade permanece bastante participativo a cada ano na amostra. Apesar de sua participação ser, em 2011, a menor observada na série, ela ainda é elevada nesse ano, com mais de 20% dos indivíduos da amostra nesse grupo. É em 2011 que o grupo de maior escolaridade atinge sua maior proporção, porém o percentual de pessoas nesse grupo em tal ano não chega a 20% da amostra.

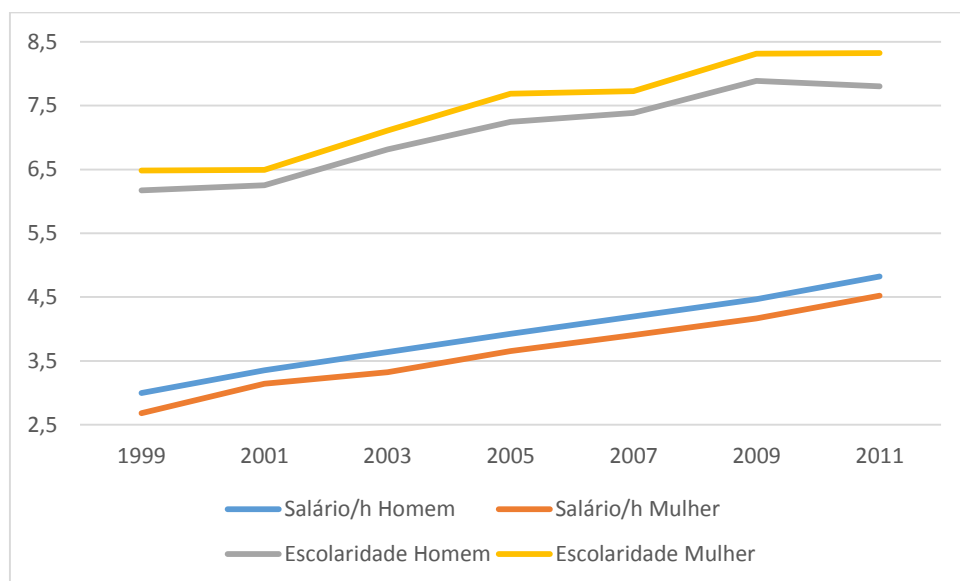
O que o Gráfico 1 informa é que, apesar de uma melhora na distribuição de escolaridade dentro da amostra, a escolaridade da maior parte da população considerada ainda é baixa. Porém, ao comparar os anos, é notável a melhora dessa proporção.

### **5.1.2 Dados Agrupados segundo Coorte de Nascimento, Raça, Gênero e Região Metropolitana**

Tendo como base a análise feita para as amostras de indivíduos, faz-se necessária uma análise semelhante para os grupos, que são as unidades de interesse no presente estudo. Os grupos foram definidos segundo o período de nascimento, o gênero, a raça, e o local de moradia do indivíduo.

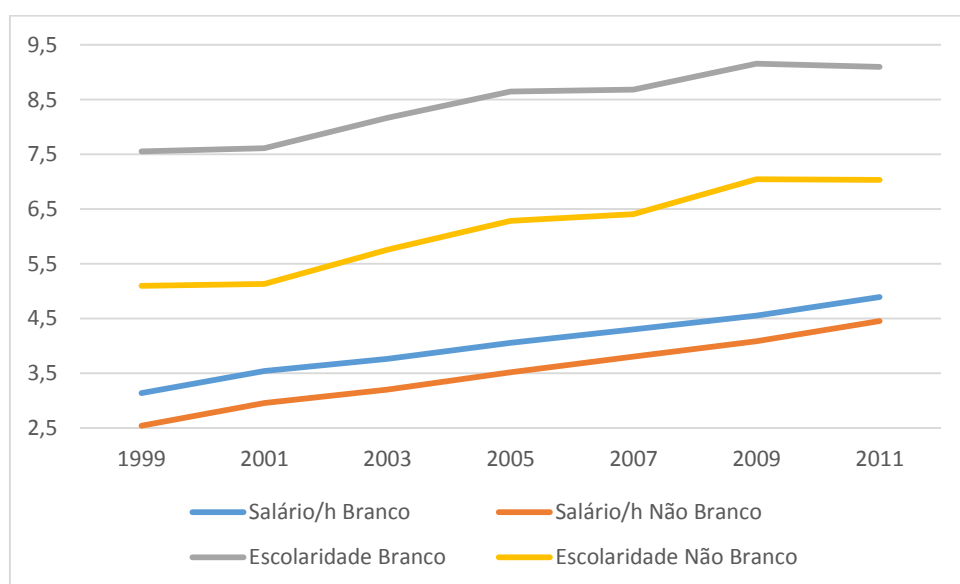
É possível dividir os grupos segundo suas características formadoras para, a partir disso, verificar se as diferenças evidenciadas nas amostras de indivíduos ano a ano são mantidas na estrutura de grupos. Para cada ano da amostra, o salário hora médio e a escolaridade de grupos femininos e masculinos são mostrados no Gráfico 2. Em seguida, o Gráfico 3 apresenta sobre as médias das mesmas variáveis por grupo, porém entre grupos brancos e não brancos. Por fim, o Gráfico 4 mostra a diferença de renda e escolaridade médias entre os grupos residentes e não residentes em regiões metropolitanas do país:

**Gráfico 2: Escolaridade e Renda Médias para Coortes Femininas e Masculinas no Brasil no período bianual de 1999 a 2011.**



Fonte: Elaborado pela autora a partir de dados das PNAD do intervalo bianual de 1999 a 2011.

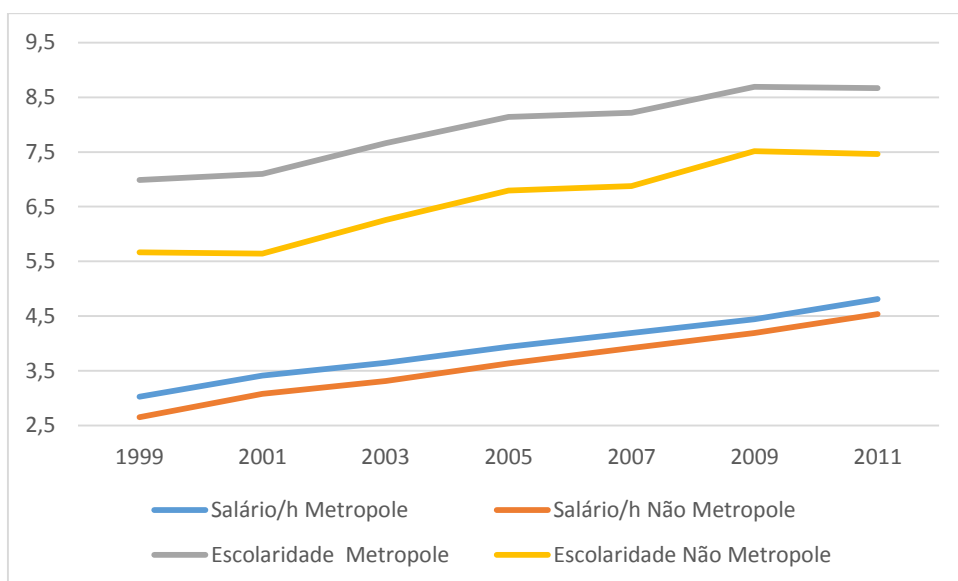
**Gráfico 3: Escolaridade e Renda Médias para Coortes Brancas\* e Não Brancas\*\* no Brasil no período bianual de 1999 a 2011.**



Fonte: Elaborado pela autora a partir de dados das PNAD do intervalo bianual de 1999 a 2011. Sendo as *coortes* brancas formadas por indivíduos que declaram-se das raças branca e amarela, e não brancos são os indivíduos que declaram-se pretos, pardos, mulatos ou não declaram a raça.



**Gráfico 4: Escolaridade e Renda Médias para Coortes residentes em Região Metropolitana e Não Metropolitana no Brasil no período bianual de 1999 a 2011.**



Fonte: Elaborado pela autora a partir de dados das PNAD do intervalo bianual de 1999 a 2011.

Os Gráficos 2 e 3 evidenciam que, como esperado, os grupos mantêm as diferenças de gênero e raça que se observa entre os indivíduos. Assim, o uso de uma amostra formada por grupos não altera a relação verificada entre variáveis como renda e escolaridade, e as características individuais. Por isso, as análises dos resultados sobre a direção dessas relações permanecem válidas para os indivíduos, apesar de não ser possível, em uma análise de grupos, definir a magnitude do efeito de uma variável sobre a outra para os indivíduos.

O Gráfico 4 também confere a possibilidade de transição das análises em grupos para a análise individual, nos mesmo termos postos acima. Mas, além disso, a comparação de escolaridade e salário/hora entre residentes e não residentes em regiões metropolitanas possui um papel ainda maior para o presente estudo. Na medida em que se assume o lugar de moradia como imutável na amostra de indivíduos aqui considerada, o resultado apresentado no Gráfico 3 corrobora a justificativa dada para essa consideração. De fato, a renda média e a escolaridade média dos grupos em diferentes regiões de residência são distintas. Portanto, controlar o lugar onde residem os

indivíduos para a formação dos grupos se mostra relevante sobre as principais variáveis de análise desse estudo.

Tendo sido analisadas as variáveis centrais deste trabalho, tanto para unidades individuais quanto para os grupos homogêneos, passa-se para a descrição e análise dos resultados da estimação dos retornos para a educação.

## **5.2 RETORNOS PARA A EDUCAÇÃO NO BRASIL**

Com o intuito de atender aos objetivos deste estudo, três formas funcionais de análise foram propostas. A primeira delas traz como medida de educação os anos de escolaridade, ou seja, a média de anos de estudo de cada grupo. A segunda usa não a escolaridade contínua, mas os grupos de estudos, ou seja, 0 a 4 anos de escolaridade, 5 a 8 anos (Esco. 5-8 anos), 9 a 10 (Esco. 9-11 anos), e 12 ou mais anos de escolaridade adquiridos (Esco. 11 ou mais).

Já a terceira considera o efeito do diploma sobre a renda. Nesse caso a medida de escolaridade é a variável *dummy*, definida na amostra de indivíduos, que assume valor 1 a cada ciclo escolar concluído. Ou seja, são quatro as variáveis qualificadoras, a primeira delas assume valor 1 quando o indivíduo possui 4 anos concluídos de estudo; a segunda quando o indivíduo possui 8 anos de estudo (Esco. 8 anos); a seguinte quando possui 11 anos concluídos de estudo (Esco. 11 anos), e a última quando possui 15 anos completos de escolaridade (Esco. 15 ou mais).

Os modelos para dados em painel foram estimados segundo o método de Efeitos Fixos, dado que os testes de Hausman e Chow identificaram esse método como o mais adequado para a amostra. O resultado desses testes são apresentado nas Tabelas A3 e A4 em Anexo.

Todos os modelos foram testados quanto à presença de Autocorrelação e Heterocedasticidade no termo de erro. O teste para a Autocorrelação de Wooldridge (2002), que tem hipótese nula a presença de Autocorrelação. O teste modificado de Wald para Heterocedasticidade, que tem como hipótese nula a presença de Heterocedasticidade (GREENE, 2008), mostraram que há autocorrelação, mas não há heterocedasticidade para o

termo de erro dos três modelos. Por essa razão, os resultados aqui apresentados são robustos pela correção da matriz de variância-covariância de White. Os resultados dos testes para cada modelo são apresentados nas Tabelas A5 e A6 da seção de Anexos.

Para cada forma funcional foram estimados os modelos segundo o Método de Mínimos Quadrados Ordinários. Essas estimações são comparadas aos resultados de Efeitos Fixos e, quanto mais próximo o parâmetro da educação estimado, por efeitos fixos, estiver de sua contraprova estimada por MQO, maior o indício de ineficiência do controle do viés de omissão da habilidade inata pelo uso de dados longitudinais. Os parâmetros estimados por MQO para educação em países em desenvolvimento, como o Brasil, tendem a superestimar o verdadeiro retorno salarial da escolaridade (CARD, 2010). Desse modo, se o estimador de efeitos fixos for inferior ao de MQO, é constatada a eficiência desse parâmetro.

Outra proposta foi a estimação de efeitos fixos sobre os dados para os grupos definidos, apenas, segundo o período de nascimento, a raça e o gênero. O que se pretende com essa estimação é comprovar a eficiência da amostra que considera a ausência de migração. Mais uma vez, há um modelo comparativo para cada forma funcional. Nesse caso, se o resultado da amostra sem a divisão por região de residência apresentar estimadores com valores próximos aos encontrados pela estimação por MQO, então a subdivisão dos grupos pela região de residência é eficiente para estimar esses parâmetros. Porém, se o estimador segundo essa amostra alternativa for inferior àquela proposta inicialmente para o presente estudo dissertativo, então a hipótese de ausência de migração causa viés na estimação, devendo não ser considerada.

Os resultados das contraprovas de MQO e com os grupos não subdivididos por região de residência são apresentados em anexo, nas Tabelas A4 e A5. Seus resultados serão comparados ainda nessa sessão com aqueles obtidos para os modelos de efeitos fixos sobre as diferentes formas funcionais propostas. A seguir, são apresentados separadamente os resultados das três formas funcionais propostas.

### 5.2.1 Retorno à Escolaridade: Modelo de Efeitos Fixos com Escolaridade Contínua

A primeira forma funcional a ser analisada é aquela que utiliza como variável de educação a escolaridade em anos de estudo. A Tabela 5 a seguir apresenta os parâmetros da estimação por Efeitos Fixos na coluna  $\beta$ , os seus desvios padrão e os seus níveis de significância.

**Tabela 5: Estimação por Efeitos Fixos da Equação de Rendimentos para o Brasil no período bianual de 1999 a 2011. Modelo 1.**

<u>Salário/hora</u>	$\beta$	Desvio Padrão	Significância
<b>Experiência</b>	0,145252	0,004736	***
<b>Experiência<sup>2</sup></b>	-0,00046	0,000073	***
<b>Escolaridade</b>	0,06057	0,015259	***
<b>Homem</b>	-0,91485	0,11843	***
<b>Branco</b>	0,036222	0,146759	-
<b>Metrópole</b>	0,266835	0,027706	***
<b>Escolaridade* Homem</b>	0,115033	0,0201	***
<b>Escolaridade*Branco</b>	0,038268	0,013345	***
<b>Escolaridade*Homem*Branco</b>	0,01174	0,015465	-
<b>periodo_1 ao periodo_15</b>	Crescente	-	
<b>1999 à 2011</b>	Decrescente	-	
<b>Const.</b>	-8,11016	2,124282	***

Fonte: Elaborado pela autora a partir de dados das PNAD do intervalo bianual de 1999 a 2011. Níveis de significância dos parâmetros: 1% (\*\*\*), 5% (\*\*), não significativo (-).

O primeiro passo na análise é verificar a eficiência do método de Efeitos Fixos. O parâmetro estimado para a variável escolaridade pelo método de MQO foi de 0,1765, como está apresentado na seção de Anexos. Já o retorno para a escolaridade determinado no método de efeitos fixos foi de 0,0606.

Essa diferença para menos sobre o modelo de Efeitos Fixos confirma a eficiência do método em questão em contornar o viés associado à omissão dos fatores constantes no tempo que são correlacionados com a escolaridade e determinam os salários/hora dos grupos.

Além disso, é necessário determinar a eficiência da amostra, dado que sua delimitação parte da forte pressuposição de que não há migração entre os períodos. Para tanto, comparam-se os retornos da escolaridade estimados pelo método de Efeitos Fixos para a amostra de grupos que considera como fixa a região de residência e para aquela que não faz essa consideração. O

resultado da estimação sobre a amostra menos restrita é apresentado na tabela A8, da seção de anexos. O Modelo 1 estimado sobre a amostra de grupos não divididos pela região de residência gerou um parâmetro para a Escolaridade de 0,1204.

A comparação desse resultado com o de MQO sugere que o parâmetro da amostra sem subdivisão pela região de residência é mais próximo do estimado por MQO. Dessa forma, é possível concluir que a amostra eficiente em supor como fixa a região de residência, e a amostra proposta inicialmente, com 126 grupos válidos, devem ser utilizadas na determinação do retorno da escolaridade no Brasil no período bianual entre 1999 e 2011

Dada a forma funcional proposta para avaliar o efeito do aumento de um ano na escolaridade de um dado grupo sobre a sua renda<sup>25</sup>, é necessário considerar o efeito diferenciado por gênero e raça. As variáveis que expressam essas diferenciações são as variáveis *Escolaridade\*Homem* e *Escolaridade\*Branco*. Como se nota no resultado do Modelo 1 (Tabela 5), tanto o parâmetro da escolaridade quanto os parâmetros das variáveis diferenciadoras foram significativas a 1%.

Desse modo, para os grupos femininos e não brancos, um ano adicional de estudo, entre dois períodos do intervalo bianual de 1999 a 2011, causa um aumento de 0,06% na renda. A elasticidade da renda em relação à educação é, para esses grupos, menor do que a unidade, indicando um comportamento inelástico da renda em relação à escolaridade na presente amostra do Brasil, no período bianual entre 1999 e 2011.

Diferenciando o retorno por gênero, pelo parâmetro da variável *Escolaridade\*Homem*, percebe-se que os grupos masculinos, tudo mais controlado, recebem um retorno maior para um ano adicional de estudo, em todos os níveis de escolaridade. Ou seja, os grupos masculinos quando elevam sua escolaridade em um ano sofrem um acréscimo de renda salarial 0,11% maior do que as mulheres com a mesma escolaridade, o que resulta em ganho total de 0,17%.

---

<sup>25</sup> Aqui, renda será considerada como sinônimo de logaritmo natural do salário/hora, por simplificação.

Quando se toma a variável diferenciadora do retorno à escolaridade por raça tem-se o seguinte: se eleva a escolaridade do grupo de 0 para 1 ano de estudo, e a renda dos grupos brancos é aumentada em 0,09% sobre a dos grupos não brancos. O ganho salarial dos brancos é de, aproximadamente, 0,1% no total.

Apesar de não ser uma medida para a existência de diferenciação de gênero e raça no mercado de trabalho, o modelo evidencia com seus resultados o mesmo que se observou na análise descritiva dos dados deste estudo. Ou seja, verifica-se um retorno à escolaridade superior entre os homens e entre brancos, ainda que se considerem grupos com o mesmo perfil de escolaridade.

A diferenciação simultânea do retorno à escolaridade por raça e gênero não foi possível, dado que a variável  $\text{Escolaridade} \times \text{Homem} \times \text{Branco}$  não foi significativa. No entanto, percebe-se que o maior retorno observado foi para os grupos masculinos e brancos que possuem um retorno total para a escolaridade de 3%, e o menor foi para os grupos de mulheres não brancas com escolaridade de 0,06%.

Esses resultados condizem com os encontrados por trabalhos anteriores para o Brasil, e indicam que os retornos para a escolaridade permanecem elevados no país. A elasticidade da renda salarial em relação à escolaridade determinada no presente estudo difere em magnitude dos resultados de trabalhos anteriores.

Apesar de terem um retorno para a escolaridade superior às mulheres, os grupos masculinos apresentam uma renda média inferior na presente amostra, dado que o parâmetro estimado para variável de gênero, Homem, foi negativo e significativo. Assim, a renda média dos grupos masculinos é 0,91% menor do que a dos grupos femininos. Esse resultado não era indicado na análise descritiva dos dados, no entanto, o tamanho dos grupos pode influenciar no resultado do modelo estimado.

Já a diferença de renda, na média, entre moradores e não moradores de região metropolitana concorda com o comportamento observado nos dados. Na amostra considerada, os moradores de região metropolitana recebem melhores remunerações médias do que aqueles que vivem nas

regiões interioranas do país. Entre os residentes de região metropolitana a renda média é 0,03% maior do que a dos residentes no restante do país.

Mais uma medida de acumulação de capital humano considerada foi a experiência. Essa variável aparece duplamente no modelo, tanto em nível quanto em sua forma quadrática. Os parâmetros foram significativos, e os seus resultados condizem com a teoria que supõe retornos marginais decrescentes para a experiência, como expressa o valor negativo do parâmetro da variável Experiência<sup>2</sup>. O resultado para o retorno à experiência mostra que de um ano a outro da amostra (que se encontra dividida em biênios), um grupo que passe de 0 para 2 anos de experiência eleva sua renda em 0,29%.

As variáveis de período de nascimento e as *dummies* de ano foram incluídas em todos os modelos que serão analisados neste trabalho como controles para os resultados, porém os parâmetros estimados são elucidativos. Como se nota, a renda média de todos os grupos amostrados foi crescente ao longo do período bianual de 1999 e 2011. Além disso, as *coortes* mais novas apresentam uma renda média superior às *coortes* mais velhas, como demonstram os parâmetros das variáveis de período de nascimento. Os parâmetros estimados, seus desvios padrão e seus níveis de significância, são apresentados em anexo na tabela A9.

A seguir, apresenta-se a estimação do Modelo 2, que considera como variável de educação as variáveis por grupo educacional.

### **5.2.2 Retorno à Escolaridade: Modelo de Efeitos Fixos com Escolaridade por Grupo**

A equação de rendimentos estimada segundo a forma funcional denominada como Modelo 2 utiliza como medida de educação as variáveis binárias de assumem valor 1 quando a escolaridade se encontra entre 0 e 4 anos (grupo controle), ou 5 e 8 anos (Esco. 5-8 anos), ou 9 e 11 anos (Esco. 9-11 anos) ou mais de 11 anos (Esco. 11 ou mais). O resultado obtido para o método de Efeitos Fixos sobre o Modelo 2 é apresentado abaixo na Tabela 6.

**Tabela 6: Estimação por Efeitos Fixos da Equação de Rendimentos para o Brasil no período bianual de 1999 a 2011. Modelo 2.**

<b>Salário/hora</b>	<b><math>\beta</math></b>	<b>Desvio Padrão</b>	<b>Significância</b>
<b>Experiência</b>	0,145666	0,00437	***
<b>Experiência<sup>2</sup></b>	-0,0004	6,87E-05	***
<b>Esco. 5-8 anos</b>	0,181946	0,16699	-
<b>Esco. 9-11 anos</b>	0,722361	0,172562	***
<b>Esco. 11 ou mais</b>	0,8305	0,282902	***
<b>Homem</b>	-0,47029	0,127689	***
<b>Branco</b>	0,380481	0,13841	***
<b>Metrópole</b>	0,247433	0,030007	***
<b>Branco*Esco. 5-8</b>	-0,07879	0,235061	-
<b>Branco*Esco.9-11</b>	-0,32533	0,209275	-
<b>Branco*Esco. 11 ou mais</b>	0,600806	0,341764	*
<b>Homem*Esco. 5-8</b>	0,439563	0,228931	*
<b>Homem*Esco. 9-11</b>	0,930058	0,242182	***
<b>Homem*Esco. 11 ou mais</b>	1,139604	0,318028	***
<b>periodo_1 ao periodo_15</b>	Crescente	-	
<b>1999 à 2011</b>	Decrescente	-	
<b>Const.</b>	-8,39318	2,088492	***

Fonte: Elaborado pela autora a partir de dados das PNAD do intervalo bianual de 1999 a 2011. O desvio padrão é representado na tabela por  $\sigma$ . Níveis de significância dos parâmetros: 1% \*\*\*, 5% \*\*, 10% \*.

A análise dos resultados apresentados na Tabela 6 segue os mesmos passos da análise do Modelo 1 (Tabela 5). Primeiro, verifica-se a eficiência do Modelo 2 comparando o retorno salarial da educação estimado segundo o método de Efeitos Fixos com o obtido pelo método de MQO, que é apresentado em anexo, na Tabela A10. Esse último foi de 0,584 para Esco. 5-8 anos contra 0,18 na estimação por Efeitos Fixos. Entretanto, o parâmetro estimado com essa amostra, e apresentado na Tabela 6, é não significativo.

Para a variável Esco. 9-11 anos, o parâmetro estimado por MQO foi de 2,25, enquanto por Efeitos Fixos o retorno salarial estimado para esse grupo foi de 0,72. Para a variável Esco.11 ou mais, o método de MQO estimou um retorno salarial de 2,05, enquanto o método com dados longitudinais proposto estimou um parâmetro de 0,83.

Como já discutido anteriormente, essa diferença revela que a escolha de dados longitudinais para a análise dos retornos salariais à educação no Brasil é coerente e os resultados com o método de Efeitos Fixos eficientes para contornar o viés associado à omissão da habilidade inata.



Comparando ainda os retornos das variáveis classificadoras de escolaridade, segundo o método de Efeitos Fixos, obtidos com as duas amostras anteriormente definidas, confirma-se a qualidade da amostra considerada no Modelo 2, que divide os grupos não só por gênero, raça e período de nascimento, mas também por região de residência.

Na seção de Anexos, Tabela A11, observa-se que os parâmetros estimados para os grupos escolares sobre a amostra foram de 1,27 para a variável Esco. 9-11 anos, e de 0,58 para a variável Esco. 11 ou mais. O parâmetro para a variável Esco. 5-8 anos também foi não significativo.

Apesar de o parâmetro de Esco. 11 ter sido inferior ao estimado com a amostra do Modelo 2, e também menor do que o estimado por MQO, ele não foi significativo sobre os grupos não subdivididos por região de residência.

Portanto, a comparação das amostras indica que o mais adequado é estimar a equação de rendimento dos grupos com a amostra definida sobre raça, sexo, período de nascimento e região de residência, como apresentado na Tabela 6.

Enquanto no Modelo 1 foram incluídas variáveis que buscaram diferenciar o rendimento da educação por raça e gênero, segundo a medida contínua de escolaridade, no Modelo 2 seis novas variáveis são incluídas partindo do mesmo princípio.

As variáveis Branco\*Esco. 5-8, Branco\*Esco.9-11, e Branco\*Esco. 11 ou mais, diferenciam o retorno salarial à educação por raça, enquanto as variáveis Homem\*Esco. 5-8, Homem\*Esco. 9-11, e Homem\*Esco. 11 ou mais diferenciam esse retorno por gênero. Assim, a análise do retorno da educação sobre os salários é feita de maneira semelhante a apresentada para o Modelo 1.

Para os grupos com escolaridade entre 5 e 8 anos, a única variável significativa de mensuração do retorno à escolaridade, que obteve apenas 10% de significância, foi aquela que diferencia o rendimento por gênero. Assim, para esse nível de escolaridade, a presente amostra permite apenas afirmar que os homens com escolaridade entre 5 e 8 anos recebem, em média, 0,44% mais do que as mulheres com a mesma escolaridade.

Quanto ao grupo com 9 a 11 anos de estudo, observa-se que os homens recebem em média 0,93% mais por sua escolaridade adicional do que as mulheres com o mesmo nível educacional. Os brancos não se diferenciam dos não brancos na amostra tomada, dado que o estimador de Branco\*Esco.9-11 é não significativo. Assim, os grupos femininos e não brancos com escolaridade entre 9 e 11 anos de estudo elevam sua renda em 0,72% acima do grupo de controle, a cada ano adicional de educação.

Entre os que possuem maior escolaridade, todos os parâmetros são significativos, pelo menos a 10%. Assim, os grupos femininos com mais de 11 anos de educação formal elevam sua renda em 0,83% a cada ano adicional de escolaridade adquirida em relação ao grupo de mulheres, não brancas e com escolaridade entre 0 e 4 anos. Já para os homens, no mesmo grupo escolar, esse crescimento é de 1,13 pontos percentuais maior, chegando a um retorno por ano adicional de educação de 1,96%. Para os grupos brancos, o crescimento do rendimento por um ano a mais de escolaridade é de 0,6% sobre os não brancos, e de 1,4% sobre o grupo de controle.

Além do retorno à educação, o Modelo 2 estimou um retorno para a experiência de 0,29%, referente ao período de um biênio amostrado.

As variáveis de período e as variáveis de ano apresentaram comportamento similar ao verificado no Modelo 2, como esperado, dado seu caráter exógeno e conjuntural. O modelo com todos os parâmetros estimados é apresentado na seção de anexos, Tabela A12.

Comparando os resultados apresentados no Modelo 1 e no Modelo 2, nota-se que o retorno para a experiência não sofreu alteração significativa. Já a comparação entre as medidas de educação evidenciam que, a cada ciclo escolar, o indivíduo se depara com uma realidade de mercado de trabalho diferenciada. No segundo modelo, de um modo geral, o retorno estimado para a educação foi estritamente superior ao estimado pelo Modelo 1. Esse resultado mostra que a diferenciação do retorno por grupo de escolaridade é relevante, e elucida fatos conclusivos sobre a realidade.

No mercado de trabalho, um ano adicional de educação nos ciclos finais de escolaridade formal é mais valorizado e melhor remunerado do que a mesma variação educacional nos anos iniciais de educação. No último

modelo, percebe-se a redução da diferenciação por gênero entre o grupo com menor escolaridade.

Passa-se agora ao último dos modelos propostos.

### **5.2.3 Retorno à Escolaridade: Modelo de Efeitos Fixos com Escolaridade medida pela Posse de Diploma**

O último modelo elaborado propõe avaliar o retorno salarial à educação segundo o efeito do diploma sobre o rendimento. Para tanto, as variáveis incluídas no modelo foram as variáveis binárias que assumem valor 1 quando a escolaridade alcançada é igual à conclusão de um ciclo escolar.

Assim, a primeira dessas variáveis se refere a obtenção do diploma de ensino primário (Esco. 4 anos), que é o controle para o Modelo 3; a segunda de ensino fundamental (Esco. 8 anos); a terceira, de segundo grau (esco. 11 anos) e a última de ensino superior (Esco. 15 ou mais). Espera-se com isso avaliar se o prêmio pago ao fim de um ciclo é diferente em relação ao retorno dentro do grupo escolar, que foi descrito no Modelo 2.

O resultado do Modelo 3 e dos retornos salariais aos diplomas no Brasil, para o período bianual de 1999 a 2011, é apresentado na Tabela 7 a seguir:

**Tabela 7: Estimação por Efeitos Fixos da Equação de Rendimentos para o Brasil no período bianual de 1999 a 2011. Modelo 3.**

<b>Salário/hora</b>	<b><math>\beta</math></b>	<b>Desvio Padrão</b>	<b>Significância</b>
<b>Experiência</b>	0,14062	0,003841	***
<b>Experiência<sup>2</sup></b>	-0,00028	7,73E-05	***
<b>Esco. 8 anos</b>	0,650191	0,21942	***
<b>Esco. 11 anos</b>	0,740125	0,172957	***
<b>Esco. 15 ou mais</b>	1,420652	0,319677	***
<b>Homem</b>	-0,26354	0,083059	***
<b>Branco</b>	0,573569	0,082084	**
<b>Metrópole</b>	0,259952	0,028745	***
<b>Branco*Esco. 8</b>	-0,72148	0,278574	***
<b>Branco*Esco.11</b>	-0,35116	0,203279	*
<b>Branco*Esco. 15 ou mais</b>	-0,01568	0,287727	-
<b>Homem*Esco. 8</b>	0,770025	0,320481	***
<b>Homem*Esco. 11</b>	0,893057	0,23343	***
<b>Homem*Esco. 15 ou mais</b>	1,124603	0,250154	***
<b>período_1 ao período_15</b>	Crescente	-	
<b>1999 à 2011</b>	Decrescente	-	
<b>Const.</b>	-9,64903	1,926982	***

Fonte: Elaborado pela autora a partir de dados das PNAD do intervalo bianual de 1999 a 2011. O desvio padrão é representado na tabela por  $\sigma$ . Níveis de significância dos parâmetros: 1% \*\*\*, 5% \*\*, 10% \*.

A validação e a eficiência<sup>26</sup> no Modelo 3 são verificadas na comparação do estimador de Efeitos Fixos com aquele obtido por MQO, e apresentadas na Tabela A13 do Anexo. Percebe-se que o parâmetro de Esco. 8 anos foi estimado em 0,65 para o primeiro, e de 0,92 para o segundo. Para a variável Esco. 11 anos esses valores foram, respectivamente, 0,74 e 2,51. Já para Esco. 15 anos ou mais, obteve-se, sobre a estimação de Efeitos Fixos, 1,4 como parâmetro, e 2,71 por MQO.

Comparando o resultado da amostra subdividida por região de residência (Tabela 7) e sua respectiva que não considera essa variável como fixa ao grupo, e apresentada na Tabela A14 do Anexo, observa-se que o único parâmetro inferior quando estimado sobre a amostra alternativa foi o da variável Esco. 8 anos. Porém, esse parâmetro não foi significativo para essa

<sup>26</sup> Quanto à correção do viés de omissão de fatores constantes no tempo correlatos com a escolaridade e determinantes da renda.

amostra. Assim como nos modelos anteriores, para o Modelo 3, a amostra que considera como fixa a região de residência é eficiente.

O retorno salarial para a escolaridade no Modelo 3 é pela valorização dada a cada diploma no mercado de trabalho. Assim como no Modelo 2, o Modelo 3 apresenta mais controles do que o Modelo 1, dado o caráter classificador de sua variável de escolaridade. As variáveis que diferenciam o retorno à escolaridade entre gêneros e raças são: Branco\*Esco.8, Branco\*Esco.11, Branco\*Esco.15 ou mais, Homem\*Esco.8, Homem\*Esco.11, Homem\*Esco.15 ou mais, Esco.11 anos, Esco.15 ou mais.

O retorno salarial estimado para mulheres não brancas que concluem o ensino fundamental de oito anos foi de 0,65% sobre aquelas com diploma do ensino fundamental de quatro anos. Já para as mulheres brancas, com essa mesma escolaridade, por seu diploma sofrer desvalorização salarial em reação às não brancas, foi de 0,72%. Os homens da mesma raça e com oito anos completos de escolaridade são remunerados com um acréscimo de 0,77% em relação à mulher da mesma raça e escolaridade.

Entre os grupos com diploma do ensino médio, os que recebem menor remuneração são aqueles compostos por mulheres não brancas, que elevam sua renda em 0,74 pontos percentuais sobre os grupos com quatro anos de estudo. Entre os homens, a obtenção deste diploma eleva sua renda em relação a das mulheres em 0,89%.

Já obtenção do título de graduado, ou um título superior a esse, confere às mulheres não brancas um acréscimo de renda de 1,42% sobre o grupo de controle. Os homens são ainda melhor remunerados quando concluem o ensino superior, elevando sua renda em relação às mulheres em 1,12%.

A experiência, o período de nascimento e o ano da amostra obtiveram resultados semelhantes aos observados nos modelos anteriores.

O que o Modelo 3 acrescenta de mais relevante aos que o antecederam é que esse modelo mostra o quanto o ensino superior é valorizado na amostra aqui trabalhada. O maior retorno observado está entre os homens com terceiro grau completo, que ganham 2,54% de acréscimo de renda por obterem o título de graduado. Enquanto isso, os grupos de mulheres não

brancas recebem apenas 1,42% a mais de renda ao obterem o mesmo diploma.

Tendo sido apresentados todos os modelos estimados sobre a presente amostra, é possível comparar os resultados obtidos neste trabalho com aqueles que o antecederam. Para tanto, serão considerados os resultados dos trabalhos que, em alguma medida, seguiram uma base de dados semelhante àquela proposta aqui. Da mesma maneira, o foco da comparação estará sobre os trabalhos que também utilizaram métodos para dados longitudinais, mais especificamente o método de efeitos fixos.

O primeiro desses trabalhos é o de Arbach, Dickerson e Green (2001), no qual os autores também formam um *pseudo-painel* sobre dados da PNAD, e o período considerado por eles compreende a década de 1990. A variável explicada pelo modelo é o salário/hora, em nível. A medida de educação por eles utilizada define seis grupos de escolaridade, e sem controles adicionais, estima um retorno de 0,379 para o grupo com escolaridade entre 1 e 3 anos, de 1,049 para o grupo com 4 a 7 anos de estudo, de 1,61 para o grupo com 8 a 10 anos de estudo, de 2,16 para o grupo com escolaridade entre 11 e 14 anos de estudo e de 3,03 para o grupo com mais de 15 anos de escolaridade.

A comparação da magnitude dos resultados de Arbach, Dickerson e Green (2001) aos que foram encontrados neste estudo dissertativo não é fidedigna aos resultados desses autores ou aos do presente estudo, pois tanto a unidade de escolaridade quanto a variável dependente não são as mesmas. Porém, a comparação dos comportamentos mostra que o perfil observado na amostra desses autores permanece presente em um *pseudo-painel* atualizado. Ou seja, os níveis mais altos de escolaridade permanecem com uma valorização salarial elevada no Brasil.

Na sequência, tem-se também o trabalho de Sachsida, Loureiro e Mendonça (2004) como referência de pressupostos adotados até aqui. Dividindo a medida de escolaridade em quatro grupos de modo a absorver o efeito do diploma de cada nível escolar sobre o rendimento, os autores encontram, em relação ao logaritmo natural do salário real, um retorno de 0,1622 para quatro anos de estudo, 0,1647 para oito anos de estudo, 0,1703 para onze, e 0,1716 para dezesseis anos completos de estudo. Mais uma vez

há uma divergência entre os resultados, em magnitude, mas também há diferenças na forma como a variável dependente e a medida de escolaridade se definem. Mas o comportamento segue o padrão no trabalho de Sachsida, Loureiro e Mendonça (2004) e neste estudo, de modo que se verifica um retorno mais elevado para os diplomas referentes à maior escolaridade.

Outro trabalho que parte de um *pseudo-painel* para determinação do retorno salarial sobre a escolaridade, e que também considera como variável dependente o logaritmo natural do salário hora, é o de França, Gasparini e Loureiro (2005). Para dados da PNAD entre os anos de 1992 e 2002, o retorno estimado para um ano adicional de educação foi de 0,1435. Esse trabalho não controla seus resultados para raças ou gêneros. O retorno médio estimado no presente estudo é inferior ao de França, Gasparini e Loureiro (2005). Essa comparação pode indicar que a presença dos controles de raça e gênero reduzem o efeito da escolaridade média, além de ser um indício de redução do retorno à escolaridade no Brasil.

Por fim, tem-se o trabalho de Teixeira e Menezes-Filho (2012), que estimam um retorno entre 0,11 e 0,05 para a escolaridade sobre o logaritmo natural do salário. O maior retorno está associado ao modelo com menor número de instrumentos utilizados, e o menor àquele com maior número de instrumentos. Esse é o trabalho que mais se assemelha aos encontrados no Modelo 1 apresentado na Tabela 5. Essa comparação com um resultado estimado sobre o método de variáveis instrumentais de Teixeira e Menezes-Filho (2012) corrobora a eficiência do método de Efeitos Fixos.

Apesar das diferenças entre os resultados aqui apresentados e aqueles encontrados por outros estudos, a relação entre escolaridade e renda permaneceu positiva e significativa em todos os modelos analisados. Além disso, os resultados obtidos na investigação desenvolvida neste documento são importantes por acrescentarem novos resultados à literatura.

## 6 CONCLUSÕES

Considerando o crescimento do investimento público em educação no Brasil, desde o início dos anos 2000, e constatada a escassez de uma avaliação atualizada do efeito da escolaridade sobre a renda no país, este trabalho dissertativo propôs mensurar e atualizar a magnitude desse retorno.

O objetivo central deste trabalho foi mensurar e analisar o retorno da educação sobre o salário/hora da população ocupada no Brasil urbano, no período bianual entre 1999 e 2001. Além de atualizar as estimativas encontradas na literatura para períodos anteriores, o presente estudo propôs ainda uma nova base de dados e um novo mecanismo de controle para o viés associado à omissão da habilidade inata que determina o nível de escolaridade e a renda simultaneamente.

O avanço apresentado sobre os trabalhos precedentes está, principalmente, no tipo de dado considerado, tendo sido priorizados dados *pseudo-painel*. O caráter longitudinal desses dados permite utilizar os indivíduos como seus próprios controles no processo de determinação do retorno salarial à escolaridade no método de Efeitos Fixos. Isso facilita o controle de todas as características constantes no tempo que podem viesar o modelo caso esse seja estimado sobre uma amostra transversal e segundo o método de mínimos quadrados.

A base de dados que se mostrou mais apropriada e eficiente foi aquela formada por grupos homogêneos de indivíduos pesquisados pelas PNADs e que foram agrupados por gênero, raça, região de residência e período de nascimento. O método que corrigiu os vieses de omissão de variáveis foi o método para dados em painel com Efeitos Fixos. Três formas funcionais foram estimadas e cada uma partiu de uma medida de escolaridade. Esse conjunto de resultados permite a comparação dos resultados do presente estudo com aqueles encontrados por trabalhos anteriores.

O primeiro modelo utiliza a medida contínua para a escolaridade e o retorno médio estimado para a educação foi de 0,06% para uma ano adicional de escolaridade, e esse resultado é semelhante ao estimado no trabalho de Teixeira e Menezes-Filho (2012). O trabalho desses autores apresenta



resultados referentes a dados de 1992 a 2002. Essa comparação sugere que, apesar do crescimento da escolaridade média dos brasileiros, como se evidenciou na análise descritiva da amostra, a elasticidade do salário/hora em relação à escolaridade pouco mudou.

Esse indício é ainda mais evidente quando tomados os resultados dos Modelos 2 e 3, que mensuram a escolaridade por variáveis *dummies* que diferenciam os grupos escolares, caso do segundo modelo, e mensuram o ganho associado aos diferentes diplomas escolares, como mostra o terceiro modelo.

No segundo modelo, observou-se que o crescimento do retorno por ano adicional de educação não é homogêneo entre os grupos escolares. Enquanto no grupo com 5 a 8 anos de escolaridade, o crescimento médio da renda por ano adicional de educação é de 0,18%, entre os que possuem 11 anos ou mais de escolaridade recebem um acréscimo de 0,83% de renda por um ano a mais de escolaridade.

Quando se observa o acréscimo salarial associado à obtenção do diploma, percebe-se uma evidência ainda mais forte da permanência de retornos divergentes entre os níveis escolares, de modo que prevalece superior o retorno associado aos níveis mais altos de escolaridade. No Modelo 3, os resultados mostraram que a obtenção do diploma de ensino fundamental de oito anos eleva a renda média em 0,65% em relação ao diploma de quatro anos de escolaridade. Já a conclusão de um nível escolar acima do ensino superior de 15 anos eleva o rendimento salarial em 1,42% em relação ao diploma de menor escolaridade considerado, o de quatro anos concluídos de educação.

Foi possível, assim, mostrar que o rendimento salarial associado à escolaridade no nível de terceiro grau permanece acima da média dos níveis escolares inferiores. Os retornos à experiência também foram elevados, significativos e com sinais coerentes. A forma funcional não afetou relevantemente o efeito da experiência sobre o salário/hora, que foi estimado em 0,14%.

A hipótese inicialmente levantada para esta pesquisa foi a de que o efeito da educação sobre a renda é menor quando o retorno é comparado

entre os anos para um mesmo grupo, do que quando se comparam indivíduos distintos em um único período de tempo. A confirmação dessa hipótese depende da comparação dos resultados aqui apresentados com resultados, para os mesmos anos, que sejam determinados por regressões sobre dados de *cross section*. O trabalho de Santos e Camillo (2011) traz resultados da equação de rendimentos de Mincer para dados de *cross section* nos anos entre 2006 e 2009. Os rendimentos da escolaridade estimados por esses autores confirmam a hipótese feita de que a análise sobre dados de seção cruzada superestima os retornos à escolaridade. Entre os anos considerados por esses autores variou 1,33 a 2,27. Apesar dos anos não serem os mesmo, o trabalho de Santos e Camillo (2011) considera um período recente, comparável aos resultados do presente estudo.

Por fim, pretendeu-se mostrar como os diferentes grupos da sociedade sofrem alterações de renda quando sua educação se altera. Nesse sentido, foram determinados os retornos à escolaridade individualizado entre raça e gênero. Como se mostrou, os resultados são robustos e confirmam os efeitos observados na análise descritiva dos dados.

Apesar de não serem suficientes para uma análise em termos de discriminação no mercado de trabalho, os resultados apresentaram uma valorização superior da escolaridade dos grupos masculinos e dos grupos brancos. O que se pode concluir dessa constatação é que, na amostra considerada, os retornos à escolaridade são discriminantes em gênero e raça, e persistem ao longo do período considerado.

Esses resultados mostram a importância de atualizar os resultados anteriores sobre a relação entre salários/hora e a escolaridade no Brasil. Além disso, o método de estimação por Efeitos Fixos se mostrou relevante e eficiente em contornar vieses associados à equação e rendimentos de Mincer.

Por fim, observou-se que os rendimentos para educação no Brasil permanecem crescentes e elevados, o que indica a necessidade de elevação do volume investido pelo Estado nesse setor. A diferença desses retornos por gênero e raça indica que o investimento deve ser feito de modo que se garanta a qualidade do capital humano acumulado, além de promover a inserção equitativa dos grupos no mercado de trabalho.

## 7 REFERÊNCIAS BIBLIOGRÁFICAS

- ACEMOGLU, D. & ANGRIST, J. How Large are the Social Returns to Education? Evidence from the Compulsory Schooling Laws. **NBER**, Working Paper No 7444, Cambridge, MA, 1999.
- ARBACHE, J. S.; DICKERSON, A. & GREEN, F. Returns to Education and Trade Liberalisation: A Pseudo-Panel Approach. **The 15th Annual Latin American and Caribbean Economic Association Meeting**, Montevideo, 2001.
- ANTMAN, F. & MCKENZIE, J. D. Earnings Mobility and Measurement Error: A Pseudo-Panel Approach. Washington: World Bank Policy Research, 2005.
- BARBOSA FILHO, F. H. & PÊSSOA, S. A. “Retorno da Educação no Brasil”. **Pesquisa e Planejamento Econômico**, Rio de Janeiro, v. 38, n. 1, p. 97-125, 2008.
- BARBOSA FILHO, F. H. & PÊSSOA, S. A. “Educação e Crescimento: O que a Evidência Empírica e Teórica mostra?”. **Revista EconomiA**, Brasília(DF), v. 11, n. 2, p.277–279, 2010.
- BARROS, R. P. & MENDONÇA, R. **Investimentos em Educação e Desenvolvimento Econômico**. Rio de Janeiro, Ipea, 1997.
- BECKER, G. S. “Investment in human capital: A theoretical analysis”. **Journal of Political Economy**, v.70, p.9-49, 1962.
- BLONDAL, S.; FIELD, S. & GIROUARD, N. Investment in human capital through upper-secondary and tertiary education. **OECD Economic Studies**, v. 34, n. 1, p. 52-111, 2002.
- CAHUC, P. Education and Human Capital. In: CAHUC, Pierre. *Labor Economics*. MIT press, p. 59-103, 2004.
- CANGUSSU, R. C.; SALVATO, M. A. & NAKABASHI, L. Uma análise do capital humano sobre o nível de renda dos estados brasileiros: MRW versus Mincer. **Estudos Econômicos (São Paulo)**, v. 40, n. 1, p. 153-183, 2010.
- CARD, D. Estimating the return to schooling: Progress on some persistent econometric problems. **Econometrica**, v. 69, n. 5, p. 1127-1160, 2001.
- COLLADO, M. D. Estimating dynamic models from time series of independent crosssections. **Journal of Econometrics**, Amsterdam, v. 82, n. 1, p.37–62, Jan. 1997.
- DEATON, A. Panel data from times series of *cross-sections*. **Journal of Econometrics**, Amsterdam, v. 30, n. 1/2, p.109–126, Oct./Nov. 1985.

- FALCÃO, N. A. & SILVEIRA NETO, R. M. "Concentração espacial de capital humano e externalidades: o caso das cidades brasileiras". Associação Nacional dos Centros de Pós-Graduação em Economia. In: XXXV Encontro Nacional de Economia. Recife – PE. **Anais...**, 2007.
- FERREIRA, S. G. The Provision of Education and its Impacts on College Premium in Brazil. **Revista Brasileira de Economia**, Rio de Janeiro, v. 58, n. 2, p. 211-233, 2004.
- FRANÇA, G. N.; GASPARINI, C. E. & LOUREIRO, P. R. de A. RELAÇÃO ENTRE ESCOLARIDADE E RENDA NO BRASIL NA DÉCADA DE 1990. Anais do 10º Encontro Regional de Economia, 2005.
- GAREN, J. The Returns to Schooling: A Selectivity Bias Approach with a Continuous Choice Variable. **Econometrica**, v. 52, n. 5, p. 1199-1218, 1984.
- GREENE, W. H. Econometric Analysis. Prentice Hall. 6ª edição, p. 233-334. 2008.
- GRILICHES, Z. Estimating the Returns to Schooling: Some Econometric Problems. **Econometrica**, v. 45, n. 1, p. 1-22, 1977.
- HADDAD, F. O Plano de Desenvolvimento da Educação: razões, princípios e programas. Brasília. **Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira**, 2008.
- HECKMAN, J. J. Shadow prices, market wages, and labor supply. **Econometrica**, v. 42, n. 4, p. 679-694, 1974.
- HECKMAN, J. J. Sample selection bias as a specification error. **Econometrica**, v.45, n.1, p.153-161, 1979
- HECKMAN, J. J. Policies to foster human capital. **Research in economics**, v. 54, n. 1, p. 3-56, 2000.
- HECKMAN, J. J.; LOCHNER, L. J. & TODD, P. E. Earnings functions and rates of return. **National Bureau of Economic Research**, Cambridge, 2003.
- INSTITUTO BRASILEIRO DE ÉTICA CONCORENCIAL – ETCO & INSTITUTO BRASILEIRO DE ECONOMIA– IBRE/FGV. Índice de Economia Subterrânea permanece em queda, apontam ETCO E IBRE/FGV. Disponível em: < <http://www.etco.org.br/destaque/indice-de-economia-subterranea-permanece-em-queda-apontam-etco-e-ibrefgv-2/>>. Acessado em: fevereiro de 2014.
- INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA - Indicadores Financeiros Educacionais. Disponível em:

<<http://portal.inep.gov.br/indicadores-financeiros-educacionais>>. Acessado em julho de 2013.

INSTITUTO DE PESQUISA ECONOMICA APLICADA-IPEA. Boletim Mercado de Trabalho: Análise do Mercado de Trabalho. v. 36, p. 3-22, Agosto de 2008.

LUCAS, R. E. On the Mechanics of Economic Development. **Journal of Monetary Economics**, Chicago, v. 22, n.1, p.18-21, 1988.

MCKENZIE, D. J. Asymptotic theory for heterogeneous dynamic pseudo-panels. **Journal of Econometrics**, v. 120, n. 2, p. 235-262, 2004.

MCKENZIE, D. J. Asymptotic theory for heterogeneous dynamic pseudo-panels. **Journal of Econometrics**, v. 120, n. 2, p. 235-262, 2004.

MILLIGAN, K.; MORETTI, E.; OREOPOULOS, P. Does education improve citizenship? Evidence from the United States and the United Kingdom. **Journal of Public Economics**, v. 88, n. 9, p. 1667-1695, 2004.

MINCER, J. Investment in Human Capital and Personal Income Distribution. **Journal of Political Economy**, v. 66, n. 4, p. 281-302, 1958.

MINCER, J. The Distribution of Labor Incomes: A Survey with Special Reference to the Human Capital Approach. **Journal of Economic Literature**, v. 8, n. 1, p. 1-26, 1970.

MINCER, J. A. Schooling, experience, and earnings. **NBER Books**, 1974.

MINCER, J. Human capital and economic growth. 1981.

MOFFITT, R. Identification and estimation of dynamic models with a time series of repeated *cross-sections*. **Journal of Econometrics**, Amsterdam, v. 59, n. 1, p.99–124, Sep. 1993.

MORA, J. J. & MURO, J. Sheepskin effects by cohorts in Colombia. **International Journal of Manpower**, v. 29, n. 2, p. 111-121, 2008.

MORETTI, E. Estimating the Social Return to Higher Education: Evidence from Longitudinal and Repeated *Cross-sectional* Data. **Journal of Econometrics**, n.121, Los Angeles, p.175-212, 2004.

MORETTI, E. Private and Social Returns to Education. **Rivista di Politica Economica**, Berkeley, v. 96, n. 3, p. 3-46, 2006.

MOURA, R. L. de. Testando as Hipóteses do Modelo de Mincer para o Brasil. **Revista Brasileira de Economia**, Rio de Janeiro, v. 64, n. 4, p. 407-449, 2008.

- OLIVEIRA, A. M. C. H. Acumulando Informações e Estudando Mudanças ao Longo do Tempo: Análises Longitudinais do Mercado de Trabalho Brasileiro. 2002. 151. Tese-Universidade Federal de Minas Gerais. Versão Digital.
- PEREIRA, V. da F et al. AVALIAÇÃO DOS RETORNOS AOS INVESTIMENTOS EM EDUCAÇÃO PARA TRABALHADORES DO SEXO MASCULINO NO BRASIL. **Revista de Economia Contemporânea**, v. 17, n.1, Rio de Janeiro, 2013.
- RESENDE, M. & WYLLIE, R. Retornos para educação no Brasil: evidências empíricas adicionais. **Economia Aplicada**, São Paulo, v. 10, n. 3, p. 349-365, 2006.
- RYDER, N. B. The cohort as a concept in the study of social change. **American Sociological Review**, v.30, n. 6, p. 843-861, 1965
- SABOIA, J. et al.. Tendências da qualificação da força de trabalho. In: KUPFER, D. et al. (Ed.). Perspectivas do investimento no Brasil: temas transversais. Projeto PIB – perspectiva do investimento no Brasil. Rio de Janeiro: Synergia, 2010. v. 4, p. 343-400.
- SACHSIDA, A; LOUREIRO, P. R. A. & MENDONÇA, M. J. C. Um Estudo Sobre Retorno em Escolaridade no Brasil. **Revista Brasileira de Economia**, Rio de Janeiro, v. 58, n. 2, p. 249-265, 2004.
- SALVATO, M. A.; FERREIRA, P. C. G. & DUARTE, A. J. M. O impacto da escolaridade sobre a distribuição de renda. **Estudos Econômicos**, São Paulo, v. 40, n. 4, p. 753-791, 2010.
- SANTOS, E. S. dos & CAMILLO, V. S. CAPITAL HUMANO E RENDA DO TRABALHO NO BRASIL: UMA INVESTIGAÇÃO EMPÍRICA. **IX CICLO DE DEBATES EM ECONOMIA INDUSTRIAL, TRABALHO E TECNOLOGIA**. Campinas, 2011.
- SCHULTZ, T. W., Capital Formation by Education. **Journal of Political Economy**, v. 58, n. 6, p.571-583, 1960.
- SCHULTZ, T. W., Investment in Human Capital. **The American Economic Review**, v. 51, n. 1, p.1-17, 1961.
- SOLOW, R. M. A contribution to the theory of economic growth. **The quarterly journal of economics**, v. 70, n. 1, p. 65-94, 1956.
- SPENCE, M. Job market signaling. **The Quarterly Journal of Economics**, v. 87, n. 3, p. 355-374, 1973.
- TEIXEIRA, W. M. & MENEZES-FILHO, N. A. Estimando o retorno à educação do Brasil considerando a legislação educacional brasileira como um

instrumento. **Revista de Economia Política**, v. 32, n. 3, p. 479-496, São Paulo, 2012.

VERBEEK, M.; VELLA, F.. Estimating dynamic models from repeated *cross-sections*. **Journal of econometrics**, v. 127, p. 83-102, 2005.

WARUNSIRI, S. & MCNOWN, R. The returns to education in Thailand: A pseudo-panel approach. **World Development**, v. 38, n. 11, p. 1616-1625, 2010.

WOOLDRIDGE, J. M. *Econometric Analysis of Cross-section and Panel Data*. The MIT Press. Cambridge. p. 247-291. 2002.

## 8 ANEXOS

### A1: Coortes por Período de Nascimento

Ano	1999	2001	2003	2005	2007	2009	2011	
Período								
1938-1938	61-63	65-63						periodo_1
1939-1941	60-58	62-60	64-62					periodo_2
1942-1944	57-55	59-57	61-59	63-61	65-63			periodo_3
1945-1947	54-52	56-54	58-56	60-58	62-60	64-62		periodo_4
1948-1950	51-49	53-51	55-53	57-55	59-57	61-59	63-61	periodo_5
1951-1953	48-46	50-48	52-50	54-52	56-54	58-56	60-58	periodo_6
1954-1956	45-43	47-45	49-47	51-49	53-51	55-53	57-55	periodo_7
1957-1959	42-40	44-42	46-44	48-46	50-48	52-50	54-52	periodo_8
1960-1962	39-37	41-39	43-41	45-43	47-45	49-47	51-49	periodo_9
1963-1965	36-34	38-36	40-38	42-40	44-42	46-44	48-46	periodo_10
1966-1968	33-31	35-33	37-35	39-37	41-39	43-41	45-43	periodo_11
1969-1971	30-28	32-30	34-32	36-34	38-36	40-38	42-40	periodo_12
1972-1974	27-25	29-27	31-29	33-31	35-33	37-35	39-37	periodo_13
1975-1977			28-26	30-28	32-30	34-32	36-34	periodo_14
1978-1980				27-25	29-27	31-29	33-31	periodo_15
1981-1983						28-26	30-28	periodo_16

Fonte: Elaborado pela autora a partir de dados das PNAD do intervalo bianual de 1999 a 2011.

### A2: Número de Indivíduos por Coorte, em cada ano da amostra.

	1999	2001	2003	2005	2007	2009	2011	Sexo	Raça	Região	Período
								*	**	***	
111	237	231	161	109	81	59	49	1	1	1	periodo_1
112	312	313	245	176	125	119	66	1	1	0	periodo_1
113	140	131	105	71	60	54	20	1	0	1	periodo_1
114	255	271	208	146	98	90	63	1	0	0	periodo_1
115	327	272	239	202	155	120	72	1	1	1	periodo_2
116	437	436	368	278	209	171	115	1	1	0	periodo_2
117	222	192	156	153	107	93	57	1	0	1	periodo_2
118	377	337	341	240	181	149	107	1	0	0	periodo_2
119	409	377	309	265	245	166	114	1	1	1	periodo_3
120	527	514	447	420	342	274	183	1	1	0	periodo_3
121	315	268	256	213	161	153	86	1	0	1	periodo_3
122	454	435	394	358	295	256	143	1	0	0	periodo_3
123	589	534	407	420	365	286	197	1	1	1	periodo_4
124	676	671	579	566	498	384	262	1	1	0	periodo_4
125	373	354	336	307	312	245	172	1	0	1	periodo_4
126	505	553	531	509	472	367	265	1	0	0	periodo_4
127	737	701	626	561	475	414	323	1	1	1	periodo_5



<b>128</b>	798	837	755	677	674	574	422	1	1	0	periodo_5
<b>129</b>	513	531	467	460	410	366	313	1	0	1	periodo_5
<b>130</b>	590	688	747	687	571	550	397	1	0	0	periodo_5
<b>131</b>	823	823	785	736	649	537	449	1	1	1	periodo_6
<b>132</b>	973	904	914	900	801	662	630	1	1	0	periodo_6
<b>133</b>	627	645	566	614	569	493	444	1	0	1	periodo_6
<b>134</b>	713	860	816	842	763	659	607	1	0	0	periodo_6
<b>135</b>	1043	972	883	811	758	716	531	1	1	1	periodo_7
<b>136</b>	1092	1180	1126	1065	1008	1002	792	1	1	0	periodo_7
<b>137</b>	770	802	688	813	729	604	592	1	0	1	periodo_7
<b>138</b>	881	1028	984	952	926	891	718	1	0	0	periodo_7
<b>139</b>	1158	1078	976	957	863	780	688	1	1	1	periodo_8
<b>140</b>	1282	1363	1317	1216	1123	1003	911	1	1	0	periodo_8
<b>141</b>	856	887	814	848	841	816	743	1	0	1	periodo_8
<b>142</b>	976	1223	1132	1174	1100	1063	909	1	0	0	periodo_8
<b>143</b>	1222	1187	1095	1081	912	934	779	1	1	1	periodo_9
<b>144</b>	1356	1522	1381	1402	1222	1179	1000	1	1	0	periodo_9
<b>145</b>	994	989	959	1080	972	928	870	1	0	1	periodo_9
<b>146</b>	1102	1348	1306	1330	1279	1206	1122	1	0	0	periodo_9
<b>147</b>	1225	1270	1156	1149	1026	990	889	1	1	1	periodo_10
<b>148</b>	1501	1647	1586	1580	1524	1351	1238	1	1	0	periodo_10
<b>149</b>	1094	1088	1109	1247	1185	1118	936	1	0	1	periodo_10
<b>150</b>	1204	1549	1491	1489	1521	1419	1285	1	0	0	periodo_10
<b>151</b>	1265	1232	1187	1080	999	938	862	1	1	1	periodo_11
<b>152</b>	1426	1578	1540	1500	1421	1420	1165	1	1	0	periodo_11
<b>153</b>	1073	1212	1121	1263	1166	1144	1028	1	0	1	periodo_11
<b>154</b>	1198	1513	1563	1589	1587	1536	1374	1	0	0	periodo_11
<b>155</b>	1242	1228	1187	1171	1057	1047	962	1	1	1	periodo_12
<b>156</b>	1480	1645	1621	1648	1491	1455	1250	1	1	0	periodo_12
<b>157</b>	1166	1256	1246	1246	1272	1297	1127	1	0	1	periodo_12
<b>158</b>	1256	1663	1685	1696	1667	1675	1513	1	0	0	periodo_12
<b>159</b>	1280	1293	1172	1221	1118	1010	953	1	1	1	periodo_13
<b>160</b>	1411	1603	1627	1641	1524	1403	1269	1	1	0	periodo_13
<b>161</b>	1179	1269	1246	1370	1396	1299	1141	1	0	1	periodo_13
<b>162</b>	1322	1729	1768	1834	1822	1746	1609	1	0	0	periodo_13
<b>163</b>	1236	1363	1298	1270	1183	1121	971	1	1	1	periodo_14
<b>164</b>	1399	1633	1582	1673	1625	1519	1360	1	1	0	periodo_14
<b>165</b>	1250	1345	1300	1526	1496	1439	1250	1	0	1	periodo_14
<b>166</b>	1395	1827	1886	1921	1904	1932	1766	1	0	0	periodo_14
<b>167</b>	1076	1330	1366	1412	1308	1277	1156	1	1	1	periodo_15
<b>168</b>	1297	1628	1733	1826	1617	1646	1458	1	1	0	periodo_15
<b>169</b>	1091	1324	1366	1651	1528	1522	1465	1	0	1	periodo_15
<b>170</b>	1368	1884	1972	2123	2051	2054	1879	1	0	0	periodo_15
<b>171</b>	612	1017	1196	1405	1327	1310	1110	1	1	1	periodo_16

<b>172</b>	857	1385	1547	1822	1742	1699	1555	1	1	0	periodo_16
<b>173</b>	650	993	1229	1544	1566	1536	1451	1	0	1	periodo_16
<b>174</b>	1023	1653	1907	2206	2178	2157	2059	1	0	0	periodo_16
<b>175</b>	123	96	81	62	51	37	21	0	1	1	periodo_1
<b>176</b>	112	145	86	76	53	39	27	0	1	0	periodo_1
<b>177</b>	89	74	70	62	40	30	17	0	0	1	periodo_1
<b>178</b>	89	115	72	36	47	35	22	0	0	0	periodo_1
<b>179</b>	168	193	132	93	78	69	30	0	1	1	periodo_2
<b>180</b>	163	181	133	114	70	74	44	0	1	0	periodo_2
<b>181</b>	125	123	105	99	87	62	33	0	0	1	periodo_2
<b>182</b>	156	159	98	96	79	83	35	0	0	0	periodo_2
<b>183</b>	222	259	190	145	129	104	54	0	1	1	periodo_3
<b>184</b>	243	249	204	159	144	102	81	0	1	0	periodo_3
<b>185</b>	185	168	143	149	125	79	63	0	0	1	periodo_3
<b>186</b>	192	222	174	158	126	99	56	0	0	0	periodo_3
<b>187</b>	344	319	274	241	183	168	98	0	1	1	periodo_4
<b>188</b>	363	356	301	255	197	163	98	0	1	0	periodo_4
<b>189</b>	245	249	218	223	196	176	102	0	0	1	periodo_4
<b>190</b>	249	265	241	226	165	154	101	0	0	0	periodo_4
<b>191</b>	466	471	413	381	324	285	175	0	1	1	periodo_5
<b>192</b>	443	431	422	374	324	284	177	0	1	0	periodo_5
<b>193</b>	373	371	322	312	269	271	182	0	0	1	periodo_5
<b>194</b>	314	375	368	335	302	248	172	0	0	0	periodo_5
<b>195</b>	570	560	502	505	433	352	285	0	1	1	periodo_6
<b>196</b>	526	571	568	528	481	396	290	0	1	0	periodo_6
<b>197</b>	447	442	441	461	399	390	285	0	0	1	periodo_6
<b>198</b>	403	444	472	444	416	390	256	0	0	0	periodo_6
<b>199</b>	756	733	675	657	567	534	401	0	1	1	periodo_7
<b>200</b>	707	730	713	665	595	550	483	0	1	0	periodo_7
<b>201</b>	550	541	567	546	512	540	423	0	0	1	periodo_7
<b>202</b>	489	567	627	586	503	536	397	0	0	0	periodo_7
<b>203</b>	794	825	748	712	655	661	501	0	1	1	periodo_8
<b>204</b>	794	922	858	810	744	746	550	0	1	0	periodo_8
<b>205</b>	626	651	604	685	646	676	507	0	0	1	periodo_8
<b>206</b>	547	710	658	681	653	644	516	0	0	0	periodo_8
<b>207</b>	831	946	866	915	782	725	652	0	1	1	periodo_9
<b>208</b>	871	931	990	987	950	805	707	0	1	0	periodo_9
<b>209</b>	743	744	744	795	790	732	640	0	0	1	periodo_9
<b>210</b>	677	754	813	813	836	762	687	0	0	0	periodo_9
<b>211</b>	937	950	946	934	923	813	752	0	1	1	periodo_10
<b>212</b>	952	1135	1073	1092	1042	976	872	0	1	0	periodo_10
<b>213</b>	788	814	826	952	904	912	679	0	0	1	periodo_10
<b>214</b>	719	907	974	1000	2202	941	814	0	0	0	periodo_10
<b>215</b>	963	935	943	943	869	860	730	0	1	1	periodo_11

216	903	1132	1032	1219	1054	1080	922	0	1	0	periodo_11
217	735	775	824	981	950	955	794	0	0	1	periodo_11
218	748	955	1011	1073	1062	1002	909	0	0	0	periodo_11
219	946	1009	901	925	892	856	816	0	1	1	periodo_12
220	897	1126	1151	1156	1137	1123	967	0	1	0	periodo_12
221	779	828	912	1031	994	1045	846	0	0	1	periodo_12
222	710	988	1048	1141	1054	1147	961	0	0	0	periodo_12
223	977	1000	1040	959	969	909	833	0	1	1	periodo_13
224	912	1101	1106	1183	1262	1162	1023	0	1	0	periodo_13
225	766	872	900	1085	1051	1098	900	0	0	1	periodo_13
226	754	978	1076	1191	1226	1264	1084	0	0	0	periodo_13
227	977	1095	1002	1063	1070	1007	903	0	1	1	periodo_14
228	1008	1174	1166	1260	1228	1244	1093	0	1	0	periodo_14
229	772	903	942	1102	1097	1088	969	0	0	1	periodo_14
230	769	1030	1180	1278	1376	1302	1175	0	0	0	periodo_14
231	844	1129	1094	1154	1069	1069	962	0	1	1	periodo_15
232	957	1232	1288	1436	1328	1318	1206	0	1	0	periodo_15
233	641	840	915	1105	1088	1218	1094	0	0	1	periodo_15
234	759	1075	1170	1274	1303	1493	1246	0	0	0	periodo_15
235	442	818	1018	1120	1155	1170	1044	0	1	1	periodo_16
236	640	1060	1206	1346	1430	1327	1182	0	1	0	periodo_16
237	378	668	808	1104	1054	1144	1038	0	0	1	periodo_16
238	523	918	1073	1327	1358	1377	1343	0	0	0	periodo_16

Fonte: Elaborado pela autora a partir de dados das PNAD do intervalo bianual de 1999 a 2011. \*Sexo: 0- mulher, 1- homem; \*\*Raça: 0-não branco, 1- branco; \*\*\*Região de Residência: 0- não metropolitana, 1- metropolitana.

### A 3: Teste de Hausman para determinação do Modelo mais adequado entre Efeitos Fixos e Aleatórios. Modelos 1, 2 e 3.

<b>Hipótese Nula: Modelo de Efeitos Aleatório é mais adequado</b>	
<b>Modelo 1</b>	
chi2=	284,74
Prob> chi2=	0,000
<b>Modelo 2</b>	
chi2=	2709,66
Prob> chi2=	0,000
<b>Modelo 3</b>	
chi2=	1468,8
Prob> chi2=	0,000

Fonte: Elaborado pela autora a partir de dados das PNAD do intervalo bianual de 1999 a 2011.

**A 4: Teste de Chow para determinação do Modelo mais adequado entre Efeitos Fixos e *Pooled*. Modelos 1, 2 e 3.**

<b>Hipótese Nula: Modelo <i>Pooled</i> é mais adequado</b>	
<b>Modelo 1</b>	
<b>F(125, 567) =</b>	10,42
<b>Prob&gt; F=</b>	0,000
<b>Modelo 2</b>	
<b>F(125, 562) =</b>	9,97
<b>Prob&gt; F=</b>	0,000
<b>Modelo 3</b>	
<b>F(125, 562) =</b>	12,42
<b>Prob&gt; F=</b>	0,000

Fonte: Elaborado pela autora a partir de dados das PNAD do intervalo bianual de 1999 a 2011.

**A 5: Teste de Autocorrelação Serial para dados em painel de Wooldridge. Modelos 1, 2 e 3.**

<b>Hipótese Nula: Ausência de Autocorrelação de 1ª Ordem</b>	
<b>Modelo 1</b>	
<b>F( 1, 110)=</b>	22,527
<b>Prob&gt; F=</b>	0,000
<b>Modelo 2</b>	
<b>F( 1, 110)=</b>	32,69
<b>Prob&gt; F=</b>	0,000
<b>Modelo 3</b>	
<b>F( 1, 110)=</b>	27,061
<b>Prob&gt; F=</b>	0,000

Fonte: Elaborado pela autora a partir de dados das PNAD do intervalo bianual de 1999 a 2011.

**A 6: Teste de Wald Modificado para Heteroscedasticidade na unidade de *cross*, para dados em painel, sobre modelos de Efeitos Fixo. Modelos 1, 2 e 3.**

<b>Hipótese Nula: Erros Homoscedásticos</b>	
<b>Modelo 1</b>	
<b>chi2 (126) =</b>	2.5e+30
<b>Prob&gt; chi2=</b>	0,000
<b>Modelo 2</b>	
<b>chi2 (126) =</b>	2.9e+29
<b>Prob&gt; chi2=</b>	0,000
<b>Modelo 3</b>	
<b>chi2 (126) =</b>	1.3e+30
<b>Prob&gt; chi2=</b>	0,000

Fonte: Elaborado pela autora a partir de dados das PNAD do intervalo bianual de 1999 a 2011.

**A 7: Modelo 1 estimado pelo Método de Mínimos Quadrados Ordinários.**

<b>Salário/hora</b>	<b><math>\beta</math></b>	<b>Desvio Padrão</b>	<b>Significância</b>
<b>Experiência<sup>2</sup></b>	-0.00079	0.000199	***
<b>Escolaridade</b>	0.176505	0.021523	***
<b>Homem</b>	-0.07981	0.094514	
<b>Branco</b>	-0.16306	0.15552	
<b>Metropole</b>	0.062228	0.046248	
<b>Escolaridade* Homem</b>	0.051234	0.014593	
<b>Escolaridade*Branco</b>	0.029162	0.010492	
<b>Escolaridade*Homem*Branco</b>	-0.00649	0.003232	

Fonte: Elaborado pela autora a partir de dados das PNAD do intervalo bianual de 1999 a 2011. O desvio padrão é representado na tabela por  $\sigma$ . Níveis de significância dos parâmetros: 1% \*\*\*, 5% \*\*, 10% \*.

**A 8: Modelo 1 estimado pelo Método de Efeitos Fixos para grupos diferenciados por: Raça, Gênero e Período de Nascimento.**

<b>Salário/hora</b>	<b><math>\beta</math></b>	<b>Desvio Padrão</b>	<b>Significância</b>
<b>Experiência</b>	0.167661	0.007352	***
<b>Experiência<sup>2</sup></b>	-0.00069	8.19E-05	***
<b>Escolaridade</b>	0.120404	0.028465	***
<b>Homem</b>	-0.67349	0.201125	
<b>Branco</b>	0.638018	0.193253	
<b>Metropole</b>	0.652998	0.20273	
<b>Escolaridade* Homem</b>	0.077577	0.027949	
<b>Escolaridade*Branco</b>	-0.04383	0.028928	
<b>Escolaridade*Homem*Branco</b>	0.03484	0.02125	

Fonte: Elaborado pela autora a partir de dados das PNAD do intervalo bianual de 1999 a 2011. O desvio padrão é representado na tabela por  $\sigma$ . Níveis de significância dos parâmetros: 1% \*\*\*, 5% \*\*, 10% \*.

### A 9: Modelo 1 Completo.

Salário/hora	$\beta$	Desvio Padrão	Significância
Experiência	0.145252	0.004735	***
Experiência <sup>2</sup>	-0.00046	0.000073	***
Escolaridade	0.06057	0.015258	***
Homem	-0.91485	0.118423	***
Branco	0.036222	0.14675	-
Metropole	0.266835	0.027704	***
Escolaridade* Homem	0.115033	0.020098	***
Escolaridade*Branco	0.038268	0.013345	***
Escolaridade*Homem*Branco	0.01174	0.015464	-
periodo_1	-2.88871	2.246953	-
periodo_2	-2.07042	0.475708	***
periodo_3	-2.20887	0.361783	***
periodo_4	-2.17321	0.301052	***
periodo_5	-2.06728	0.265765	***
periodo_6	-1.76979	0.232532	***
periodo_7	-1.57302	0.203152	***
periodo_8	-1.37967	0.167634	***
periodo_9	-1.3806	0.182559	***
periodo_10	-0.67993	0.15044	***
periodo_11	-0.3326	0.217315	-
periodo_12	-0.36853	0.158947	**
periodo_13	0.300708	0.077504	***
periodo_14	0.462395	0.09291	***
periodo_15	0.642545	0.129263	***
1999	-14.5053	1.580547	***
2001	-11.8529	1.262652	***
2003	-9.38811	1.07595	***
2005	-7.69727	0.768861	***
2007	-6.0913	0.487998	***
2009	-5.13847	0.273651	***
Const.	-8.09344	2.131398	***

Fonte: Elaborado pela autora a partir de dados das PNAD do intervalo bianual de 1999 a 2011. O desvio padrão é representado na tabela por  $\sigma$ . Níveis de significância dos parâmetros: 1% \*\*\*, 5% \*\*, 10% \*.

**A 10: Modelo 2 estimado pelo Método de Mínimos Quadrados Ordinários.**

Salário/hora	$\beta$	Desvio Padrão	Significância
Experiência	0.138276	0.010862	***
Experiência <sup>2</sup>	-0.00079	0.000169	***
Esco. 5-8 anos	0.584202	0.170876	***
Esco. 9-11 anos	2.252824	0.315582	***
Esco. 11 ou mais	2.054799	0.421814	***
Homem	-0.06714	0.071777	
Branco	0.065881	0.102048	
Metropole	0.040601	0.043741	
Branco*Esco. 5-8	-0.18059	0.170341	
Branco*Esco.9-11	-0.21083	0.184107	
Branco*Esco. 11 ou mais	0.735134	0.299684	
Homem*Esco. 5-8	0.7108	0.142772	
Homem*Esco. 9-11	0.577914	0.259779	
Homem*Esco. 11 ou mais	0.206978	0.119317	

Fonte: Elaborado pela autora a partir de dados das PNAD do intervalo bianual de 1999 a 2011. O desvio padrão é representado na tabela por  $\sigma$ . Níveis de significância dos parâmetros: 1% \*\*\*, 5% \*\*, 10% \*.

**A 11: Modelo 2 estimado pelo Método de Efeitos Fixos para grupos diferenciados por: Raça, Gênero e Período de Nascimento.**

Salário/hora	$\beta$	Desvio Padrão	Significância
Experiência	0.162385	0.007341	***
Experiência <sup>2</sup>	-0.00077	9.43E-05	***
Esco. 5-8 anos	-0.11153	0.326574	-
Esco. 9-11 anos	1.279026	0.342558	***
Esco. 11 ou mais	0.584183	0.437444	-
Homem	-0.4546	0.206945	
Branco	0.451197	0.169805	
Metropole	0.724097	0.204546	
Branco*Esco. 5-8	0.660822	0.338388	
Branco*Esco.9-11	-0.6951	0.312643	
Branco*Esco. 11 ou mais	0.720935	0.62481	
Homem*Esco. 5-8	0.348818	0.34527	
Homem*Esco. 9-11	0.832544	0.33558	
Homem*Esco. 11 ou mais	0.866313	0.515239	

Fonte: Elaborado pela autora a partir de dados das PNAD do intervalo bianual de 1999 a 2011. O desvio padrão é representado na tabela por  $\sigma$ . Níveis de significância dos parâmetros: 1% \*\*\*, 5% \*\*, 10% \*.



**A 12: Modelo 2 Completo.**

<b>Salário/hora</b>	<b><math>\beta</math></b>	<b>Desvio Padrão</b>	<b>Significância</b>
<b>Experiência</b>	0.145666	0.00437	***
<b>Experiência<sup>2</sup></b>	-0.0004	6.87E-05	***
<b>Esco. 5-8 anos</b>	0.181946	0.16699	-
<b>Esco. 9-11 anos</b>	0.722361	0.172562	***
<b>Esco. 11 ou mais</b>	0.8305	0.282902	***
<b>Homem</b>	-0.47029	0.127689	***
<b>Branco</b>	0.380481	0.13841	***
<b>Metropole</b>	0.247433	0.030007	***
<b>Branco*Esco. 5-8</b>	-0.07879	0.235061	-
<b>Branco*Esco.9-11</b>	-0.32533	0.209275	-
<b>Branco*Esco. 11 ou mais</b>	0.600806	0.341764	*
<b>Homem*Esco. 5-8</b>	0.439563	0.228931	*
<b>Homem*Esco. 9-11</b>	0.930058	0.242182	***
<b>Homem*Esco. 11 ou mais</b>	1.139604	0.318028	***
<b>periodo_1</b>	-3.29191	2.116923	-
<b>periodo_2</b>	-2.37897	0.466523	***
<b>periodo_3</b>	-2.56239	0.356833	***
<b>periodo_4</b>	-2.47258	0.282786	***
<b>periodo_5</b>	-2.29128	0.240675	***
<b>periodo_6</b>	-1.95256	0.224506	-
<b>periodo_7</b>	-1.68789	0.192812	***
<b>periodo_8</b>	-1.51692	0.169187	***
<b>periodo_9</b>	-1.37744	0.166519	***
<b>periodo_10</b>	-0.66942	0.1528	***
<b>periodo_11</b>	-0.31294	0.216046	-
<b>periodo_12</b>	-0.3774	0.130646	-
<b>periodo_13</b>	0.322186	0.073253	***
<b>periodo_14</b>	0.476808	0.083041	***
<b>periodo_15</b>	0.665827	0.119036	***
<b>1999</b>	-12.4056	1.463214	***
<b>2001</b>	-10.1637	1.168708	***
<b>2003</b>	-7.85522	0.98161	***
<b>2005</b>	-6.48553	0.702938	***
<b>2007</b>	-5.18934	0.458711	***
<b>2009</b>	-4.56681	0.260241	***
<b>Const.</b>	-8.39318	2.088492	***

Fonte: Elaborado pela autora a partir de dados das PNAD do intervalo bianual de 1999 a 2011. O desvio padrão é representado na tabela por  $\sigma$ . Níveis de significância dos parâmetros: 1% \*\*\*, 5% \*\*, 10% \*.

**A 13: Modelo 3 estimado pelo Método de Mínimos Quadrados Ordinários.**

Salário/hora	$\beta$	Desvio Padrão	Significância
Experiência	0.139047	0.011787	***
Experiência <sup>2</sup>	-0.00088	0.000216	***
Esco. 8 anos	0.925131	0.314971	***
Esco. 11 anos	2.510721	0.427628	***
Esco. 15 ou mais	2.713586	0.500195	***
Homem	0.165283	0.051241	
Branco	0.214033	0.099597	
Metropole	0.04147	0.047023	
Branco*Esco. 8	-0.77551	0.342035	
Branco*Esco.11	-0.2882	0.171778	
Branco*Esco. 15 ou mais	0.222648	0.362932	
Homem*Esco. 8	0.706303	0.286571	
Homem*Esco. 11	0.52657	0.276592	
Homem*Esco. 15 ou mais	-0.08073	0.169285	

Fonte: Elaborado pela autora a partir de dados das PNAD do intervalo bianual de 1999 a 2011. O desvio padrão é representado na tabela por  $\sigma$ . Níveis de significância dos parâmetros: 1% \*\*\*, 5% \*\*, 10% \*.

**A 14: Modelo 3 estimado pelo Método de Efeitos Fixos para grupos diferenciados por: Raça, Gênero e Período de Nascimento.**

Salário/hora	$\beta$	Desvio Padrão	Significância
Experiência	0.158901	0.00591	***
Experiência <sup>2</sup>	-0.00066	0.000082	***
Esco. 8 anos	0.225706	0.516037	-
Esco. 11 anos	1.373686	0.37676	***
Esco. 15 ou mais	1.694582	0.450003	***
Homem	-0.27588	0.105345	
Branco	0.870557	0.106503	
Metropole	0.683799	0.201755	
Branco*Esco. 8	-0.80679	0.521738	
Branco*Esco.11	-1.17448	0.320081	
Branco*Esco. 15 ou mais	-0.66396	0.503851	
Homem*Esco. 8	0.910939	0.452646	
Homem*Esco. 11	0.610039	0.298007	
Homem*Esco. 15 ou mais	1.076329	0.39322	

Fonte: Elaborado pela autora a partir de dados das PNAD do intervalo bianual de 1999 a 2011. O desvio padrão é representado na tabela por  $\sigma$ . Níveis de significância dos parâmetros: 1% \*\*\*, 5% \*\*, 10% \*.

**A 15: Modelo 3 Completo.**

<b>Salário/hora</b>	<b><math>\beta</math></b>	<b>Desvio Padrão</b>	<b>Significância</b>
<b>Experiência</b>	0.1406198	0.0038406	***
<b>Experiência<sup>2</sup></b>	-0.0002842	0.0000773	***
<b>Esco. 8 anos</b>	0.6501912	0.2194197	***
<b>Esco. 11 anos</b>	0.7401245	0.1729566	***
<b>Esco. 15 ou mais</b>	1.420652	0.3196771	***
<b>Homem</b>	-0.2635402	0.0830585	***
<b>Branco</b>	0.573569	0.0820836	**
<b>Metropole</b>	0.2599521	0.0287446	***
<b>Branco*Esco. 8</b>	-0.7214834	0.2785742	***
<b>Branco*Esco.11</b>	-0.3511606	0.2032792	*
<b>Branco*Esco. 15 ou mais</b>	-0.0156778	0.2877268	-
<b>Homem*Esco. 8</b>	0.7700253	0.3204811	***
<b>Homem*Esco. 11</b>	0.8930567	0.2334295	***
<b>Homem*Esco. 15 ou mais</b>	1.124603	0.2501536	***
<b>periodo_1</b>	-3.519987	2.242915	-
<b>periodo_2</b>	-2.506621	0.4659678	***
<b>periodo_3</b>	-2.767163	0.3507302	***
<b>periodo_4</b>	-2.753851	0.253113	***
<b>periodo_5</b>	-2.516599	0.2141892	***
<b>periodo_6</b>	-2.114831	0.1983078	***
<b>periodo_7</b>	-1.760044	0.1883395	***
<b>periodo_8</b>	-1.584838	0.1735932	***
<b>periodo_9</b>	-1.424297	0.1814547	***
<b>periodo_10</b>	-0.7998616	0.1527607	***
<b>periodo_11</b>	-0.434321	0.1659639	***
<b>periodo_12</b>	-0.2861106	0.138023	***
<b>periodo_13</b>	0.4104863	0.0673728	***
<b>periodo_14</b>	0.5479298	0.0740481	***
<b>periodo_15</b>	0.7200406	0.1082552	***
<b>1999</b>	-10.46226	1.454457	***
<b>2001</b>	-8.344861	1.190671	***
<b>2003</b>	-6.190608	0.9424738	***
<b>2005</b>	-5.069981	0.6755211	***
<b>2007</b>	-4.189857	0.4323951	***
<b>2009</b>	-3.701032	0.2258939	***
<b>Const.</b>	-9.649027	1.926982	***

Fonte: Elaborado pela autora a partir de dados das PNAD do intervalo bianual de 1999 a 2011. O desvio padrão é representado na tabela por  $\sigma$ . Níveis de significância dos parâmetros: 1% \*\*\*, 5% \*\*, 10% \*.